THERMODYNAMICS – PHYSICAL CHEMISTRY OF AQUEOUS SYSTEMS

Edited by **Juan Carlos Moreno-Piraján**

INTECHWEB.ORG

Thermodynamics – Physical Chemistry of Aqueous Systems
Edited by Juan Carlos Moreno-Piraján

Published by InTech
Janeza Trdine 9, 51000 Rijeka, Croatia

Publishing Process Manager Viktorija Zgela
Technical Editor Teodora Smiljanic
Cover Designer Jan Hyrat
Image Copyright Andrejs Pidjass, 2010. Used under license from Shutterstock.com

First published September, 2011
Printed in Croatia

A free online edition of this book is available at www.intechopen.com
Additional hard copies can be obtained from orders@intechweb.org

Thermodynamics – Physical Chemistry of Aqueous Systems,
Edited by Juan Carlos Moreno-Piraján
p. cm.
ISBN 978-953-307-979-0

Contents

Preface IX

Chapter 1 **Thermodynamics of Molecular Recognition by Calorimetry 1**
Luis García-Fuentes, Ramiro, Téllez-Sanz, Indalecio Quesada-Soriano and Carmen Barón

Chapter 2 **Theory and Application of Thermoelectrochemistry 27**
Zheng Fang

Chapter 3 **Thermodynamics and the Glass Forming Ability of Alloys 49**
Chengying Tang and Huaiying Zhou

Chapter 4 **Information Thermodynamics 73**
Bohdan Hejna

Chapter 5 **Mesoscopic Thermodynamics in the Presence of Flow 105**
I. Santamaría-Holek, R. Lugo-Frías, R. F. Rodríguez and A. Gadomski

Chapter 6 **Non-Instantaneous Adiabats in Finite Time 131**
Delfino Ladino-Luna and Ricardo T. Páez-Hernández

Chapter 7 **Heterogeneous Melting in Low-Dimensional Systems and Accompanying Surface Effects 157**
Dmitry G. Gromov and Sergey A. Gavrilov

Chapter 8 **Pressure Effects on Thermodynamics of Polymer Containing Systems 191**
Shichun Jiang and Hongfei Li

Chapter 9 **Potential-pH Diagrams for Oxidation-State Control of Nanoparticles Synthesized via Chemical Reduction 223**
Shunsuke Yagi

Chapter 10 **On the Extremum Properties of Thermodynamic Steady State in Non-Linear Systems 241**
Gy. Vincze and A. Szasz

Chapter 11 **Thermodynamic Study of Grinding-Induced Loratadine Inclusion Complex Formation Using Thermal Analysis and Curve-Fitted FTIR Determination 317**
Shan-Yang Lin, Hong-Liang Lin, Chih-Cheng Lin, Cheng-Hung Hsu, Tieh-kang Wu and Yu-Ting Huang

Chapter 12 **Three-Dimensional Constitutive Viscoelastic Model for Isotropic Materials 327**
Donald Picard and Mario Fafard

Chapter 13 **Hydrogen Bond Interactions Between Water Molecules in Bulk Liquid, Near Electrode Surfaces and Around Ions 351**
Abhishek Rastogi, Amit K. Ghosh and SJ Suresh

Chapter 14 **The Stability of a Three-State Unfolding Protein 365**
Yang BinSheng

Chapter 15 **Phase Diagram and Waterlike Anomalies in Core-Softened Shoulder-Dumbbell Complex Fluids 391**
Paulo A. Netz, Guilherme K. Gonzatti, Marcia C. Barbosa, Juliana Z. Paukowski, Cristina Gavazzoni and Alan Barros de Oliveira

Chapter 16 **Effect of Magnetic and Mechanical Fields on Phase Liquid Crystalline Transitions in Solutions of Cellulose Derivatives 407**
S. A. Vshivkov

Preface

Thermodynamics is one of the most exciting branches of physical chemistry which has greatly contributed to the modern science. Since its inception, great minds have built their theories of thermodynamics. One should name those of Sadi Carnot, Clapeyron Claussius, Maxwell, Boltzman, Bernoulli, Leibniz etc. Josiah Willard Gibbs had perhaps the greatest scientific influence on the development of thermodynamics. His attention was for some time focused on the study of the Watt steam engine. Analysing the balance of the machine, Gibbs began to develop a method for calculating the variables involved in the processes of chemical equilibrium. He deduced the phase rule which determines the degrees of freedom of a physicochemical system based on the number of system components and the number of phases. He also identified a new state function of thermodynamic system, the so-called free energy or Gibbs energy (G), which allows spontaneity and ensures a specific physicochemical process (such as a chemical reaction or a change of state) experienced by a system without interfering with the environment around it. The essential feature of thermodynamics and the difference between it and other branches of science is that it incorporates the concept of heat or thermal energy as an important part in the energy systems. The nature of heat was not always clear. Today we know that the random motion of molecules is the essence of heat. Some aspects of thermodynamics are so general and deep that they even deal with philosophical issues. These issues also deserve a deeper consideration, before tackling the technical details. The reason is a simple one - before one does anything, one must understand what they want.

In the past, historians considered thermodynamics as a science that is isolated, but in recent years scientists have incorporated more friendly approach to it and have demonstrated a wide range of applications of thermodynamics.

These four volumes of applied thermodynamics, gathered in an orderly manner, present a series of contributions by the finest scientists in the world and a wide range of applications of thermodynamics in various fields. These fields include the environmental science, mathematics, biology, fluid and the materials science. These four volumes of thermodynamics can be used in post-graduate courses for students and as reference books, since they are written in a language pleasing to the reader.

They can also serve as a reference material for researchers to whom the thermodynamics is one of the area of interest.

Juan Carlos Moreno-Piraján
Department of Chemistry
University of the Andes
Colombia

1

Thermodynamics of Molecular Recognition by Calorimetry

Luis García-Fuentes, Ramiro, Téllez-Sanz,
Indalecio Quesada-Soriano and Carmen Barón
University of Almería, Almería
Spain

1. Introduction

When Otto von Guericke, stimulated by the previous work of Galileo and Torricelli, constructed the world's first-ever vacuum pump in 1650 to disprove Aristotle's supposition that "nature abhors a vacuum", he could not imagine the newly-born scientific field would get us closer to understand one of the oldest questions in the history of mankind: what is life?. Even though nobody is able to answer this question correctly yet, thermodynamics helps us to address another one, equally important: how does it work? The cellular machinery is a highly complex system, probably the most complex ever created by nature. A perfect gear with thousands of chemical reactions taking place synchronously requiring high efficiency enzymes, which are responsible for providing the cell in time with the products it needs. One of the basic aims of the biophysical research is to be able to control how enzymes work. But there's no possible control if you don't previously understand how the molecular recognition between ligand and protein occurs and how favorable it is.

Thermodynamics is the only scientific field allowing to address the matter. Any molecular recognition process, as a chemical reaction, is associated with a change in the molecular properties of the reactants. Understanding the molecular recognition processes between small ligands and biological macromolecules takes a complete characterization of the binding energetic, as well as the correlation between thermodynamic data and chemical structure. Techniques such as fluorimetry, spectrophotometry or circular dichroism are convenient, fast and low sample-consuming, but their application is not universal.

However, there is such a universal technique, the Isothermal Titration Calorimetry (ITC), standing above the others. Modern isothermal titration calorimeters (e.g. VP-ITC or iTC-200 from Microcal (http://www.microcal.com/) and nano ITC (http://www.tainstruments.com/) are able to measure the energetic of ligand binding (for example, a drug) in a highly reliable, fast and accurate way, using relatively small amounts of material. Typically, these calorimeters require less than 500 μg of protein per complete calorimetric titration and can measure heat effects as small as 0.1 μcal, thus allowing the determination of binding constants as large as 10^8 to 10^9 M^{-1}. Chemical interaction changes are always associated with a heat energy exchange with the environment. This fact turns ITC, among the possible choices, in the safest bet to address these studies. In an ITC experiment the heat

evolved when two reactants are mixed is monitored as a titration curve where one of them, frequently the macromolecule, is titrated at constant temperature against the ligand. Planning and careful performing the experiments is crucial to get quality data from which extracting reliable thermodynamic parameters and interpretations. ITC is currently used in a large number of molecular recognition studies, such as antigen-antibody, protein-peptide, protein-protein, sugar-protein, DNA-protein and protein-ligand studies, as well as enzyme kinetics.

The quantitative analysis of the molecular association driving forces between a biological macromolecule and a ligand requires the determination of thermodynamic parameters. The suitability of ITC lays on its ability to not only providing the affinity, usually expressed in terms of the association constant, but also the enthalpic (ΔH) and entropic (ΔS) contributions to the Gibbs free energy of association ($\Delta G=\Delta H-T\Delta S$). Under right conditions, a single ITC experiment is able to give the values for these changes along with the stoichiometry or number of binding sites (n). Moreover, in the cases in which more than one binding site is present, it is also possible to examine the sites for cooperativity.

Each interaction, either hydrophobic, pi-stacking, electrostatic, proton release or uptake... has its own energetic fingerprint. Splitting the global energetics into individual contributions is the key to know, with a great deal of reliability, details such as which enzyme residue is involved in a proton uptake or which other locks the ligand into position through a pi-stacking. This assignment of individual residue roles would not be possible without the calorimetric study of some protein mutants, obtained by directed mutagenesis. It is also important to have structural information about the complexes. When X-ray crystallographic data is not available, molecular docking studies can replace it as long as it is used with enough precaution and the user has previous structural knowledge from similar ligands.

The combination of different protein-ligand binding experiments performed under different solution conditions allows for other parameters to be calculated. For instance, the heat capacity change (ΔC_p) can be determined through a temperature change. ΔC_p is closely related to, among other factors, changes in the solvent Accessible Surface Areas (ASA's) upon complex formation. Or, when a change in the protonation state of one of more groups accompanies the complex formation, the number of protons uptaken or released can be determined from a series of experiments carried out with different buffers or at different pH values. ITC is, thus, a key tool to elucidate which chemical structures a ligand must possess to bind a protein with high affinity and specificity. Since this is the main requirement for a rational drug design against a biological target, ITC is a valuable technique for identifying and optimizing molecules with therapeutic properties.

The chapter will deal, through experimental results, with the necessary requirements for a complete thermodynamic protein-ligand binding study allowing for the maximal amount of information to be obtain about the molecular recognition basis: how to plan the experiments, data analysis, strategies to follow to overcome difficulties and the splitting of energetic parameters into individual contributions.

2. Background of binding thermodynamic

For a simple reversible bimolecular binding reaction between a target macromolecule (M) and a ligand (X), represented as:

$$M + X \rightleftarrows MX \tag{1}$$

the change in the Gibbs free energy (ΔG), for the ligand-macromolecule complex formation of the complex (MX) is related to the standard Gibbs free energy change (ΔG^o), by the equation:

$$\Delta G = \Delta G^0 + RT \ln\left(\frac{[MX]}{[M][X]}\right) \tag{2}$$

At equilibrium, under standard conditions, when ΔG=0, this becomes:

$$\Delta G^0 = -RT \ln\left(\frac{[MX]}{[M][X]}\right) = -RT \ln K_a = RT \ln K_d \tag{3}$$

where K_a is the equilibrium association constant, commonly named as affinity, and K_d is the dissociation constant. Moreover, the binding parameter, ν, is defined as the ratio between the concentrations of bound ligand, $[X]_b$ and the total macromolecule, $[M]_t$:

$$\nu = \frac{[X]_b}{[M]_t} = \frac{[MX]}{[M]_t} = \frac{[MX]}{[MX]+[M]} = \frac{K_a[X]}{1+K_a[X]} = \frac{[X]}{K_d+[X]} \tag{4}$$

K_a or K_d can be measured using a great variety of experimental techniques (fluorescence, circular dichroism, equilibrium dialysis, surface plasmon resonance, etc.). However, a complete thermodynamic characterization requires the enthalpy change, which accounts for the heat exchange during the association reaction, to be measured. When this is done, the entropic contribution to the overall observed Gibbs free energy can be calculated through the relationship $\Delta G^0=\Delta H-T\Delta S^0$ (assuming that $\Delta H=\Delta H^0$). The sign and value of the observed enthalpy are the global result of the interaction changes taking place at binding time: their type and number, bond length and angle changes… but perhaps the most important contribution, enthalpically speaking, is the hydrogen bonding. Thus, the sign indicates if there is a net favorable (negative) or unfavorable (positive) redistribution of the hydrogen bond network between the reacting species, including the solvent. The entropy change can be related to the relative degree of disorder after binding. For instance, the release of water molecules to the bulk solvent is a source of favorable entropy. Thus, hydrophobic interactions are characterized by a small enthalpy (negative or positive), and a favorable entropy. Thus, two interactions with similar affinities and structures can have different enthalpic and entropic contributions to the Gibbs free energy of binding.

Enthalpy changes can be measured in an indirect way through the integrated form of the van't Hoff equation. However, this is done under the assumption that ΔH^o is constant within the studied temperature range, which is seldom the case. ITC is by far the preferred method, since it provides a direct and accurate measurement at every temperature. There are reported discrepancies between calorimetric and van't Hoff enthalpies (Horn et al., 2001), proving the advantage of using ITC to determine enthalpy changes.

3. General aspects of ITC

3.1 Instrumentation

The basic design of ITC instruments has scarcely changed over the last 10 years. The most modern instruments operate a differential cell feedback system, where the reference cell is filled with water or buffer and the sample cell usually contains the macromolecule. A syringe that also serves as the stirrer adds the ligand in a stepwise fashion at preset intervals during the course of the experiment. Heat produced or absorbed during the binding reaction is monitored as a temperature change. Any temperature difference between the sample and reference cells triggers a feedback system which modulates the applied thermal power in order to keep the temperature difference between both cells as low as possible. The instrument slowly increases the temperature of both cells during each titration (less than 0.1 ^{o}C per hour), in a way that approximates isothermal conditions. Usually cell volumes are around 1.5 mL, the thermostat temperature can be set between 5 and 80 ^{o}C, and heats as small as 0.1 μcal can be measured.

3.2 Experimental planning

The setup of an ITC experiment is largely dependent on the thermodynamic characteristics of the system of interest, i.e., the expected binding affinity and the heat effect of the interaction. To obtain high quality data, an appropriate protocol has to be established by optimizing ligand and protein concentrations, the injection volume and the values of K, ΔH, and n (or a larger set of parameters for binding models other than the n equal and independent binding sites). The shape of the binding curve is dependent on the C-value, defined this, as product of the association constant K_a and the sites molar concentration of the macromolecule $[M]_T$ being titrated (Wiseman et al, 1989). This value is crucial for an accurate determination of the binding parameters. Experience shows that for a good ITC experimental design (sigmoidal termogram) a C-value in the 10-100 range should be chosen. However, in many cases, the intrinsic properties of the system avoid reaching a good C-value, and it is up to the user to choose the more adequate experimental conditions. Clearly, simulations are important in optimizing an ITC experiment and in achieving a balance between detectable heats and thermogram curvature.

As an example, we will use the binding of dUDP to trimeric dUTPase from *Plasmodium falciparum* (PfdUTPase) in glycerophosphate buffer at pH 7 and 25 ^{o}C (Quesada-Soriano et al., 2007). Fig. 1 shows a typical calorimetric titration. What is needed to reach such an experimental outcome?. As indicated above, the appropriate concentration range for the macromolecule placed in the cell depends on the binding constant. Since in this case the approximate value for the binding constant at 25 ^{o}C is $K_a = 6 \cdot 10^5$ M^{-1}, with a stoichiometry of 3 mol of ligand per mol of trimeric enzyme, a concentration of macromolecule of approximately 20 μM yields a C-value of 36, within its ideal 10-100 range. The actual value was 22.7 μM.

The macromolecule in the cell is titrated with a series of small injections of the ligand solution from the syringe, the concentration of which must be much higher than that for the macromolecule in the cell since the titration experiment is planned to approach or reach complete saturation of the binding sites at the end (Fig. 1). The number and volume of the ligand injections should be chosen so that the sigmoidal shape is as well defined as possible, usually with a large number of small aliquots, between 5 and 10 μL. Only if the heat signal is small it will be necessary to choose larger injection volumes.

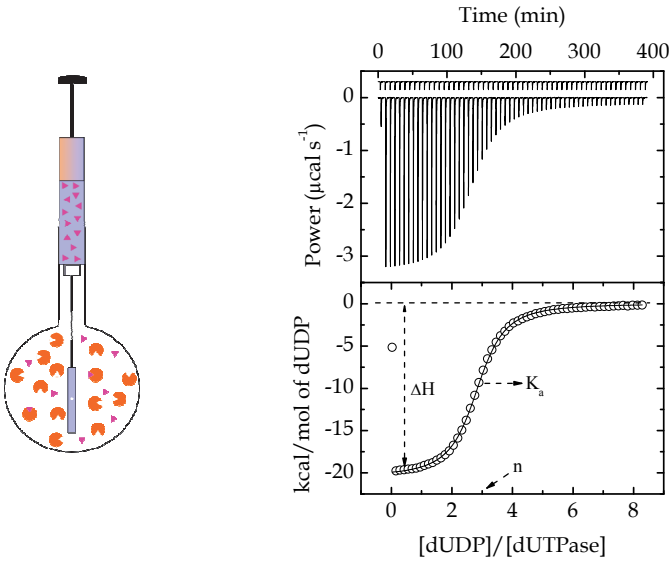


Fig. 1. Scheme of the calorimeter reaction cell (left) and results of a typical ITC experiment (right). The sigmoidal thermogram in the upper panel on the right side corresponds to the binding of a ligand (dUDP) to a trimeric protein (dUTPase). The small linear thermogram above the sigmoidal one comes from the so-called "ligand dilution experiment", where the ligand is injected into the sample cell containing just the buffer. Each thermogram consists of the heat peaks generated by a series of 5 µL injections from the syringe (containing the ligand solution) into the sample cell (containing the macromolecule or plain buffer solutions). The bottom panel shows the non-linear least squares analysis of the thermodynamic data to a suitable model, yielding the values for K_a, ΔH and n.

These requisites define the titration protocol, and it is up to the user to find the ideal compromise. In this particular example, 58 injections, 5 µL-each (preliminary 1 µl injection), of a ligand solution (dUDP) with a concentration 40 times higher than that of the protein solution will give an adequate binding isotherm. If association is fast compared to the response time of the calorimeter, and the heat signal is not very large, the instrument baseline will be recovered in a short time. In those cases, four or five minutes are usually enough to reach baseline again after injection. In Fig. 1, this time was set to about five minutes. In contrast, heat signals of slow processes, such as covalent reactions or enzymatic kinetics, require much more time to reach thermal equilibrium.

Finally, other issues, also related to experimental design, should be taking into account. It is very important that ligand and macromolecule solutions are pure and exactly each other regarding pH and solution conditions. For this reason, the macromolecule and the ligand should be are preferably dissolved in the same buffer. It is a good practice to dialyze the protein prior to the experiment and dissolve the ligand in the last dialysis buffer change. Furthermore, air bubbles have to be avoided in the sample cell. Thus, it is very important to degas, all solutions prior to the experiment during a short time. Also, any air bubble left in the syringe after filling it can cause variation in the injected volume or lead to additional heat signals. Finally, in most experiments the heat effect of the first injection of a series of

injections is obviously too small. This results from diffusion while equilibrating the system. Even if care is taken to avoid this leakage, the problem may persist. Therefore it is common practice to make a small first injection of 1 μL and then to remove the first data point before data analysis.

The result from a titration is a plot of the recorded power, dQ/dt, *vs.* time, as shown in the right upper panel of Fig. 1. Each peak represents the thermal effect associated with an injection. The right lower panel shows the integrated areas as a function of the ligand/protein molar ratio.

The heat effects after every ligand injection arise from four main sources: binding interaction, ligand dilution, macromolecule dilution and a mixing heat effect. Generally, the dialysis/dialysate approach will virtually eliminate the mixing. The dilution heat of both the macromolecule and the ligand must be measured in separate experiments. For the first, buffer is injected from the syringe into the macromolecule solution in the sample cell, whereas for the latter the ligand is injected into the sample cell containing just buffer. Since the macromolecule concentration placed in calorimetric cell is usually in the micromolar range, its dilution heat is negligible. Thus, this titration can be skipped. However, the ligand dilution heat is not always negligible and it needs to be measured in an independent titration and substrated from the injection heats measured in the binding titration (Fig. 1).

There are situations where the general procedure above is not the best choice, like when the ligand is poorly soluble. In these cases a so-called "reverse titration" may be preferred, where the macromolecule is inside the syringe and the ligand in the sample cell. The analysis procedure has to be modified accordingly, especially if the macromolecule has several binding sites.

4. Data analysis

4.1 Equal and independent binding sites model

The equal and independent binding sites model describes the simplest way a macromolecule can interact with a ligand. The system described above will be used as an example (i.e. dUDP/PfdUTPase). Structurally, *Pf*dUTPase is a trimer with three identical active sites located at the subunit interfaces. Each active site is made up by residues from all three subunits, five or which are highly conserved. For such a system the binding parameter, υ, is related to the fractional saturation, Y, by

$$\nu = n \cdot Y \tag{5}$$

where n is the number of binding sites, in this case n=3. The concentration of free ligand is related to the total ligand, $[X]_t$, and the bound ligand, $[X]_b$, by the mass conservation law:

$$[X] = [X]_t - [X]_b \tag{6}$$

By using Eqs. 4 and 5, Eq. 6 can be represented by the relationship

$$[X] = [X]_t - nY[M]_t \tag{7}$$

On the other hand, the binding constant, K_a, is given by,

$$K_a = \frac{Y}{(1-Y)[X]} \tag{8}$$

The combination of Eqs. 7 and 8 gives the quadratic equation

$$Y^2 - Y\left(1 + \frac{1}{nK_a[M]_t} + \frac{[X]_t}{n[M]_t}\right) + \frac{[X]_t}{n[M]_t} = 0 \tag{9}$$

where the only root with physical meaning is,

$$Y = \frac{1}{2}\left(1 + \frac{1}{nK[M]_t} + \frac{[X]_t}{n[M]_t} - \sqrt{\left(1 + \frac{1}{nK[M]_t} + \frac{[X]_t}{n[M]_t}\right)^2 - \frac{4[X]_t}{n[M]_t}}\right) \tag{10}$$

The accumulated or integral binding heat of the process after the i^{th} injection is given by

$$Q = n[M]_t V_0 \Delta H_t Y_i \tag{11}$$

where V_0 is the cell volume and ΔH_t is the molar enthalpy change of the binding reaction. The heat of the i^{th} injection (differential heat) is,

$$q_i = V_0 \Delta H_t \cdot \Delta[L]_b = V_0 \Delta H_t \cdot n[M]_t (Y_i - Y_{i-1}) \tag{12}$$

with $\Delta[L]_b$ being the difference in the bound ligand concentration between the i^{th} and $(i-1)^{th}$ injections. It is very important to underline that the functional form of $\Delta[L]_b$ depends on the specific binding model. Thus, for this simplest model, when the protein has n binding sites, Eq. 12 becomes

$$q_i = V_0 \Delta H_t \cdot n[M]_t \left(\frac{K_a[X]_i}{1 + K_a[X]_i} - \frac{K_a[X]_{i-1}}{1 + K_a[X]_{i-1}}\right) \tag{13}$$

The experimental titration data from Fig. 1 can be non-linearly fitted to the sigmoidal curve defined by Eqs. 7 and 13 (q_i *vs.* $[X]_t$, or *vs.* $[X]_t/[M]_t$). The model yields the values for its parameters (K_a, ΔH_t, and n) from a single experiment. In the example in Fig. 1, the parameter values obtained were n=2.85, $K_a = 5.7 \cdot 10^5$ M^{-1} and ΔH_t = -20.4 kcal/mol. It is worth noticing that the resulting stoichiometry differs somewhat from three (three binding sites). This discrepancy between the calorimetric determined stoichiometry and the real number of binding sites in the enzyme is very frequent and there are two main reasons for it to appear: concentration errors (ligand and/or protein) and the presence of a small fraction of damaged macromolecule unable to bind ligand. These small errors are acceptable and within experimental error, although a usual procedure is to remove the stoichiometry parameter from the fitting session by fixing it at a constant value (only if its value is known and trusted). This way only ΔH_t and K_a are calculated by the fitting procedure.

4.2 Equal and interacting binding sites model

When good quality data has been obtained but the simplest model above is unable to yield a successful fit, then it is not valid to describe the macromolecule-ligand interaction. If the macromolecule is composed of identical subunits, the next logical step is trying an equal and interacting binding sites model. This model makes the assumption that a ligand molecule binds the macromolecule with a different affinity than the previous one, i.e, there

is cooperativity. When the complexity of the model increases it is common practice to use a statistical thermodynamic approach to deduce the binding equations. However, the fitting success strongly depends on the quality of the experimental data.

To describe this model we are using experimental data from the binding of the Pi class human glutathione S-transferase enzyme to two glutathione conjugates. Human glutathione transferase P1-1 (hGSTP1-1), a homodimeric protein of ≈46 kDa, has been extensively studied for its potential use as a marker during chemical carcinogenesis and its possible role in the mechanism of cellular multidrug resistance against a number of anti-neoplastic agents (Hayes et al., 2005). S-nitroglutathione (GSNO) binds to wild-type hGSTP1-1 with negative cooperativity, whereas the C47S mutation induces positive cooperativity towards both GSNO and (ethacrynic and glutathione conjugate) EASG binding (Tellez-Sanz et al., 2006; Quesada-Soriano et al., 2009).

The equilibrium between a ligand and a macromolecule with two ligand binding sites can be described in terms of two different sets of association constants: the macroscopic association constants (overall, β_i, or stepwise, K_i), or the microscopic or intrinsic constants:

$$\beta_1 = K_1 = \frac{[MX]}{[M][X]}; \quad \beta_2 = \frac{[MX_2]}{[M][X]^2} = K_1 \cdot K_2 \tag{14}$$

The microscopic binding constants, K_i^0, are related to the intrinsic ligand binding to a site, and therefore reflect the intrinsic binding affinities to each site. The relationship between macroscopic and microscopic binding constants is a statistical factor given by

$$K_i = \frac{n-i+1}{i} K_i^0 \tag{15}$$

Therefore, for the two interacting sites case, there are two microscopic constants, one per site ($K_1^0=1/2K_1$ and $K_2^0=2K_2$), and the binding parameter or Adair´s equation, ν, will be given by

$$\nu = \frac{2K_1^0[X] + 2K_1^0K_2^0[X]^2}{1 + 2K_1^0[X] + K_1^0K_2^0[X]^2} \tag{16}$$

The denominator in Eq. 16 is called the binding polynomial, P, or binding partition function and it represents the sum of the different macromolecular species concentrations relative to that of the free macromolecule that is taken as the reference:

$$P = \sum_{i=0}^{n} \frac{[MX_i]}{[M]} \tag{17}$$

which with two sites (n=2), and using Eqs. 14 and 15, the expression shown in the denominator of Eq. 16 is deduced.

The free ligand concentration is related to the total ligand $[X]_t$ and the bound ligand, $[X]_b$ by

$$[X] = [X]_t - [X]_b = [X]_t - [M]_t \left(\frac{2K_1^0[X] + 2K_1^0K_2^0[X]^2}{1 + 2K_1^0[X] + K_1^0K_2^0[X]^2} \right) \tag{18}$$

The accumulated or integral binding heat of the titration after the ith injection is given by

$$Q = V_0[M]_t \left[\frac{2K_1^0 \Delta H_1[X] + K_1^0 K_2^0 (\Delta H_1 + \Delta H_2)[X]^2}{1 + K_1^0[X] + K_1^0 K_2^0[X]^2} \right] \tag{19}$$

where ΔH_1 and ΔH_2 will be the binding enthalpy changes for the first and second site, respectively. These expressions are completely general for any macromolecule with two interacting ligand-binding sites, irrespective of positive or negative cooperativity.
Consequently, the heat of the ith injection is,

$$q_i = Q_i - Q_{i-1} \tag{20}$$

When a system behaves according to this model, a nonlinear fit using Eqs. 18, 19 and 20 can fit the titration data (q_i vs. $[X]_t$, or vs. $[X]_t/[M]_t$). Fig. 2 (left panel) shows a typical ITC profile for the binding of GSNO (12.7 mM) to dimeric wt-hGSTP1-1 (43.7 μM) in phosphate buffer at pH 7.0 and 25°C. Control experiments were also carried out in order to measure the ligand dilution heat. A noncooperative model is unable to fit these calorimetric data properly.

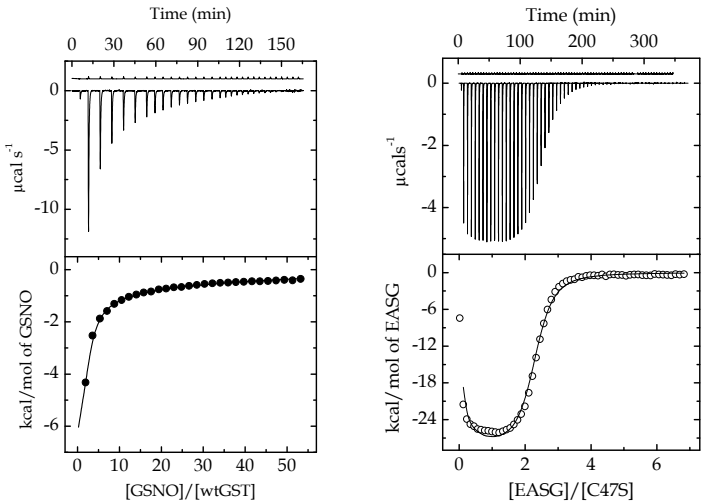


Fig. 2. Representative isothermal titration calorimetry measurements for the binding at 25.1 °C of GSNO (left panel) and EASG (right panel) to wt-hGSTP1-1 and its C47S mutant, respectively. Solid lines in the bottom plots represent the fit to an equal and interacting binding sites model.

The calculated parameters for the left panel in Fig. 2 by an iterative fitting procedure were: (K_1^0=7.0±0.4) ·10^4 and K_2^0= (2.3±0.2) ·10^3 M^{-1}; ΔH_1=-7.9±0.8 and ΔH_2=-24.8±0.7 kcal/mol. Thus, the intrinsic binding constant value for the first site, K_1^0, is higher than that for the second site, K_2^0, and consequently ligand binding induces negative cooperativity in the enzyme. Right panel in Fig. 2 shows the binding of EASG (1 mM) to the C47S mutant (33 μM) of hGSTP1-1. From a closer inspection of both thermograms, it is clear the ligand binds in a different way to the wild-type (left panel) than to the C47S mutant (right panel). For the right panel, the calculated parameter values are: K_1^0=(4.9±0.3) ·10^4 and K_2^0= (1.2±0.1) ·10^6 M^{-1}; ΔH_1=-1.8±0.1 and ΔH_2=-58.5±0.2 kcal/mol. In this case, the microscopic binding constant for

the first site is lower than that for the second site. This result indicates the binding of EASG to the first subunit induces a favorable conformational change on the second one, which in turn displays an increased affinity for EASG (positive cooperativity).

4.3 Others binding models

The procedure described above is suitable for deducing the equations for other models such as multiple sets of independent sites, multiple sets of interacting sites,… However, in all of these cases the model to fit the data to must be known or suspected in advance.

There exists a general formalism to analyze many experimental biological systems involving ligand binding that can be used when the user has no hint about which model to choose: the theory of the binding polynomial (Freire et al., 2009). It allows the analysis of experimental data by employing a general model-free methodology that provides essential information about the system, such as whether there exists binding cooperativity, if it is positive or negative or the magnitude of the cooperativity energy. The binding polynomial contains all the information about the system and allows derivation of all the thermodynamic experimental observables. Consequently, when the model to fit the data is not known, this method is the preferred starting point for the analysis of complex binding situations.

Another situation requiring a special approach to analyze the data is when the ligand binding to the protein is so tight that the high association constant value cannot be measured directly by one ITC experiment. In this case, a displacement titration can solve the issue. It consists of a regular binding experiment where the protein in the cell is premixed with a weaker competitive ligand. Three titrations must be carried out: a direct titration of the high-affinity ligand to the target protein, a direct titration of the weak ligand to the target protein and a displacement titration of the high-affinity ligand to the weak ligand-target protein complex (Velazquez-Campoy & Freire, 2006).

4.4 Data kinetic analysis

ITC also provides a direct and accurate method for determining the kinetic parameters of a chemical reaction when its observed kinetic constant is smaller than the calorimeter response time, from the heat absorbed or released during the reaction. Thus, it is a valid method for some enzyme catalyzed processes.

There are two different scenarios: when the binding heat evolved is negligible compared to that of the chemical reaction, and when both are of a similar magnitude.

4.4.1 Measurement of kinetic parameters by ITC when the binding heat is negligible

This is the case when the enzyme is present at catalytic concentrations. ITC provides a direct and accurate assay for determining kinetic parameters of an enzyme catalysed reaction, based on the measurement of the heat absorbed or released during the catalytic reaction. Frequently, two procedures are employed: multiple injections or the continuous method (Todd & Gomez, 2001). Although the two methods provide analogous results, the use of one or the other depends on the particular characteristics of the reaction under study. However, as a rule of thumb, the multiple injection assay is recommended when the K_m value is greater than 10-15 μM, leaving the continuous assay for when K_m < 10 μM.

As an example we describe the methodology and data analysis to obtain the kinetic parameters for the PfdUTPase enzyme. dUTPase catalyzes the hydrolysis of α-β-pyrophosphate bond of dUTP to yield dUMP and inorganic pyrophosphate (PP_i). This

hydrolysis reaction releases protons and typically is studied spectrophotometrically in a coupled assay by using a pH indicator in a weak buffered medium with similar pK_a. However, the ITC method proved to be more sensitive, even detecting the activity in the case of some mutants where the spectrophotometric method was unable to detect it.
Briefly, the ITC method for enzyme assays is based on the fact that the thermal power reflects the heat flow (dQ/dt). This thermal power or heat flux ($\mu cal \cdot s^{-1}$) is directly proportional to the rate of product formation or substrate deflection and can be described as

$$v = d[P]/dt = \frac{dQ}{dt} / \Delta H_{obs} V_0 \tag{21}$$

where V_0 is the effective volume of the calorimetric cell, and Q and t are values measured during the experiment. ΔH_{obs} is the observed molar enthalpy for the conversion of substrate to product. Fig. 3 shows a typical experimental thermogram (using the continuous method also termed as "single injection assay") for the titration of PfdUTPase with dUTP in the presence of 25 mM $MgCl_2$ at pH 7 and 25.2 °C (Quesada-Soriano et al., 2008).

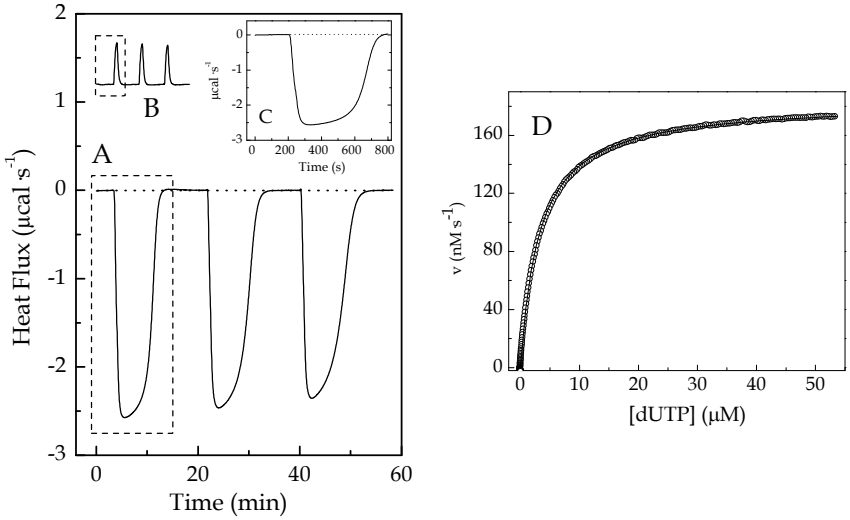


Fig. 3. PfdUTPase-catalyzed hydrolysis of dUTP in 25 mM MES, 100 mM NaCl, 25 mM $MgCl_2$, 1 mM β-mercaptoethanol at pH 7 and 25 °C. (A) Typical calorimetric trace (µcal/s versus time) obtained after addition of three injections of 5 mM dUTP (20 µl) to the calorimetric cell containing PfdUTPase (5.3 nM). (B) dUTP dilution thermogram. (C) Calorimetric trace (first injection) resulting after subtracting the first peak of the dilution experiment. (D) Net thermal power was converted to rate and fitted to the Michaelis–Menten equation, giving $\Delta H_{obs} = -10.4$ kcal/mol, $K_m = 3.2$ µM, $k_{cat} = 11.7\ s^{-1}$.

It is very important to observe that the response time of the calorimetric signals when a kinetic process occurs is much higher than in a typical binding process. In this way, the catalyzed reaction progress can be followed from analysis of the calorimetric peaks (first or second injections). Thus, the reaction rate can be calculated since the heat flow (dQ/dt) is directly proportional to the rate of reaction (Eq. 21). The area under each peak gives the

total heat for the reaction (Q_T) which is converted to ΔH_{obs} (cal/mol) by means of the expression

$$\Delta H_{obs} = \left(-\int_0^{\infty} (dQ / dt) dt \right) / V_0 [S]_0 \tag{22}$$

where $[S]_0$ is the initial substrate concentration, and in this example $[dUTP]_0$. The downward displacement of the baseline after the substrate injection indicates the exothermic nature of the reaction. On the other hand, the substrate concentration in any time $[S]_t$ was calculated by the expression

$$[S]_t = [S]_0 - \left(\int_0^{t} (dQ / dt) dt \right) / V_0 \Delta H_{obs} \tag{23}$$

Fig. 4, shows the raw data for thermal power change as a function of time in the multiple injections assay for the hydrolysis process of dTTP by human dUTPase (Quesada-Soriano et al., 2010).

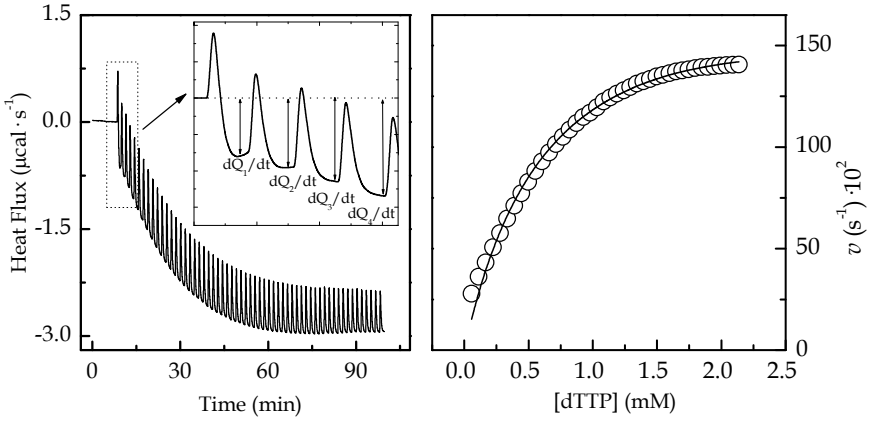


Fig. 4. Raw data for thermal power change as a function of time in the multiple injections assay for the hydrolysis process of dTTP by human dUTPase (left panel). The non-linear least squares fits of the experimental data to the Michaelis–Menten equation (right panel). Kinetic assays were performed in 25 mM MES, 5 mM NaCl, 1 mM β-mercaptoethanol and 25 mM $MgCl_2$ (pH 7, 25 °C), making 5 μL injections of 15 mM substrate into calorimetric cell containing 15–25 μM enzyme.

Briefly, this method measures the rate of enzyme catalysis during stepwise increase of substrate concentration in an enzyme solution. After initial equilibration, as a result of the deoxynucleoside triphosphate injections (5 μL each), an initial endothermic peak corresponding to the heat dilution was generated. The baseline then stabilises at a lower power level than those of the previous injections as a consequence of the heat generated by the enzymatic reaction and due to the fact that the higher the substrate accumulation, the faster the reaction occurs. The drop in baseline also indicates that this particular reaction proceeds with a negative (exothermic) enthalpy. Values of rate determined in this way will yield data

into units of μcal per second. These can be readily converted in units of molar per second using Eq. 21 and the ΔH_{obs} value. Thus, in the kinetic experiments using the stepwise injection method is need to perform an additional experiment using a higher enzyme concentration in the calorimetric cell, in order to determine ΔH_{obs} (Eq. 22). The non-linear least squares fit of the experimental data to the Michaelis-Menten equation for determining the hydrolysis parameters (K_m= 687 μM, k_{cat}= 0.12 s^{-1}) is shown in the right side of thermogram.

4.4.2 Evaluation of kinetic reactions associated to a binding process

Generally, the binding of a ligand to a macromolecule occurs due to different types of non-covalent interactions such as hydrogen bonds, van der Waals, π-stacking, electrostatic …. Thus, the affinity and the energetic for the binding process are intrinsically related with the nature, strength and number of those interactions. The observed heat in a thermogram is a global value of all the contributions taking place simultaneously. It is of outmost importance to determine which processes are contributing to the observed heat and to correct for them, if needed, in order to get the value of the intrinsic binding enthalpy. A particular case occurs when the binding process is followed by a covalent reaction. In those cases, the calorimetric signals will be the result of the two processes: binding and covalent reaction. The analysis of the calorimetric thermograms results in those cases more complicated and for this reason is avoided in the interaction studies. As a result it is difficult to find study cases in the literature. Our description will be based in the study of the interaction of diuretic ethacrynic acid (EA) with the enzyme hGSTP1-1 in phosphate buffer at pH 7 and 25 ºC (Quesada-Soriano et al., 2009). We demonstrated that EA (inhibitor and substrate of hGSTP1-1) binds irreversibly to the α2 loop Cys 47. Fig. 5 shows a representative thermogram.

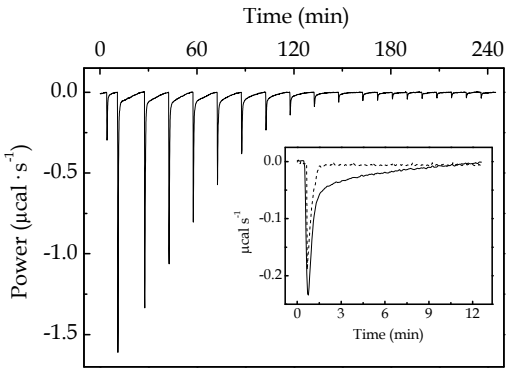


Fig. 5. Calorimetric thermogram for the titration of 23 μM wt GSTP1-1 with 5 μL injections (1 μL first injection) of 2.1 mM EA in 20 mM sodium phosphate, 5 mM NaCl and 0.1 mM EDTA at pH 7.0 and 25ºC. Inset plot: Comparison between a peak from a calorimetric thermogram for the titration of wt enzyme with EA, reflecting the slow kinetic process caused by covalent modification (solid line) and a representative calorimetric trace of a typical binding peak in the absence of a kinetic process (dashed line).

Fig. 5 reveals important differences when comparing the thermogram corresponding to a typical binding titration to one with a concomitant kinetic process. The calorimetric responses after each injection of ligand (peaks in the thermogram) show at least two exothermic phases clearly differentiated by their response times. The response time for the

first phase, ~2 min (comparable to the response time either of a typical binding or a dilution experiment) is smaller than that obtained for the second phase (~12 min). Therefore, the presence of these two phases reveals the occurrence of two sequential and separable steps in the binding reaction. The first process (fast) could be the binding of EA to the protein and the second one could correspond to a chemical reaction (slow) occurring concomitant to the binding process.

The analysis of this thermogram can be done by applying a theoretical model that adequately describes the results from an ITC instrument with feedback system. The capacity to correctly predict the response of the calorimeter to any given thermal effect can allow us to satisfactorily analyze the obtained thermogram. We demonstrated that for an isothermal calorimeter with dynamic power compensation (feedback system) and a similar configuration as used here (VP-ITC from Microcal), the response function to a power pulse (W_0) of finite duration (ξ) can be expressed by the Eq. 24,

$$W(t) = \frac{W_0}{\tau_S - \tau_Q}\left[\tau_S e^{-(t-\xi)/\tau S} - \tau_Q e^{-(t-\xi)/\tau Q} - \left(\tau_S e^{-t/\tau S} - \tau_Q e^{-t/\tau Q}\right)\right] \quad (24)$$

where, if $t<\Delta t$, then $\xi=t$, and if $t\geq \Delta t$, then ξ becomes constant and equal to Δt (García-Fuentes et al., 1998). τ_S and τ_Q are the response times for the sample cell and the assembly sample cell-feedback system, respectively. It is also worth noting that each injection (peak) displayed in the thermogram in Fig. 5, includes ligand dilution, which always occurs in any binding experiment. Moreover, the possible proton exchange should be also included in each injection. However, since the buffer used is phosphate which has a small ionization enthalpy, the possible protonation/deprotonation effects may be negligible. If the behaviour for each individual process (binding, chemical and dilution) is represented by Eq. 24, the signal generated by each ligand injection will be composed by three contributions:

$$W_g(t) = W_{binding}(t) + W_{chemical}(t) + W_{dilution}(t) \quad (25)$$

By subtracting the dilution trace by a blank experiment, the resultant thermogram will include now only two processes and the generated global power ($W_g(t)$) in each injection will be:

$$W_g(t) = W_{binding}(t) + W_{chemical}(t) \quad (26)$$

where,

$$W_{binding}(t) = \frac{W_{0,b}}{\tau_{S,b} - \tau_Q}\left[\tau_{S,b} e^{-(t-\xi)/\tau S,b} - \tau_Q e^{-(t-\xi)/\tau Q} - \left(\tau_{S,b} e^{-t/\tau S,b} - \tau_Q e^{-t/\tau Q}\right)\right] \quad (27)$$

$$W_{chemical}(t) = \frac{W_{0,c}}{\tau_{S,c} - \tau_Q}\left[\tau_{S,c} e^{-(t-\xi)/\tau S,c} - \tau_Q e^{-(t-\xi)/\tau Q} - \left(\tau_{S,c} e^{-t/\tau S,c} - \tau_Q e^{-t/\tau Q}\right)\right] \quad (28)$$

Eq. 26 has six parameters to fit: W_0 and τ_S for the binding process (denoted as $W_{0,b}$ and $\tau_{S,b}$ in Eq. 27);W_0 and τ_S for the chemical process (denoted as $W_{0,c}$ and $\tau_{S,c}$ in Eq. 28); ξ (injection duration) and τ_Q (characteristic time constant for the instrument and therefore, independent of the process occurring in the cell). $\tau_{S,b}$ and τ_Q should be equal to those obtained in the

dilution experiment. This is also true for ξ, because the volume and duration time of the injections are the same in both experiments. On the basis of this empirical procedure, we developed a computational algorithm to fit the individual calorimetric peaks contained in the thermograms obtained from titration at each temperature. Firstly, the peaks from the dilution experiment were used to obtain those three parameters (i.e. $\tau_{S,dilution}=\tau_{S,b}$; $\tau_{Q,\ dilution}=\tau_Q$; ξ), which are then included as fixed values in Eq. 26. This way, each peak in the calorimetric thermogram originated from titrating the protein with the ligand is fitted to Eq. 26 after subtracting the corresponding dilution peak in the reference experiment, obtaining the remaining three parameters: $W_{0,c}$, $W_{0,b}$ and $\tau_{S,c}$.
Fig. 6 shows, as an example, the theoretical deconvolution for a peak from Fig. 5. The two contributions included in Eq. 27, binding and chemical reaction, are visualized as dashed and dotted lines, respectively.

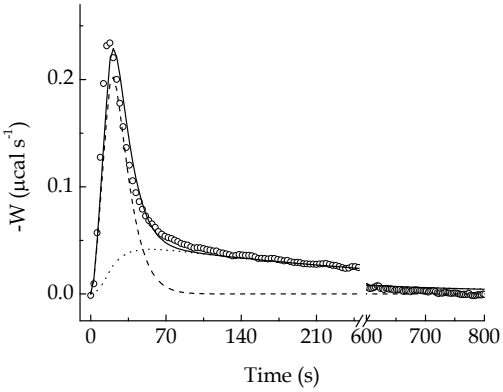


Fig. 6. Schematic example of the theoretical deconvolution procedure for a calorimetric peek of Fig. 5. Solid, dash and dot lines correspond to global signal, binding contribution and kinetic contribution, respectively.

The same deconvolution was applied at all the peaks in the calorimetric trace, and as a result two thermograms are generated. The binding thermogram was used to calculate the binding parameters with the proper model (in this case, two equal and independent sites). Lastly, the kinetic constant for the chemical reaction at this temperature can be calculated from the response time, $\tau_{S,c}$, as $k=1/\tau_{S,c}$.

5. Evaluation of protonation effects in binding processes

The calorimetric enthalpy measured (ΔH_{obs}) is the sum of different heat effects taking place during any reaction. Thus, for instance, if a kinetic reaction involves the release (or uptake) of protons, ΔH_{obs} will be a combination of the reaction intrinsic enthalpy, ΔH_{rxn}, and the protonation (or ionization) enthalpy for each proton absorbed (or released) by the buffer. For instance, the hydrolysis reaction of dUTPase is accompanied by a release of protons, which will be absorbed by the buffer. If the same experiment is done under the same solution conditions but using different buffers with different a protonation enthalpies, the peaks in the traces will have different areas. Fig. 7 shows such a case for the PfdUTPase/dUTP reaction in glycerophosphate, Pipes, Mes, Hepes and TES buffers.

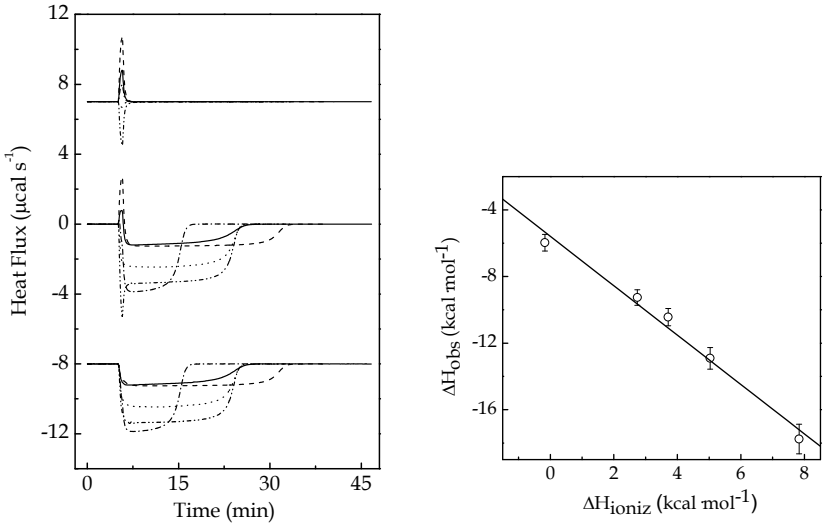


Fig. 7. Protonation effect in the dUTP hydrolysis by PfdUTPase at pH 7 and 25 °C. The calorimetric thermograms correspond to one 20 μL injection of 9.98 mM dUTP to the calorimetric cell containing PfdUTPase (5.5 nM) in (solid line) glycerophosphate, (dash line) Pipes, (dash dot line) Mes, (dot line) Hepes and (dash dot dot line) TES. Upper, medium and bottom panels correspond to the calorimetric traces for dUTP dilution, protein titration and net (protein titration minus dUTP dilution), respectively. Right figure correspond to the fitting to Eq. 29 of the observed reaction enthalpy change, ΔH_{obs}, obtained in each buffer system, *versus* the ionization enthalpy.

$$\Delta H_{obs} = \Delta H_{rxn} + n_H \Delta H_{ioniz} \tag{29}$$

Thus, the enthalpy change of this hydrolysis reaction, ΔH_{rxn} can be calculated from the linear relationship between ΔH_{obs} and the ionization enthalpy, ΔH_{ioniz} (Eq. 29). The intercept gives a ΔH_{rxn} of -5.58 ± 0.52 kcal/mol, and the slope being the number of protons exchanged during the hydrolysis reaction, n_H: -1.48 ± 0.12. In many cases these linked effects are crucial for understanding an enzyme mechanism.

Similarly to a kinetic reaction, whenever binding is coupled to changes in the protonation state of the system, the measured heat signal will contain the heat effect due to ionization of buffer (Baker & Murphy, 1996). Therefore, the observed enthalpy changes (ΔH_{obs}) derived from binding isotherms, are not solely contributed to by the physical forces governing the protein–ligand interactions, but they often contain contributions from the ionization enthalpy of the buffer species and/or changes in the protein conformations. Although the enthalpic contributions from protein conformational changes can be taken as an integral component of the overall binding process, the enthalpic contributions due to protonation changes of the buffer species must be subtracted from the observable. Repeating the calorimetric experiment at the same pH in buffers of different ΔH_{ioniz} allows to determine the number of protons n_H that are released ($n_H > 0$) or taken up ($n_H < 0$) by the buffer, and thus to calculate the intrinsic binding enthalpy corrected for protonation heats using a similar relationship to Eq. 29.

6. Case studies

6.1 Thermodynamics of binding to a protein and mutant

A great amount of molecular recognition studies are directed to investigate the binding or affinity of a series of ligands or drugs, with similar structure, to a target protein. The comparative studies with mutants of that protein help to find the role of individual residues both in the catalytic mechanism and in the binding mode. Two examples will illustrate this. Complete thermodynamic profiles consisting of free energy, enthalpy and entropy changes can be obtained for the reactions of interest by calorimetric techniques. The thermodynamic parameters calculated allow to know both the nature of binding site and the functional groups of the ligand that are important for the interaction to occur. This information is very valuable in drug design and cannot be obtained from structural or computational methods alone.

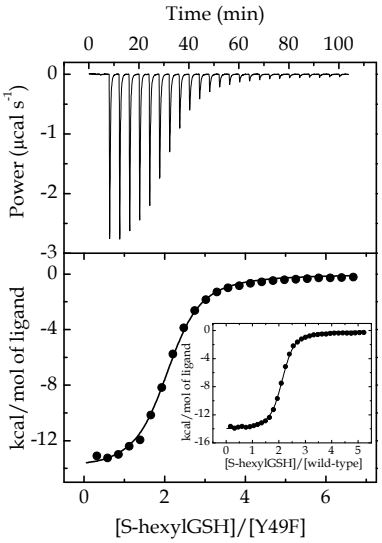


Fig. 8. Representative isothermal titration calorimetry measurements of the binding of S-hexylGSH to the Y49F mutant of hGST P1-1. A 15.32 µM mutant enzyme solution was titrated with 1.12 mM S-hexylGSH. (inset) Integrated heats per mol of S-hexylGSH injected into a solution of wild-type enzyme. 32.66 µM wild type was titrated with 1.31 mM S-hexylGSH.

In the first example we show the thermodynamics of binding of the substrate glutathione (GSH) and the competitive inhibitor S-hexylglutathione to the wt-enzyme and the Y49F mutant of the human glutathione S-transferase (hGST P1-1) (Ortiz-Salmerón et al., 2003). Structural studies revealed that two residues (Cys 47 and Tyr 49) located in a mobile helix, denoted as α2, could participate in intersubunit communication between active sites of the dimer. Fig. 8 shows a representative titration of the Y49F mutant with S-hexylGSH in phosphate buffer at pH 6.5.

6.1.1 Protonation state change

Calorimetric titration experiments were repeated in various buffers of different ΔH_{ioniz} (phosphate, MOPS, MES and ACES) at pH 6.5 and 25 °C. The binding enthalpy, ΔH and the number of exchanged protons were obtained from the intercept and slope of a linear plot

according to Eq. 29 (Table 1). A negative slope was obtained ($n_H < 0$), with n_H ~-0.44 and n_H~-0.11 for the binding of GSH to the wild type enzyme and its Y49F mutant, respectively (Table 1).

		Y49F mutant		Wild-type	
		GSH	S-hexylGSH	GSH	S-hexylGSH
Buffer	ΔH_{ioniz} (kcal mol^{-1})	$-\Delta H_{obs}$ (kcal mol^{-1})		$-\Delta H_{obs}$ (kcal mol^{-1})	
Phosphate	1.22	13.04 ± 0.31	17.14 ± 0.33	9.91 ± 0.19	16.13 ± 0.26
Mes	3.72	-	-	11.27 ± 0.31	15.58 ± 0.42
Mops	5.27	13.48 ± 0.33	16.55 ± 0.41	11.67 ± 0.23	-
Aces	7.53	13.74 ± 0.40	15.41 ± 0.25	12.71 ± 0.32	16.20 ± 0.45
n_H		-0.11 ± 0.01	0.24 ± 0.09	-0.44 ± 0.08	-0.02 ± 0.01

Table 1. Protonation effect.

Hence, upon enzyme-GSH complex formation, the number of protons released for the wild-type enzyme binding was higher than that for the Y49F mutant. This means that one or more pK_a values, corresponding to some donor proton groups in the ligand and/or enzyme, decrease (i.e. become more acidic). Although n_H indicates the global number of protons uptaken or released upon ligand binding, it is usually the result of only one or two residues changing their protonation state. Since ~-0.45 H^+/monomer are released in the wild-type/GSH binding, but the number is practically zero for the binding of the S-hexylGSH inhibitor to the wild-type hGSTP1-1, the protons released during the binding of substrate (GSH) to the wild-type enzyme might come from the thiol group of the sulfhydryl in GSH. However, there exists a net concomitant uptake of protons for the formation of the S-hexylGSH-Y49F complex (n_H~0.25) (Table 1). The results for the Y49F mutant were explained by assuming that the mutation induces slight changes in the environment of the binding site, producing a pK_a shift in one or more groups of the ligand and/or enzyme upon complex formation. In this case, it cannot be assumed that only the sulfhydryl group in GSH is responsible for the proton exchange at this pH, since the binding of S-hexylGSH to Y49F should therefore take place without a proton exchange, which was contrary to our results. Overall the following could be deduced: 1) the proton of the thiol group of GSH is released upon binding to both the Y49F mutant and wild-type anzymes; and 2) upon binding of GSH to the mutant, at least two groups participate in the proton exchange: the sulfhydryl group in GSH and a second group with a low pK_a capable of increasing its pK_a as a consequence of binding. This second group takes up ~0.25 protons from the buffer media. Whereas an unambiguous assignment of the specific residue(s) responsible for the binding-induced uptake of protons is not possible, the side chains of Asp and Glu are likely candidates as ionizing groups at pH 6.5. Crystallographic studies showed that the Asp 98 of the adjacent subunit is located at the active site. This residue (pK_a=4.8), involved in a hydrogen-bonding network around the γ-glutamate of GSH or S-hexylGSH, could increase its pK as a consequence of a small local conformational change arising from the mutation, thus explaining the number of protons taken up in the association with these ligands.

6.1.2 Temperature dependence

When an ITC experiment is carried out at several temperatures the heat capacity change of the reaction can be obtained from the enthalpy dependence on temperature. Fig. 9 shows the

dependency of the thermodynamic parameters on temperature for the wild-type enzyme and its Y49F mutant. The binding of these ligands to both enzymes is noncooperative within the temperature range analyzed, which suggests that the interaction does not induce a conformational change affecting the binding of the second ligand molecule to the other site of the dimer.

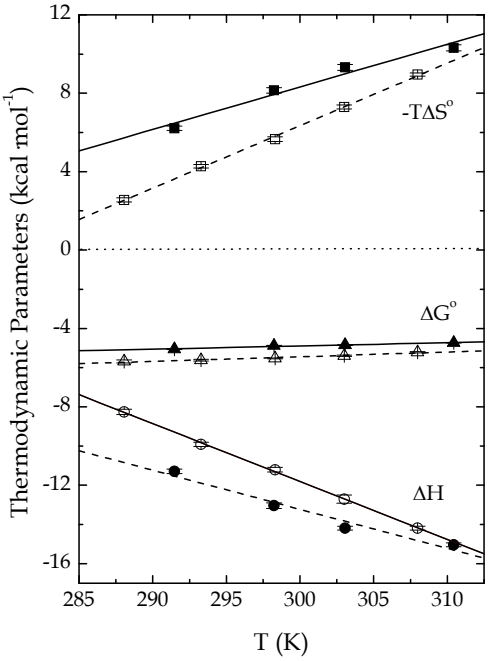


Fig. 9. Temperature dependence of the thermodynamic parameters for the binding of GSH to both the Y49F mutant (filled symbols) and the wild-type (open symbols) enzyme.

ΔG^0 is almost insensitive to the temperature change, in both cases, but increases as a consequence of the mutation, so the ligand affinity (GSH and S-hexylGSH) for Y49F is lower than for the wild-type enzyme. This behavior was observed at all the temperatures studied (Fig. 9). This result alone could also be obtained from other techniques, such as fluorescence. However, ITC is able to split the affinity values into the enthalpic and entropic contributions, allowing for a deeper understanding of the binding process.

Protein	Ligand	K	ΔH	TΔS°	$\Delta C_p°$
		M^{-1}	kcal mol^{-1}		cal K^{-1} mol^{-1}
WT	GSH	11630±302	-11.21±0.12	-5.65±0.12	-294.2±2.7
Y49F	GSH	3883±83	-13.04±0.14	-9.17±0.14	-199.5±26.9
WT	S-hexylGSH	(8.1±0.3) ·10^5	-16.13±0.07	-8.04±0.07	-441.6±48.7
Y49F	S-hexylGSH	(4.3±0.1) ·10^5	-17.14±0.08	-9.45±0.08	-333.6±28.8

Table 2. Thermodynamic parameters for the interaction of GSH and S-hexylGSH to the Y49F mutant and wild-type enzymes of hGST P1-1, at 25.2°C and pH 6.5.

As shown in Fig. 9, although ΔG^0 is almost insensitive to the change in temperature, ΔH and $T\Delta S^0$ strongly depend on it, for both enzymes. This feature is known as enthalpy-entropy compensation, and it is very common in most of the thermodynamic binding studies of biological systems.

The enthalpy-entropy compensation is related to the properties of the solvent as a result of perturbing weak intermolecular interactions. In Fig. 9 a linear dependence of ΔH with the temperature was observed, from the slope of which the heat capacity change (ΔC_p^o) upon ligand binding was obtained. The absolute values of ΔC_p^o calculated for the binding of either substrate or S-hexylGSH to the mutant Y49F decreased ~100 cal K^{-1} mol^{-1} when compared to those for the wild-type enzyme. This is most probably due to a different release of water molecules from both complexes, being the number of water molecules released by the wt-ligand complex higher than in the mutant-ligand case. The release of water molecules is accompanied by an increased entropy change. This suggestion for the ΔC_p^o difference is supported by a comparison between the entropy change values for both enzymes (Table 2), since it is slightly higher for the wild-type at all temperatures (Fig. 9). The data reported here indicate that, thermodynamically, the mutation leads to increased changes in negative enthalpy and negative entropy (Table 2 and Fig. 9). Although the interaction between the Y49F mutant and GSH is enthalpically more favorable than that for wild-type enzyme, the entropic loss due to binding is also increased, indicating that the mutation is both enthalpically favorable and entropically unfavorable (Table 2). The same tendency was obtained for the binding of S-hexylGSH. This means that the thermodynamic effect of this mutation is to decrease the entropic loss due to binding. The unfavorable entropy change outweighs the enthalpic advantage, resulting in a 3-fold lower binding constant for the binding of GSH to the wild-type enzyme. Also, compared to GSH, S-hexylGSH shows more negative values for the two contributions (enthalpic and entropic), maybe as a consequence of the higher apolarity in the inside of the active site provided by the hexyl chain of this inhibitor. This also could explain its higher affinity.

6.1.3 Correlation between ΔC_p and the buried surface area

Works carried out during the last decade supports the view that the changes in the thermodynamic quantities associated with the folding and ligand-binding processes can be parametrized in terms of the corresponding changes in nonpolar (ASA_{ap}) and polar (ASA_p) areas exposed to the solvent (in $Å^2$ units). Thus, Freire and co-workers (Murphy & Freire, 1992) have suggested the following equations for the ΔC_p and the enthalpy change (at the reference temperature of 60 °C (this temperature is taken as a reference because it is the mean value of the denaturation temperatures of the model proteins used in the analysis).

$$\Delta C_p = 0.45\Delta ASA_{ap} - 0.26\Delta ASA_{pol} \tag{30}$$

$$\Delta H_{60} = -8.44\Delta ASA_{ap} + 31.4\Delta ASA_{pol} \tag{31}$$

Therefore, from an experimentally determined ΔC_p and a measured enthalpy change at 25 °C, the corresponding enthalpy change at 60°C is calculated by $\Delta H_{60}=\Delta H_{25} +\Delta C_p$ (60-25). If we admitted Eqs. 30 and 31 to be valid for binding processes, we would have a two equation system with two unknowns from which we could calculate changes in accessible surface

areas. Changes in apolar (ΔASA_{ap}) and polar (ΔASA_p) solvent-accessible surface areas upon complexation have been estimated by those relationships mentioned above. On the basis of the X-ray crystallographic data of several proteins, the changes in the water-accessible surface areas of both nonpolar (ΔASA_{ap}) and polar (ΔASA_p) residues on protein folding have been calculated. Such calculations reveal that the ratio $\Delta ASA_{ap}/\Delta ASA_p$ varies between 1.2 and 1.7. This range is comparable with the medium value for the ratio of $\Delta ASA_{ap}/\Delta ASA_p$ of ~1.2, calculated for the interactions described in this study. The application of Murphy's approach (Murphy & Freire, 1992) to the experimentally determined values indicates that the surface areas buried on complex formation comprise ~54% of the nonpolar surface. Therefore, as Spolar and Record (1992) indicate, these values can be taken as the "rigid body" interactions, and therefore, no large conformational changes occur as a consequence of the association with these ligands. On the other hand, Eq. 30 can also be used to obtain an estimation of the ΔC_p value for a generic macromolecule-ligand interaction, under the assumption that this parameterization is valid for binding processes. In those cases, structural information should be available for the complex as well as for the interacting species. In general, changes in solvent-accessible surface area (ΔASA) are determined as the difference in ASA between the final and initial states. For a molecular interaction process, this is the difference between the ASA for the complex and the sum of the ASA for the macromolecule and the ligand, resulting in negative values of ΔASA. ΔASA is further subdivided into nonpolar and polar contributions by simply defining which atoms take part in the surface. The original description is based on the Lee and Richards (1971) algorithm implemented in software NACCESS (Hubbard & Thornton, 1996), using a sphere (of solvent) of a particular radius to 'probe' the surface of the molecule. There is a high number of other parameterizations used also to determine ASA. However, since each implementation yields slightly different results, it is very important to assure the implemented parameters when performing calculations.

6.2 Application of calorimetry to predict the binding mode of a ligand

Frequently, as it has been described before, negative ΔC_p values are usually attributed to the burial of apolar groups from water. However, in many systems it is also thought to be associated with hydrophobic stacking interactions, presumably resulting from the dehydration of highly ordered water molecules surrounding hydrophobic surfaces. In order to describe in more detail this observation we will show two examples.

6.2.1 Binding of EASG to wt and Y108V mutant of hGSTP1-1

The EASG is a conjugate of ethacrynic acid (with aromatic groups) and reduced glutathione (tripeptide). The thermodynamic parameter values obtained are shown in Table 3. As it can be observed, the thermodynamic parameter values are very similar, the main difference being the absolute ΔC_p values, which were considerably larger (approximately twice as large) for the binding to the Y108V mutant when compared to the wt enzyme.

These results might be used to make an assessment on the predilection of the Y108V mutant to adopt a particular 3D structure, when it is complexed with EASG, where there is a high drop in the centroid distance between the ligand EA moiety and Phe 8. This analysis is based on the assumption that the overall conformation of the Y108V mutant complexed with

EASG is relatively unchanged from that of the wt protein and that the main effect is the π-stacking between the EA and flanking aromatic amino acids such as Phe 8 located in the active site. Thus, the more negative value of ΔC_p was interpreted as coming from a strengthening of the π-stacking between the EA moiety and Phe 8 in the absence of Tyr 108. This increase in the strength of a particular π-π interaction in the Y108V mutant explained the larger negative ΔC_p value for the interaction with it compared to the wt enzyme. These results were corroborated by the X-ray crystallography of this free mutant and its complex with EASG.

Protein	K_d (μM)	ΔH (kcal mol^{-1})	TΔS^0 (kcal mol^{-1})	ΔC^0_p (cal mol^{-1}K^{-1})
WT	0.5 ± 0.1	-14.17 ± 0.37	-5.59 ± 0.37	-264 ± 24
Y108V	0.3 ± 0.1	-13.34 ± 0.46	-4.55 ± 0.31	-415 ± 17

Table 3. Thermodynamic parameters of the interaction of EASG with wt GST P1-1 and theY108V mutant at 25 ^{o}C and pH 7.0

Therefore, these results demonstrate that ITC measurements can provide a thermodynamic fingerprint of drug–protein interactions, which can be related to the molecular mode of binding in multimeric enzymes. Calorimetric data can therefore be a useful method to complement X-ray crystallographic studies. However, even in the absence of structural information, thermodynamics can give a guideline to improve drug potency. Reports of comparisons of drug binding energies with relevant structural studies are still rather scarce in the literature. Thus, studies such as ours are of great general interest.

6.2.2 Binding of S-benzylglutathione to GST from *Schistosoma japonicum*

When X-Ray data are not available for a particular macromolecule-ligand complex, but there is enough structural knowledge about the ligand binding modes of similar ligands to the same or similar proteins, the combination of ITC thermodynamic data with docking studies may predict the preferred binding mode or pose of the ligand under study in the complex.

As an example, we have carried out the thermodynamic studies of the binding of S-benzylglutathione to the GST from *Schistosoma japonicum* (wt-SjGST) and three important Tyr 111 mutants (Y111L, Y111T and Y111F), because this residue has an active role in the enzyme activity (to be published).

The ΔC_p values for the interaction of this ligand with wt-SjGST and the Y111F, Y111L and Y111T mutants were -733, -745, -576 and -499 cal mol^{-1}K^{-1}, respectively, within a 20-40 ^{o}C range.

From these ΔC_p values it is clear the interaction of the ligand with the wt and the Y111F mutant must proceed in a similar fashion, whereas the thermodynamic data show the interaction with the Y111L and Y111T mutants has been altered compared to the wt.

Following the previous reasoning relating ΔC_p values to a change in π-stacking interactions, and taking into account the structural knowledge and ligand binding modes of this protein (Cardoso et al., 2003), it is not illogical to think the difference in the ΔC_p values might come from a lacking π-stacking interaction between the benzyl moiety of the ligand and the aromatic ring of the 111 residue in the active site in the case of the Y111L and Y111T mutants. To address this question we carried out a series of docking studies with AutoDock

Vina (Trott & Olson, 2010), where the ligand was docked to the wt and its three mutants, which were built with the "Mutagenesis Wizard" in PyMOL from the wt structure (The PyMOL Molecular Graphics System, Version 1.3.1_pre3925, Shrödinger, LLC). Fig. 10 (PyMOL) superimposes the two best docked poses for S-benzylglutathione in the active site for the four enzymes, as well as the binding mode of the very similar S-2-iodobenzylglutathione as a reference (pdb structure 1M9B, Cardoso et al., 2003).

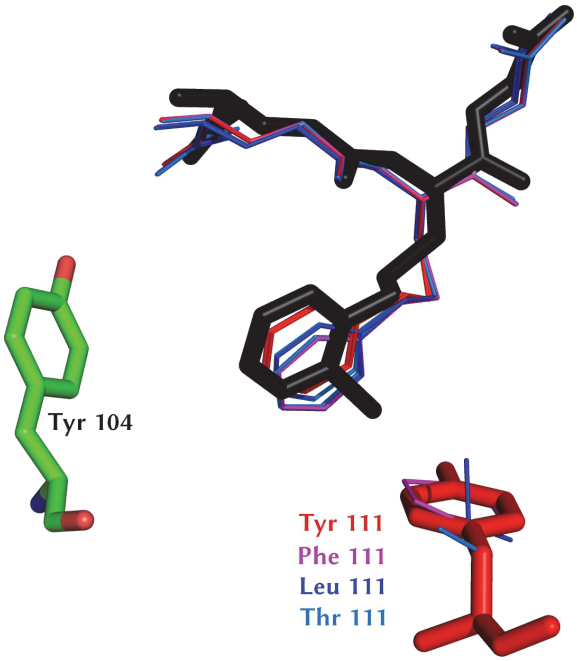


Fig. 10. Two best docked poses of S-benzylglutathione in the active site of wt SjGST (red) and its three mutants: Y111F (magenta), Y111L (blue) and Y111T (light blue). Only the relevant 104 and 111 residues are displayed. S-2-iodobenzylglutathione (black) from the PDB structure 1M9B is shown as a reference.

	Angle (Tyr 104)	Centroid distance (Tyr 104)	Angle (res. 111)	Centroid distance (res. 111)
Reference [a]	50.0°	6.6 Å	80.8° [b]	6.6 Å
wt	58.2°	6.9 Å	73.9° [b]	6.5 Å
Y111F	59.2°	7.4 Å	73.5° [c]	6.3 Å
Y111L	60.0°	7.3 Å	-	-
Y111T	59.3°	7.2 Å	-	-

Table 4. Mean geometric data for the interaction between the S-benzylglutathione ring plane and the Tyr 104 and residue 111 ring planes. [a]S-2-iodobenzylglutathione from the PDB structure 1M9B. [b] wt Tyr 111. [c]Phe 111

The predicted S-benzylglutathione binding modes agree with the crystallographic structure for this reference ligand (Table 4), where the aromatic ring of the benzyl moiety stacks between the side-chains of Tyr 104 and Tyr 111 in the case of the wt and the Y111F mutant, but stacks only against Tyr 104 in the case of the Y111L and Y111T mutants. We propose the lacking π-stacking interaction against the side-chain of residue 111 in these two mutants is responsible for the difference in the ΔC_p measured.

7. Conclusions

In the beginning of this chapter we stated thermodynamics can help us understand how life works. We hope the reader has now a better understanding why we did so. Among the available thermodynamic techniques to address the question, Isothermal Titration Calorimetry is perhaps the most powerful tool at our disposal. Apart from being a universal technique, not only does it provide us with the crucial affinity between two interacting species, but it also allows us to know the reason for it. By splitting the affinity into enthalpic and entropic contributions, and after combining the thermodynamic parameters with the structural knowledge available or molecular modeling predictions, the individual interactions responsible for the recognition between a ligand and a macromolecule emerge. This information is of the utmost importance for a rational drug design against a desired target. Last but not least, ITC can even be used for kinetic studies. It is able to detect weak enzyme activity under situations where the traditional spectrophotometric method fails, as we demonstrated. In fact, it is even possible to detect a covalent bonding in the cases where the binding process under study is followed by an unexpected chemical reaction, a situation impossible to properly address with other titration techniques. We have shown how to find out the thermodynamic and kinetic parameters from a single ITC trace in this case.

8. Acknowledgments

Research in the authors' laboratory is supported by grants CTQ2010-17848 from Spanish Plan Nacional , Ministerio de Ciencia e Innovación (co-financied by FEDER) and FQM-3141 and CVI-6028 from the Andalusian Region, Junta de Andalucía (Spain).

9. References

Baker, B.M. & Murphy, K.P. (1996) Evaluation of linked protonation effects in protein binding reactions using isothermal titration calorimetry. *Biophys. J.* 71, 2049 – 2055.

Cardoso, R.M.F.; Daniels, D.S.; Bruns, C.M. & Tainer, J.A. (2003) Characterization of the electrophile binding site and substrate binding mode of the 26-kDa glutathione S-transferase from *Schistosoma japonicum*. *Proteins*, 51, 137-146.

Freire, E. ; Schön, A. & Velazquez-Campoy, A. (2009) Isothermal titration calorimetry: general formalism using binding polynomials. *Methods Enzymol*. 455, 127-55.

García-Fuentes, L. ; Barón, C. & Mayorga. O.L. (2008) Influence of dynamic power compensation in an isothermal titration microcalorimeter. *Anal. Chem*. 70, 4615-23.

Hilser, V.J. ; Gómez, J. & Freire, E. (1996) The enthalpy change in protein folding and binding: Refinement of parameters for structure-based calculations. *Proteins* 26, 123–133.

Horn, J. R. ; Russell, D.M. ; Lewis, E.A. & Murphy, K.P. (2001) Van' t hoff and calorimetric enthalpies from isothermal titration calorimetry: Are there significant discrepancies?. *Biochemistry* 40, 1774 – 1778.

Hubbard, S.J. & Thornton, J.M. (1996) NACCESS Computer Program. 2.1.1 ed. London: Department of Biochemisry & Molecular Biology, University College.

Lee, B. & Richards, F.M. (1971) The interpretation of protein structures: estimation of static accessibility. *J. Mol. Biol.*; 55, 379–400.

Murphy, K. P. & Freire, E. (1992) Thermodynamics of Structural Stability and Cooperative Folding Behavior in Proteins. *Adv. Protein Chem.* 3, 313–361.

Ortiz-Salmerón, E. ; Nuccetelli, M. ; Oakley, A.J. ; Parker, M.W. ; Lo Bello, M. & García-Fuentes, L. (2003) Thermodynamic description of the effect of the mutation Y49F on human glutathione transferase P1-1 in binding with glutathione and the inhibitor S-hexylglutathione. *J. Biol. Chem.* 278, 46938-48.

Quesada-Soriano, I. ; Casas-Solvas, J.M. ; Recio, E. ; Ruiz-Pérez, L.M. ; Vargas-Berenguel, A. ; González-Pacanowska, D. & García-Fuentes, L. (2010) Kinetic properties and specificity of trimeric *Plasmodium falciparum* and human dUTPases. *Biochimie* 92,178-86.

Quesada-Soriano, I. ; Leal, I. ; Casas-Solvas, J.M. ; Vargas-Berenguel, A. ; Barón, C. ; Ruíz-Pérez, L.M. ; González-Pacanowska, D. & García-Fuentes, L. (2008) Kinetic and thermodynamic characterization of dUTP hydrolysis by *Plasmodium falciparum* dUTPase. *Biochim. Biophys. Acta.* 1784, 1347-55.

Quesada-Soriano, I. ; Musso-Buendia, J.A. ; Tellez-Sanz, R. ; Ruíz-Pérez, L.M. ; Barón, C., González-Pacanowska, D. & García-Fuentes, L. (2007) *Plasmodium falciparum* dUTPase: studies on protein stability and binding of deoxyuridine derivatives. *Biochim. Biophys. Acta* 1774, 936-45.

Quesada-Soriano, I. ; Parker, L.J. ; Primavera, A. ; Casas-Solvas, J.M. ; Vargas-Berenguel, A. ; Barón, C. ; Morton, C.J. ; Mazzetti, A.P. ; Lo Bello, M. ; Parker, M.W. & García-Fuentes, L. (2009) Influence of the H-site residue 108 on human glutathione transferase P1-1 ligand binding: structure-thermodynamic relationships and thermal stability. *Protein Sci.* 18, 2454-70.

Spolar, R.S. ; Livingstone, J.R. & Record, M.T. (1992) Coupling of local folding to site-specific binding of proteins to DNA. *Biochemistry* 31, 3947–3955.

Téllez-Sanz, R. ; Cesareo, E. ; Nuccetelli, M. ; Aguilera, A.M. ; Barón, C. ; Parker, L.J. ; Adams, J.J. ; Morton, C.J. ; Lo Bello, M. ; Parker, M.W. & García-Fuentes, L. (2006) Calorimetric and structural studies of the nitric oxide carrier S-nitrosoglutathione bound to human glutathione transferase P1-1. *Protein Sci.* 15, 1093-105.

Todd, M.J. & Gomez, J. (2001) Enzyme kinetics determined using calorimetry: a general assay for enzyme activity?. *Anal. Biochem.* 296, 179–187.

Trott, O. & Olson, A.J. (2010) AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* 31, 455-461.

Velazquez-Campoy, A. & Freire, E. (2006) Isothermal titration calorimetry to determine association constants for high-affinity ligands. *Nature Protocols*, 1, 186-191.

Wiseman, T. ; Williston, S. ; Brandts, J. F. &Lin, L.-N. (1989) Rapid measurement of binding constants and heats of binding using a new titration calorimeter. *Anal. Biochem.* 179, 131 – 137.

2

Theory and Application of Thermoelectrochemistry

Zheng Fang
Chemistry & Chemical Engineering College,
Central South University, Changsha
People's Republic of China

1. Introduction

Thermoelectrochemistry (TEC) is a subject that combines the theories and techniques of both thermo- and electro-chemistry to investigate the cell and electrode reactions [1]. That is, the parameters of thermodynamics [2-4] and kinetics [5] of the electrochemical reactions can be obtained by the simultaneous measurements and analysis of heat flow, electrode potential, electric current and time signals under the various conditions. Therefore, TEC can provide the availably and expansively additional information more for electrochemical reactions. It compensates the insufficiency for a single electrochemical study or a single thermochemical research to some extent. In earlier period, a lot of techniques and instruments used to research the heat effects of cell and half-cell was set up [6-30], such as thermoelectric power measurements [6,7], electrolytic calorimeter [8], controlled-potential and controlled-current polarizations [9], Kinetic method on the stationary heat effect [10], non-stationary temperature wave method [11], cyclic-voltammo-thermometry[12], Lumped-heat-capacity analysis [13], steady state electrolysis [14], differential voltammetric scanning thermometry [15], acoustic calorimetry[16], thermistor probe determination[17], potentiodynamic and galvanostatic transient techniques [18], non-isothermal cell [19], etc to obtain the electrochemical Peltier heat (EPH) of the electrode reactions.

In these researches, a mainly purpose is to acquire EPHs of cell or half-cell reactions. The EPH could be considered as a basic issue of TEC. Before the identification of this problem there had been two puzzled questions, one is that the heat effects for a reversible reaction, Q can be calculated by the formula $Q = T\Delta S$ where ΔS is the entropy change of this reaction and T temperature in Kelvin. However, this formula that is valid for most reactions is not viable at least for a reversible single electrode reaction in aqueous solution. For a reversible single electrode reaction, the experimental value of the heat effect is not in agreement with that calculated on the current thermodynamic databank of ions, that is, with which, the product of the calculated entropy change and the temperature of the electrode reaction always differs from the experimental measurements [2]. For example, for the electrode reaction at the standard state:

$$Cu^{2+}\ (aq., a_{Cu^{2+}}=1) + 2e^- = Cu\ (Pure\ crystal) \quad (1)$$

where a_{Cu2+} is activity of copper ions, and metal copper and its ions lie to the each standard state. Its change in entropy is calculated to be about $\Delta S^{\varnothing} = S^{\varnothing}_{Cu} - 2S_{e^-} - S^{\varnothing}_{Cu2+} = 2.2\ J{\cdot}K^{-1}{\cdot}mol^{-1}$. The heat effects, Q should be $T\Delta S^{\varnothing}$=0.65 kJ· mol^{-1} at 298.15K, but it was evaluated by an experimental as 52.8 kJ· mol^{-1}. The difference of both is bigger. Another problem is that there had been no workable method that could be used to calculate or predict the "real" heat effect of a standard reversible electrode reaction by means of the current thermodynamic knowledge. For example, we did not know how to get the value of heat effects, 52.8 kJ·mol^{-1} for reaction (1) except the experiment at least up to now. These two problems should be resolved in TEC discipline.

In order to identify EPHs of the cell or electrode reactions from the experimental information, there had been two principal approaches of treatments. One was based on the heat balance under the steady state or quasi-stationary conditions [6, 11, 31]. This treatment considered all heat effects including the characteristic Peltier heat and the heat dissipation due to polarization or irreversibility of electrode processes such as the so-call heats of transfer of ions and electron, the Joule heat, the heat conductivity and the convection. Another was to apply the irreversible thermodynamics and the Onsager's reciprocal relations [8, 32, 33], on which the heat flux due to temperature gradient, the component fluxes due to concentration gradient and the electric current density due to potential gradient and some active components' transfer are simply assumed to be directly proportional to these driving forces. Of course, there also were other methods, for instance, the numerical simulation with a finite element program for the complex heat and mass flow at the heated electrode was also used [34].

2. Electrochemical Peltier heat and the absolute scale

2.1 The electrochemical Peltier heat of cell reaction

The terminology of EPH originated from the thermoelectric phenomena in Physics. Dated back to more than 100 years ago, such as the Seebeck effect, the Peltier effect and the Thomson effect were successionally discovered. The Peltier heat was first found by the French physicist Peltier in 1834. The Peltier effect shows that the heat flow would be generated on the junction between two different metals in an electric current circumstance. The junction acts as a heat sink or as a heat source, which depends on the direction of the electric current. And the strength of the heat was found to be proportional to the current intensity. The Peltier effect can express as [35]

$$dQ(T)/dt = p_{I\text{-}II} i\pi \tag{2}$$

where i is electric current, $Q(T)$, Peltier heat dependent on temperature, T in Kelvin, t, time and π_{I-II}, the Peltier coefficient which, sometimes, is considered as the difference of the "heats of evaporation" of electrons in the dissimilar metals, I and II.

The Peltier effect is a reverse one of the Seebeck effect that was discovered by the German physicist Seebeck at earlier period (1822). Seebeck discovered that a potential difference will be resulted between two connection points in a loop composed of two dissimilar metals, if the two junctions are maintained at different temperatures. Thereafter, in 1854, the English physicist Lord Kelvin (W. Thomson) was to discover that a uniform conductor with electric current passing through will suck heat up from the surrounding when there has a temperature gradient in the conductor, which is called as the Thomson effect.

In 1877, Mill called the heat effects in the electrochemical reactions as the electrochemical Peltier heats. Later, Bouty and Jahn demonstrated experimentally the EPH effects. Subsequently, a few of experimental studies on the heat effects for the electrochemical cells had also been presented. However, the heat effects in physics are different from those in electrochemical reactions. No distinct definition for EPH was given in history, except it was defined as the heat effects observed when electric current passes through. Obviously, this definition is not precise.

Vetter has elegantly defined the EPH effect to be the heat arising out or the heat consumption in a reversible cell reaction[36], i.e.

$$-T\,(\partial\,(\Delta G)/\partial T)_P = z\,\mathbf{F}\,T\,(\partial\varepsilon_0/\partial T)_P \tag{3}$$

where ΔG is the free energy change of the reaction, z electron transfer number, $\mathbf{F}$ Faraday constant, ε_0 cell voltage, T absolute temperature and P pressure.

Two features of this definition are worth noting. One is that EPH is defined as the heat of a reversible reaction, which essentially eliminates the various uncertainties arising from the irreversible factors such as overvoltage, Joule heat, thermal conductivity, concentration gradient and forced transfer of various particles like ions and electrons in electrical field, and makes the physical quantity more definite and comparable. This indicates that EPH is a characteristic measure of a cell reaction, because the term $(\partial\,(\Delta G)/\partial T)_P$ is an amount independent on reaction process, and only related to changes in the function of state. That is to say, EPH is determined only by the initial and the final states of the substances taking part in the reaction that occurs on the electrode-electrolyte interfaces, although other heats due to irreversible factors are accompanied. EPH is, unlike the heat of dissipation (Joule heat and the heats due to irreversibility of electrode processes and transfer processes), one of the fundamental characteristics of the electrode process.

Another one is that the thermodynamic functions of the standard hydrogen electrode (SHE) are taken as the reference of ΔG and ε_0 in this definition. This reference is called the "conventional scale" where all thermodynamic functions of SHE reaction including its free energy change, enthalpy change and entropy change as well as its electrode potential are designated to zero at any temperature.

This definition, where ε_0 emerges, also reflects that EPH differs from the physically Peltier heat, although the terminology of "Peltier heat" originated from the physical phenomena. The physically Peltier heat mostly concerns with the electron transfer from an energy level (material I) to another level (material II), there being no electric potential difference and no change in composition of substances on the contact interfaces between two conductors, so that it is called as the difference of the "heats of evaporation" of electrons in the materials, I and II. But, EPH mostly concerns with the atom reorganization and the change of valence of the active element by redox of substances under the action of the cell voltage, ε_0. Compared with EPH, the physically Peltier heat is so small that it generally can be ignored. Like the contact point between copper and silver, the Peltier coefficient is about $100\text{x}10^{-6}$ volt [35] according to Fermi-Dirac statistics used to free electron gas in metal, while the electrochemical Peltier coefficient of the redox between copper and its ions is 0.274 volt [1].

2.2 The electrochemical Peltier heat of electrode reaction and the absolute scale

When applying the Vetter's definition to a reversible electrode (or half-cell) reaction, it is no longer able to use the conventional scale as the reference of the free energy change and the

electrode potential. Otherwise, for the SHE reaction itself, we will draw a conclusion that the heat effects of the reaction are always zero in all temperatures. Obviously this is not true, because even the SHE reaction, certainly there are the "old" chemical bond fracturing and a "new" chemical bond constructing process accompanied by the emergence of the heat effect [1]. Then where does the problem come from? Look at the SHE reaction:

$$H^+ (aq., a_{H^+}=1) + e^- = 1/2\ H_2 \text{ (hypothetical ideal gas at unit fugacity, Pt)} \quad (4)$$

In the conventional scale, the entropy of the hydrogen ion and the change in entropy of this reaction are all arbitrarily set at zero[37], which would result in a bigger difference between these quantities and the "real" values. Just this entropy that is arbitrarily specified to the hydrogen ion is taken as reference point of the other ion entropies once again. This will also make the calculated entropy change differ from the "real" value for other electrode reaction. Therefore, the heat effect calculated based on the change in entropy of the reaction must differ from that obtained experimentally. In this case, in order to make the calculated result much approximate to the experimental data, we should adopt a new reference scale, i.e. the "absolute scale" as the reference, in which the enthalpy change, the entropy change, the free energy and its standard electrode potential for the SHE reaction are not able to be arbitrarily specified as zero. Define the reversible electrode potential of any electrode, ϕ^*, in the absolute scale as follows [38]:

$$\phi^* = \phi\ (vs.\ \text{SHE}) + \phi^*(H^+/H_2), \quad (5)$$

where the amount marked with an asterisk is on the "absolute scale" (the same below), $\phi^*(H^+/H_2)$ the electrode potential of SHE on this scale and ϕ (*vs.* SHE) on the conventional scale. And both ϕ^* and $\phi^*(H^+/H_2)$ are designated to zero when $T \to 0$ [2].
On this scale, the entropy change for a single-electrode reaction, ΔS^*_T will be characterized as:

$$\Delta S^*_T = z\ \mathcal{F} (\partial \phi^*_T / \partial T)_P, \quad (6)$$

When integrating Eq. (6), the integral constant, ΔS^*_0 is appointed to zero at $T \to 0$ based on the third law of thermodynamics.
The resultant EPH of the electrochemical reaction, Π is defined as

$$\Pi = z\ \mathcal{F}\ T\ (\partial \phi^* / \partial T)_{P,\ i \to 0}, \quad (7)$$

or

$$\Pi = -T\ (\partial\ (\Delta G^*) / \partial T)_P = T(\Delta S^*)_{p,\ i \to 0}\ , \quad (8)$$

where $i \to 0$ indicates that the considered electrode reaction progresses in the reversible manner. It should be noted that in this scale, the electronic entropy is specified to be zero. It will be seen below that this is a designation close to truth.
The definition of EPH for the electrode reaction given by Eq. (7) or Eq. (8) is all similar to that of a cell reaction except on the absolute scale. These equations indicate that EPH of a half-cell, just like that of the cell reaction, is also a characteristic quantity that only relates to changes in the function of state, i.e. the entropies on the absolute scale, of substances taking part in the reaction. The heat effect occurs on the electrode-electrolyte interfaces. Evidently, when Eq. (7) or Eq. (8) is applied to a cell reaction, the terms, $\phi^*(H^+/H_2)$ in Eq. (5), common to both electrodes of the cell, does not appear explicitly because they are deleted ultimately.

The ϕ^*in Eq.(5) will be replaced by the corresponding function on the conventional scale; the expression is reduced to Vetter's definition. It is obvious that the definition based on Eq. (3) is a special case of Eq.(7) or Eq. (8) which could be more widely used.
The establishment of the absolute scale is dependent on determination of ΔS^*_T according to Eq.(6) which can be rewritten as

$$\phi^*_T = \int \Delta S^*_T \, d\,T \,/\, (z\,\mathcal{F}) \tag{9}$$

with the integral range from zero to a designated temperature T.
It should be pointed out that a specified reference is $\phi^*_T \to 0$ and $\Delta S^*_T \to 0$ when $T \to 0$ on the absolute scale. This merely is for expedience to identify other values at other than zero in Kelvin experimentally. According to Nernst, for an isothermal process of the condensed system, the third law of thermodynamics can be expressed as

$$\lim_{T\to 0} (\Delta S)_T = 0 \tag{10}$$

Although the formula is confirmed to be correct in many cases, for the electrode reaction where the hydrated ions and the electrons would take part in it, the validity is to be confirmed. Even, ϕ^*_T might have an own individual value for each electrode when $T \to 0$.

3. The basic equations for thermoelectrochemistry and experiments for determination of the entropy change of SHE on absolute scale

3.1 The basic equations for thermoelectrochemistry

It has been mentioned above that two methods, the heat balance under the steady state or quasi-stationary conditions, and the irreversible thermodynamics and Onsager's reciprocal relations, had been used to treat the heat effects in the electrochemical reactions. Although these methods can determine EPH of electrode reaction under some assumption, they are helpless to answer those problems presented in introduction.
Here a method based on the equilibrium thermodynamics will be introduced. In Eq.(5), a relationship between the electrode potentials on the absolute scale and the conventional scale is given. According to the relationship, the thermodynamic functions such as the entropy S, the free energy G and the enthalpy H, can be written as:

$$\Delta S^* = \Delta S + z\,\Delta S^*(H^+/H_2), \tag{11}$$

$$\Delta G^* = \Delta G + z\,\Delta G^*(H^+/H_2), \tag{12}$$

$$\Delta H^* = \Delta H + z\,\Delta H^*(H^+/H_2), \tag{13}$$

where $\Delta S^*(H^+/H_2)$, $\Delta G^*(H^+/H_2)$ and $\Delta H^*(H^+/H_2)$ are the corresponding thermodynamic functions of the SHE reaction with a single electron transfer on the absolute scale, z is the electron transfer number of the considered electrode reaction.
For a reversible electrode reaction, on the absolute scale, we still have the following relationship

$$\Delta G^* = \Delta H^* - T\Delta S^*, \tag{14}$$

Combined with the equations (8), (12), (13) and (14), and noting that $\Delta G = -W_e$ (*vs.* SHE) = $-z\,\mathcal{F}\phi$, we have

$$\Pi\text{-}W_e\,(vs.\mathrm{SHE}) = \Delta H^{\square}, \tag{15}$$

or

$$\Pi\text{-}z\,\mathcal{F}\phi\,(vs.\mathrm{SHE}) = \Delta H^{\square}, \tag{16}$$

where W_e is the reversible electric work done on the conventional scale, $\Delta H^{\square} = \Delta H + z\,T\Delta S^*(H^+/H_2)$, which is called as the apparent enthalpy change. It should be noted that $\Delta S^*(H^+/H_2)$ at a given temperature is a constant.

According to Faraday law, for more than or less than one mole change, Eq. (8) and (15) can be, respectively, rewritten as

$$\Pi = T\Delta S^*(\int i\,dt/(z\,\mathcal{F}))_{i\to 0} \tag{17}$$

$$\Pi\text{-}W_e\,(vs.\ \mathrm{SHE}) = \Delta H^{\square}\int i\,dt/\ (z\,\mathcal{F}), \tag{18}$$

where i is electric current, t time and the integral range from zero to t. Eq. (17) and (18) are the fundamental equations for the electrode reaction.

When a small electric current passes through, Eq. (17) can be approximately written, as

$$Q = T\Delta S^*\int i\,dt/\ (z\,\mathcal{F}) \tag{19}$$

Being differential on both sides of Eq. (19) and letting $i\to 0$, we get

$$\Delta S^* = \mathcal{k}\,(q/i)_{\,i\to 0} \tag{20}$$

where $\mathcal{k} = z\,\mathcal{F}/T$ and $q = dQ/dt$ which is called as the heat flow. The ratio of the heat flow to the electric current at $i\to 0$, $(q/i)_{\,i\to 0}$ is called as the Peltier coefficient and represented by the symbol, π. The definition of the Peltier coefficient is completely similar to one in physics in format. The Peltier heat effect can be obtained by the product of the electric quantity and the Peltier coefficient, *i.e.* $z\,\mathcal{F}\,(q/i)_{\,i\to 0} = z\,\mathcal{F}\pi$. The formula (20) is also a fundamental equation to obtain the entropy change on the absolute scale and the Peltier heat for an electrode reaction.

3.2 An experiment for evaluation of the entropy change of SHE on the absolute scale

As shown by Eq. (9), if ΔS^*_T at various temperatures are experimentally obtained, ϕ^*_T at a designated temperature T can be determined. The problem is how to obtain ΔS^*_T experimentally at any temperature except zero in Kelvin. According to Eq. (11), if $\Delta S^*(H^+/H_2)$ can be determined experimentally, ΔS^*_T, subsequently, can be obtained. Take the evaluation of $\Delta S^*(H^+/H_2)$ at 298.15K as an example. A work has done on the thermoelectrochemical experiment on the ferro-ferricyanide couple and from this the result of the entropy change of SHE on the absolute scale has been evaluated.

The heat effects were determined by temperature-rise calorimetry. The experimental setup is shown by Fig. 1.

A three-electrode system with a platinum working (**H**), a platinum auxiliary (**K**) and a reference saturated calomel electrode (**F**) was placed in a U-type compartment (**M**) in an isothermal water bath (**E**). Two platinum wires jointing with the working and auxiliary electrodes, respectively, straight reached to electrolyte outside and connected to two copper wires which linked with the electrochemical workstation (**B**). The reference electrode was used to track the change in potential of electrode (**H**). A thermo-sensitive resistor (**L**)

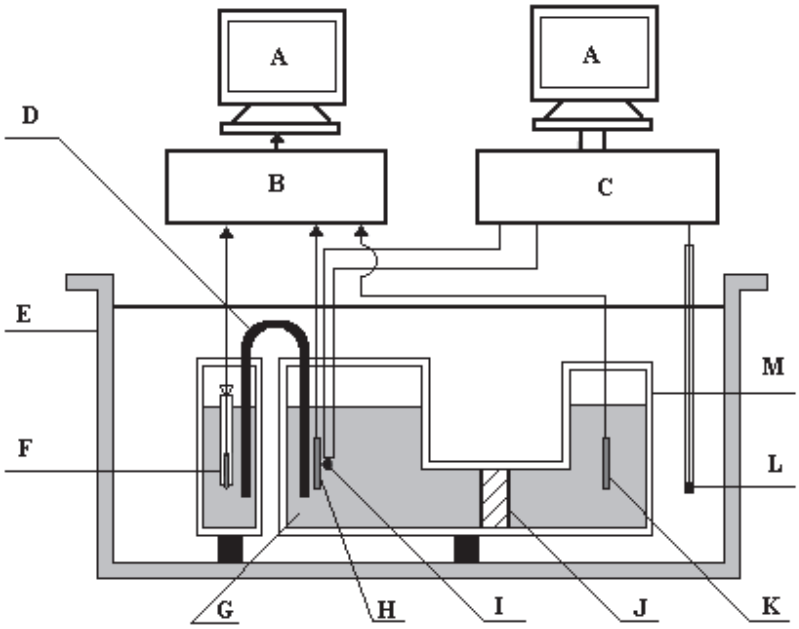


Fig. 1. Schematic diagram of experimental apparatus for measuring $\Delta S^*(H^+/H^2)$ (A) computer, (B) electrochemical workstation, (C) solution-reaction isoperibol calorimeter, (D) salt bridge, (E) isothermal water bath, (F) saturated calomel electrode, (G) electrolyte, (H) working electrode, (I) thermo-sensitive resistor, (J) porous ceramics, (K) auxiliary electrode, (L) thermo-sensitive resistor, and (M) U-type compartment.

controlled the isothermal surrounding with 0.001 K fluctuations. Another thermo-sensitive resistor (**I**) was cemented to the back of working electrode (**H**) immersed in electrolyte (**G**). A precision resistor inside the solution-reaction calorimeter (**C**) with ±0.001K was the reference. The electric bridge of the calorimeter measured change in temperature of the working electrode. At the beginning of every run, the electric bridge was balanced by adjusting the inner reference resistor. The potential signals for temperature of the working electrode (**H**) were output when a small electric current passed through the loop consisting of electrodes (**H**) and (**K**). The relation between output potentials and temperature change was predetermined by a calibration experiment. The potential signals of temperature change, electric currents and electrode potentials as a function of time were simultaneously collected by calorimeter (**C**) and the electrochemical workstation (**B**), and were recorded by means of computers (**A**). The temperature change was used for determining Q and the electric currents and the potential for We'.

The $[Fe(CN)_6]^{-3/-4}$+1mol.dm^{-3} KCl solutions were prepared with equal molar concentrations of the two negative ions being 0.075 mol.dm^{-3}, 0.15 mol.dm^{-3}, 0.2 mol.dm^{-3}, 0.25 mol.dm^{-3}and 0.3 mol.dm^{-3}, respectively. The experiments were done at 298.15K. The experimental data are in table 1, and the typical experimental curves for electrode potentials against time and the potential signals for temperature difference against time for the 0.2 mol.dm^{-3} $[Fe(CN)_6]^{-3/-4}$ system are shown in Fig. 2.

c /mol.dm^{-3}	i /mA	$\int_0^t idt$ /A.S	$\frac{\int_0^t (\Delta V+6.9\times10^{-6})^{*} dt}{\int_0^t idt}$ /V.A^{-1}	$\frac{\int_0^t i(\phi+\eta)(\mathrm{SHE})\,dt}{\int_0^t idt}$ /V	$-k$ /A
0.075	$i\to0$		0.732	0.484	0.566
	0.5	0.060	0.761	0.470	
	0.8	0.096	0.763	0.463	
	1.0	0.120	0.770	0.459	
	1.3	0.156	0.775	0.451	
	1.5	0.180	0.788	0.448	
	1.7	0.204	0.812	0.443	
	2.3	0.276	0.829	0.423	
0.15	$i\to0$		0.675	0.489	0.715
	1.0	0.12	0.689	0.471	
	1.5	0.18	0.716	0.466	
	2.5	0.30	0.726	0.451	
	3.5	0.42	0.747	0442	
	4.0	0.48	0.748	0.436	
	4.5	0.54	0.764	0.421	
	5.0	0.60	0.785	0.406	
0.20	$i\to0$		0.698	0.494	0.750
	1.0	0.12	0.707	0.481	
	2.0	0.24	0.733	0.471	
	2.5	0.30	0.735	0.467	
	3.5	0.42	0.750	0.457	
	4.0	0.48	0.752	0.453	
	5.0	0.60	0.760	0.443	
	6.0	0.72	0.778	0.430	
	7.0	0.84	0.803	0.414	
0.25	$i\to0$		0.724	0.491	0.783
	1.0	0.12	0.734	0.484	
	1.5	0.18	0.741	0.480	
	2.5	0.30	0.743	0.473	
	3.0	0.36	0.746	0.470	
	3.5	0.42	0.753	0.466	
	4.0	0.48	0.763	0.462	
	4.5	0.54	0.760	0.458	
	5.0	0.60	0.771	0.454	
	5.5	0.66	0.775	0.451	
0.30	$i\to0$		0.814	0.497	0.751
	1.0	0.12	0.826	0.486	
	2.0	0.24	0.829	0.483	
	3.0	0.36	0.845	0.474	
	4.0	0.48	0.858	0.467	
	4.5	0.54	0.864	0.464	
	6.0	0.72	0.871	0.451	
	8.0	0.96	0.895	0.433	

* : 6.9×10−6 (V) originated from experimental calibration ΔV(mV) = 170.199 ΔT(K) -0.0069. From Ref. [2]

Table 1. The experimental data and results for $[Fe(CN)_6]^{-3/-4}$ system.

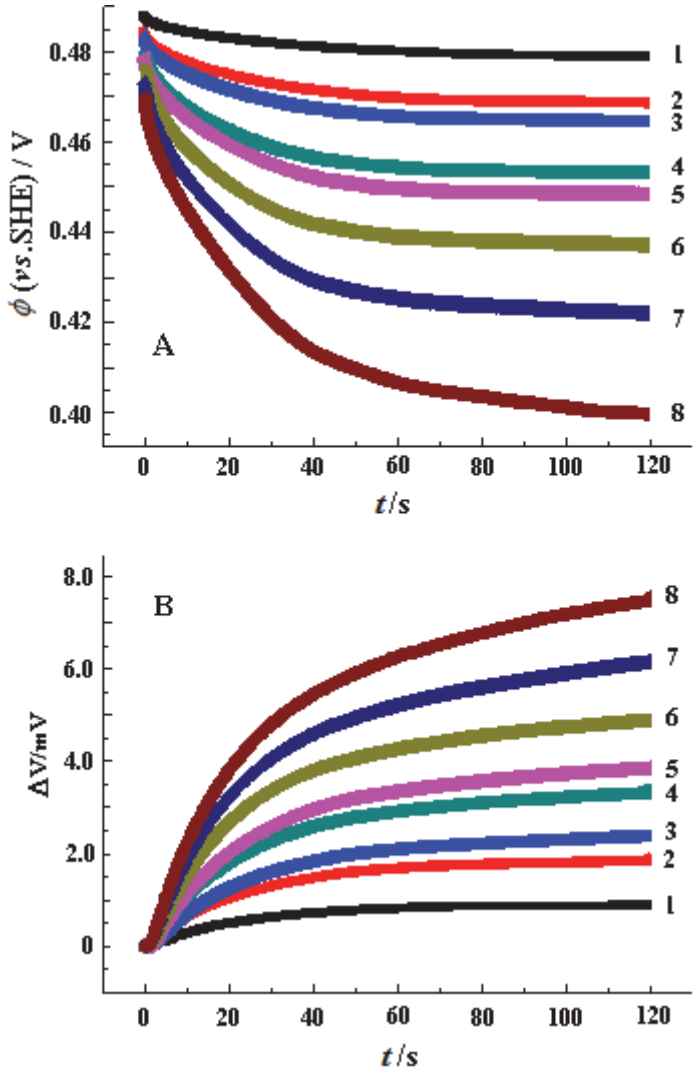


Fig. 2. Plots of the electrode potentials and the potential signals for temperature difference against time for the 0.2 mol.dm^{-3} $[Fe(CN)_6]^{-3/-4}$ system at various electric currents, (A) the electrode potentials *vs.* time, (B) the potential signals for temperature difference *vs.* time; the curves from No.1 to No.8 correspond to currents 1.0, 2.0, 2.5, 3.5, 4.0, 5.0, 6.0 and 7.0 mA, respectively

3.3 Experimental data-processing and evaluation of entropy change of SHE on the absolute scale

The redox equation considered in this experiment is as follows:

$$Fe(CN)_6^{-3} + e^- = Fe(CN)_6^{-4} \tag{21}$$

When a small electric current passes through, Eq.(18) can be approximately written as

$$Q - W_e' \ (vs.\ \text{SHE}) = \Delta H^{\square\prime} \int i dt / \ (z\ \mathcal{F}), \tag{22}$$

where Q denotes the observed heats including all heat dissipation due to the various polarizations, W_e' $(vs$ .SHE) = $\int i\ (\phi + \eta)$ $(vs.$ SHE) dt with equilibrium potential ϕ and overpotential η, and $\Delta H^{\square\prime}$, a quantity corresponding to $\Delta H^{\square}$ but includes the change due to irreversible effects. Apparently, $Q \rightarrow \Pi$, $W_e' \rightarrow W_e$ (or $\eta \rightarrow 0$) and $\Delta H^{\square\prime} \rightarrow \Delta H^{\square}$ and Eq. (22) reduces to Eq.(18) when $i \rightarrow 0$. The both sides of Eq.(22) being simultaneously divided by the term, $\int i dt / (z\mathcal{F})$ gives

$$Q / \ (\int i dt / \ (z\ \mathcal{F})) - W_e' (vs.\ \text{SHE}) / (\int i dt / (z\ \mathcal{F})) = \Delta H^{\square\prime}, \tag{23}$$

Noting that $Q/(\int i dt/(z\ \mathcal{F})) \rightarrow \Pi$ and $W_e'(vs$ .SHE)$/(\int i dt/(z\ \mathcal{F})) \rightarrow W_e$ $(vs.$ SHE) at $i \rightarrow 0$, and extrapolating the electric current to zero, Π can be determined from a series values of $Q/(\int i dt/(z\ \mathcal{F}))$ at different currents. Similarly, W_e $(vs.$ SHE) can be also obtained. Therefore, according to Eq. (15) $\Delta H^{\square}$ is identified. The $\Delta H^{\square}$ obtained by the extrapolation treatment is merely at a given concentration of active electrolyte, and denoted by $\Delta H^{\square}(c)$. When further extrapolating the concentration to zero, we will acquire $\Delta H^{\square}$ at the infinite dilution, i.e. $\Delta H^{\square}(c \rightarrow 0)$. Based on the definition of the apparent enthalpy change, we have

$$\Delta H^{\square}(c \rightarrow 0) = \Delta H(c \rightarrow 0) + z\, T\Delta S^{*}(H^{+}/H_2), \tag{24}$$

Considering a thermodynamic principle which holds that the enthalpy of the solute at unlimited dilution is always equal to that at the standard state, when the standard state is designated to a hypothetical solution which obeys Henry's law at unit molal concentration of solute [39], we can write

$$\Delta H\ (c \rightarrow 0) = \Delta H^{\varnothing} \tag{25}$$

where $\Delta H^{\varnothing}$ is the standard enthalpy change of the considered electrode reaction. Consequently, Eq.(24) can be rewritten as

$$\Delta H^{\square}(c \rightarrow 0) = \Delta H^{\varnothing} + z\, T\Delta S^{*}(H^{+}/H_2), \tag{26}$$

The values of $\Delta H^{\square}$ (c) for $[Fe(CN)_6]^{-3/-4}$ redox system at various concentrations have been calculated from the experimental data, as shown in table 2.

c**/mol.dm^{-3}**	0.075	0.15	0.20	0.25	0.30
$-\Delta H^{\square}(c)$**/ kJ. mol^{-1}**	86.67	93.75	98.17	102.07	106.96
$-\Pi(c)$**/ kJ. mol^{-1}**	39.97	46.57	50.51	54.70	58.98

Table 2. Some thermodynamic functions for $[Fe(CN)6]^{-3/-4}$ system. From Ref. [2].

From table 2, plot $\Delta H^{\square}$ (c) against concentration, as shown in Fig.3 where it can be seen that $\Delta H^{\square}(c)$ is good linearly related to concentration. $\Delta H^{\square}$ $(c \rightarrow 0)$ can be obtained to be -80.16 kJ.mol^{-1} by the linear extrapolation of concentration to zero. From the literature, $\Delta H^{\varnothing} = \Delta H^{\varnothing}(Fe(CN)_6^{-4}) - \Delta H^{\varnothing}(Fe(CN)_6^{-3}) =$ -106.3kJ.mol^{-1} can be obtained, then the entropy change of SHE reaction on the absolute scale, $\Delta S^{*}(H^{+}/H_2)$ is calculated as 87.6±1.0 J.K^{-1}.mol^{-1} at 298.15K according to Eq.(26).

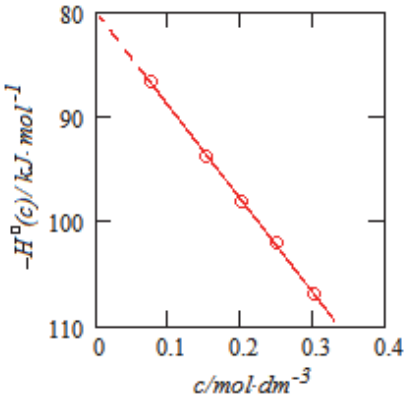


Fig. 3. Plot of $\Delta H^{\square}(c)$ *vs*. concentration for the $[Fe(CN)_6]^{-3}/^{-4}$ system.

3.4 Relationship between the quantities determined by TEC technology and those calculated with the current thermodynamic databank of ions

A mathematical expression of the first law of thermodynamics for an infinitesimal process is as follows

$$d\,U = \delta\,Q - \delta\,W \tag{27}$$

where $d\,U$ is the change in thermodynamic energy, $\delta\,Q$ and $\delta\,W$ are, respectively, the heat transferred and work done to the system which includes the expansion work done due to volume change of the system and the other work done except expansion due to the change in surface area, charge in electrochemical cell, magnetic moment in the magnetic field, etc. Considering a system where there are only the electric work done and the expansion work at constant pressure, Eq. (27) can be written as

$$\delta Q = dU + \delta W_e + p\,dV \tag{28}$$

or

$$\delta\,Q = d\,H + \delta\,W_e \tag{29}$$

where H is the enthalpy function, p, pressure and V, volume. For a system where a certain amount of change occurs, Eq. (29) can be written as

$$Q\text{-}W_e\,(vs\ .\mathrm{SHE}) = \Delta H \tag{30}$$

where the electric work done is appointed to the reference of SHE. This is an expression of the first law of thermodynamics.

When the law is used to an electrode reaction with the given electron transfer number at the given temperature, the enthalpy change calculated based on the electric work done and the heat effect obtained experimentally differs from that calculated by Eq.(30) on the current thermodynamic databank including the ion data. The difference between them is almost a constant. For this phenomenon it has been not yet explained reasonably so far, and this greatly influenced the development of thermoelectrochemistry.

It should be seen from Eq.(15) that the value on TEC experiments is the apparent enthalpy change, while the value calculated by Eq.(30) on the current thermodynamic databank is the

enthalpy change of the electrode reaction on the conventional scale. They just differ by a constant that is z $T\Delta S^*(H^+/H_2)$. If adding z $T\Delta S^*(H^+/H_2)$ to both sides of Eq. (30) at same time and noting that $Q = T\Delta S$ and $\Pi = T\Delta S^*$, we will get Eq. (15). The difference between the two terms about the enthalpy changes, $\Delta H^\square$ and ΔH, in Eq. (15) and Eq.(30) is just a constant that is EPH of the SHE reaction at temperature T, z $T\Delta S^*(H^+/H_2)$.
In Eq. (30), Q is a product of temperature T and the entropy change derived from the current thermodynamic databank including the ion data which is constructed on the conventional scale, and can be named as "the traditional heat effect"; while in Eq. (15), Π is the heat effect identified by the experiments, called as "the measured heat effect". The difference between them is z $T\Delta S^*(H^+/H_2)$. Consequently, the first problem mentioned above, why is this formula, $Q = T\Delta S$, unsuitable for a reversible single electrode reaction, is answered.

3.5 About the absolute value of thermodynamic function on the absolute scale

When these thermodynamic functions on the absolute scale are applied to a cell reaction, all terms concerning with SHE on the absolute scale, such as $\Delta S^*(H^+/H_2)$, $\Delta G^*(H^+/H_2)$ and $\Delta H^*(H^+/H_2)$ in Eq.(11), (12) and (13), common to both electrodes of the cell, are cancelled ultimately. It will be not any impact to the normal thermodynamic analysis of the cell reaction.
Here an issue is whether or not the potential of single electrode on the absolute scale is a "real" value? It should be said that the potential of electrode is not its "real" value in absolute sense, and the thermodynamic functions on the absolute scale like ΔS^*, ΔG^* and ΔH^* of an electrode reaction are not the "real" values either, although the words, "real" and "absolute", are used to this text. The scale is called as "absolute scale" that is only compared with "conventional scale". They are also some relative quantities obtained based on this reference, i.e. $\phi^*_T \to 0$, $\Delta S^*_T \to 0$ when $T \to 0$, and the assumption of zero electronic entropy [40, 41]. It, called as "absolute scale", is merely that these values on the scale much more close to "reality" than on the conventional scale. In the conventional scale, the entropy of the hydrogen ion and the entropy change of the SHE reaction are all arbitrarily set at zero, which would result in a bigger difference between these quantities and the "verity". However, based on the "absolute scale", the partial molal entropy of the hydrogen ion is calculated as -22.3 $J.mol^{-1}.K^{-1}$ (see below), the entropy change of SHE reaction (Eq. (4)), 87.6 $J.mol^{-1}.K^{-1}$ at 298.15K, and the electronic entropy is specified as zero. All of these quantities are much more close to "verity" than those on the "conventional scale". In the conventional scale the electronic entropy is calculated as 65.29.$mol^{-1}.K^{-1}$ [41] according to a principle which states that the algebraic sum of the entropies of all substances taking part in the reaction (4) equals to zero. This result, obviously, is far from the "really" electronic entropy that can approximately be considered to be zero according to Fermi-Dirac statistics theory applied to the electron gas in metal. Therefore, we call the scale that is more close to the "verity" as the "absolute scale".

4. Some application of the basic equations for thermoelectrochemistry

4.1 The partial molal entropies of hydrogen ion and some other hydrated ions on the absolute scale

According to Eq.(4),

$$\Delta S^*(H^+/H_2) = S^\varnothing(1/2H_2) - \tilde{S}^*(e^-) - \tilde{S}^*(H^+) \qquad (31)$$

where $\widetilde{S}^*_{(e^-)}$ and $\widetilde{S}^*_{(H^+)}$ are the partial molal entropy of electron and hydrogen ion on the absolute scale, respectively. Substituting the corresponding data, 87.6, 65.29 and 0 for $\Delta S^*(H^+/H_2)$, $S^{\varnothing}(1/2H_2)$ and $\widetilde{S}^*_{(e^-)}$, respectively, $\widetilde{S}^*_{(H^+)}$ is calculated to be -22.3 J.K^{-1}.mol^{-1} which is very good in agreement with the value reported previously on some other experimental methods. Therefore, the entropies of other hydrated ions on the conventional scale can be recalculated to be those on the absolute scale with $\widetilde{S}^*_{(H^+)}$. The formula used for conversion is as follows:

$$\widetilde{S}^*_{(N^z)} = \widetilde{S}_{(N^z)} - 22.3.z \tag{32}$$

where N^z represents ion with the z valence of positive (z >0) or negative (z<0) ion. Table 3 gives the entropies for some hydrated ions on absolute scale based on the data for ion entropies from Ref. [42].

Ion	$\widetilde{S}_i$	$\widetilde{S}_i^*$	Ion	$\widetilde{S}_i$	$\widetilde{S}_i^*$	Ion	$\widetilde{S}_i$	$\widetilde{S}_i^*$
H^+	0	-22.3	Be^{2+}	-129.7	-174.3	Al^{3+}	-325.0	-391.9
Li^+	12.24	-10.0	Mg^{2+}	-137.4	-182.0	Fe^{3+}	-315.9	-382.8
Na^+	58.45	36.2	Ca^{2+}	-56.2	-100.8	S_2^{2-}	-14.6	30.0
K^+	101.2	78.9	Sr^{2+}	-32.6	-77.2	HS^-	67.0	89.3
Rb^+	121.75	99.5	Ba^{2+}	9.6	-35.0	F^-	-13.8	8.5
Cs^+	132.1	109.8	Fe^{2+}	-137.7	-182.3	Cl^-	56.6	78.9
Ag^+	72.45	50.2	Co^{2+}	-113.0	-157.6	Br^-	82.55	104.8
Cu^+	40.6	18.3	Ni^{2+}	-128.9	-173.5	I^-	106.45	128.7

Table 3. The partial molal entropies for some hydrated ions on the conventional scale and on absolute scale (J.K^{-1}.mol^{-1}).

4.2 The electrochemical Peltier heats and the Peltier coefficients of the standard electrode reactions

Similar to the $[Fe(CN)_6]^{-3/-4}$ redox couple, EPH of an electrode reaction can directly be measured by the thermoelectrochemical experiments. However, it is hard to measure EPH of an electrode reaction at the standard state directly, because the standard state chosen usually in thermodynamics is even physically unrealizable in most cases. According to Eq.(8), EPH of a standard electrode reaction can be determined provided that ΔS^* of the reaction is known. For example, for the $[Fe(CN)_6]^{-3/-4}$ couple at each standard state of the components, its entropy change is calculated as $\Delta S^{\varnothing}([Fe(CN)_6]^{-3/-4})$ = -240.6 J.K^{-1}.mol^{-1}, therefore, $\Delta S^*([Fe(CN)_6]^{-3/-4})$ = -153.0 J.K^{-1}.mol^{-1}, and $\Pi^{\varnothing}([Fe(CN)_6]^{-3/-4})$ = -45.6 kJ.mol^{-1}. The EPHs of some standard electrode reactions at 298.15K are given in table 4. Similarly, the electrochemical Peltier coefficient, a characteristic quantity of electrode reaction, can be also determined by Eq.(20).

Like this, the second problem mentioned in introduction, that is, how to seek a feasible method to calculate or predict the "real" heat effect of a standard reversible electrode reaction, is also resolved.

Reaction	ΔS^* (J.mol^{-1}.K^{-1})	π (V)	$\Pi^{\varnothing}$ (kJ.mo l^{-1})	Reaction	ΔS^* (J.mol^{-1}.K^{-1})	π (V)	$\Pi^{\varnothing}$ ((kJ.mol^{-1})
$H^+ + e^- = 0.5H_2$	87.6 ±1.0	0.27_1	26.1±0.3	$Am^{3+}+3e^-=Am$	288.6±3.0	0.29_7	86.1±0.9
$Ag^+ + e^- = Ag$	-8.6±1.0	-0.02_7	-2.6±0.3	$Ce^{4+}+4e^-=Ce$	462.0±4.0	0.35_7	137.8±1.2
$Cu^+ + e^-=Cu$	14.9±1.0	0.04_6	4.4±0.3	$S(orth)+2e^-=S^{2-}$	-1.5±2.0	-0.00_2	-0.5±0.6
$Be^{2+}+2e^-=Be$	183.8±2.0	0.28_4	54.8±0.6	$Cl_2+2e^-=2Cl^-$	-75.3±2.0	-0.11_6	-22.44±0.6
$Mg^{2+}+2e^-=Mg$	214.3±2.0	0.33_1	63.9±0.6	$Br_2(l)+2e^-=2Br^-$	57.5±2.0	0.08_9	17.2±0.6
$Ca^{2+}+2e^-=Ca$	142.4±2.0	0.22_0	42.5±0.6	$I_2(c)+2e^-=2I^-$	141.4±2.0	0.21_8	42.2±0.6
$Sr^{2+}+2e^-= Sr$	132.2±2.0	0.20_4	39.4±0.6	$Fe(CN)_6^{-3}+e^-= Fe(CN)_6^{-4}$	-153.0±1.0	-0.47_3	-45.6±0.3
$Ba^{2+}+2e^-=Ba$	97.5±2.0	0.15_1	29.1±0.6	$Co^{3+}+e^-=Co^{2+}$	214.3±1.0	0.66_2	63.9±0.3
$Ra^{2+}+2e^-=Ra$	61.6±2.0	0.09_5	18.4±0.6	$Fe^{3+}+e^-=Fe^{2+}$	200.5±1.0	0.62_0	59.8±0.3
$Mn^{2+}+2e^-= Mn$	150.2±2.0	0.23_2	44.8±0.6	$MnO_4^-+e^-=MnO_4^{2-}$	-109.9±1.0	-0.34_0	-32.8±0.3
$Co^{2+}+2e^-=Co$	187.6±2.0	0.29_0	55.9±0.6	$CuCl+e^-=Cu+Cl^-$	25.9±1.0	0.08_0	7.7±0.3
$Ni^{2+}+2e^-=Ni$	203.4±2.0	0.31_4	60.7±0.6	$AgCl+e^-=Ag+Cl^-$	25.2±1.0	0.07_8	7.5±0.3
$Cu^{2+}+2e^-=Cu$	177.4±2.0	0.27_4	52.9±0.6	$AgBr+e^-=Ag+Br^-$	40.3±1.0	0.12_5	12.0±0.3
$Al^{3+}+3e^-=Al$	420.3±3.0	0.43_3	125.3±0.9	$AgI+ e^-=Ag+I^-$	55.8±1.0	0.17_2	16.6±0.3
$Co^{3+}+3e^-=Co$	401.9±3.0	0.41_4	119.8±0.9	$AuCl+e^-=Au+Cl^-$	33.4±1.0	0.10_3	9.9±0.3
$Sc^{3+}+3e^-=Sc$	356.6±3.0	0.36_7	106.3±0.9	$Au(CN)_2^-+e^-=Au+2CN^-$	85.9±1.0	0.26_5	25.6±0.3
$Y^{3+}+3e^- =Y$	362.3±3.0	0.37_3	108.0±0.9	$AuCl_4^{2-}+3e^-=Au+4Cl^-$	73.8±3.0	0.07_6	22.0±0.9
$La^{3+}+3e^-=La$	341.4±3.0	0.35_2	101.8±0.9	$PtCl_4^{2-}+2e^-=Pt+4Cl^-$	141.9±2.0	0.21_9	42.3±0.6
$Ce^{3+}+3e^-=Ce$	343.9±3.0	0.35_4	102.5±0.6	$Hg_2Cl_2+2e^-=2Hg(l)+2Cl^-$	118.0±2.0	0.18_2	35.2±0.6
$Nd^{3+}+3e^-=Nd$	345.2±3.0	0.35_6	102.9±0.9	$Zn(OH)_2(\beta)+2e^-=Zn+2OH^-$	-16.8±2.0	-0.02_6	-5.0±0.6
$Sm^{3+}+3e^-=Sm$	348.2±3.0	0.35_9	103.8±0.9	$Cd(CN)_4^{2-}+2e^-=Cd(\gamma)+4CN^-$	150.7±2.0	0.23_3	44.9±0.6
$Eu^{3+}+3e^-=Eu$	366.7±3.0	0.37_8	109.3±0.9	$AsO_2^-+2H_2O+3e^-=As(\alpha)+4OH^-$	-122.1±3.0	-0.12_6	-36.4±0.9
$Gd^{3+}+3e^-=Gd$	340.9±3.0	0.35_1	101.6±0.9	$AsO_4^{3-}+2H_2O+2e^-=AsO_2^-+4OH^-$	64.5±2.0	0.10_0	19.2±0.6
$Tb^{3+}+3e^- =Tb$	366.2±3.0	0.37_7	109.2±0.9	$Ba(OH)_2.8H_2O+2e^-=Ba+8H_2O+2OH^-$	217.9±2.0	0.33_7	65.0±0.6
$Dy^{3+}+3e^-=Dy$	373.5±3.0	0.38_5	111.4±0.9	$S(orth)+2H^++2e^-=H_2S(g)$	218.4±2.0	0.33_7	65.1±0.6
$Ho^{3+}+3e^-=Ho$	369.0±3.0	0.38_0	110.0±0.9	$H_3BO_3(aq)+3H^++3e^-=B+3H_2O$	120.3±3.0	0.12_4	35.9±0.9
$Er^{3+}+3e^-=Er$	384.4±3.0	0.39_6	114.6±0.9	$WO_3+6H^++6e^-=W+3H_2O$	300.4±6.0	0.15_5	89.6±1.8
$Tm^{3+}+3e^-=Tm$	383.9±3.0	0.39_5	114.5±0.9	$Al(OH)_3+3e^-=Al+3OH^-$	-8.5±3.0	-0.00_9	-2.5±0.9
$Yb^{3+}+3e^-=Yb$	364.8±3.0	0.37_6	108.8±0.9	$O_2+2H_2O(l)+4e^-=4OH^-$	-299.4±4.0	-0.23_1	-89.3±1.2
$Lu^{3+}+3e^-=Lu$	381.9±3.0	0.39_3	113.9±0.9	$O_2+4H^++4e^-=2H_2O$	24.0±4.0	0.01_9	7.2±1.2

Table 4. The entropy change on absolute scale, EPHs and Peltier coefficients for some standard electrode reactions in aqueous solution at 298.15K. From Ref. [1], but recalculating a part of data according to Ref. [42]

4.3 Determination of the electric potentials of the standard electrode by $\Delta H^{\square}$ (c→0)

The standard electrode potential of an electrode is a very important electrochemical quantity. Conventionally, it is determined by the extrapolating the electrode potentials of extremely dilute solution along the line predicted by the Debye-Hückel theory. Nevertheless, in the thermoelectrochemistry, it can be obtained by the measurement of the apparent enthalpy change. Based on a thermodynamic principle mentioned above (see Eq. (25)), Eq.(16) can be rewritten as

$$\Pi^{\varnothing} - z\,\mathcal{F}\phi^{\varnothing}(vs.\ \mathrm{SHE}) = \Delta H^{\square}(c\rightarrow 0) \tag{33}$$

where the superscript "∅" represents the standard state. Therefore, if $\Delta H^{\square}(c\rightarrow 0)$ and $\Pi^{\varnothing}$ are identified, then the standard electrode potential, $\phi^{\varnothing}(vs.\mathrm{SHE})$ can be determined.

For the $[Fe(CN)_6]^{-3}/^{-4}$ couple, when $\Pi^{\varnothing}$= -45.6 kJ.mol^{-1}(see section 4.2) and $\Delta H^{\square}(c\rightarrow 0)$ = -80.2kJ.mol^{-1} (see section 3.3) are substituted to Eq.(33), its standard potential is identified to be +0.3580 volt at 298.15K.

4.4 The enthalpy change and entropy change at designated concentration *c*, and diluted heats and diluted entropies of ions

The enthalpy change, $\Delta H\,(c_2\rightarrow c_1)$ of an electrode reaction from a designated concentration c_2 to another concentration c_1, can be evaluated from the measurable quantity $\Delta H^{\square}(c)$. Based on the definition of $\Delta H^{\square}$ (=ΔH + z $T\Delta S^*(H^+/H_2)$), the dilute enthalpy from concentrations c_2 to c_1 can be computed by the following cycle (Fig. 4) [43] where Me^{n+} is a specified metallic ion.

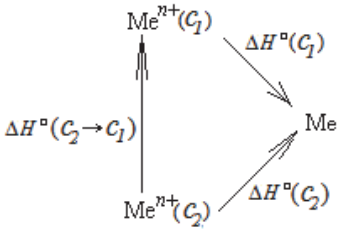


Fig. 4. The thermal cycle for getting dilute enthalpy from concentration c_2 to c_1

According to the thermal cycle, the corresponding computing equations are following

$$\Delta H^{\square}(c_2\rightarrow c_1) = \Delta H^{\square}(c_2) - \Delta H^{\square}(c_1) \tag{34}$$

Noting that $\Delta S^*(H^+/H_2)$ in the expression of $\Delta H^{\square}$ at a designated temperature is constant, the dilute enthalpy, $\Delta H\,(c_2\rightarrow c_1)$ is

$$\Delta H\,(c_2\rightarrow c_1) = \Delta H^{\square}(c_2)-\Delta H^{\square}(c_1)$$

$$= \Delta H(c_2)-\Delta H(c_1) \tag{35}$$

where $\Delta H\,(c)$ is the enthalpy change of the considered electrode reaction at the electrolyte concentration c. For example, a thermoelectrochemical experiment similar to the $[Fe(CN)_6]^{-3}/^{-4}$ couple was done on the Cu^{2+}/Cu couple in the $CuCl_2$+1mol.dm^{-3}KCl solution. The experimental data for the apparent enthalpies of the Cu^{2+}/Cu couple are shown in table 5. The dilute enthalpies of Cu^{2+} in the solution are calculated from table 5 , and listed in table 6.

c /mol.dm^{-3}	0.01	0.02	0.03	0.04	0.05
$-\Delta H^{\square}(c)$ / kJ. mol^{-1}	11.27	9.52	8.25	6.46	5.23

Table 5. Some thermodynamic functions for Cu^{2+}/Cu system. From Z.Yang, *an unpublished work*.

c /mol.dm^{-3}	0.05	0.04	0.03	0.02	0.01
0.05	0	1.23±0.28	3.02±0.28	4.29±0.28	6.04±0.28

Table 6. Dilute enthalpies $\Delta H\,(c)$ (kJ. mol^{-1}) for Cu^{2+} in $CuCl_2$+1 mol.dm^{-3} KCl electrolyte solution.

Similarly, the dilute entropy, $\Delta S(c_2 \rightarrow c_1)$ of the electrolyte solution from concentration c_2 to c_1 at the designated temperature can be evaluated from the measurable quantity $\Delta \Pi$ (c) according to Eq.(36),

$$\Delta S\ (c_2 \rightarrow c_1) = \Delta S\ (c_2) - \Delta S\ (c_1)$$

$$= (\Pi\ (c_2) - \Pi\ (c_1))/T \tag{36}$$

In a word, because the apparent enthalpy, ΔH □ and the Peltier heat, Π can be measured by the thermoelectrochemical experiments as a function of the function of state, all thermodynamic functions related to these quantities of the considered electrode reaction could be evaluated.

5. Application to hydrometallurgy

The Seebeck effect has been used for the power generation practice. This is called as the thermoelectricity. Now, an application of thermoelectrochemistry to hydrometallurgy has been explored. The hydrometallurgy is a kind method that the valuable metals are separated and extracted from the corresponding ores or concentrates mainly by aqueous solution treatment. The leaching is a very important process in this method, by which the valuable metals are transformed into metallic or complex hydrated ions. By the subsequent purified and separated from impurity, the valuable components are extracted. In the direct leaching process, usually, the ores or concentrates are mixed with the acidic or the basic solution or the solution with the redox. A large amount of heats is suck up from surrounding or released out during the leaching. The heats are completely wasted except heating the leaching solution sometimes. The leaching heat effect, Q, is generally considered to be corresponding to the enthalpy change, ΔH of the leaching reaction at the constant pressure. If the leaching is designed as a process that takes place in a galvanic cell, according to Eq.(30), a part of the energy of the leaching reaction would be released as the electric work done. This treatment that takes a galvanic cell to complete the leaching process is named to be the "power generation leaching".

For the wet-extraction of copper, the chalcopyrite ($CuFeS_2$) is usually used as the raw materials, and the leaching reagent is the ferric sulphate. Its leaching reaction is

$$CuFeS_2 + 4Fe^{3+} \rightarrow Cu^{2+} + 5Fe^{2+} + 2S(s) \tag{37}$$

In the direct leaching, the released energy due to the disaggregation of the raw materials is educed as the heats. The energy that can do useful work could not be effectively released and functioned. If the power generation leaching is chosen, the galvanic cell could be constructed as shown by the following diagram,

$$-)\ CuFeS_2 \mid CuSO_4(c_1) \parallel Fe_2SO_4(c_2) \mid \text{graphite}\ (+ \tag{38}$$

The salt bridge or a kind of ionic membrane suited to the leaching system is located between the anodic and cathodic compartments. The electrode reactions are as follows,

$$-)\ CuFeS_2 - 4e^- \rightarrow Cu^{2+} + Fe^{2+} + 2S\ (s)$$

$$+)\ 4Fe^{3+} + 4e^- \rightarrow 4Fe^{2+} \tag{39}$$

The cell reaction is corresponding to Eq.(37). When the reaction progresses reversibly, the electric work could achieve about 100 kJ.mol^{-1} at room temperature.

Apparently, in the power generation leaching, a part of the energy due to the disaggregation of the ores or concentrates will be released as the electric work. Bockris had defined a coefficient as $\varepsilon_{max} = \Delta G/\Delta H$ [44] that represents the maximum transfer efficiency of the chemical energy to the electric work done, and it could be a measure as choosing leaching reaction. This also may be a criterion of the efficiency of the power generation leaching.

In the leaching with the galvanic cell, the element sulfur produced due to its insolubility would cover the surface of ores or the concentrate particles, inhibiting the further progress of the chalcopyrite leaching. When a kind of microorganism, *Acidithiobacillus ferrooxidans* (*A. ferrooxidans*) would be added to the leaching system, the insoluble simple substance could be oxidized to the positive six-valence sulfur, and further forming sulfate group, SO_4^{2-} in the oxygen-containing condition [45]. This will reinforce the leaching result, increase the leaching ratio of the mineral and make the amount of the electric work done more. The leaching reaction could be written as,

$$CuFeS_2+4Fe^{3+}+3O_2+2H_2O \rightarrow Cu^{2+}+5Fe^{2+}+ 2SO_4^{2-} +4H^+ \quad (40)$$

The maximum electric work done could achieve about 1100 kJ.mol^{-1}, when the reaction would be reversible at room temperature. That is to say, the power generation leaching would create a probability that the chemical energy of an electrochemical reaction transfers to the useful electric work.

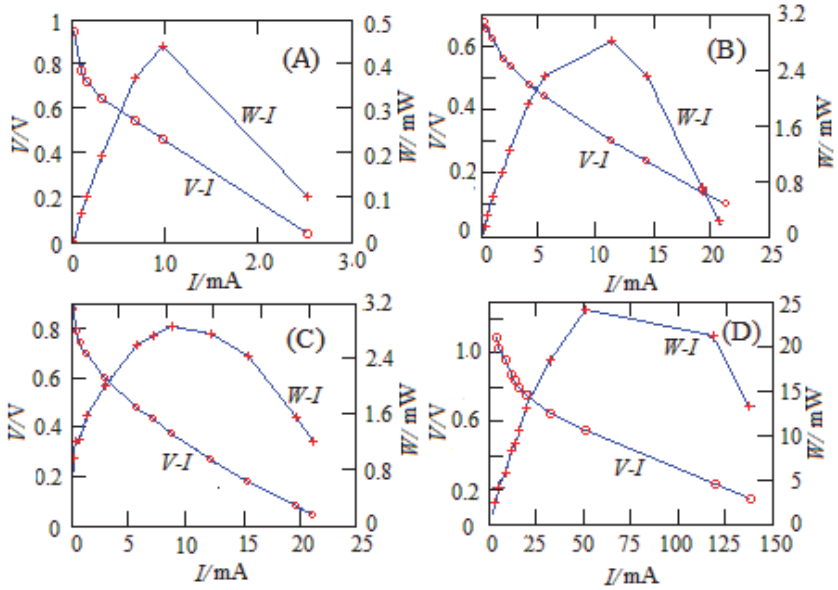


Fig. 5. Potential (V)–power (W)–current (I) curves for (A) Ni_3S_2-$FeCl_3$ system with a salt bridge, (B) PbS (beforehand steeped in Ag^+ solution)-$FeCl_3$ system, (C) Ni_3S_2 (beforehand steeped in La^{3+} solution)-$FeCl_3$ system and (D) ZnS (containing acetylene black)-$FeCl_3$ system with PE anion-selective membrane instead of salt bridge. The figures, (A), (B), (C) and (D) are from H. Zhang, et al, *Nonferrous Met.* 44 (1992) 69; S. Wang, et al, *J. Changsha Univ. Sci. & Technol.* 1 (2004), 87; *ibid* ; and S. Wang, et al, *Electrochem.* 11 (2005) 77, respectively.

In order to make the reaction progress under a high oxidation potential, the bacteria, *A. ferrooxidans* are also added into the cathodic room sometimes because they could be cultivated on the Fe^{2+} solution and grow[46]. The bacteria can promote the ferrous ions to oxidize to the ferric ones with the rates 10^5 to 10^6 times faster than those of inorganic oxidation at the corresponding pH value. This indicates that the high ratio of Fe^{3+}/Fe^{2+} is always kept in catholyte, so that the leaching reaction is also kept under a high strength of the oxidizing potentials to encourage the reaction progress with a fast rate. Leaching experiments have done on such the minerals as Cu_2S, Ni_3S_2, PbS, FeS and ZnS [47]. The researches of the power generation leaching also concern with the simultaneous leaching of two minerals, for example, Cu_2S or other metallic sulfide like Ni_3S_2, PbS, FeS or ZnS (as anode) and MnO_2(as cathode) [48]. The curves of electric currents-potentials-electric power for the power generation leaching for some metallic sulfides are shown in Fig.5, from where it can be seen that the resulting currents and the electric power obviously increase with improvement of the conditions of leaching.

In a word, this power generation leaching not only can generate electricity, but also reduce the purified steps of the leaching solution and the reagents consumption as well as is able to adjust the reaction evolution by control of the power output. From the point of view of the economy and ecological protection, this is a worthy advocating method.

6. Summary

In this chapter, a very important thermodynamic quantity, the electrochemical Peltier heat of a single electrode process, and some concepts related to this quantity have been discussed. They include the definition about EPH and the Peltier coefficient for the electrode process, the absolute scale, and the fundamental equations on this scale for thermoelectrochemistry. The equations on this scale, actually, are a special depiction of the first and second laws of the thermodynamics used to the electrode reactions. Firstly, the energy conservation equation with a new form based on the classical equilibrium thermodynamics is set up. A new reference point for the heat effects and the enthalpy change is designated in the equation, that is both ϕ^*_T and ΔS^*_T are zero when $T \to 0$ and an assumption of zero electronic entropy. For SHE reaction, the designation of $\phi^*_{T\to 0} = 0$ and $\Delta S^*_{T\to 0} = 0$ on the absolute scale is as the same as that of $\phi_{T\to 0} = 0$ and $\Delta S_{T\to 0} = 0$ on the conventional scale. Secondly, EPH on electrode–electrolyte interfaces is specially emphasized to be a quantity related to the function of state. The reversible heat effect, or EPH on the electrode–electrolyte interfaces, can be determined by the change in entropy of reaction on the absolute scale. This is an extension of the second law of thermodynamics. In order to acquire ϕ^*_T, the quantity of ΔS^*_T, dependent on temperature, needs to be beforehand identified by thethermoelectrochemistry experiments at different temperatures, and sequentially it is determined as a function of temperature. From this the electrode potential on the absolute scale would be evaluated by means of the relationship of ϕ^*_T and ΔS^*_T. Additionally, in the third law of thermodynamics, the change in entropy for a constant temperature process of a condensed system will approach to zero with the temperature moving toward zero according to Eq.(10). If further assumed that the constant temperature process could include the electrode reaction with the hydrated ions and the electrons taking part in, then

$$\lim_{T\to 0} (\Delta S^*)_T = \lim_{T\to 0} (\Delta S + z\Delta S^*(H^+/H_2))_T = 0 \tag{41}$$

That is, the stipulation of $\Delta S^*_{T\to 0} = 0$ would be also fit to the third law of thermodynamics. To sum up, the set up of the absolute scale and the evaluation of the change in entropy of SHE on this scale are useful to resolve those two problems which are mentioned at the beginning of this chapter. From this, one can get the relationship between the thermodynamic functions on the conventional and the absolute scale, and obtain the changes in entropy on the absolute scale and EPHs for electrode reactions, as a result greatly enriching the thermodynamic database. This also provides a new scale to study the electrode reaction, being valuable to the further development of the thermoelectrochemistry. Thermoelectrochemical applications have now been extended in many areas, especially in the surface-electrochemical treatment of the functional materials, electrode modifying, and the charge and discharge-control of the batteries. The application to hydrometallurgy has also been explored.

7. Acknowledgement

The author would like express his gratitude to the National Natural Science Foundation of China (No.50874119, No. 50374077) for the financial support.

8. References

[1] Z. Fang (2011). Some basic matters on the heat effects at electrode–electrolyte interfaces. *Thermochim. Acta* Vol. 516, 1-7.

[2] Z. Fang, S. Wang, Z. Zhang, G. Qiu,(2008). The Electrochemical Peltier heat of the standard hydrogen Electrode reaction. *Thermochim. Acta* Vol. 473, 40–44.

[3] H. Zhang, P. Zhang, Z Fang (1997), Coupling microcalorimeter with electrochemical instruments for thermoelectrochemical research. *Thermochem. Acta* Vol. 303, 11–15.

[4] H. Nakajima, T. Nohira, Y. Ito,(2004). The single electrode Peltier heats of Li^+/Li, H_2/H^- and Li^+/ Pd–Li couples in molten LiCl–KCl systems. *Electrochim. Acta*, Vol. 49, 4987-4991

[5] D. Bedeaux, S. K. Ratkje, (1996). The dissipated energy of electrode surfaces: Temperature jumps from coupled transport processes. *J. Electrochem. Soc.* Vol. 143, 767-779.

[6] Y. Ito, R. Takeda, S. Yoshizawa,Y. Ogata, (1985). Eletrode heat balances of electrochemical Cells: Application to NaCl electrolysis. *J. Appl. Electrochem.* Vol.15, 209–215.

[7] M. Kamata, Y. Ito, J. Oishi, (1987). Single electrode Peltier heat of a hydrogen electrode in H_2SO_4 and NaOH solution, *Electrochim. Acta*, Vol.32, 1377-1381.

[8] Y. Ito, H. Hayashi, N. Hayafuji, S. Yoshizawa, (1985). Single electrode heat of molten NaCl electrolysis: Measurement by electrolytic calorimeter and heat flux transducer (HFT). *J. Appl. Electrochem*. Vol. 15, 671–674.

[9] S. Shibata, M.P. Sumino, (1985). The electrochemical Peltier heat for the adsorption and desorption of hydrogen on a platinized platinum electrode in sulfuric acid solution. *J. Electroanal. Chem*. Vol. 193, 135–143.

[10] Y.V. Kuz'minskii, A.V. Gorodyskii, (1988). Thermal analysis of electrochemical reactions: Part I. Kinetic method of determining Peltier heats. *J. Electroanal. Chem.* Vol. 252, 21-37.

[11] Y.V. Kuz'minskii, A.A. Andriiko, (1988). Thermal analysis of electrochemical reactions: Part II. The non-stationary temperature wave method — A method for the determination of Peltier heats at the electrode/molten electrolyte interface. *J. Electroanal. Chem*. Vol. 252, 39–52.

[12] T. Ozeki, I. Watanabe, S. Ikeda, (1983). Analysis of copper (I) ion chloride solution with cyclic-voltammo-thermometry. *J. Electroanal. Chem*. Vol. 152, 41–45.

[13] H. Wang, D. Wang, B. Li, S. Sun, (1995). Improved methods to determine the electrochemical Peltier heat using a thermistor I: Improved heat-sensor electrodes and lumped-heat-capacity analysis. *J. Electroanal. Chem.* Vol. 392, 13-19.

[14] Y. Ito, F.R. Foulkes, Sh. Yashizawa, (1982). Energy analysis of a steady-state electrochemical reactor. *J. Electrochem. Soc*. Vol.129. 1936–1943.

[15] B.B. Graves, (1972). Differential voltammetric scanning thermometry of Thenth formal formal formaldehyde solution in formal perchloric acid. *Anal. Chem*. Vol. 44, 993–1002.

[16] F. Decker, M. Fracastoro-Decker, N. Cella, H. Vargas, (1990). Acoustic detection of the electrochemical Peltier effect. *Electrochim. Acta*., Vol. 35, 25-26.

[17] P. Boudeville, (1994). Thermometric determination of electrochemical Peltier heat (thermal effect associated with electron transfer) of some redox couples. Vol.226, 69-78

[18] S. Shibata, M.P. Sumina, A. Yamada, (1985). An improved heat-responsive electrode for the measurement of electrochemical Peltier heat. *J. Electroanal. Chem*. Vol. 193, 123–134.

[19] S. K. Ratkje, T. Ikeshoji, K. Syverud, (1990). Heat and internal energy changes at electrodes and junctions in thermocells. *J. Electrchem. Soc*. Vol.137, 2088–2095.

[20] T. Ozeki, N. Ogawa, K. Aikawa, I. Watanabe, S. Ikeda, (1983). Thermal analysis of electrochemical reactions: Influence of electrolytes on peltier heat for Cu(0)/Cu(II) and Ag(0)/Ag(I) redox systems, *J. Electroanal. Chem. & Interfacial Electrochem*., Vol. 145, 53-65

[21] P. Boudeville, A. Tallec, (1988). Electrochemistry and calorimetry coupling: IV. Determination of electrochemical peltier heat, *Thermochim. Acta*, Vol. 126, 221-234.

[22] T. Ozeki, I. Watanabe, S. Ikeda, (1979). The application of the thermistor-electrode to peltier heat measurement : Cu/Cu^{2+} system in aqueous perchlorate solution, *J. Electroanal. Chem*., Vol. 96, 117-121.

[23] Masahiro Kamata, Yasuhiko Ito, Jun Oishi, (1988). Single electrode peltier heat of a hydrogen electrode in NaOH solutions at high concentration, *Electrochim. Acta*, Vol.33, 359-363.

[24] R. Tamamushi, (1973). An experimental study of the electrochemical peltier heat, *J. Electroanal. Chem. & Interfacial Electrochem*., Vol. 45, 500-503.

[25] R. Tamamushi, (1975). The electrochemical Peltier effect observed with electrode reactions of Fe(II)/Fe(III) redox couples at a gold electrode, *J. Electroanal. Chem.& Interfacial Electrochem*., *Vol. 65,* 263-273

[26] E. Lange, Th. Hesse, (1933). Concerning the existence of the so-called heats of transfer (q^* values) in peltier heats, *J. Am. Chem. Soc*., Vol. 55, 853–855.

[27] J. Newman, (1995).Thermoelectric effects in electrochemical systems, *Ind. Eng. Chem. Res*., Vol.34, 3208–3216.

[28] B. B. Graves, (1972). Differential voltammetric scanning thermometry of tenth formal formaldehyde solution in formal perchloric acid. *Anal. Chem.*, Vol. 44, 993–1002.

[29] S. Kjelstrup , E. Olsen, J. Qian, (2001). The Peltier heating of aluminium, oxygen and carbon-carbon dioxide electrodes in an electrolyte of sodium and aluminium fluorides saturated with alumina. *Electrochim. Acta,* Vol. 46, 1141–1150.

[30] Y. Maeda, T. Kumagai, (1995). Electrochemical Peltier heat in the polypyrrole4ectrolyte system. *Thermochim. Acta,* Vol.267, 139-148.

[31] Z. Jiang, W. Zhang, Xi Huang, (1994). Ac electrochemical-thermal method for investigating hydrogen adsorption and evolution on a platinised platinum electrode. *J. Electroanal. Chem*. Vol. 367, 293–296.

[32] Q. Xu, S. Kjelstrup, B.Hafskjold, (1998). Estimation of single electrode heats. *Electrochim. Acta,* Vol.43, 2597-2603.

[33] Y. Ito, H. Kaiya, S. Yoshizawa, S.K. Ratkje, T. Forland, (1984). Electrode heat balances of electrochemical Cells. *J. Electrochem. Soc*. Vol. 131, 2504–2509.

[34] F. Qiu, R.G. Compton, B.A. Coles, F. Marken, (2000). Thermal activation of electrochemical processes in a Rf-heated channel flow cell: experiment and finite element simulation. *J. Electroanal. Chem*. Vol. 492 ,150-155.

[35] T.Y. Wu, (1983) *Thermodynamics, Gas Kinetic Theory and Statistical Mechanics*, Science Press, Beijing (in Chinese).

[36] K. J. Vetter, (1967). *Electrochemical Kinetics: Theoretical and Experimental Aspects*, Academic Press, New York.

[37] R.S. Berry, S.T. Rice, J. Ross, (1980). *Physical Chemistry*, John Wiley & Sons, New York.

[38] Z. Fang, H. Zhang, P. Zhang, S. Huang, L.Guo, G. Hu, (1996). Basic equations for Thermo- electrochemistry and the entropy change of the standard hydrogen electrode reaction. *Acta Metall. Sinica*. Vol. 9, 189-192.

[39] I. M. Klotz, R. M. Rosenberg, (1972). *Chemical Thermodynamics, basic theory and methods,*3rd ed. Benjamin Inc., Menlo Park CA.

[40] J. W. Cobble, (1964). The Thermodynamic Properties of High Temperature Aqueous Solutions. VI. *J. Am. Chem. Soc*. Vol. 86, 5394–5401.

[41] Z. Jiang , J. Zhang, L. Dong , J. Zhuang, (1999). Determination of the entropy change of the electrode reaction by an ac electrochemical–thermal method. *J. Electroanal. Chem*. Vol. 469, 1–10.

[42] J .A. Dean (Ed.), (1999). *Lange's Handbook of Chemistry*, 15th ed., McGraw-Hill, New York.

[43] Z. Fang, L. Guo, H. Zhang, P. Zhang, (1998). Determination of the entropy change for the electrode reaction and dulute enthalpy of some ions by thermo-electrochemical technology. *J. Cent. South Univ. Technol*. Vol. 5, 38-40.

[44] J. O'M. Bockris, A. K. N. Reddy, (1973). *Modern Electrochemistry*, Vol.2, Plenum Publishing. New York.

[45] L. Xiao, Z. Fang，G. Qiu, J. Liu, (2007).Electro-generative mechanism for simultaneous leaching of pyrite and MnO_2 in presence of *A. ferrooxidans*, *Trans. Nonferrous Met. Soc. China.* Vol.17, 1373-1378

[46] H. Tributsch, (1979). Solar bacterial biomass bypasses efficiency limits of photosynthesis, *Nature,* Vol.281, 555-556

[47] S. Wang, Z. Fang，Y. Wang， Y. Chen, (2004). Electrogenerative leaching of nickel sulfide concentrate with ferric chloride, *J. Cent. South Univ. Technol.* Vol. 11, 405-409.

[48] S. Wang, Z. Fang, Y. Tai (2006). Application of thermo-electrochemistry to simultaneous leaching of sphalerite and MnO_2, *J. Thermal Analysis and Calorimetry* .Vol. 85, 741–743.

3

Thermodynamics and the Glass Forming Ability of Alloys

Chengying Tang and Huaiying Zhou
Guilin University of Electronic Technology
P. R. China

1. Introduction

Bulk metallic glasses (BMGs) have received a great deal of attention due to scientific and technological interest ever since the first successful synthesis of an amorphous phase in the Au–Si system in 1960 (Klement et al., 1960). There has been a lot of interest to identify parameters to assess the glass forming ability (GFA) of various alloy systems and compositions. A great deal of scientific efforts for quantification of GFA of alloys has been devoted to investigation of the GFA of alloys. There have been a lot of parameters to assess the glass forming ability (GFA) of various alloy systems and compositions. As a result, many criteria, including the confusion rule and the deep eutectic rule, for evaluating the glass forming ability (GFA) of an amorphous alloy have been proposed. Among them, the criteria used usually are the supercooled liquid region $\Delta T_x(=T_x-T_g$, where T_g and T_x are the glass transition temperature and the crystallization temperature, respectively) (Inoue et al., 1993), the reduced glass transition temperature $T_{rg}(=T_g/T_l$, where T_l is the liquidus temperature) (Turnbull, 1969) and the recently defined parameters $\gamma(=T_x/(T_g+T_l))$ (Lu & Liu, 2002), $\delta(=T_x/(T_l-T_g))$ (Chen, et al., 2005), $\beta[=T_xT_g/(T_l+T_x)^2]$ (Yuan, et al., 2008), $\phi(=\Delta T_{rg}(T_x/T_g)^{0.143})$ (Fan, et al. 2007), $\omega[=T_l(T_l+T_x)/(T_x(T_l-T_x))]$ (Ji & Pan, 2009), $\gamma_c[=(3T_x-2T_g)/T_l)$ (Guo, 2010), and so on. These criteria have generally proved useful parameters for evaluating the GFA of an amorphous alloy. In order to guide the design of alloy compositions with high GFA, Inoue et al. (Inoue et al., 1998) and Johnson (Johnson, 1999) have proposed the following empirical rules: (I) multicomponent systems, (II) significant atomic size ratios above 12%, (III) negative heat of mixing and (IV) deep eutectic rule based on the T_{rg} criterion. However, Al-based metallic glasses with rare earth metal additions (Guo et al., 2000), rare earth (RE) based glasses and some binary BMGs such as Zr-Cu, Ni-Nb binary alloy (Xia et al., 2006), provide important exception from this generality, because most of above mentioned GFA parameters and rules capable of searching metallic glasses with high GFA are not applicable to these Al–based and RE-based amorphous systems. Furthermore, all the above parameters need the alloy to be first prepared in glassy form to be able to measure the crystallization temperature T_x, the liquidus temperature T_l, and/or the glass transition temperature T_g. Hence, the above parameters are not predictive in nature, as they cannot predict a good glass forming composition without actually making that alloy and rapidly solidifying it into the glassy state. It is well known that crystallization is the only event that prevents the formation of an amorphous phase. Metallic glass formation is always a competing process between the undercooled melt and the resulting

crystalline phases. The GFA of a melt is thus virtually determined by the stability of the undercooled melt and the competing crystalline phases. Thermodynamic analysis could be useful in evaluating the stability of the undercooled melt and the formation enthalpies of crystalline phases. So far, several attempts have been made successfully to investigate the GFA and predict glass forming range (GFR) in several binary and ternary amorphous alloy systems, using a pure thermodynamic approach or a combined thermodynamics and kinetics approach.

From a thermodynamic point of view, there are generally following methods for calculating the GFA and predicting glass forming range (GFR) of an alloy system. The first approach is based on the T_0 curve, which has been used to predict the GFR on several binary and some ternary systems. The quality of these predictions depends critically on the accuracy of the thermodynamic description. The second method is based on the semi-empirical Miedema's model, which has been successfully applied to calculate and predict the glass forming range of some binary or ternary systems. The third consideration is directly employed on the calculation of the driving forces of crystalline phases (minimum driving force criterion) in a supercooled melt using calculation of phase diagram (CALPHAD) database. By employing driving force criterion with the obtained thermodynamic description for the investigated system, the GFA and predicted GFR of an alloy system were determined by comparing the driving force of crystalline phases precipitated from an undercooled melt. This evaluation has been successfully used to evaluate the GFA of several binary or ternary systems. Especially, it can be used to analyze the GFA of some alloy systems with unique glass forming ability, such as Al-based system. The other thermodynamic considerations, such as suppression of the formation of intermetallic phases, have been introduced.

From a combined thermodynamics and kinetics approach, the GFA of the alloys were evaluated by introducing thermodynamic quantities obtained from CALPHAD method into Davies–Uhlmann kinetic formulations. In this evaluation, by assuming homogeneous nucleation without pre-existing nuclei and following the simplest treatment based on Johnson-Mehl-Avrami's isothermal transformation kinetics, the time–temperature-transformation (TTT) curves were obtained, which are a measure of the time t for formation of the phase Φ with a minimum detectable mass of crystal as a function of temperature. The critical cooling rates (R_c) for the glass formation calculated on the basis of the TTT curves was used to evaluate the glass-forming ability of this binary or ternary alloy. The calculated GFA results show good agreement with the experimental data in the compositional glass formation range of the investigated systems.

This chapter is intended to present systematically the methods and progress on the glass forming ability investigated by a thermodynamic approach or a combined thermodynamics and kinetics approach.

2. Calculation of GFA based on thermodynamics analysis

Usually, it is regarded the formation of metallic glasses is controlled by two factors, i.e., the cooling rate and the composition of the alloy. The critical cooling, which is the most effective gauge for GFA of the alloys, is hard to be measured experimentally. Hence, a great deal of efforts has devoted to the investigation on the correlation between the GFA and the composition of glass forming alloys. Inoue et al. (Inoue et al., 1998) and Johnson (Johnson, 1999) proposed the empirical rules to predict the element selection and compositional range of glass forming alloy. These rules have played an important role as a guideline for synthesis of BMGs for the last decade. However, recent experimental results have shown that the "confuse

principal" and "deep eutectic rule" cannot be applicable to the Cu-Zr, Ni-Nb binary system (Xia et al., 2006) and Al-based ternary system (Guo et al., 2000). From a thermodynamic point of view, it is well known that crystallization is the only event that prevents the formation of an amorphous phase. During a melt–quenching process for metallic glass formation, the glass formation is exposed to crystallization competition of other crystalline phases from the undercooled melt between liquidus temperature T_l and glass transition T_g. The GFA of a melt is thus virtually determined by the stability of the undercooled melt and the competing crystalline phases, which can be analyzed by thermodynamic analysis. In this section, several GFA calculation based on thermodynamics analysis were introduced.

2.1 Calculation of the GFA of alloys based on T_0 curve

2.1.1 Method

Generally, a glass can be formed during cooling when crystallization is avoided up to the occurrence of the glass transition. Thus, in order to predict the tendency to glass formation in a system and the composition regions where it is most probable, nucleation of crystals in the undercooled melt must be considered. The GFR will be the region of composition where nucleation of crystalline phases is less likely. Various models have been developed to analyze the GFA of alloys in the literature, as will be discussed in the following section, with different levels of approximation. T_0 curve is one of the approaches used to estimate the GFA of the alloys.

A T_0 curve is the locus of the compositions and temperatures where the free energies of two phases are equal. Thus, T_0 curves can be calculated provided that their Gibbs free energy is known, i.e. an assessment of the system is available. The T_0 curve between the liquid and a solid phase determines the minimum undercooling of the liquid for the partitionless formation of a crystalline solid with the same composition (Boettinger & Perepezko, 1993). Fig. 1 showed one example for a simple eutectic system. Alloys with T_0 curves plunge steadily at low temperatures (dashed line in Fig. 1a), there will be no driving force for partitionless transformation in the composition region between them. If the equilibrium crystalline phases are not prone to nucleation, the glass can thus form. On the contrary, if T_0 curves that are only slightly depressed below the stable liquidus curves are good candidate for partionless transformation of crystalline phases in the entire composition range (dashed line in Fig. 1b).

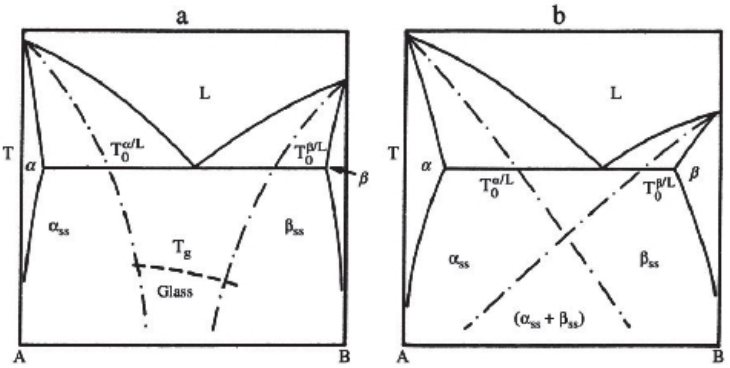


Fig. 1. Hypothetical T_0 curves for a binary eutectic A–B system. (a) T_0 curves drop to low temperature: glass formation is possible. (b) T_0 curves intersect at low temperature: partitionless crystalline phase formation occurs (redrawn from Boettinger & Perepezko, 1993).

2.1.2 Application of T_0 curve

Predictions of GFR based on T_0 curves have been performed on several binary and some ternary systems. The construction of T_0 curves for alloy glass needs a precise knowledge of thermodynamic properties of the supercooled liquid alloy and the introduction of the transition to the glassy state (Kim, et al., 1998). The quality of these predictions depends critically on the accuracy of the thermodynamic description and the introduction of the excess specific heat contribution is expected to improve the quality of results (Palumbo & Battezzati, 2008). However, as pointed out by Schwarz and co-workers (Schwarz et al., 1987), some discrepancies have been observed between the prediction and experimental results. For example, even when using the most recent thermodynamic assessment (Kumar, 1996) to calculate T_0 curves in the Cu–Ti system, the results are not agreement with the reported experimental GFR. In fact, T_0 curves for terminal solid solutions do not plunge at low temperatures as expected for glass forming systems (Kumar et al. 1996). Battezzati and co-workers (Battezzati, et al., 1990) have shown that in the Cu–Ti system the contribution of the excess specific heat is essential for describing the glass forming ability. An excess specific heat contribution has also been considered in the Al–Ti system (Cocco, et al., 1990) and the Fe–B system (Palumbo, et al., 2001).

2.2 Calculation of the GFA of alloys based on Miedema's model

2.2.1 Method

Miedema's model is an empirical theory for calculating heat of mixing in various binary systems both for the solid state (Miedema et al., 1975) and liquid (Boom et al., 1976). This model involves the calculations of the formation enthalpy of metallic glasses (amorphous phase) (ΔH^{amor}), solid solutions (ΔH^{SS}), and intermetallic compounds (ΔH^{inter}) according to the following equations (Bakker 1988; Boer et al. 1988).

$$\Delta H^{amor} = \Delta H^{chem}(amor) + \Delta H^{topo} \tag{1}$$

$$\Delta H^{ss} = \Delta H^{chem}(SS) + \Delta H^{elastic} + \Delta H^{structure} \tag{2}$$

and

$$\Delta H^{\text{int}er} = \Delta H^{chem}(\text{int}er) \tag{3}$$

where ΔH^{chem}(amor) is the chemical mixing enthalpy of the amorphous state, ΔH^{topo} is the topology enthalpy of a glass, ΔH^{chem}(SS) is the chemical mixing enthalpy of a solid solution, $\Delta H^{\text{elastic}}$ is the elastic enthalpy of the solid solution calculated based on the continuous elastic model proposed by Friedel (Friedel, 1954) and Eshelby (Eshelby, 1954 & 1956), $\Delta H^{\text{structure}}$ is the structure enthalpy induced by the structural changes, and ΔH^{chem}(inter) is the chemical mixing enthalpy of an intermetallic compound. The formation enthalpy ΔH^{inter} of a composition between two adjacent intermetallic compounds can be calculated using the level principle.

The chemical contribution of enthalpy of mixing of solid solution can be written as

$$\Delta H^{chem} = x_A x_B [x_A \Delta H^{SS}_{BinA} + x_B \Delta H^{SS}_{AinB}] \tag{4}$$

where x_A and x_B represent the mole fraction of A and B atoms and ΔH^{SS} is the enthalpy of solution of one element in another at infinite dilution. The data have been taken from Niessen et al. (Niessen, et al., 1983).
The elastic term in the enthalpy of formation originates from the atomic size mismatch, which can be expressed as

$$\Delta H^{elastic} = x_A x_B [x_A \Delta H^{elastic}_{BinA} + x_B \Delta H^{elastic}_{AinB}] \tag{5}$$

The $\Delta H^{elastic}_{iinj}$ has been obtained by using the formalism by Simozar and Alonso (Simozar & Alonso, 1984) as

$$\Delta H^{elastic}_{iinj} = \frac{2\mu_j (V_i - V_j)^2}{V_j (3 + 4\mu_j K_i)} \tag{6}$$

where μ_j is the shear modulus of the solvent, V_i and V_j are the molar volumes of the solute and the solvent, respectively and K_i is the compressibility of the solute.
The structural contribution of enthalpy for solid solution originates from the valence and the crystal structure of the solute and the solvent atom. It is found to have a very minor contribution and it is difficult to calculate. Hence, the structural contribution to enthalpy has been usually neglected (Basu, et al., 2008). In the case of the elastic and structural contributions are absent, thus the formation enthalpy of glasses can be calculated as

$$\Delta H^{amor} = \Delta H^{chem}(amor) + 3.5 \sum_{1}^{n} x_i T_{m,i} \tag{7}$$

Where x_i represents the mole fraction of component *i* atom, $T_{m,i}$ is the melting temperature of the component *i*.
According to the Miedema's model, an amorphous phase can be formed if the enthalpy of formation of the amorphous phase is less than that of the solid solution phase. The heat of formation in alloys generally arises from the interactions among the constituent atoms where the interfacial energy plays a major role. The interfacial energy mainly comes from the atomic size difference. It has also been postulated that the number of intermetallic phases appearing in an alloy system is a strong function of the heat of mixing. The number of intermetallic phase in an alloy system increases with the increase in the heat of mixing. This model can be directly used to determine the glass forming range in binary alloy systems and can be extended to ternary systems by neglecting the ternary interactions.

2.2.2 Calculation of the GFA for the binary alloy systems

Since the metallic glass formation process is controlled by thermodynamic factors, Miedema's model was firstly used to predict the composition range of amorphous binary transition metal alloys (Kolk et al., 1988; Coehoorn et al., 1988; Murty, et al., 1992; Basu, et al., 2008). It is found that the predicted glass forming composition ranges are in good agreement with the experimental results. In the work of Takeuchi and Inoue (Takeuchi & Inoue, 2000), this approach has been used to calculate the mixing enthalpy and mismatch entropy of a number of bulk metallic glass alloy systems. It has been observed that the mixing enthalpy and normalised mismatch entropy for glass forming alloys vary within a certain range.

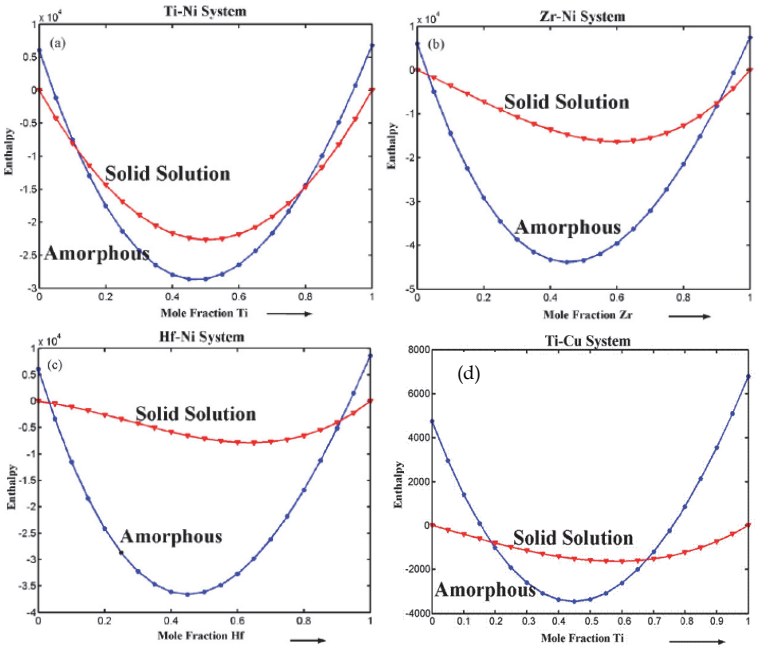


Fig. 2. Enthalpy–composition curves for binary Ti–Ni, Zr–Ni, Hf–Ni, Ti-Cu alloy systems (a–d). The curve with (•) and the curve with (Δ) represent amorphous and solid solution phase, respectively. The enthalpy values are in J/mol (From Basu, et al., 2008).

As shown in Fig. 2, in the work of Basu et al. (Basu, et al., 2008), glass forming range (GFR) has been determined for different binary (Ti–Ni, Zr–Ni, Hf–Ni, Ti–Cu, Zr–Cu, Hf–Cu) in (Zr, Ti, Hf)–(Cu, Ni) alloys based on the mixing enthalpy and mismatch entropy calculations. Though copper and nickel appear next to each other in the periodic table, the glass forming ability of the copper and nickel bearing alloys is different. Thermodynamic analysis reveals that the glass forming behaviour of Zr and Hf is similar, whereas it is different from that of Ti. The smaller atomic size of Ti and the difference in the heat of mixing of Ti, Zr, Hf with Cu and Ni leads to the observed changes in the glass forming behaviour. Enthalpy contour plots can be used to distinguish the glass forming compositions on the basis of the increasing negative enthalpy of the composition. This method reveals the high glass forming ability of binary Zr–Cu, Hf–Cu, Hf–Ni systems over a narrow composition.

In the recent work performed by Xia (Xia, et al., 2006), the GFA of an alloy is considered that the formation of the meta-stable amorphous state should include two aspects: (1) the driving force for the glass formation, i.e., $-\Delta H^{amor}$, and (2) the resistance of glass formation against crystallization, i.e. the difference between the driving force for glass phase and for the intermetallic compound formation $\Delta H^{amor}-\Delta H^{inter}$. When two glass forming alloys have the same $-\Delta H^{amor}$ but different $\Delta H^{amor}-\Delta H^{inter}$, their GFA can then be dominated by $\Delta H^{amor}-\Delta H^{inter}$. The lower the value of $\Delta H^{amor}-\Delta H^{inter}$, the higher the GFA of the alloy. On the other hand, when two glass forming alloys have the same $\Delta H^{amor}-\Delta H^{inter}$ but different ΔH^{amor}, their GFA is dominated by $-\Delta H^{amor}$. The higher the value of $-\Delta H^{amor}$, the better the GFA. Since the

contribution from entropies is much smaller as compared with that from the formation enthalpy of solid compounds (Delamare, et al., 1994), the GFA is expressed in terms of formation enthalpy alone. Based on this thermodynamic consideration, a new parameter γ* to evaluate GFA for glass formation was proposed by Xia et al. (Xia, et al., 2006) and expressed as

$$\gamma^* = GFA \propto \frac{\Delta H^{amor}}{\Delta H^{inter} - \Delta H^{amor}} \tag{8}$$

where ΔH^{amor} and ΔH^{inter} are the enthalpies for glass and intermetallic formation, respectively. Both ΔH^{amor} and ΔH^{inter} are calculated by Miedema's macroscopic atom model. This parameter has been successfully used to predict the GFR and the best GFA alloy compositions in Zr-Cu and Ni-Nb system by comparing the value of γ* of various alloy systems, respectively (Xia, et al., 2006).

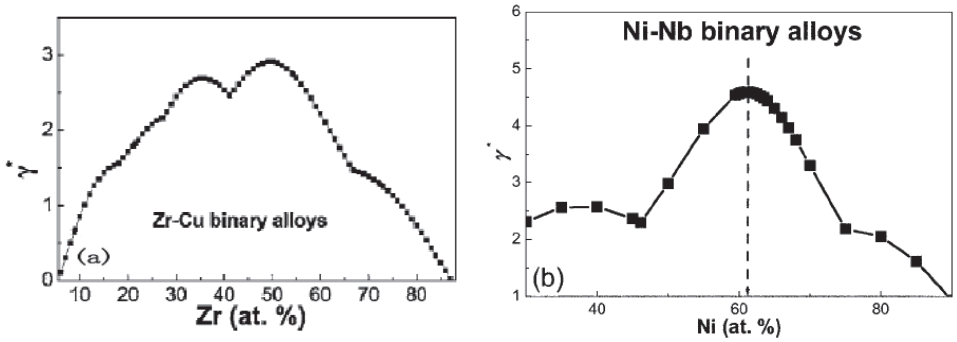


Fig. 3. Calculated dependence of the parameter γ* on Zr and Ni concentration in Cu-Zr (a) and Ni-Nb (b) binary alloys, respectively (from Xia, et al., 2006).

Fig. 3 shows the calculated dependence of the parameter γ* on Zr and Ni concentration in Cu-Zr (a) and Ni-Nb (b) binary alloys, respectively, suggesting that the alloys $Cu_{64}Zr_{36}$ and $Cu_{50}Zr_{50}$ in Cu-Zr system, and $Ni_{61.5}Nb_{38.5}$ in Ni-Nb system are the best glass former, respectively. These predicted results are in good agreement with the experimentally reported $Cu_{64.5}Zr_{35.5}$ and $Cu_{50}Zr_{50}$, and $Ni_{62}Ni_{38}$ that could be made into bulk metallic glass rods with 2 mm in diameter, indicating that γ* is an effective parameter in identifying the best glass former in the Zr-Cu and Ni-Nb binary system.

Similarly, considering both the stability of liquid employing $\Delta H^{liq}/\Delta H^{inter}$, and the competition of glass and crystal using $\Delta H^{amor}/\Delta H^{inter}$, Ji et al. (Ji, et al., 2009) proposed a new parameter γ′ of GFA as

$$\gamma' = GFA \propto \frac{\Delta H^{liq} \bullet \Delta H^{amor}}{(\Delta H^{inter})^2} \tag{9}$$

As Ji et al. described, this parameter γ′ is not only verified in five different binary bulk metallic glasses (Cu–Hf, Ni–Nb, Cu–Zr, Ca–Al, Pd–Si) but also showed wider application range comparing with the former model, but also have a better GFA estimation on the different composition than the parameter γ* because it is including ΔH^{liq} in the evaluation expression. The predicated results are in good agreement with the experiments in all five different kinds of binary BMG systems and the biggest deviation of the peak of γ′ from the

best current GFA composition is only about 6 at.% in Ca–Al alloy. Comparing with former GFA parameter γ*, γ′ takes account of liquid stability and shows more universal for evaluation GFA in different kinds of binary alloys (Ji, et al., 2009). Recently, Wang et al. also made a modification to Xia's proposal and it works more convenient to describe the GFA of transition metal systems (Wang, et al., 2009).

2.2.3 Calculation of the GFA for the multicomponent alloy systems

Miedema's approach has been extensively used by Nagarajan and Ranganathan (Nagarajan & Ranganathan, 1994), Takeuchi and Inoue (Takeuchi & Inoue, 2001 &2004) and other researchers (Murty, et al., 1992; Rao, et al.; 2007; Basu, et al, 2008; Wang & Liu, 2009; Sun, et al., 2010) to determine the glass forming composition range (GFR) in a number of ternary and multicomponent systems. In the work performed by Takeuchi and Inoue (Takeuchi & Inoue, 2001), the amorphous-forming composition range (GFR) was calculated for 338 ternary amorphous alloy systems on the basis of the database given by Miedema's model in order to examine the applicability of the model, to analyze the stability of the amorphous phase, and to determine the dominant factors influencing the ability to form an amorphous phase. The mixing enthalpies of amorphous and solid solution phases were expressed as a function of alloy compositions on the basis of chemical enthalpy. The GFR was calculated for 335 systems except for the Al-Cu-Fe, Al-Mo-Si and Au-Ge-Si systems. The calculated results are in agreement with the experimental data for Cu-Ni- and Al-Ti-based systems. For typical amorphous alloy systems exemplified by the Zr-, La-, Fe- and Mg-based systems, it was recognized that the calculated GFR had been overestimated as a result of the model being simplified. It is found that the elastic enthalpy term arising in a solid solution phase stabilizes the amorphous phase, and the stabilization mechanism is particularly notable in Mg-based amorphous alloy systems. Short-range order plays an important role in the formation of Al-, Fe- and Pd-metalloid based systems (Takeuchi & Inoue, 2001).

Based on Miedema's model and Alonso's method, the glass forming ability/range (GFA/GFR) of the Fe–Zr–Cu system was studied by thermodynamic calculation. It is found that when the atomic concentration of Zr is between 34% and 56%, no matter what the atomic concentrations of Fe and Cu are, amorphous phase could be obtained, thus the atomic mismatch playing a dominating role in influencing the GFA. While the atomic concentration of Zr is out of the above range, the GFA is highly influenced by the immiscibility between Fe and Cu (Wang & Liu, 2009).

Glass forming composition range for ternary Zr–Ti–Ni, Zr–Hf–Ni, Ti–Hf–Ni, Zr–Ti–Cu, Zr–Hf–Cu and Ti–Hf–Cu systems has been determined by extending the Miedema's model to ternary alloy systems and by neglecting the ternary interaction parameter (Basu, et al., 2008). In their calculations, solid pure metals have been chosen to be the standard state and their enthalpy has been assigned to be zero. It is seen that the glass forming composition range for Ni bearing alloys is higher than that of the Cu bearing alloys, as heat of mixing of Ni is higher than that of Cu with Ti, Zr and Hf. In these ternary (Zr, Ti, Hf)–(Cu, Ni) alloys mixing enthalpy and mismatch entropy varies between (−13) and (−42) kJ/mol and 0.13 and 0.25, which is within the range predicted for glass formation (Basu, et al., 2008).

In the work of Oliveira et al., the γ* parameter proposed by Xia et al. was extended to the ternary Al–Ni–Y system. The calculated γ* isocontours in the ternary diagram are compared with experimental results of glass formation in that system. Despite some misfitting, the best glass formers are found quite close to the highest γ* values, leading to the conclusion that

this thermodynamic approach can be extended to ternary systems, serving as a useful tool for the development of new glass-forming compositions (Oliveira et al., 2008).
Rao et al. (Rao et al. 2007) identified the composition with highest glass forming ability in Zr-Ti-Ni-Cu-Al quinary systems with the Gibbs-energy change between the amorphous and solid solution phases as the thermodynamic parameter by calculating the Gibbs-energy change with the help of Miedema, Miracle, mismatch entropy and configurational entropy models. ΔG shows the strong correlation with the reduced glass transition temperature (T_g/T_l) in Zr-based metallic glasses. Thus, ΔG can be used as a predictive GFA parameter to identify compositions with the highest GFA. The compositions with the highest GFA have been identified in a number of quinary systems by iso-free energy contour maps by representing quandary systems as quasiternary systems (Rao et al. 2007). The best glass forming composition has been identified by drawing iso-Gibbs-energy change contours by representing quinary systems as pseudo-ternary ones. Attempts have been made to correlate the Gibbs-energy change with different existing glass forming criteria and it is found that the present thermodynamic parameter has good correlation with the reduced glass transition temperature. Further, encouraging correlations have been obtained between the energy required for amorphization during mechanical alloying to the Gibbs-energy change between the amorphous and solid solutions.

2.3 Calculation of the GFA of alloys based on driving force criterion

2.3.1 Method

The basic underlying concept to predict the compositions of alloys having high GFA using the thermodynamic approach is that the compositions exhibiting the local melting minimum points favour amorphous phase formation. Thermodynamic approach of driving force criterion is based on a different concept. During a melt–quenching process for metallic glass formation, the glass formation is exposed to crystallization competition of other crystalline phases from the undercooled melt between liquidus temperature T_l and glass transition T_g. It is well known that the crystallization is the only event that prevents the formation of amorphous phase. Considering that crystallization is usually through the nucleation and growth process, the high GFA can be inversely predicted by searching a condition where the nucleation and growth of crystalline phases can be retarded. There are three dominating factors for the kinetics, (i) chemical driving force, (ii) interfacial energy, as an energy barrier, between the amorphous phase and the crystalline phases, (iii) the atomic mobility for rearrangement or transport of the partitioning atoms. According to the classical nucleation theory, the driving force of formation of the crystalline phases and interfacial energy, among other things, affects the nucleation rate of product phases. The interfacial energy between liquid and crystalline phases is known to be small compared to surface energy or grain boundary energy (Porter & Easterling, 1992), and therefore the role of interfacial energy in the nucleation kinetics of crystalline phases would be small. Then, the driving force of formation becomes the major factor that affect the nucleation kinetics of crystalline phases from amorphous alloy melts. It is believed that alloys with lower driving force for the formation of crystalline phases under the supercooled liquid state suggest higher GFA in the glass forming range. Therefore, Kim and co-workers proposed the minimum driving force criterion as a new thermodynamic calculation scheme to evaluate the composition dependence of the GFA (Kim et al., 2004). The driving force for the crystalline phases can be calculated using the critical assessed thermodynamic parameters by the CALPHAD method (Kaufman & Bernstein, 1970). In the CALPAHD method, the Gibbs energies of individual phases are described using

thermodynamic models. Then, the model parameters are optimized considering relevant experimental information on phase equilibria or the other thermodynamic properties. The calculation of phase equilibrium is performed based on the minimum Gibbs energy criterion.

2.3.2 Application of the driving force criterion

Driving force criterion has been successfully used to explain the composition dependence of GFA in several glass forming alloys with unique GFA, such as Cu-Zr-Ti (Kim, et al., 2004), Mg-Cu-Y (Kim, et al., 2005), Al-Ce-Ni (Tang, et al. 2010), and Al-Cu-Zr (Bo, et al, 2010) systems, by calculating the driving force of formation of crystalline phases under metastable supercooled liquid states and by searching the local minima of the driving forces for crystallization. The calculated results are in good agreement with the experimental results. It has been indicated that the driving force criterion can be used as a new thermodynamic scheme to estimate the composition dependence of GFA in multicompoent alloy systems for the development of bulk amorphous alloys.

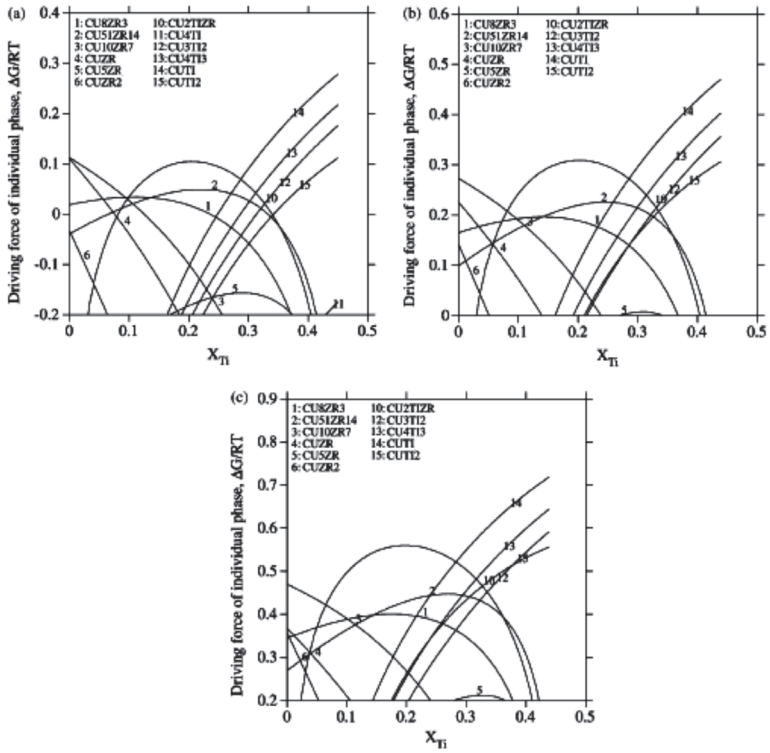


Fig. 4. Calculated driving forces of crystalline phases for $Cu_{55}Zr_{45-x}Ti_x$ alloys, versus Ti content at (a) 1073 K, (b) 973 K, and (c) 873 K (from Kim, et al., 2004).

For the Cu-Zr-Ti system, among a series of ternary alloys $Cu_{60}Zr_{40-x}Ti_x$ (x = 10, 20, 30), the alloy with the highest GFA should be the alloy $Cu_{60}Zr_{20}Ti_{20}$ according to the maximum T_{rg} criterion (Turnbull, 1969), while experiments (Inoue, et al., 2001) show it is $Cu_{60}Zr_{30}Ti_{10}$. Although the other alloys, $Cu_{55}Ti_{35}Zr_{10}$ (Lin & Johnson, 1995) and $Cu_{47}Ti_{33}Zr_{11}Ni_8Si_1$ (Choi, et

al., 1998) based on the Cu–Ti–Zr ternary system but at different region, have been published as alloys with high GFA, there is no empirical rule or factor that can explain why the high GFA is obtained at certain compositions in the Cu–Ti–Zr system (roughly Zr: Ti=3:1 and Zr: Ti=1:3). As already the thermodynamic parameters for all phases in this system obtained by the CALPHAD method, the GFA is estimated by calculating the driving forces of all crystalline phases under the undercooled liquid state. Fig. 4 shows the calculated driving forces of individual crystalline phases as a function of Ti content in a temperature range (600–800 °C) where the alloys correspond to supercooled liquids state. As shown in this figure, along the composition line $Cu_{55}Zr_{45-x}Ti_x$ with varying Ti content, the driving forces of crystalline phases show two local minimums, one at Zr-rich region (x = 7–10) and the other at Ti-rich region (x= 28–29). According to the driving force criterion, the two local minimum points in Fig. 4 are the compositions where the GFA is expected to be higher than other compositional region. In a sense that the Zr:Ti ratios in the two local minimum points are toughly close to 3:1 and 1:3, it can be said that the former is close to the Inoue's composition and the latter is close to Johnson's composition (Kim et al., 2004). This finding indicates that the composition dependency of the GFA in the Cu-Zr-Ti ternary alloy system can be explained by calculating the driving forces of formation of crystalline phases under metastable supercooled liquid states and by searching the local minima of the driving forces for crystallization (Kim, et al. 2004).

Similarly, Al–based amorphous, which was discovered in 1988 (He, et al., 1988 & Inoue, et al., 1988), is also of particular interest because of its low density, good bending ductility and high tensile strength. It was found that, however, most of above mentioned parameters and rules capable of searching metallic glasses with high GFA are not applicable to Al–based amorphous (Guo, et al., 2000, Hackenberg, et al., 2002, Gao, et al., 2003, Zhu, et al. 2004).Al–Ce–Ni system is an unique Al–based system, which can be synthesized into a strong, flexible metallic glass with the widest GFR covering 2–15 at.% Ce and 1–30 at.% Ni (Inoue, 1998, Kawazoe et al., 1997). The alloys with high GFA are situated away from the eutectic point. Experimental results of the Al–Ce–Ni bulk amorphous alloys prepared with copper mold casting indicate that the amorphous sheets with 5 mm width and 0.2 mm thickness are obtained in $Al_{86}Ce_4Ni_{10}$ and $Al_{88}Ce_6Ni_6$ alloys without appreciable glass transition. On contrary, alloys $Al_{82}Ce_8Ni_{10}$ and $Al_{80}Ce_6Ni_{14}$ with $\triangle$Tx values of 20 and 21K consist mainly of crystalline phases (Inoue, 1998). After a thermodynamic assessment of the Al-Ce-Ni system in the Al-rich corner was performed, a set of consistent thermodynamic parameters were obtained, and the thermodynamic properties of the Al-Ce-Ni amorphous alloys were calculated. The calculated results indicated that the alloys with high GFA in the Al–Ce–Ni system are far from the eutectic point, and the heats of mixing are from –15 to –49 kJ/mol of atom for the observed amorphous alloys (Tang et al., 2010).

As shown in Fig. 5, the relatively smaller nucleation driving forces for the formation of crystalline phases for the Al–10Ce based alloys (Fig. 5a) are generally indicative of their higher GFA with a reportedly wider GFR (1–30at.% Ni) (Kawazoe et al., 1997). In contrast, the relatively larger driving forces in the Al–10Ni based alloys (Fig. 7b) are associated with their poorer GFA and narrower GFR (2–10 at.% Ce) (Kawazoe et al., 1997). This finding is further confirmed by the melt spinning (Tang, et al., 2010) and the copper mold casting experimental results (Inoue, 1998).

Based on the experimental enthalpies of mixing of ternary liquid and undercooled liquid alloys as well as the evaluated isothermal sections, the Al–Cu–Zr ternary system has been assessed using the CALPHAD method. Most of the calculated results show good agreement

with the experimental thermodynamic data and the reported phase diagrams. By employing the driving force criterion with the present thermodynamic description, the observed glass-forming ability in the Al–Cu–Zr system can be accounted for satisfactorily (Bo, et al., 2010).

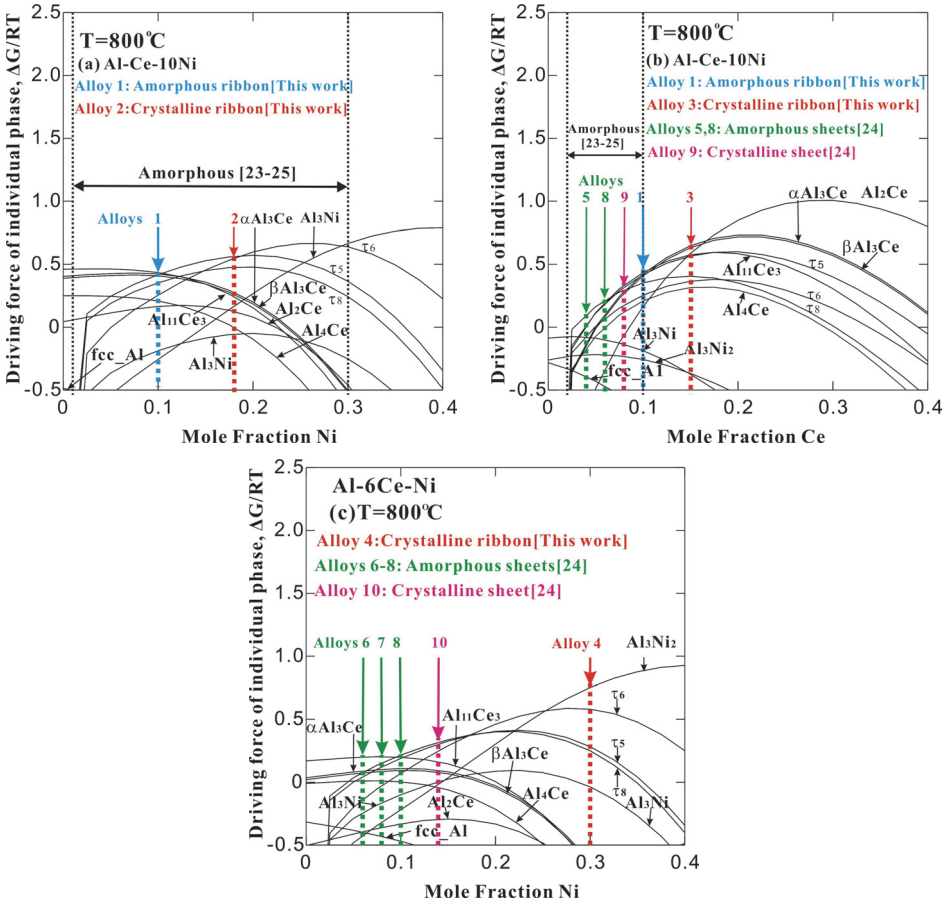


Fig. 5. Calculated normalized nucleation driving force (per mole of atoms) for crystalline phases from undercooled (a) Al–10Ce–Ni, (b) Al–Ce–10Ni and (c) Al-6Ce-Ni metastable liquid at 800 °C (from Tang, et al., 2010).

2.4 Calculation of the GFA of alloys based on the other thermodynamic approach

By treating the glass transition as a second-order phase transformation from liquid phase (Palumbo, et al., 2001, Shao et al., 2005), which can give good predictions of all important GFA indicators such as the reduced glass transition temperature and the thermodynamic stability of the amorphous phase, Shao et al. established a full thermodynamic database for glass forming ability (GFA). The resultant thermodynamic database can be used to produce all major temperature-related GFA indicators such as T_g/T_l, T_g/T_m and $T_x/(T_g+T_l)$. It is indicated that together with phase diagram prediction, such an extensive CALPHAD approach is a powerful tool for designing alloys with large GFA (Shao et al., 2005).

By using the computational thermodynamic approach exhibiting low-lying liquidus surfaces coupled with the reduced glass transition temperature criterion of Turnbull, regions of alloy composition suitable for experimental tests for glass formation of Zr–Ti–Ni–Cu–Al system were identify rapidly by Cao et al. (Cao, et al., 2006). The glass forming ability of the alloys we studied can be understood in terms of the relative liquidus temperature in a thermodynamically calculated temperature vs. composition section through a multicomponent phase diagram. It does not follow several other proposed thermodynamic or topological criteria.

A thermodynamic parameter ($\Delta H^{chem} \times S_\sigma/k_B$) in the configuration entropy (S_{config}/R) range of 0.8-1.0 has been developed to identify excellent BMG composition using enthalpy of chemical mixing (ΔH^{chem}), the mismatch entropy normalized by Boltzmann's constant (S_σ/k_B) and the configurational entropy (S_{config}/R) by Bhatt et al. and it has been demonstrated for the Zr-Cu-Al based ternary system. It is found that this approach can be used to predict the best BMG composition more closely than the earlier models (Bhatt, et al., 2007).

Based on the undercooling theory resulting from the existence of multicomponent chemical short-range order (CSRO) domains, the glass forming range (GFR) in Zr-Ni-Ti alloy system was predicted by thermodynamic calculation. The GFR predicted by the thermodynamic calculation is consistent with the experiment results (Liu, et al. 2008).

One of the ways to predict the possible bulk glass formation composition is the phase diagram calculation with suppression of the formation of intermetallic phases. The formation of stoichiometric intermetallic compounds which have the ordered structure of atoms into specific lattice sites can take a time for the rearrangement of atoms from liquid state. Thus, the formation of intermetallic compounds can be suppressed during the fast solidification process normally applied to the bulk glass production. Combining the obtained the thermodynamic database and the above concept, the amorphous formation diagram of the Cu–Zr–Ag system with the suppression of all binary and ternary intermetallic phases has been proposed by Kang and Jung (Kang & Jung, 2010)

3. Calculation of the GFA of alloys based on a combined thermodynamics and kinetics approach

As discussed above, the thermodynamic approach is useful since the thermodynamic parameters can be used to calculate the GFR in binary alloys and can also be used to predict the GFR in ternary systems based on the constituent binaries. One of the limitations of a purely thermodynamic approach is that it does not give the critical cooling rates for the glass formation. A combined thermodynamic and kinetic treatment, based on time-temperature-transformation curves (TTT) in the manner of Uhlmann and Davies has been presented (Saunders & Miodownik, 1986 & 1988). This combined approach takes the thermodynamic parameters obtained from the phase diagram calculations and derives values for the free energy barrier for nucleation, free energy driving forces, and melting points used in kinetic equations. The combined approach has been successfully used to calculate the glass forming ability (GFA) of a wide range of binary and ternary alloy systems (Saunders & Miodownik, 1988). The calculated glass forming ranges for a wide number of binary and ternary alloy systems are in good agreement with experiment. A significant advantage of the combined approach is that data from binary alloy systems, often with little or no ternary modification, can be used to calculate the necessary thermodynamic input for the kinetic equations in higher order systems. This section outlines briefly the combined thermodynamic and kinetics method used for the calculation of the GFA for alloy systems.

3.1 Method

Critical cooling rates for glass formation can be obtained by Johnson-Mehl-Avrami isothermal transformation kinetics using the equation

$$X = 1 - \exp[(-\pi / 3) I_v U_c^3 t^4] \tag{10}$$

where X is the volume fraction of material transformed, I_v is the nucleation frequency, U_c is the crystal growth rate, and t is the time taken to transform X. In the early stages of transformation the value of X approximates to

$$X \cong \pi I_v U_c^3 t^4 \tag{11}$$

For homogeneous nucleation without pre-existing nuclei, the nucleation frequency I_v^h is given by

$$I_v^h = \frac{D_n N_v}{a_0^2} \exp(-\Delta G^* / kT) \tag{12}$$

where D_n is the diffusion coefficient necessary for crystallisation, N_v is the number of atoms per unit volume. a_0 is an atomic diameter. k is Boltzmann's constant. T is the transformation temperature, and ΔG^* is the free energy barrier for nucleation of a spherical nucleus given by the expression

$$\Delta G^* = \frac{16\pi}{3} (\sigma^3 / G_v^2) \tag{13}$$

where σ is the liquid/crystal interfacial energy and G_v is the change in free energy per unit volume on solidification. An equation for U_c can be written as

$$U_c = \frac{f D_g}{a_0} [1 - \exp(-\Delta G_m / RT)] \tag{14}$$

where D_g is the diffusion coefficient for the atomic motion necessary for liquid to crystal growth, ΔG_m is the molar free energy driving force for liquid to crystal growth, and R is the universal gas constant, and f is a structural constant denoted the fraction of sites on the interface where atoms may preferentially be added or removed, and is given by the following expression (Uhlmann, 1972)

$$f = 0.2(T_m - T) / T_m \tag{15}$$

where T_m is the liquidus temperature. By assuming that D_n=D_g=the bulk liquid diffusivity and invoking the Stokes-Einstein relationship between diffusivity and viscosity η, equation (12) and (14) can be derived to give the time t needed to form a volume fraction X of transformed crystalline phase in an undercooled liquid as following

$$t \approx \frac{9.3\eta}{kT} \left\{ \frac{a_0^9 X}{f^3 N_v} \frac{\exp(\Delta G^* / kT)}{[1 - \exp(-\Delta G_m / RT)]^3} \right\}^{1/4} \tag{16}$$

where t is the time taken to transformation volume fraction X of crystalline solid. η is the viscosity of liquid, a_0 is an atomic diameter, f is a structural constant, N_v is the number of atoms per unit volume, ΔG^* is the Gibbs energy barrier to nucleation and ΔG_m is the Gibbs energy driving force for the liquid-crystal transformation. The constants have been typically taken as $X=10^{-6}$, $a_0=0.28\times10^{-9}$m, $f=0.1$ and $N_v=5\times10^{28}$ atoms/m^3. In order to apply this equation to a real alloy system, it is necessary to derive or estimate the parameters η, ΔG^*, and ΔG_m (Saunders & Miodownik, 1988).

3.2 Estimation of η, ΔG^*, and ΔG_m

Since it is very difficult to measure experimentally the viscosity of supercooled liquid, there have been few measurements of it. In this case, the viscosity can be generally described as being between the liquidus temperature T_m and the glass transition T_g using a Doolittle-type expression involving the relative free volume f_T (Ramachandrarao, et al., 1977) as

$$\eta = A\exp(B / f_T) \tag{17}$$

where

$$f_T = C\exp(-E_H / RT) \tag{18}$$

E_H is the hole formation energy and A, B, and C are constants. Because of the lack of experimental data and the E_H value was estimated by means of a direct relationship from T_g (Ramachandrarao, et al., 1977). Assuming B is unity and f_T and η are 0.03 and 10^{12}Ns/m^2, respectively at Tg, A and C have been approximated at 3.33×10^{-3} and 10.1, respectively. If T_g values are not available, crystallisation temperatures T_x are used as a first approximation.

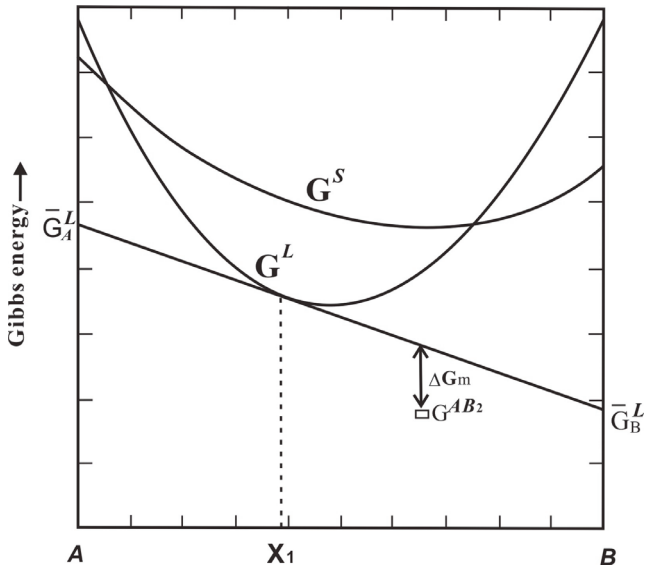


Fig. 6. The construction used in calculating the driving force, ΔG_m, for the crystallization of compound AB_2 from a liquid of composition x_1 in the A-B system.

For the crystallization of compounds AB_2 from a liquid of composition x_1 in the A-B system (Fig. 6), the molar free energy driving force for liquid to crystal growth, ΔG_m, represents the Gibbs energy required to form one mole of crystalline phase from the liquid of composition x_1, which can be obtained from thermodynamic phase diagram calculations to give explicitly molar heats of fusion H_m^f and driving force that can be used to ΔG^*. Values of H_m^f and ΔG_m are calculated from partial molar Gibbs energies of elements A and B and free energy values. Therefore, ΔG_m is expressed by the following equation

$$\Delta G_m = x_A \overline{G}_A^L + x_B \overline{G}_B^L - G_{cryst} \tag{19}$$

where x_A and x_B are the mole fractions of elements A and B in the precipitating crystalline phase, respectively. $\overline{G}_A^L$ and $\overline{G}_B^L$ are the partial molar Gibbs energies of elements A and B in the liquid phase, respectively, and G_{cryst} is the integral free energy of the precipitating crystalline phase. The Gibbs energy functions in Eq. (19) can be obtained from the thermodynamic model parameters evaluated in the literature. In an A-B alloy system, a liquid composition x_1 becomes unstable with respect to the compounds AB_2 at the liquidus temperature T_m. At a given temperature T_1, there is a driving force for the precipitation of the compound AB_2 given by G_1 (Fig. 6), where G_1 is defined as the driving force to form one mole of compound AB_2 in a liquid of composition x_1. In all cases here ΔG_m is equal to ΔG_1. By using heats of formation in place of free energy values, H_m^f can be similarly evaluated.

The Gibbs energy barrier to nucleation of a spherical nucleus ΔG^* can be described as

$$\Delta G^* = \frac{16\pi}{3N}(\sigma_m^3 / \Delta G_m^2) \tag{20}$$

where N is Avogadro's number and the σ_m the molar liquid/crystal interfacial energy. σ_m is directly related to the molar enthalpy of fusion H_m^f and expressed as

$$\sigma_m = \alpha H_m^f \tag{21}$$

where α is a proportional constant. H_m^f can be obtained in a similar way to evaluate ΔG_m based on bond energy values across the interface (Turnbull, 1950). Saunders and Miodownik empirically evaluated the constant α to be 0.41 (Saunders & Miodownik, 1988).

3.3 Calculation of critical cooling rates below T_0 of disordered solid phases

The expression for t in equation (16) is derived assuming that the kinetics of the liquid to crystal transformation are limited by the bulk diffusivity, which is appropriate when the crystal composition differs from that of the liquid, or at compound compositions where substantial diffusion is necessary before the correct spatial relationships that define the ordered structure of the compound are achieved. However, at the temperatures below the T_0 temperature of a disordered solid solution phase, the liquid becomes unstable with respect to a molecularly simple phase of the same composition. Consequently, no long range diffusion is necessary for the liquid to crystal and the kinetics are governed by atom motions of less than one atom in diameter. Then, the transformation is considered to be extremely difficult to suppress and this forms the T_0 criteria for GFA. In such cases, it has been suggested that the rate limiting step for crystal growth is proportional to the rate at which

atoms collide at the liquid/crystal interface, and an expression for the crystal growth rate is then given (Boettinger et al., 1984) by

$$U_c = fV_0[1-\exp(-\Delta G_m / RT)] \tag{22}$$

where V_0 is the velocity of sound in the liquid metal. This is the same form as equation (14), but with V_0 replacing the D_g/a_0. Replacing the D_n/a_0 in equation (12) with V_0 and rearranging equation (12) and (14), an expression for t is derived as

$$t \approx \frac{1}{V_0}\left\{\frac{Xa_0}{\pi f^3 N_v}\frac{\exp(\Delta G^* / kT)}{[1-\exp(-\Delta G_m / RT)]^3}\right\}^{1/4} \tag{23}$$

The value for V_0 has been taken as 1000 m/s by Saunders and Miodownik (Saunders & Miodownik, 1988), close to the a value used by Boettinger et al. (Boettinger et al., 1984) and no transformation is considered to occur below T_g.

From equations (10) to (23), the time-temperature-transformation (TTT) curve can be obtained. The critical cooling rate R_c necessary for amorphous phase formation with a melt quenching method can be evaluated from TTT curve calculated and approximated as follows

$$R_c = \frac{T_m - T_n}{5t_n} \tag{24}$$

where T_m and t_n are the temperature and time at the nose of the TTT curve, respectively, since the cooling rate calculated directly from the isothermal transformation curve is somewhat overestimated compared with that from the CCT (continuous cooling transformation) curve, the right side of Equation (24) has divided by a factor of 5 to emulate continuous cooling. In the composition range with $R_c<1\times10^{-7}$ K/s, which has been generally known to be a maximum available cooling rate for melting quenching, the amorphous phase formation may be possible.

3.4 Evaluation of glass forming ranges in alloy systems

The combined thermodynamic and kinetic approach has been undertaken to evaluate the GFA of a wide number of binary and ternary alloy systems since the pioneering work performed by Sanders and Miodownik (Saunders & Miodownik, 1988; Shim et al., 1999; Clavaguera-Mora, 1995; Tokunaga, et al., 2004; Abe, et al., 2006; Ge, et al., 2008; Palumbo, & Battezzati, 2008; Mishra & Dubey, 2009). They calculated the free energy driving forces from the thermodynamic databases, free energy barrier for nucleation and melting points, and employed this data to kinetic calculation. There is in good agreement between the predicted glass forming ranges and those experimentally observed. It is indicated that the approach has the potential to predict glass forming ability in multicomponent alloys using mainly binary input data.

The first attempts to couple kinetic models with reliable thermodynamic data using the CALPHAD methodology was performed by Saunders and Miodownik (Saunders & Miodownik, 1988). In his work, the combined thermodynamics and kinetics approach was presented in detail and undertaken to evaluate the GFA of a wide range of binary (Au-Si, Pd-Si, Ti-Be, Zr-Be, Hf-Be, Cu-Ti, Co-Zr, Ni-Zr, Cu-Zr, Ni-P, Pd-P) and ternary (Ni-Pd-P, Cu-Pd-P,

Co-Ti-Zr, Zr-Be-Hf, Ti-Be-Hf) alloy systems (Saunders & Miodownik, 1988). The TTT curves and the critical cooling rate for glass formation R_c were estimated. There is excellent agreement between the predicted and observed GFRs of binary systems, apart from the discrepancies in the Ti-Be and Cu-Ti systems. The approach was then extended to give predications for critical cooling rates in ternary and multicomponent alloys using mainly binary information. The results would appear to indicate that the combined approach takes into account a number of the major effects that govern glass formation and has the potential to predict GFA in multicomponent systems (Saunders & Miodownik, 1988).

In the work performed by Ge et al., (Ge, et al., 2008), the glass forming ability (GFA) of nine compositions of Cu-Zr and thirteen of Cu-Zr-Ti alloys in terms of critical cooling rate and fragility were evaluated by combining CALPHAD technique with kinetic approach. The driving forces for crystallization from the undercooled liquid alloys were calculated by using Turnbull and Thompson-Spaepen (TS) Gibbs free energy approximate equations, respectively. As shown in Fig. 7, time-temperature-transformation (TTT) curves of these alloys were obtained with Davies-Uhlmann kinetic equations based on classical nucleation theory. With Turnbull and TS equations, the critical cooling rates are calculated to be in the range of 9.78 $\times 10^3$-8.23$\times 10^5$ K/s and 4.32 $\times 10^2$-3.63$\times 10^4$ K/s, respectively, for Cu-Zr alloys, and 1.38$\times 10^2$-7.34$\times 10^5$ K/s and 0.64-1.36$\times 10^4$ K/s, respectively, for Cu-Zr-Ti alloys (Ge, et al., 2008).

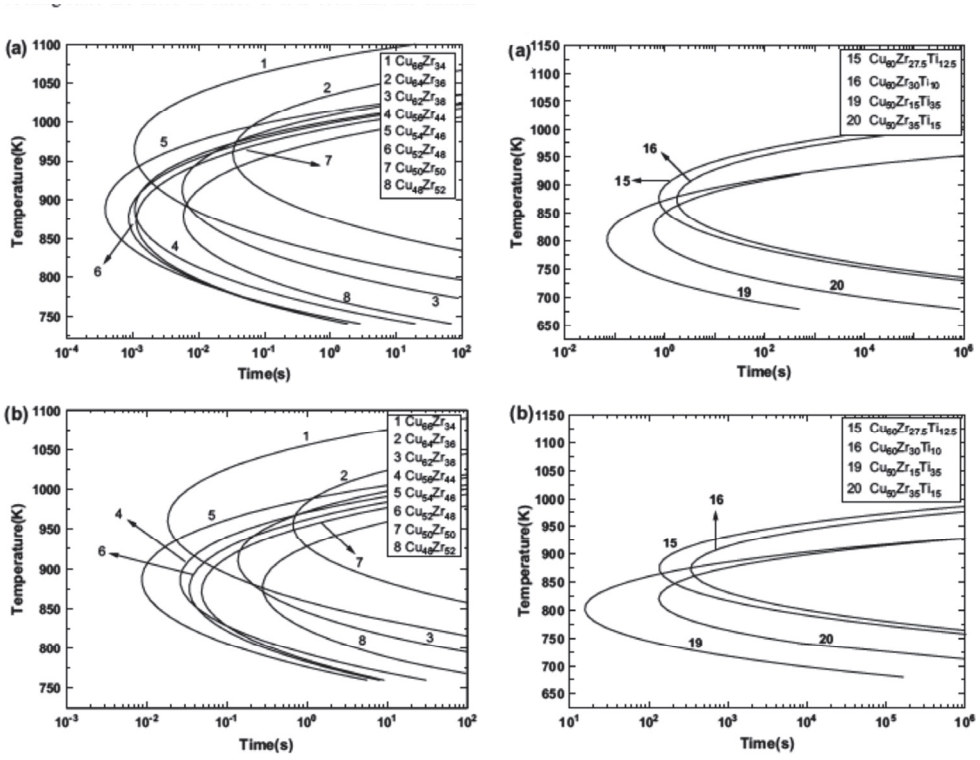


Fig. 7. Calculated TTT curve of Cu-Zr (left) and Cu-Zr-Ti (right) alloys by (a) Turnbull model and (b) TS model (from Ref. Ge, et al., 2008).

Based on topological, kinetic and thermodynamic considerations, Yang et al. (Yang, et al., 2010) have discussed the existence of the multiple maxima in GFA in a single eutectic system in Al–Zr–Ni system. It is apparent that, when taken alone, none of the factors seemed to be able to fully explain this phenomenon we have observed. It is suggested that glass formation is an intricate balance of kinetic, thermodynamic and also topological factors. Perhaps in good glass formers, all factors could come to a consensus at one composition or one compositional zone, where the best glass former(s) are located. However, for marginal glass formers like Al-based alloys, each of these factors could point to a different alloy composition, where conditions are best suited for glass formation.
Recently, Considering chemical short-range ordering and metastability of undercooled melts, Zhu and co-workers have applied a simplified quasi-kinetic approach in order to predict the GFR in binary Al-rare earth (Zhu, et al., 2004) and Al-based Al-Gd-Ni(Fe) ternary system (Zhu, et al., 2004), using CALPHAD databases. They derived the derived an expression for the reduced time $t' = t/t_{\min}$ for the formation of a minimal quantity of crystalline solid, where t represents the composition-dependent time needed for the transformation and $t_{\min}$ the minimum transformation time at a certain optimum composition:

$$t' \approx \left(\frac{exp(\Delta G^* / kT)}{[1-exp(-\Delta G_m / RT)]^3}\right)^{1/4} \tag{25}$$

This formula is in fact equivalent to Eqs. (10) to (23), except that the effect of parameters related to atomic transport is neglected. The calculated reduced times for various solid crystalline phases are then used to predict the GFR, i.e. the region of the composition space where these times are higher. A qualitative satisfactory agreement can be observed (Zhu, et al., 2004). The ability to predict the GFR of candidate metallic glass systems indicates a simple but effective approach for reducing reliance on extensive experimental trial and error in the search for new metallic glass systems (Zhu, et al., 2004).

4. Conclusion

Search for new bulk metallic glasses (BMGs) system or composition by predicting the GFA of an alloy system is of interesting and theoretical and practical significance. In this chapter, the progress on the calculation or predication the glass forming ability by thermodynamics approach or a combined thermodynamics and kinetics approach have been reviewed. It is found that a good agreement between the predicated glass forming ability and those experimentally observed has been obtained. It is indicated that the thermodynamic approach developed in the literature has proved useful to predict the glass forming ability of a number of alloys system. It has revealed that the combined thermodynamics and kinetics approach has the advantage to predict the glass forming ability of the multicomponent alloys using the reliable database of binary system assessed by CALPHAD method. It has been accepted that the thermodynamic approach and/or the combined thermodynamic and kinetic approach are effective ways for the prediction of the GFA of metallic glass alloys.

5. Acknowledgment

The financial support from the National Natural Science Foundation of China (NSFC) (Grant No. 50861006) and the Natural Science Foundation of Guangxi (Grant No. 0991002Z, 2011GXNSFC018002) are greatly acknowledged.

6. References

Abe, T.; Shimono M.; Ode, M. & Onodera, H. (2006). Thermodynamic modeling of the undercooled liquid in the Cu–Zr system. *Acta Materialia,* Vol. 54, No. 4, (February 2006), pp. 909–915, ISSN 1359-6454

Bakker, H. (1988). Enthalpies in Alloys: Miedema's Semi-Empirical Model. Trans Tech, ISBN 978-087-8497-83-6, Uetikon-Zurich, Switzerland

Basu, J. ; Murty, B. S & Ranganathan, S. (2008). Glass forming ability: Miedema approach to (Zr, Ti, Hf)–(Cu, Ni) binary and ternary alloys. *Journal of Alloys and Compounds*. Vol. 465, No. 1/2, (October 2008), pp. 163–172, ISSN 0925-8388

Battezzati, L.; Baricco, M.; Riontino, G. & Soletta, I. (1990). Thermodynamic evaluation of the Cu-Ti system in view of solid state amorphization reactions. *Journal de Physique Colloques*. Vol, 51, No, C4, (1990), pp. 79- 85, ISSN 0449-1947

Bhatt, J.; Jiang, W.; Xia, J.; Wang, Q. Chuang, D. & Murty, B.S. (2007). Optimization of bulk metallic glass forming compositions in Zr-Cu-Al system by thermodynamic modeling. *Intermetallics*, Vol. 15, No. 5/6, (May/June 2007), pp. 716-721, ISSN 0966-9795

Bo, H.; Wang J.; Jin, S.; Qi, H. Y.; Yuan, X. L.; Liu L. B. & Jin, Z.P. (2010). Thermodynamic analysis of the Al–Cu–Zr bulk metallic glass system. *Intermetallics*, Vol. 18, No. 12, (December 2010), pp. 2322-2327, ISSN 0966-9795

Boom, R. ; De Boer, M. R. & Miedema, A. R. (1976). On the heat of mixing of liquid alloys. *Journal of Less-Common Metetals*, Vol. 45, No. 2, (March 1976), pp. 237-245, ISSN 0925-8388

Boettinger, W.J. & Perepezko, J.H. (1993) in: *Rapidly Solidified Alloys*, Liebermann, H. H., (Ed.), Dekker M., p. 17, ISBN 978- 082-4789-51-2, New York,

Clavaguera-Mora, M.T. (1995). Glass materials: thermodynamic and kinetic quantities. *Journal of Alloys and Compounds*, Vol. 220, No. 1/2, (April 1995), pp. 197-205, ISSN 0925-8388

Cocco, G.; Soletta, I.; Battezzati, L.; Baricco, M. & Enzo, S. (1990). Mechanical alloying of the Al-Ti system. *Philosophical Magazine B*, Vol. 61, No. 4, (April 1990), pp. 473-479, ISSN 1364-2812

Coehoorn, R.; Van der Kolk, G. J. ; Van den Broek, J. J. ; Minemura, T. & Miedema, A. R. (1988). Thermodynamics of the stability of amorphous alloys of two transition metals. *Journal of Less-Common Metetals*, Vol. 140, No. 1/2, (June 1988), pp. 307-316, ISSN 0925-8388

Chen, Q.J.; Shen, J.; Fan H.B. ; Sun, J.F. ; Huang, Y.J. & Mccartney, D.G. (2005). Glass-forming ability of an Iron-based alloy enhanced by Co addition and evaluated by a new criterion. *Chinese Physics Letters* , Vol. 22, No.7, (July 2005), pp.1736–1738, ISSN 0256-307X

Choi,Y. H. ; Busch, R. & Johnson, W. L. (1998). The effect of silicon on the glass forming ability of the $Cu_{47}Ti_{34}Zr_{11}Ni_8$ bulk metallic glass forming alloy during processing of composites. *Journal of Applied Physics*, Vol. 83, No. 12, (June 1998), pp. 7993-7997. ISSN 0021-8979

Delamare, J. ; Lemarchand, D. & Vigier P. (1994), Structural investigation of the metastable compound A1 in an as-cast Fe---Nd eutectic alloy. *Journal of Alloys and Compouds*, Vol. 216, No. 2, (January 1994), pp. 273-280, ISSN 0925-8388

De Boer, F. R. ; Boom, R. ; Mattens, W. C. M ; Miedema, A. R. & Niessen, A. K. (1988). Cohesion in Metals: Transition Metals Alloys. North-Holland, ISBN 978-044-4870-98-8, Amsterdam, The Netherlands

Eshelby, J. D. (1954). Distortion of a Crystal by Point Imperfections. *Journal of Applied Physics*, Vol. 25, No. 2, (February 1954), pp. 255-261. ISSN 0021-8979

Eshelby, J. D. (1956), The continuum theory of lattice defects. *Solid State Physics : Advances in Research and Applications*. *Frederick* Seitz and David Turnbull, (Ed.), Vol. 3, (1956), pp. 79-144, Elsevier, ISBN 978-0-12-374292-6, Amsterdam, the Netherlands

Friedel, J. (1954). Electronic structure of primary solid solutions in metals. *Advances in Physics*, Vol. 3, No. 12, (October 1954), pp. 446-507, ISSN 0001-8732

Fan, G. J.; Choo, H. & Liaw, P. K. (2007). A new criterion for the glass-forming ability of liquids. *Journal of Non-Crystalline Solids*, Vol. 353, No.1, (January 2007), pp. 102−107, ISSN 0022-3093

Gao, M. C.; Hackenberg, R.E. & Shiflet, G.J. (2003). *Journal of Alloys and Compounds*, Vol. 353, No. 1/2, (April 2003), pp. 114–123, ISSN 0925-8388

Ge, L.; Hui, X.; Wang, E.R.; Chen, G.L.; Arroyave, R. & Liu, Z.K. (2008). Prediction of the glass forming ability in Cu-Zr binary and Cu-Zr-Ti ternary alloys. *Intermetallics*, Vol. 16, No. 1, (January 2008), pp. 27-33, ISSN 0966-9795

Guo, F. Q.; Poon, S.J. & Shiflet, G.J. (2000). Investigation of glass formability in Al-based multinary alloys. *Scripta Materialia*, Vol. 43, No. 12, (November 2000), pp. 1089–1095, ISSN 1359-6462

Guo, F. Q.; Poon, S. J. & Shiflet, G. J. (2000), Glass formability in Al-based multinary alloys, *Materials Science Forum*, Vol. 331-337, (2000), pp. 31–42, ISSN 1662-9752

Guo, S. & Liu, C.T. (2010). New glass forming ability criterion derived from cooling consideration. *Intermetallics*, Vol.18, No. 11, (November 2010) pp. 2065-2068, ISSN 0966-9795

Hackenberg, R. E.; Gao, M. C.; Kaufman, L. & Shiflet, G.J. (2002). Thermodynamics and phase equilibria of the Al-Fe-Gd metallic glass-forming system. *Acta Materialia*, Vol. 50, No. 9, (May 2002), pp. 2245–2258, ISSN 1359-6454

He, Y.; Poon, S. J. & Shiflet, G. J. (1988). Synthesis and properties of metallic that contain aluminum. *Science,* Vol. 241, No. 4873, (September 1988), pp. 1640–1642. ISSN 0036-8075

Inoue, A.; Zhang, T. & Masumoto, T. (1993). Glass-forming ability of alloys. *Journal of Non-Crystalline Solids*, Vol.156-158, No. 2, (May 1993), pp.473−480, ISSN 0022-3093

Inoue, A.; Zhang, T. & Takeuchi, A. (1998). Ferrous and Nonferrous Bulk Amorphous Alloys, *Materials Science Forum*, Vol. 269-272, (1998), pp. 855-864, ISSN 0255-5476

Inoue, A. ; Zhang, W. ; Zhang, T. & Kurosaka K. (2001). High-strength Cu-based bulk glassy alloys in Cu–Zr–Ti and Cu–Hf–Ti ternary systems, *Acta Materialia*, Vol. 49, No.14, (August 2001), pp. 2645–2652, ISSN 1359-6454

Inoue, A.; Ohtera, K.; Tsai, A. P. & Masumoto, T. (1988). Aluminum-based amorphous alloys with tensile strength above 980 MPa (100 kg/mm2). *Japanese Journal of Applied Physics*, Vol. 27, No. 4, (April 1988), pp. L479–L482, ISSN 0021-4922

Inoue, A. (1998). Amorphous, nanoquasicrystalline and nanocrystalline alloys in Al-based systems. *Progress in Materials Science*, Vol. 43, No. 5, (December 1998), pp. 365–520, ISSN 0079-6425

Ji, X.L.; & Pan, N.Y. (2009). A thermodynamic approach to assess glass–forming ability of bulk metallic glasses. *Transaction of Nonferrous Metals Society of China*, Vol. 19, No. 5 (October 2009), pp.1271– 1279, ISSN 1003-6326

Ji, X.L.; Pan, Y. & Ni, F. (2009). A thermodynamic criterion for predicting glass-forming ability in binary metallic glasses. *Materials and Design*, Vol. 30, No. 3, (March 2009), pp. 842–845, ISSN 0261-3069

Jonson, W. L. (1999). Bulk glass-forming metallic alloys: science and technology. *MRS Bulletin*. Vol. 24, No. 10, (October 1999), pp. 42-56, ISSN 0883-7694

Kaufman, L. & Bernstein H. (1970). Computer calculation of phase diagrams. Academic Press; ISBN 978-1-84569-215-5, NY, USA

Kang, D. H. & Jung, I. H. (2010). Critical thermodynamic evaluation and optimization of the Ag–Zr, Cu–Zr and Ag–Cu–Zr systems and its applications to amorphous Cu–Zr–Ag alloys. *Intermetallics*, Vol. 18, No. 5, (May 2010), pp. 815–833, ISSN 0966-9795

Kawazoe, Y.; Masumoto, T.; Suzuki, K.; Inoue, A.; Tsai, A. P.; Yu, J. Z.; Aihara, J.T. & Nakanomyo, T. (1997). Nonequilibrium phase diagrams of ternary amorphous alloys. In: *Landolt– Bönstein, New Series III: Condensed Matter,* Y. Kawazoe, J. Z. Yu, A. P. Tsai & T. Masumoto, (Ed.), vol. 37A, p. 112, ISBN 978-3-540-65664-7

Kim, D.; Lee B. J. & Kim, N. J. (2004). Thermodynamic approach for predicting the glass forming ability of amorphous alloys. *Intermetallics*. Vol. 12, No. 10-11, (October-November 2004), pp. 1103–1107, ISSN 0966-9795

Kim, D. ; Lee, B. J. & Kim, N. J. (2005). Prediction of composition dependency of glass forming ability of Mg–Cu–Y alloys by thermodynamic approach. *Scripta Materialia*, Vol. 52, (2005), pp. 969–972, ISSN 1359-6462

Kim, Y. K.; Soh, J. R.; Kim H. S. & Lee, H. M. (1998). Thermodynamic prediction of glass forming range in Al-Mg-REM ternary system. *CALPHAD*. Vol. 22, No. 2, (June 1998). pp. 221-2301, ISSN 0364-5916

Kumar, K. C. H.; Ansara, I.; Wollants, & P. Delaey, L. (1996). Thermodynamic optimisation of the Cu-Ti system. *Zeitschrift für Metallkunde*. Vol. 87, No. 8, (August 1996), pp. 666-672, ISSN 0044-093

Klement, W. ; Willens, R.H. & Duwez, P. (1960). Noncrystalline structure in solidified gold-silicon alloys. *Nature*, Vol.187, No.4740, (September 1960), pp. 869-870, ISSN 0028-0836

Lin, X. H. & Johnson, W.L. (1995). Formation of Ti-Zr-Cu-Ni bulk metallic glasses. *Journal of Applied Physics*, Vol. 78, No. 11, (December 1995), pp. 6514-6519. ISSN 0021-8979

Lu, Z.P. & Liu, C.T. (2002). A new glass-forming ability criterion for bulk metallic glasses. *Acta Materialia*, Vol.50, No.13, (August 2002), pp.3501–3512, ISSN 1359-6454

Liu, X.J.; Hui, X.D. & Chen, G.L. (2008). Thermodynamic calculation and experimental investigation of glass formation in Zr-Ni-Ti alloy system. Intermetallics, Vol. 16, No. 2, (February 2008), pp. 262-266, ISSN 0966-9795

Miedema, A. R. ; Boom, R. & De Boer, M. R. (1975). On the heat of formation of solid alloys. *Journal of Less-Common Metals*, Vol. 41, No. 2 (July 1975), pp. 283-298, ISSN 0925-8388

Mishra, R. K. &. Dubey S. (2009). Glass forming ability of materials: A thermodynamic approach. *Journal of Non-Crystalline Solids.* Vol. 355, No. 45-47 (November 2009), pp. 2199–2204, ISSN 0022-3093

Murty, B. S.; Ranganathan, S. & Rao, M. M. (1992). Solid state amorphization in binary Ti-Ni, Ti-Cu and ternary Ti---Ni---Cu system by mechanical alloying. *Materials Science and Engeering A*, Vol. 149, No. 2, (January 1992), pp. 231-240, ISSN 0921-5093

Nagarajan, R. & Ranganathan, S. (1994). A study of the glass-forming range in the ternary Ti-Ni-Al system by mechanical alloying. *Materials Science and Engeering A*, Vol, 179/180, No. 1, (May 1994), pp. 168–172, ISSN 0921-5093

Niessen, A. K. ; Boer, F.R. de ; Boom, R. ; Chatel, P.F. de ; Mattern, W.C.M. & Miedema, A.R. Model predictions for the enthalpy of formation of transition metal alloys II. *CALPHAD*, Vol. 7, No. 1, (Januay-March 1983), pp. 51–70, ISSN 0364-5916

Oliveira, M. F.; Aliaga, L.C. ; Bolfarini, R.C. ; Botta, W.J. & Kiminami, C.S. (2008). Thermodynamic and topological instability approaches for forecasting glass-forming ability in the ternary Al–Ni–Y system. *Journal of Alloys and Compounds,* Vol. 464, No. 1/2, (September 2008), pp. 118–121, ISSN 0925-8388

Porter, D. A. & Easterling K. E. (1992). Phase transformation in metals and alloys. Chapman & Hall, ISBN 978- 074-8757-41-1 , London, UK

Palumbo, M.; Cacciamani, G.; Bosco, E. & Baricco, M. (2001). Thermodynamic analysis of glass formation in Fe-B system, *CALPHAD*, Vol. 25, No. 4, (December, 2001), pp. 625-637. ISSN 0364-5916

Palumbo, M. & Battezzati, L. (2008). Thermodynamics and kinetics of metallic amorphous phases in the framework of the CALPHAD approach. CALPHA, Vol. 32, No. 2, (June 2008), pp. 295–314, ISSN 0364-5916

Rao, B. S.; Bhatt, J. & Murty, B. S. (2007). Identification of compositions with highest glass forming ability in multicomponent systems by thermodynamic and topological approaches. *Materials Science and Engineering A*, Vol. 449-451, No. 1, (March 2007), pp. 211-214, ISSN 0921-5093

Ramachandrarao, P.; Cantor, B. & Cahn, R.W. (1977). Free volume theories of the glass transition and the special case of metallic glasses. *Journal of Materials Science*, Vol. 12, No. 12, (December 1977), pp. 2488-2502, ISSN 0022-2461

Saunders, N. & Miodownik, A. P. (1988). Evaluation of glass forming ability in binary and ternary metallic alloy systems - an application of thermodynamic phase diagram calculation. *Materials Science and Technology*, Vol. 4, No. 9, (September 1988), pp. 768-777, ISSN 0267-0836

Saunders, N. & Miodownik, A.P. (1986). Thermodynamic aspects of amorphous phase formation. *Journal of Materials Research*, Vol. 1, No. 1, (January 1986), pp. 38-46. ISSN 0884-2914

Saunders, N. & Miodownik, A.P. (1985). The use of free energy vs composition curves in the prediction of phase formation in codeposited alloy thin films. *CALPHAD*, Vol. 9, No. 3, (September 1985), pp. 283-290, ISSN 0364-5916

Schwarz, R. B.; Nash, P. & Turnbull, D. (1987). The use of thermodynamic models in the prediction of the glass-forming range of binary alloys. *Journal of Materials Research*. Vol. 2, No. 4, (July 1987), pp. 456-460, ISSN 0884-2914

Shao, G.; Lu, B.; Liu, Y.Q. & Tsakiropoulos P. (2005). Glass forming ability of multi-component metallic systems. *Intermetallics*, Vol. 13, No. 3/4, (March/April 2005), pp. 409–414, ISSN 0966-9795

Shim, J. H.; Chung, H.-J. & Lee, D. N. (1999). Calculation of phase equilibria and evaluation of glass-forming ability of Ni–P alloys. *Journal of Alloys and Compounds*, Vol. 282, No. 1/2, (January 1999), pp. 145-181, ISSN 0925-8388

Simozar, S. & Alonso, J.A. ; On the Factors Controlling Glass Forming Ability of Metallic Alloys Formed by Fast Liquid Quenching. *Physica Status Solidi (a)*, Vol. 81, No.1, (January 1984), pp. 55–61, ISSN 1862-6319

Sun, S.P.; Yi, D. Q.; Liu, H.Q.; Zang B. &Jiang, Y. (2010). Calculation of glass forming ranges in Al–Ni–RE (Ce, La, Y) ternary alloys and their sub-binaries based on Miedema's model. *Journal of Alloys and Compounds*, Vol. 506, No. 1, (September, 2010) 377–387, ISSN 0925-8388

Takeuchi, A. & Inoue, A. (2000). Calculations of Mixing Enthalpy and Mismatch Entropy for Ternary Amorphous Alloys. *Materials Transactions JIM*, Vol. 41, No. 11, (November 2000), pp. 1372-1378, ISSN 0916-1821

Takeuchi, A. & Inoue, A. (2001). Calculations of Amorphous-Forming Composition Range for Ternary Alloy Systems and Analyses of Stabilization of Amorphous Phase and Amorphous-Forming Ability. *Materials Transactions JIM*, Vol, 42, No. 7 (July 2001) 1435–1444, ISSN 0916-1821

Takeuchi, A. & Inoue, A. (2001). Thermodinamic and kinetic analyses of formation of amorphous and nanocrystalline alloys with the aid of computer and database. *Journal of Optoelectronics and Advanced Materials*, Vol, 6, No. 2, (June 2004), pp. 533–540, ISSN 1454-4164

Tang, C. Y.; Du, Y.; Wang, J.; Zhou, H. Y.; Zhang, L.J.; Zheng, F.; Lee, J. & Yao, Q. (2010). Correlation between the thermodynamics and the glass forming ability in the Al-Ce-Ni system, *Intermetallics*, Vol. 18, No. 5, (May 2010), pp. 900- 906, ISSN 0966-9795

Tokunaga, T.; Ohtani, H. & Hasebe, M. (2004). Thermodynamic evaluation of the phase equilibria and glass-forming ability of the Fe–Si–B system. *CALPHAD*, Vol. 28, No. 4, (December 2004), pp. 354–362. ISSN 0364-5916

Turnbull, D. (1950). Formation of crystal nuclei in liquid metals. Journal of Applied Physics, Vol. 21, No. 11, (November 1950), pp. 1022-1028, ISSN 0021-8979

Turnbull, D.(1969). Under what conditions can a glass be formed? *Contemporary Physics*, Vol.10, No. 5, (October 1969), pp.473-488, ISSN 0010-7514

Uhlmann, D.R. (1972). A kinetic treatment of glass formation. *Journal of Non-Crystalline Solids*, Vol. 7, No. 4 (May 1972), pp.337-348, ISSN 0022-3093

Van der Kolk, G. J. ; Miedema A. R. & Niessen, A. K. (1988). On the composition range of amorphous binary transition metal alloys. *Journal of Less-Common Metetals*, Vol. 145, No. 1/2, (December 1988), pp. 1-17, ISSN 0925-8388

Wang, T.L.; Wang, W.C.; Li, J.H. & Liu, B.X. (2010). Glass formation of the Fe–Hf system studied by thermodynamic calculation and ion beam mixing. *Journal of Alloys and Compounds,* Vol. 493, No 1/2. (March 2010), pp. 154–157, ISSN 0925-8388

Wang, T.L.; Li, J.H. & Liu, B.X. (2009). Proposed thermodynamic method to predict the glass formation of the ternary transition metal systems. *Physical Chemistry Chemical Physics*, Vol. 11, No, 14, (February 2009), pp. 2371-2373, ISSN 1463-9076

Wang, T.L. & Liu B.X. (2009). Glass forming ability of the Fe–Zr–Cu system studied by thermodynamic calculation and ion beam mixing. *Journal of Alloys and Compounds*, Vol. 481, No. 1/2, (July 2009), pp. 156–160, ISSN 0925-8388

Xia, L.; Fang S. S.; Wang, Q.; Dong, Y. D. & Liu, C.T. (2006). Thermodynamic modeling of glass formation in metallic glasses. *Applied Physic Letters*, Vol. 88, No. 17, (April 2006), pp. 171905 -171905-3, ISSN 0003-6951

Xia, L.; Li, W. H.; Fang, S. S.; Wei B. C. & Dong, Y. D. (2006). Binary Ni-Nb bulk metallic glasses. *Journal of Applied Physics*, Vol. 99, No. 2, (January 2006), pp. 026103-1-3, ISSN 0021-8979

Yang, H.; Lim, K.Y. & Li Y. (2010). Multiple maxima in glass-forming ability in Al–Zr–Ni system. *Journal of Alloys and Compounds*, Vol. 489, No.1, (January 2010), pp. 183–187, ISSN 0925-8388

Yuan, Z.Z.; Bao, S.L. ; Lu, Y.; Zhang, D.P. & Yao, L. (2008). A new criterion for evaluating the glass–forming ability of bulk glass forming alloys. *Journal of Alloys and Compouds*, Vol. 459, No. 1/2, (July 2008), pp. 251–260, ISSN 0925-8388

Zhu, A.; Shiflet, G. J. & Miracle, D. B. (2004). Glass forming ranges of Al–rare earth metal alloys: thermodynamic and kinetic analysis. *Scripta Materialia*, Vol. 50, No. 7, (April 2004), pp. 987–991, ISSN 1359-6462

Zhu, A.; Poon, S. J. & Shiflet, G. J. (2004). On glass formability of Al–Gd–Ni (Fe). *Scripta Materialia*, Vol. 50, No. 12, (June 2004), pp. 1451–1455, ISSN 1359-6462

4

Information Thermodynamics*

Bohdan Hejna
Institute of Chemical Technology Prague
Department of Mathematics, Prague
Czech Republic

1. Introduction

A working medium of any heat cycle can be considered as a thermodynamic, average-value model or, as such, as a realization of an information transfer channel. More generally we could use some kind of *directly shared energy* (6). The cycle models an information transfer process in the channel, which enables input messages to be transferred *repeatedly* and uses transformations of such energies.
Our thermodynamic-information derivation based on a heat cycle demonstrates the fact that it is impossible, in such a type of channel, for the *bound* (2) information contained in an input message to be transferred without its (average) loss, *even when the ideal case of a noiseless channel* is considered. This loss of information is the *necessary* condition for such a repeatable transfer of messages. Such information transfer can be worsened only by heat dissipation of energies, which means by a *noise* heat generated by the irreversible processes in the channel, *subtractive* in this case. This channel is described by a transformer of input heat, which has non-ideal properties (inner friction).[1]
To explain the Gibbs paradox and to prove the II. P.T. we use the concept of *bound* information (2; 15). This method is identical with introducing the *Boltzman* function of statistical physics. Its negative value, determined by a *detailness* of our description of an observed system, is proved to be a value of Clausius entropy (in a certain substitute equivalent equilibrium thermodynamic way (19)). We show that a physical realization of such observation is equivalent to a scheme of a relevant (reversible) heat cycle (7). Its properties are expressible in terms of the Gibbs paradox.
The bound entropies of our realized observation, the input, the output and the conditional are, as the *free* ones (2; 15), associated by the *channel equation* (3; 20). This equation is, in this sense, an *information* description of a cyclical transformation of heat energy of an observed, measured system (7). In this way, its derivation proves the II. P.T. and gives its information forms as well (7). It is a *most general* formulation of the II. P.T. and yields in the *Equivalence Principle* of the I., II. and III. P.T.
In this article we are thinking, in a *free way*, about possible applications of the Information Thermodynamics point of view in *biology*.
We can consider a principle loss of information (structure) within the process of generation of cells by their duplication (dividing). Within any dividing of a cell (*the predecessor*) its *follower*

*Supported by the grant of Ministery of Education of Czech Republic MSM 6046137307.
[1] This text is a composite of the papers (7; 8); also, it is the very core of the book (9).

(equivocant) is generated and the a distortion of the *duplicated* (copied) structure (information) of the parent cell arises. This loss is measurable by the quantity of (average) information amount. It is a loss of part of message being copied (transferred), a loss of information within this process of such an information transfer. (The whole *structure of the cell is a message*, also including 'a program' for its functionality in a texture.) It is a mechanism of oldering by a 'tooth of time'. This mechanism can be described in a functional way by *'carnotized' model of information transfer - direct Carnot cycle viewed informationally*.
Within the pathological proliferation of cells the opposite situation arises. In this case the cells with a precise structure (but, of an another type in comparison with the normal structure of its own original and 'normal' type is). But, this grow of structure in a certain locality in the whole organism is paid by a pumping off energy from an environment of this locality. This situation is describable again by our information-thermodynamic model, but by a *reverse* this time.

2. Notion of entropy

2.1 Information entropy

With a stochastic phenomenon ξ, the probability of which is $p(\xi)$, its "own" *information*, *information amount* $\mathcal{I} \stackrel{\text{Def}}{=} -K \cdot \log_z p(\xi)$ is associated. The quantity $\mathcal{I}$ is measured in *information* units *bit* when $K = 1$, $z = 2$ or *nat* when $K = 1$, $z = e$ or *Hartley* when $K = 1$, $z = 10$ is used. Also it is possible to measure it in *thermodynamic* units: *Boltzman* when $K = k$, k is *Boltzman constant*[2] and $z = 10$; for $K = k$ and $z = e$ it is *Clausius*. Then it is valid that $k \cdot Hartley = Boltzman$ and $k \cdot nat = Clausius$.
Let the stochastic phenomenon $\xi \in \mathcal{X}$ be a *realization* of a *discrete* stochastic quantity X with a *probability distribution* $[p(\cdot)]_{\mathcal{X}}$, $\mathcal{X}$ is its *selecting space*,

$$X \stackrel{\text{Def}}{=} [\mathcal{X},\ p(\cdot)],\ \xi \in \mathcal{X},\ \sum_{\xi \in \mathcal{X}} p(\xi) = 1,\ 0 \leq p(\xi) \leq 1. \tag{1}$$

The *expectation value* $H(X)$ of the stochastic quantity $[\{\mathcal{I}\}_{\mathcal{X}},\ p(\cdot)]$,

$$H(X) \stackrel{\text{Def}}{=} E[\{\mathcal{I}\}_{\mathcal{X}},\ p(\cdot)], \tag{2}$$

is called *Shannon*, *(information) entropy* (3) of X. With the denotation $\xi \stackrel{\triangle}{=} \xi_j$ and $p(\xi_j) \stackrel{\triangle}{=} p_j$, we write

$$H(X) = -K \cdot \sum_j p_j \log_z p_j,\ \sum_j p(j) = 1. \tag{3}$$

For a *continuous* stochastic quantity X the *differential* entropy is defined (15),

$$H(X) \stackrel{\text{Def}}{=} -K \cdot \int_{\mathcal{X}} w(\xi) \log_z w(\xi) \mathrm{d}\xi;\ \int_{\mathcal{X}} w(\xi) \mathrm{d}\xi = 1,\ w(\xi) \geq 0; \tag{4}$$

$\mathcal{X} \subseteq \Re$ is a *non-degenerated* interval, $w(\xi)$ is a *density* of probability of X, $\xi \in \mathcal{X}$.

[2] *Gay-Lussac* experiment, $k = \dfrac{R}{N_A}$, R is *gas constant*, N_A is *Avogadro number* (11)

2.1.1 Channels of information theory

A *discrete transfer channel* $\mathcal{K}$ is defined as the tripartite structure

$$\mathcal{K} \stackrel{\text{Def}}{=} [(X,\ p(\cdot|\cdot),\ Y)],\ \text{where,} \tag{5}$$

X is an *input* discrete stochastic quantity, a *source of messages*, with its selecting space $\mathcal{X} = \{x_1, x_2, ...\}$ and a probability distribution $[q(\cdot)]_{\mathcal{X}} \stackrel{\triangle}{=} [q(\cdot)]_{\{i\} \subseteq \{1,2,\ ...\}}$;
Y is an *output* discrete stochastic quantity, a *receiver* of messages. Its selecting space is $\mathcal{Y} = \{y_1, y_2, ...\}$ and its probability distribution is $[p(\cdot)]_{\mathcal{Y}} \stackrel{\triangle}{=} [p(j)]_{\{j\} \subseteq \{1,2,\ ...\}}$;
the quantity $p(\cdot|\cdot)$ is a probability distribution of an *error* in channel $\mathcal{K}$, the distribution of probability of receiving (measuring, observing) $y_j \in \mathcal{Y}$ on the channel output when the input $x_i \in \mathcal{X}$ has been sent (output j for input i).
Shannon entropy of X or Y respectively, is called the *input* or the *output entropy*, and, following the definition (3) with $K = 1$ and $z = e$,

$$H(X) = -\sum_i q(i) \ln q(i),\ \ H(Y) = -\sum_j p(j) \ln p(j). \tag{6}$$

The information amount $\mathcal{I}_{i|j}$, contained in a phenomenon with the conditional probability $p(i|j)$, is defined by the equality $\mathcal{I}_{i|j} \stackrel{\text{Def}}{=} -\ln p(i|j)$, and, is called the *loss information*. For its average information, its (*conditional*) entropy, we figure:

$$H(X = x_i|Y) = -\sum_j p(i|j) \ln p(i|j), \tag{7}$$
$$H(X|Y) = \sum_i q(i) \sum_j p(i|j) \ln p(i|j) = -\sum_i \sum_j p(i,j) \ln p(i|j),$$

where $p(i,j)$ is the *simultaneous* probability of a phenomenon i and j. The conditional entropy $H(X|Y)$ is called the *loss* (*residual*) entropy.
Similarly, $\mathcal{I}_{j|i} \stackrel{\text{Def}}{=} \ln p(j|i)$. This quantity is called the *noise* information. Its average information, the conditional entropy, is called the *noise* entropy in channel $\mathcal{K}$,

$$H(Y|X) = -\sum_i \sum_j p(i,j) \ln p(j|i). \tag{8}$$

The *usable information* $\mathcal{I}_{i;j}$ in an output message y_j about an input x_i is defined by the equality

$$\mathcal{I}_{i;j} \stackrel{\text{Def}}{=} \mathcal{I}_i - \mathcal{I}_{i|j} = -\ln q(i) + \ln p(i|j) = \ln \frac{p(i|j)}{q(i)}$$

The average information amount of $\mathcal{I}_{i;j}$ in the message j about the message i is the quantity

$$T(X;Y) \stackrel{\text{Def}}{=} \sum_i \sum_j \mathcal{I}_{i;j} p(i,j) = \sum_i \sum_j p(i,j) \ln \frac{p(i|j)}{q(i)} = -H(X|Y) + H(X), \tag{9}$$

called the average usable *transferred information*, *transinformation* (15). Similarly,

$$\mathcal{I}_{j;i} = \mathcal{I}_j - \mathcal{I}_{j|i} = -\ln p(j) + \ln p(j|i) = \ln \frac{p(j|i)}{p(j)},$$

and then

$$T(Y;X) \stackrel{\text{Def}}{=} \sum_i \sum_j \mathcal{I}_{j;i} p(i,j) = \sum_i \sum_j p(i,j) \ln \frac{p(j|i)}{p(j)} = -H(Y|X) + H(Y). \tag{10}$$

For a simultaneous distribution $p(i,j)$ $[= p(j,i)]$, $i,j \in \{1,2,\ ...\}$ we define the *simultaneous* entropies

$$H(X,Y) \stackrel{\text{Def}}{=} -\sum_i \sum_j q(i)p(j|i) \ln[q(i)p(j|i)] = H(X) + H(Y|X), \tag{11}$$

$$H(Y,X) \stackrel{\text{Def}}{=} -\sum_i \sum_j p(j)p(i|j) \ln[p(j)p(i|j)] = H(Y) + H(X|Y)$$

and then $H(Y,X) = H(X,Y)$.
A *continuous transfer channel* $\mathcal{K}$ is defined in a similar formal way as a discrete one, but the quantities X, Y, $(X|Y)$, $(Y|X)$ are continuous. Adequate differential entropies are defined (15).
It follows from equalities (9)-(11) that, both in a discrete and in a continuous transfer channel, the *law of entropy (information) conservation*, the *channel equation* is valid,

$$H(X) + H(Y|X) = H(Y) + H(X|Y), \ \ H(X) - H(X|Y) = H(Y) - H(Y|X). \tag{12}$$

Or, also, the *symmetry* of transinformation, $T(X;Y) = T(Y;X)$, is valid.

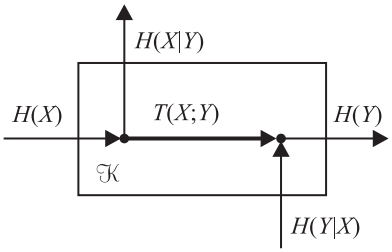


Fig. 1. The relations among entropies of an information tranfer channel

2.2 Thermodynamic entropy

2.2.1 Definition of equilibrium macroscopic thermodynamics

With an arbitrary thermodynamic system $\mathcal{A}$, in the *thermodynamic equilibrium*, the *macroscopic (global, extensive)* and thus the *additive* quantity, called the *thermodynamic (Clausius) entropy*, denoted as S, is associated. But, the *phenomenologic (macroscopic, classic)* thermodynamics defines its change ΔS only, generated by a reversible exchange of heat Δq at an *absolute* constant temperature (in $^\circ$K) between the system $\mathcal{A}$ and its *environment*. Or, it should be defined by another change of heat energy Δq, expressible by an equivalent *substitute* reversible way[4] at a certain constant temperature $\Theta > 0$ (19). This change is defined by the *Clausius equation*

$$\Delta S \stackrel{\text{Def}}{=} \frac{\Delta q}{\Theta}, \ \ \Theta > 0. \tag{13}$$

[4] or, by a possibility of such a type of changes

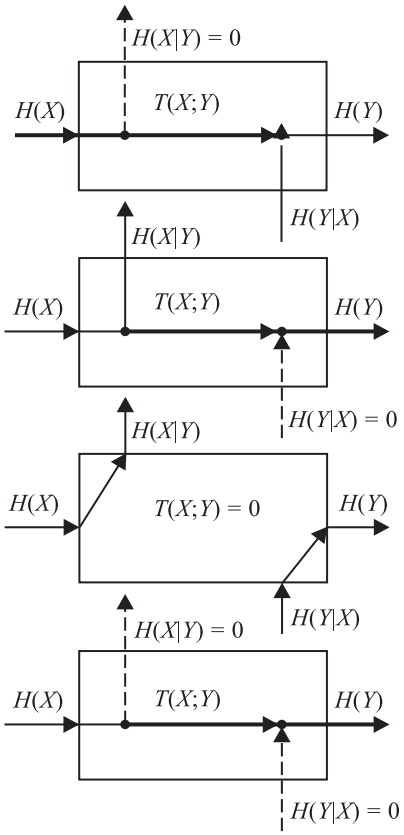


Fig. 2. The information transfer channel: without losses, without noise, disrupted, without noise and losses

By integration of $\mathrm{d}S \stackrel{\triangle}{=} \Delta S$, when $\delta q \stackrel{\triangle}{=} \Delta Q$, the entropy S can be figured, but, with such an exception of an additive integration constant S_0. For the entropy S, as a function of a temperature $\Theta > 0$, it is to be valid that

$$S = S(\Theta) = \int \frac{\delta q}{\Theta} = \sigma(\Theta) + S_0, \tag{14}$$

where Θ^{-1} is *Pfaffian* integrating factor (5; 16).
The quantity S_0 is a non-zero constant, independent from *state variables* of system $\mathcal{A}$, but dependent on *matter units* (*mol*, *kmol*) of $\mathcal{A}$. Then, $S_0 \stackrel{\triangle}{=} S_0(n) \neq 0$, where n is a number of matter units of $\mathcal{A}$. Not respecting this fact leads to the known *Gibbs paradox* (13; 14; 18). A certain *mathematical* solution of this phenomenon will now be revised shortly.

2.2.2 Gibbs paradox

Just by the (thought) "dividing" of an equilibrium system $\mathcal{A}$ by *diaphragmas* (18), without any influence on its thermodynamic (macroscopic) properties, a non-zero difference of its entropy,

before and after its "dividing", is evidenced.
Let us consider a thermodynamic system $\mathcal{A}$ in volume V and with n matter units of ideal gas in the thermodynamic equilibrium. The *state equation* of $\mathcal{A}$ is $pV = nR\Theta$. For an elementary change of the *internal* energy U of $\mathcal{A}$ we have $\mathrm{d}U = nc_v\mathrm{d}\Theta$.
From the state equation of $\mathcal{A}$ and from the general *law of energy conservation* [for a (substitute) reversible exchange of heat δq between the system and its environment] we formulate the *I. Principle of Thermodynamics*, $\delta q = \mathrm{d}U + p\mathrm{d}V$.
From this principle, and, from the Clausius equation (13), it follows that

$$S = n\int\left(c_v\frac{\mathrm{d}\Theta}{\Theta} + R\frac{\mathrm{d}V}{V}\right) = n\left(c_v\ln\Theta + R\ln V\right) + S_0(n) = \sigma(\Theta, V) + S_0(n). \qquad (15)$$

Let us "divide" the equilibrial system $\mathcal{A}$ in a volume V and at a temperature Θ, or, better said, the whole volume V (or, its whole given state space) occupiable, and now just occupied by all its constituents (particles, matter units), with diaphragmas (thin infinitely, or, "thought" only), not affecting the thermodynamic properties of $\mathcal{A}$ supposingly, to m parts $\mathcal{A}_i$, $i \in \{1, ..., m\}$, $m \geq 1$ with volumes V_i with matter units n_i. Evidently,

$$n = \sum_{i=1}^{m} n_i \text{ and } V = \sum_{i=1}^{m} V_i$$

Now let $S_0(n) = 0$ and $S_{0i}(n_i) = 0$ for all i. For the entropies S_i of $\mathcal{A}_i$ considered individually, and for the change ΔS, when volumes V, V_i are expressed from the state equations, and for $p = p_i$, $\Theta = \Theta_i$ it will be gained that $\sigma_{[i]} = Rn_{[i]}\ln n_{[i]}$. Then, for $S_i = \sigma_i = n_i\left(c_v\ln\Theta + R\ln V_i\right)$ is to be valid, we have

$$\sum_{i=1}^{m} S_i = \sum_{i=1}^{m}\sigma_i = nc_v\ln\Theta + R\ln\left(\prod_{i=1}^{m} V_i^{n_i}\right), \qquad (16)$$

$$\Delta S = S - \sum_{i=1}^{m} S_i = \sigma - \sum_{i=1}^{m}\text{œ}_i = \Delta\sigma = R\ln\frac{V^n}{\prod_{i=1}^{m} V_i^{n_i}} = -nR\sum_{i=1}^{m}\frac{n_i}{n}\ln\frac{n_i}{n} > 0.$$

Let us denote the last sum as B further, $B < 0$.
The result (16), $\Delta S = -nRB \neq 0$, is a *paradox*, a contradiction with our presumption of not influencing a thermodynamic state of $\mathcal{A}$ by diaphragmas, and, leads to that result, that the heat entropy S (of a system in equilibrium) *is not* an extensive quantity. But, by the definition of the differential $\mathrm{d}S$, this *is not* true.
Due to this contradiction[6] we must consider non-zero integrating constants $S_0(n)$, $S_{0i}(n_i)$, in such a way, that $\Delta S = (\sigma + S_0) - \sum_{i=1}^{m}(\sigma_i + S_{0i}) = 0$. This equation is solvable for the system $\mathcal{A}$ and all its parts $\mathcal{A}_i$ by solutions $S_{0[i]}(n_{[i]}) = -n_{[i]}R\ln\frac{n_{[i]}}{\gamma_{[i]}}$.

Then $S_{[i]} \overset{\triangle}{=} S_{Claus,[i]}$, and we write and derive that

$$S_{\mathrm{Claus}} = \sum_{i=1}^{m} S_{\mathrm{Claus,i}} = \sum_{i=1}^{m} n_i R\ln\gamma_i = nR\ln\gamma \;\Rightarrow\; \gamma = \gamma_i;\; \Delta S = 0. \qquad (17)$$

[6] Quantity $-B$ expressed in (16) is information entropy of a source of messages with an alphabet $[n_1, n_2, ..., n_m]$ and probability distribution $\left[\frac{n_i}{n}\right]_{i=1}^{m}$. Such a division of system to m parts defines an information source with information entropy of maximum $\ln m$.

2.2.3 Definition of statistical thermodynamics

The term *(macro)state* of a system $\mathcal{A}$, both *equilibrium* and *non-equilibrium*, means a *class of equivalence*[7] on the set of all its possible *microstates* - microscopic *arrangements* of all its constituents (particles, matter units) within all (by us distinguished, defined) parts, *cells*, of its whole, by them occupiable, volume V, or, better said, of a given *(state) space* of $\mathcal{A}$[8].

Let us denote the cardinality of such a class as *thermodynamic probability* $\tilde{P}$ of a given macrostate.

Let m be a number of cells of the whole volume, (state) space, of $\mathcal{A}$, occupiable by a set of all its constituents, let M be a number of all its constituents, and, let m_l, $0 \leq m_l \leq M$, be a number of constituents in a cell l of a given (state) space, $1 \leq l \leq m$. Then, $\tilde{P} = \frac{M!}{\prod_{l=1}^{m} m_l!}$, $\sum_{l=1}^{m} m_l = M$.

Then the *Boltzman definition* (11) of the physical (thermodynamic) entropy S of $\mathcal{A}$ (of, the whole volume V, or, of a given state space) in a macrostate with thermodynamic probability $\tilde{P}$ [could be non-equilibrium too, composed by (equilibrium) subsystems $\mathcal{A}_i$ in states $\boldsymbol{\theta}_i$, $i \in \{1, 2, ..., m\}$, not interacting mutually (at the same temperature Θ, or also, at various tempertures Θ_i) with thermodynamic probabilities $\tilde{P}_i$] is valid;

$$S_{\text{Boltz}} \stackrel{\text{Def}}{=} k \cdot \ln \tilde{P}, \;\; S_{\text{Boltz}} = \sum_{i=1}^{m} S_{\text{Boltz},i} = k \cdot \sum_{i=1}^{m} \ln \tilde{P}_i, \;\; \tilde{P} = \prod_{i=1}^{m} \tilde{P}_i. \tag{18}$$

Let us denote by N and N_j numbers of particles of system $\mathcal{A}$ and of its subsystems $\mathcal{A}_j$, $j = 1, \ldots, m \leq N$,

$$N_j = n_j N_A, \;\; \sum_j N_j = N = nN_A$$

By Stirling's formula and (18) we have, that

$$S_{\text{Boltz}} = -kN \sum_j \frac{N_j}{N} \ln \frac{N_j}{N} \stackrel{\triangle}{=} -kNB$$

The last sum B defines *Boltzman function* B_{Boltz} of statistical physics, $B^{Bolts} \stackrel{\triangle}{=} B$. Evidently, $\ln \tilde{P} = -NB_{\text{Boltz}}$.

For Boltzman entropy per one particle we write $\frac{S_{\text{Boltz}}}{N} = -kB_{\text{Boltz}}$.

The equilibrium system $\mathcal{A}$ (with N particles) is possible to be identified with the (state) space it occupies, being defined by "one-particle" railing of cells [*"one-particle"* (state) space], with the uniform probability distribution of occupation these cells by particles, $\left[\frac{N_j}{N}\right]_j = \left[\frac{1}{N}\right]_j$, where $j = 1, \ldots, m$ and $m = N$. In this space the quantities $B*$, $\tilde{P}*$, $S*$ are defined,

$$B* \stackrel{\triangle}{=} \sum_{j=1}^{N} \frac{1}{N} \ln \frac{1}{N} = \ln N = \frac{-1}{N} \ln \tilde{P}*, \;\; S* = -kNB*. \tag{19}$$

[7] Due to individal particles being undistinguishable mutually by means of macroscopic thermodynamics, microstates of a given macrostate are undidisguishable too

[8] State space of $\mathcal{A}$ can be *three-dimensional* or *general configuration* space, *impulse* or the whole *phase* space

3. Clausius, Boltzman and Shannon entropy

Let us explore the entropy S_{Boltz} during the Gay-Lussac experiment.
The value S_{Boltz} is dependent on a time development of $\mathcal{A}$ within the whole volume V divided into cells, and, in the maximum level of detail (the railing with "one-partical" cells) we have $S* = \max\limits_{\{m,\, p(\cdot)\}} \{S_{\mathrm{Boltz}}\}$ for the equlibrium state.

In the final state of $\mathcal{A}$ (equilibrium) we have: $\Theta = const.$, $n = \dfrac{N}{N_A}$ and $n_i = \dfrac{N_i}{N_A}$, where $i = 1, 2, \ldots, m$ is given by a railing of our diaphragmas.

Probability distribution $\dfrac{N_i}{N}$, $i = 1, \ldots, m$ (defined in such railing) states a probability of an *observer's choice* of i-th cell with N_i particles.
The uncertainity $-B$ of this choice is given by Shannon entropy at the value $-B_{\mathrm{Boltz}}$,

$$-B_{\mathrm{Boltz}} = -\sum_{i=1}^{m} \frac{N_i}{N} \ln \frac{N_i}{N} \stackrel{\triangle}{=} -B_{\mathrm{Gibbs}}$$

Its maximum $-B*$ is achieved for the most possible detailed description ($m = N$) of the observed equilibrium system $\mathcal{A}$ and with the uniform probability distribution $\dfrac{N_i}{N} = \dfrac{1}{N}$, $i = 1, \ldots, N$.
Till this moment, the expression $-B_{\mathrm{Boltz}}$ (18) defines an uncertainity of probability distribution too, but, defined by a *time propagation* of the set of particles of $\mathcal{A}$, now not equilibrium, within the all cells of the given railing.
When $\mathcal{A}$ achieves its final, equilibrium state, each "its own" "one-particle" cell is occupied by one particle and the thermodynamic entropy of $\mathcal{A}$ (of its entire volume, or, of a given state space) per one cell, is $-kB*$.
But when the observer's railing is considered, the index $i = j = 1, \ldots, m$ is used, and, the value $-B_{\mathrm{Boltz}} = -B_{\mathrm{Gibbs}} \leq B*$ is defined. Till this moment, the relation $-B_{\mathrm{Boltz}} < -B_{\mathrm{Gibbs}} \leq -B*$ is valid.
During the time development of $\mathcal{A}$ such probability distributions are defined that are impossible for observing its equilibrium state (for which our railing defines the value $-B_{\mathrm{Gibbs}}$); for the sets of all values of $-B_{\mathrm{Gibbs}}$ and $-B_{\mathrm{Boltz}}$ we have

$$\{-B_{\mathrm{Gibbs}}\} \subset \{-B_{\mathrm{Boltz}}\},\ \text{but, } \max\{-B_{\mathrm{Gibbs}}\} = \max\{-B_{\mathrm{Boltz}}\} = -B* = \frac{S*}{kN}.$$

The quantity $-B*$, introduced in (19), has a meaning of the maximal uncertainity of distribution of particles, and, also, of their energies within all "one-particle" cells.
Thus, using a measure r of the accuracy, detail of our observation, we write

$$-B_{\mathrm{Gibbs}} = -\sum_{i=1}^{m} \frac{N_i}{N} \ln \frac{N_i}{N} = -\frac{1}{r} \sum_{j=1}^{N} \frac{1}{N} \ln \frac{1}{N} = \frac{1}{r} \ln N = -\frac{1}{r} B*,\ \ r \geq 1, \tag{20}$$

$$\ln N = -r \sum_{i=1}^{N} \frac{N_i}{N} \ln \frac{N_i}{N} = -r \cdot B_{\mathrm{Gibbs}} = -r \cdot B_{\mathrm{Boltz}} = -B*,$$

$$S* = kN \ln N = -kNB* = S_{\mathrm{Claus}} = -r \cdot kNB = rS_{\mathrm{Gibbs}} = rS_{\mathrm{Boltz}}.$$

Following (20), for the system $\mathcal{A}$ with "its own" "one-particle" railing, $j = 1, ..., N$, and, for an observer's railing, $i = 1, ..., m$, it is valid that

$$\ln\gamma = -s\cdot B,\ \ s \stackrel{\triangle}{=} \frac{r\cdot Q}{\Theta\cdot kN\ln N},\ \ s>0,\ \ \frac{s}{r} = \frac{\ln\gamma}{\ln N} = const. \tag{21}$$

It is provable that $s = r$ for all $r \geq 1$, and thus, $\gamma = N > 1$ (8).
For $\mathcal{A}$ with volume V, in the thermodynamic equilibrium at a temperature Θ, its heat $Q = \int_V \delta q(\Theta, V)$ and its entropy $S_{\text{Claus}} = \int_V \frac{\delta q(\Theta,V)}{\Theta} = \frac{Q}{\Theta}$. Then, following note 3, we have,

$$S_{\text{Claus}} = \frac{Q}{\Theta} = kN\ln\gamma = kN\ln N = S*,\ \ \gamma = e^{\frac{\varepsilon}{k\Theta}},\ \varepsilon \stackrel{\triangle}{=} \frac{Q}{N} = \frac{Q_i}{N_i}. \tag{22}$$

For the (final) equilibrium state of $\mathcal{A}$ and for our most possible detailed observation of $\mathcal{A}$ [$max(i) = m = N$ with the uniform probability distribution of our choice] $r = s = 1$ is valid. Otherwise $s = r > 1$.
Considering a general observer's railing, $i = 1, ..., m \leq N$, any time development of $\mathcal{A}$ is expressed by a sequence of values from the set $\{S_{\text{Boltz}}\} - \{S_{\text{Boltz}}\}\bigcap\{S_{\text{Gibbs}}\}$, growing, defined by a time propagation of its N particles (bearing heat Q at temperature Θ) through its whole occupiable volume (a state space). Any of values of such a sequence defines Clausius entropy, relevant to a certain equivalent substitute equilibrium state, at value $\frac{S_{\text{Claus}}}{r'} \stackrel{\triangle}{=} \frac{Q}{T''} \stackrel{\triangle}{=} \frac{Q}{r'\Theta},\ \ r' \geq r \geq 1$.
On the set $\{S_{\text{Boltz}}\}\bigcap\{S_{\text{Gibbs}}\}$ the equality $-B_{\text{Boltz}} = -B_{\text{Gibbs}}$ is valid. It expresses the observer's uncertainity only (as $r' = r$). It can be changed by a various organization of railing, up or down respectively. Thus, the best possible knowledge about the equilibrium $\mathcal{A}$ should be achieved just when both the observer and the system $\mathcal{A}$ are in equilibrium and "one-particle" railing is considered.
Then, $r' = r = 1$, otherwise, $\frac{S_{\text{Boltz}}}{S_{\text{Gibbs}}} = \frac{r}{r'} \leq 1$.
The entropy S_{Claus} is the maximum of the physical entropies S_{Boltz} and S_{Gibbs} which are bound Shannon entropies and figurred in physical units.

4. Elements of heat transformation

4.1 Reversible Carnot cycle

A reversible Carnot cycle, here denoted as $\mathcal{O}$, is the most simple heat *cyclical process* ending with a gain of mechanical work, denoted as ΔA. It consists of: 4 *reversible* changes; 2 isothermal changes related to two constant absolute temperatures, T_W and T_0, $T_W \geq T_0 > 0$; and two adiabatic changes effecting the transposition between these two temperatures in a *working medium*. The working medium, a *thermodynamic system* denoted as $\mathcal{L}$, is the transformer of the input energy, denoted as ΔQ_W, and as it passes through the cycle it acquires this *input* heat ΔQ_W. This occurs during the reversible isothermal expansion phase, during which the temperature T_W of the system is equal to that of the heater $\mathcal{A}$. Within the reversible isothermal compression phase, while the temperature T_0 of the system is equal to that of the cooler $\mathcal{B}$, the system $\mathcal{L}$ is giving out "residual" heat, denoted as ΔQ_0, to $\mathcal{B}$.
To achieve a positive gain in ΔA it must be valid that $T_W > T_0$. The difference between

the values of ΔQ_W and ΔQ_0 is equal to that of ΔA gained within the cycle; $\Delta A = \Delta Q_W + (-\Delta Q_0)$ or $\Delta A = \Delta Q_W + \Delta Q^*{}_0$, when $\Delta Q^*{}_0 \stackrel{\triangle}{=} -\Delta Q_0$.
The sign $+$ denotes the heats delivered (directly) to the system $\mathcal{L}$ and the sign $-$ denotes the heats drained off (directly) from the system $\mathcal{L}$ during the cycle. The *transformation efficiency* of

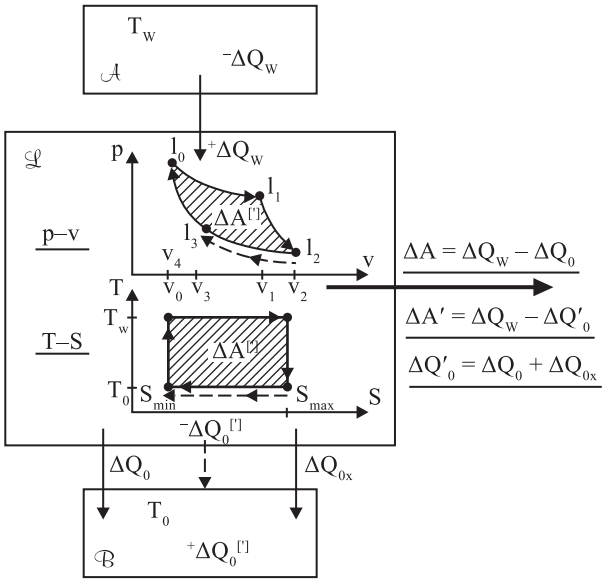


Fig. 3. Carnot Cycle

a reversible Carnot cycle, denoted as η_{max}, is defined by the equation

$$\eta_{max} \stackrel{\text{Def}}{=} \frac{\Delta A}{\Delta Q_W} = \frac{\Delta Q_W + (-\Delta Q_0)}{\Delta Q_W}. \tag{23}$$

Following from both the definition (23) and the condition of the cyclical form of the whole thermodynamic change $\mathcal{O}$ (11), we have

$$\eta_{max} = \frac{\Delta Q_W - \Delta Q_0}{\Delta Q_W} = \frac{T_W - T_0}{T_W}, \quad \eta_{max} < 1. \tag{24}$$

Carnot's theorem (see later) states that η_{max} is the *maximum* of a set of efficiencies from all those heat cycles with those (extreme) working temperatures. Following from (24) in a reversible Carnot cycle it is valid that

$$\frac{\Delta Q_W}{T_W} = \frac{\Delta Q_0}{T_0}, \quad \frac{\Delta Q_W}{T_W} + \left(-\frac{\Delta Q_0}{T_0}\right) = 0 \text{ or } \frac{\Delta Q_W}{T_W} + \frac{\Delta Q^*{}_0}{T_0} = 0, \tag{25}$$

or more generally that

$$\sum_{i \in [W,0]} \frac{\Delta Q_i}{T_i} \stackrel{\triangle}{=} \oint_{\mathcal{O}} \frac{\delta Q(\Theta)}{\Theta} = 0. \tag{26}$$

The expressions ΔQ_i and $\delta Q(\Theta)$ denote all the heats shared reversibly (isothermally) when the temperature of the system $\mathcal{L}$ is either T_i or Θ respectively. The ratios $\frac{\Delta Q_i}{T_i}$ and $\frac{\Delta Q(\Theta)}{\Theta}$ are known as changes of *heat (Clausius') entropy S*.
The integral in (26) is named *Clausius' integral* and this relation itself is the mathematical integral form of the II. *Principle of Thermodynamics* in **Thomson-Planck's** formulation: *It is impossible to construct a heat cycle transforming all heat delivered to a system ($\mathcal{L}$, going through this cycle) into an equivalent amount of mechanical work (ΔA).*
Therefore in the case of a reversible Carnot cycle, Clausius' integral, the algebraic sum of all the *by-temperature-reduced* heats [both delivered (directly) to the cycle and drained off (directly) from the cycle], equals 0. According to Thomson-Planck's formulation there must be both heat $\Delta Q_{i=W}$ delivered into the cycle $\mathcal{O}$ and heat $\Delta Q_{i=0}$ drained off from the cycle. As a consequence of the existence of these heats (the input ΔQ_W and the "residual" ΔQ_0), and in accordance with the I. Principle of Thermodynamics, it must be valid that $\eta_{max} < 1$.
Another formulation of the II. Principle of Thermodynamics is **Carnot's** theorem, the **first part** of which states: *The efficiencies of all reversible Carnot cycles with the working temperatures T_W and T_0 are equal.*
A reversible Carnot cycle is an abstract construction working without any restriction on the values T_W and T_0 (except in the case that $T_W \geq T_0$) and any restriction on the medium $\mathcal{L}$. Therefore we can consider that it runs in both an *ideal* medium and a *non-ideal* medium.
Reversibility of a (thermodynamic) change of status in a thermodynamic system means that, in such a system, there is no resistance that this change must overcome. Thus there is no friction in a system going through such a change, and consequently there is no (positive) *production* of heat within it. We can denote such heat as ΔQ_{0x}. This situation, where $\Delta Q_{0x} = 0$, stands in any ideal medium $\mathcal{L}$ (without viscosity, i.e. in *ideal* gas), both when a Carnot cycle runs finitely slowly and when it runs infinitely slowly. We can call this an *ideal* cycle.
However, a reversible Carnot cycle can run in a non-ideal, *real* system $\mathcal{L}$ (with viscosity) when the kinetic energy of the system is constant and, as a consequence, no heat $\Delta Q_{0x} > 0$ is generated. This situation occurs when the change of states in a (real) system is *quasi-stationary*, namely when it is infinitely small and runs infinitely slowly, so that it is valid that $\Delta t \to \infty$ (where t denotes time). In this case $\Delta Q_{0x} = 0$ also holds true.

4.2 Irreversible Carnot cycle

When the system $\mathcal{L}$ is non-ideal, for instance it is a *real* gas and the heat cycle operates with a finite, but not zero, speed,

$$0 < \Delta t < \infty, \tag{27}$$

some inner friction caused by the viscosity of $\mathcal{L}$ generates *noise* heat ΔQ_{0x}. The fact that the medium $\mathcal{L}$ is non-ideal is only discovered when the changes the system undergoes are not reversible (e.g. they are not the cycle $\mathcal{O}$ or the cycle $\mathcal{O}'$ where $\Delta t \to \infty$, $\Delta Q_{0x} = 0$ is valid), namely when it undergoes *non-quasi-stationary* changes lasting a finitely long period of time (27). Then the kinetic energy of $\mathcal{L}$ varies and, due to inner friction on the passive resistants in $\mathcal{L}$, (positive) production of the heat ΔQ_{0x} arises,

$$\Delta Q_{0x} > 0. \tag{28}$$

The **second part** of **Carnot's** theorem states: *The efficiency of any irreversible heat cycle with (extreme) working temperatures T_W and T_0, $T_W \geq T_0 > 0$, is less than the efficiency of a reversible Carnot cycle with those same working temperatures.*

The *substitute* thermodynamic path (19) is denoted as $\mathcal{O}'$, expressing the irreversible cycle itself. For this path to be closed, the heat $\Delta Q'_0$ must be drained off from the medium $\mathcal{L}$ when its temperature is T_0. This heat is greater than the heat ΔQ_0 when only the reversible cycle $\mathcal{O}$ is considered. Thus we have

$$\Delta Q'_0 = \Delta Q_0 + \Delta Q_{0x}. \tag{29}$$

Consequently for the output work $\Delta A'$ of an irreversible Carnot cycle we have

$$\Delta A' = \Delta Q_W + (-\Delta Q'_0) = \Delta A - \Delta Q_{0x}. \tag{30}$$

For an irreversible Carnot cycle with the working temperatures T_W and T_0, where $T_W \geq T_0 > 0$, it follows from the definition of the (transformation) efficiency, denoted as η, that

$$\eta = \frac{\Delta A'}{\Delta Q_W} < \frac{\Delta A}{\Delta Q_W} = \frac{T_W - T_0}{T_W} = \eta_{max}. \tag{31}$$

The equality (30) and the inequality (31) implies that

$$\frac{\Delta Q'_0}{\Delta Q_W} > \frac{T_0}{T_W}. \tag{32}$$

Within the reversible part $\mathcal{O}$ of the irreversible cycle $\mathcal{O}'$ the equality (25) is valid and thus, following (29) and (32), for the whole irreversible $\mathcal{O}'$ we have

$$\frac{\Delta Q_W}{T_W} + \left(-\frac{\Delta Q'_0}{T_0}\right) = -\frac{\Delta Q_{0x}}{T_0} < 0, \tag{33}$$

or, more generally (*Clausius' inequality*),

$$\sum_{i \in [W,0]} \frac{\Delta Q(T_i)}{T_i} \triangleq \oint_{\mathcal{O}'} \frac{\delta Q(\Theta)}{\Theta} < 0, \quad \Delta Q(T_i) \triangleq \Delta Q_i, \tag{34}$$

where ΔQ_i and $\delta Q(\Theta)$ denote all the (elementary) heats shared reversibly (isothermally) when the temperature of the system $\mathcal{L}$ is respectively either T_i or Θ. The relations (26) and (34) are also known as **Kelvin's** formulation of the II. Principle of Thermodynamics.

In the case of an irreversible Carnot cycle $\mathcal{O}'$, Clausius' integral, the algebraic sum of all the *by-temperature-reduced* heats (34) [both delivered (directly) to the cycle and drained off (directly) from the cycle], is less than 0. This is caused by the heat ΔQ_{0x} being drained off from the medium $\mathcal{L}$ (into $\mathcal{B}$) when its temperature is T_0. This is a consequence of the requirement that the thermodynamic path $\mathcal{O}'$ in $\mathcal{L}$ must be cyclical, and, thus, that the whole process be *repeatable*.

Thus with respect to the work $\Delta A'$ produced within one run of an irreversible Carnot cycle $\mathcal{O}'$, this cycle can be considered to be a reversible cycle $\mathcal{O}$ accompanied by the generation of noise heat $\Delta Q_{0x} > 0$. The "additional" heat ΔQ_{0x} existing in an irreversible cycle then enters $\mathcal{B}$ during the compression phase, with the compression work having a value equivalent to the noise heat ΔQ_{0x}. This compression work diminishes the mechanical work ΔA acquired during the reversible part $\mathcal{O}$ of the $\mathcal{O}'$ cycle. Thus we can envisage one "additional" isothermal compression running in parallel with the "original" compression . Therefore, with regards to the heat $\Delta Q'_0$ being drained off from the medium $\mathcal{L}$, when $\mathcal{L}$ is going through the irreversible cycle $\mathcal{O}'$, this cycle can be considered to be an *additive* superposition of the reversible part $\mathcal{O}$ and the irreversible part; the latter simply being this "additional" isothermal compression.

Irreversibility is the consequence of both the medium $\mathcal{L}$ being non-ideal and its state changes being of finite and non-zero speed. Therefore the equation $\Delta Q_W = \Delta A = \Delta A'$ is the limit, but *not achievable* in reality.

4.3 Reverse reversible Carnot cycle

This cycle is a reverse Carnot cycle in which the states of the system $\mathcal{L}$ are changing along the curve $\mathcal{O}$, but, in the opposite way, starting with the isothermal expansion at the temperature T_0 when the *diathermal* contact between $\mathcal{L}$ and $\mathcal{B}$ is maintained. It is a cycle run of the medium $\mathcal{L}$ through a sequence of (equilibrium) states along the curve $\mathcal{O}$ oriented in the opposite way than it is in the case of a direct Carnot cycle. This run is called the *cooling* run too. The Carnot machine with this running through of its medium $\mathcal{L}$ along the heat cycle $\mathcal{O}$ is working as a *heat pump, cooling machine*. The functionality of such a cooling machine is that:
Within the isothermal expansion at the temperature T_0 the system $\mathcal{L}$ is draining off the heat ΔQ_0 being *pumped out, transferred* from the cooler $\mathcal{B}$. This heat is, within the isothermal compression at the temperature T_W, delivered from $\mathcal{L}$ into the heater $\mathcal{A}$. We call it the *output* heat ΔQ_W,

$$\Delta Q_W = \Delta Q_0 + \Delta A \tag{35}$$

The quantity ΔA is the *input* mechanical energy (work) delivered into $\mathcal{L}$ within the isothermal compression at the temperature T_W.
Following the first part of Carnot's theorem (*the efficiencies of all reversible Carnot cycles with the working temperatures T_W and T_0, $T_W \geq T_0 > 0$, are equal*) we must conclude that in the opposite case we would be able to couple two Carnot machines with different efficiencies η_1, η_2, $\eta_1 \neq \eta_2$ in such a way that the resulting machine would be the *perpetuum mobile of the II. order* (machine changing cyclically, permanently the whole input heat ΔQ_W in the output work ΔA; $\Delta Q_W = \Delta A$). So it must be valid that $\eta_1 = \eta_2$.
The first part of Carnot's theorem is another variant of the Thomson-Planck formulation of the II. Principle of Thermodynamics.

5. Heat transformation and information transfer

It follows from section 3 and also from Brillouin (2) and (15), Landauer (17) and Gershenfeld (4) that when an average amount of information (denoted as ΔJ) is being *recorded, transmitted* or *computed*, etc. while the temperature is Θ, there is a need for a (minimal) average energy ΔW, $\Delta W \geq k \cdot \Theta \cdot \Delta J$. In this paper $\Delta W \stackrel{\triangle}{=} \Delta Q_W$.
We term the following ratios [expressed in *information* units (*Hartley*, *nat*, *bit*)] *changes of the thermodynamic entropies of system $\mathcal{L}$ in cycles $\mathcal{O}$ or $\mathcal{O}'$*:

$$\frac{\Delta Q_W}{kT_W}\ input,\ \frac{\Delta A^{[\prime]}}{kT_W}\ output\ (\stackrel{\triangle}{=} \Delta I^{[\prime]}),\ \frac{\Delta Q_0}{kT_W}\ loss,\ \frac{\Delta Q_{0x}}{kT_W}\ noise, \tag{36}$$

where k is Boltzman's constant.
These changes are the absolute values of $H(\cdot)$, $H(\cdot|\cdot)$ contained in any message on inputs and outputs of a "*carnot" (thermodynamically)* described transfer channel $\mathcal{K}$.

5.1 Reversible Carnot cycle and noiseless channel

A reversible Carnot cycle $\mathcal{O}$ running in $\mathcal{L}$ (producing noise heat $\Delta Q_{0x} = 0$) can be considered to be a *thermodynamic, average-value* realization or, as such, as a model of an information transfer process running in a channel $\mathcal{K}$ *without noise*. For the average noise information (entropy) $H(Y|X)$ defined in (8) it is valid that $H(Y|X) = 0$.
On the channel $\mathcal{K}$ the average information amounts $H(X)$, $H(Y)$, $H(Y|X)$, which have been

expressed in (6) and (8), and with their values satisfying (36), are defined as:

$$H(X) \stackrel{\text{Def}}{=} \frac{\Delta Q_W}{kT_W} \quad \left[= \frac{\Delta Q_0}{kT_0}, \text{ following (25)}\right], \tag{37}$$

$$H(Y) \stackrel{\text{Def}}{=} \frac{\Delta A}{kT_W} = \frac{\Delta Q_W - \Delta Q_0}{kT_W} = \frac{\Delta Q_W}{kT_W} \cdot \eta_{max} = H(X) \cdot \eta_{max} \stackrel{\triangle}{=} \Delta I,$$

$$H(Y|X) \stackrel{\text{Def}}{=} 0.$$

Thus we assume the medium $\mathcal{L}$ going through a reversible Carnot cycle $\mathcal{O}$ works as a

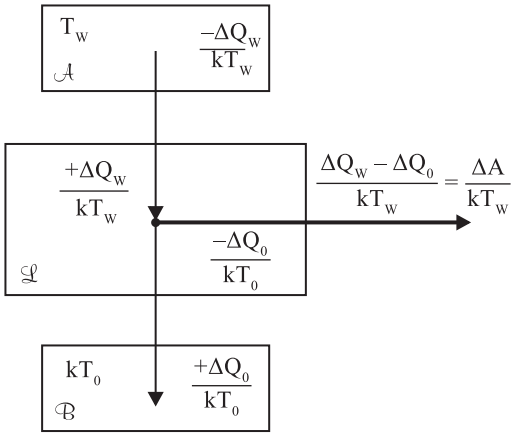


Fig. 4. The schema of a reversible Carnot Cycle

thermodynamic, average-value model of an information channel $\mathcal{K}$ ($\mathcal{K} \cong \mathcal{L}$) and that the equation (12) is valid for the values of the quantities in (37).
Following from (25), the definition (37) and the assumption (12) we then have

$$\frac{\Delta Q_W}{kT_W} \cdot \eta_{max} - 0 = \frac{\Delta Q_W}{kT_W} - H(X|Y) \text{ and thus} \tag{38}$$

$$H(X|Y) = \frac{\Delta Q_W}{kT_W} \cdot (1 - \eta_{max}) = \frac{\Delta Q_W}{kT_W} \cdot \beta = \frac{\Delta Q_0}{kT_W}, \ \beta = \frac{T_0}{T_W}.$$

Thus for the transinformation defined in (9), (10) and using the definitions (37), we have

$$T(X;Y) = H(X) \cdot (1 - \beta) = H(X) \cdot \eta_{max}, \text{ then} \tag{39}$$

$$T(Y;X) = \frac{\Delta Q_W}{kT_W} \cdot \eta_{max} = \frac{\Delta A}{kT_W} \text{ and thus } T(X;Y) = \Delta I = T(Y;X).$$

As a consequence of Thomson-Planck's formulation of the II. Principle of Thermodynamics the result (38) and the relations (39) imply the sharp inequality in the relation

$$\Delta I = T(X;Y) < H(X). \tag{40}$$

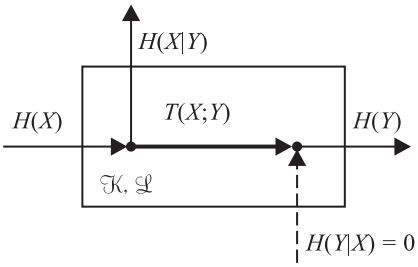


Fig. 5. The information transfer channel modelling a reversible Carnot Cycle

The equality in the relation (40) is valid for all reversible Carnot cycles (with temperatures T_W and T_0) viewed informationally, and can be considered to be an *information* formulation of the first part of the Carnot's theorem.
When the medium $\mathcal{L}$ has gone through the cycle $\mathcal{O}$ once [and in agreement with (26)], we have the following for the change $\Delta S_{\mathcal{L}}$ of the heat entropy $S_{\mathcal{L}}$ of the whole reversible Carnot engine:

$$\Delta S_{\mathcal{L}} = \oint_{\mathcal{O}} \frac{\delta Q}{T} = \frac{\Delta Q_W}{T_W} - \frac{\Delta Q_0}{T_0} = 0. \tag{41}$$

As a consequence of the additivity of (substitute reversible (19)) changes of heat entropy, when the medium $\mathcal{L}$ has gone through the reversible cycle $\mathcal{O}$ once, we have the following for the change $\Delta S_{\mathcal{AB}}$ of the heat entropy $S_{\mathcal{C}}$ in the system $(\mathcal{AB})$ (consisting of the heater $\mathcal{A}$ and the cooler $\mathcal{B}$):

$$\Delta S_{\mathcal{AB}} = -\frac{\Delta Q_0}{T_W} + \frac{\Delta Q_0}{T_0} = \frac{\Delta Q_0}{T_0} \cdot \eta_{max} = \frac{\Delta Q_W}{T_W} \cdot \eta_{max}. \tag{42}$$

A reversible Carnot engine is an isolated system, in which the transformation of heat energy $\Delta Q_W \sim x$ to mechanical energy $\Delta A \sim y$ is being performed. As a consequence of the additivity of (substitute reversible) changes of heat entropy, and when (41) and (42) are used, we have the following for the *result* change $\Delta S_{\mathcal{C}}$ of the heat entropy $S_{\mathcal{C}}$ of the whole reversible Carnot engine:

$$\Delta S_{\mathcal{C}} = \Delta S_{\mathcal{L}} + \Delta S_{\mathcal{AB}} = \frac{\Delta Q_W}{kT_W} \cdot \eta_{max}. \tag{43}$$

The derivation (39) and the equation (43) together then imply that

$$\begin{gathered} \Delta S_{\mathcal{C}} - k{\cdot}T(X;Y) = k{\cdot}H(X) \cdot (\eta_{max} - \eta_{max}), \text{ then} \\ \Delta S_{\mathcal{C}} - k{\cdot}\Delta I = 0 \text{ or also } \Delta(S_{\mathcal{C}} - k{\cdot}I) = 0. \end{gathered} \tag{44}$$

Therefore the result change $\Delta S_{\mathcal{C}}$ (43) of the heat entropy $S_{\mathcal{C}}$ of the whole reversible Carnot engine, together with the output information ΔI defined in (37), satisfies **Brillouin's** (2; 15) ***extended*** formulation of the II. Principle of Thermodynamics[13]

$$\Delta(S_{\mathcal{C}} - k{\cdot}I) \geq 0 \text{ or also } \mathrm{d}(S_{\mathcal{C}} - k{\cdot}I) \geq 0. \tag{45}$$

We assume the relation (44), incorporating the equality in (45), is valid for any information transfer, such as a *watching* process (i.e. transmitting, measuring, recording, computing, etc.),

[13] The *information* member I does not exist in the traditional (differential) formulation of this theorem, $\mathrm{d}S \geq 0$.

using energy transformations[14] expressible in a *reversible* thermodynamic way, such as a cycle $\mathcal{O}$ running in $\mathcal{L}$.
This process can transfer a message $x \sim \Delta Q_W$ from a source of messages with the information entropy $H(X) = \frac{\Delta Q_W}{kT_W}$. It is transmitting a message with the same average amount of information $H(X)$. It also could be measuring the status (x) of the thermodynamic system $\mathcal{A}$ or $(\mathcal{AB})$ observed through the process $\mathcal{O}$. With an information transfer process of this type we obtain an output message $y \sim \Delta A$ with the average amount of information $\Delta I = H(Y) < H(X)$. This information gain is expressed by the by-temperature-reduced output work ΔA performed, while the system $\mathcal{L}$ underwent the cycle $\mathcal{O}$ once, at the temperature T_W. This expresses the definition (37) and the equation (12).
Consequently on receiving the output information $\Delta I = \frac{\Delta A}{kT_W}$ in $\mathcal{O}$, the *distinguishability* of the structure of the mechanical output of this cycle increases in volume, and is realized by the addition ΔA to the potential output energy; the distinguishability being measured by ΔI. But at the same time the *thermodynamic* distinguishability of the parts $\mathcal{A}$ and $\mathcal{B}$ of the system $(\mathcal{AB})$, the structure of which we express by the mutually different volumes of heats in $\mathcal{A}$ and $\mathcal{B}$, has decreased in volume by exactly the value $\Delta S_{\mathcal{C}}$ (43). Thus we have $\Delta S_{\mathcal{C}} = k{\cdot}\Delta I$, stated in (44), and in (45) with the equals sign.
The increasing ΔI of the *structural* distinguishability of the mechanical output models the receiver's average information increase by exactly the average information addition ΔI contained in any message y on the output of $\mathcal{K}(\cong \mathcal{L})$.
The value ΔS_C in $\mathcal{O}$ thus represents the increase in the whole *extensity* (see remark 3) of energy used for coding the input message, and thus the indistinguishability of this energy within the whole isolated transfer chain. In agreement with (44) and (45) we can say that any such *measuring influences what is measured* (7).
In the following section we will show that the measured subject [in our heat example $\mathcal{A}$ or $(\mathcal{AB})$] is not only influenced by the way measuring is *organized* (here described by the cycle $\mathcal{O}$ with $\Delta Q_{0x} = 0$), but also that the result [y or $H(Y)$] of measuring is influenced by its *construction* in a *real* environment (further described by the irreversible cycle O' with $\Delta Q_{0x} > 0$) (7).

5.2 Irreversible Carnot cycle and noise channel

We will now consider an irreversible Carnot cycle $\mathcal{O}'$ running in the medium $\mathcal{L}$ [its irreversibility is caused by a non-ideal working medium $\mathcal{L}$ undergoing such a cycle with a finite, but not zero, speed (27) and (28)], to be a thermodynamic, average-value realization or, as such, as a model of an information transfer process running in a channel $\mathcal{K}$ *with noise*. For the noise information entropy $H(Y|X)$ defined in (8) it is valid that $H(Y|X) \neq 0$. Thus we can assume that the medium $\mathcal{L}$ going through an irreversible cycle works as an average-value, thermodynamic realization, model, of an information channel $\mathcal{K}$ ($\mathcal{K} \cong \mathcal{L}$).
On the channel $\mathcal{K}$ the average information amounts $H(X)$, $H(Y)$, $H(X|Y)$, which have been expressed in (6) and (7), and $H(X)$, $H(X|Y)$ with their values satisfying (36), (37) and (38), are

[14] Generally we could consider any type of so-called *directly-shared-energy* other than heat (e.g. electricity), and instead of the term *heat entropy* we could then speak about *extensity* and *Extensity Grow Principle* of the energy considered (6). Naturally thermodynamic irreversibility would again reveal itself by the generation and dissipation of noise heat, i.e. on the electrical resistants in the case of electric energy

defined as:

$$H(X) \stackrel{\text{Def}}{=} \frac{\Delta Q_W}{kT_W}, \tag{46}$$

$$H(Y) \stackrel{\text{Def}}{=} \frac{\Delta A'}{kT_W} = H(X)\cdot\eta_{max} - \frac{\Delta Q_{0x}}{kT_W} = \frac{\Delta Q_W}{kT_W}\cdot\eta = H(X)\cdot\eta \stackrel{\triangle}{=} \Delta I'$$

and following the result (39), for the reversible case we define

$$H(X|Y) \stackrel{\text{Def}}{=} \frac{\Delta Q_0}{kT_W},$$

because, in respect to the mechanical work $\Delta A'$ gained when $\mathcal{L}$ has gone through the cycle once, an irreversible Carnot cycle $\mathcal{O}'$ can be considered to be an additive superposition of its reversible part $\mathcal{O}$ [(25) is valid], and its irreversible part (with noise given by the production of the heat $\Delta Q_{0x} > 0$ in the medium $\mathcal{L}$). The noise heat is drained off into $\mathcal{B}$ within the "added" isothermal compression, (33) and (34).

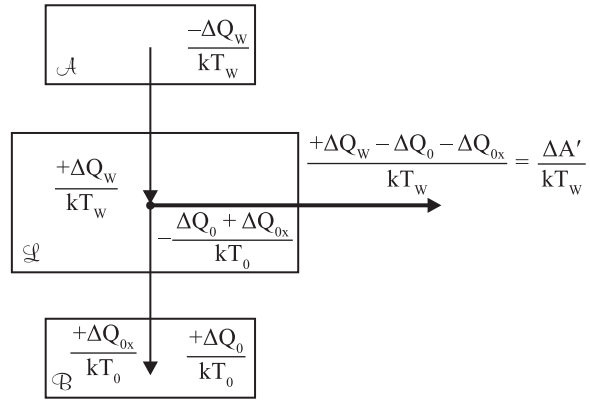


Fig. 6. The schema of an irreversible Carnot Cycle

We can assume that for the values of the quantities $H(X)$, $H(Y)$, $H(X|Y)$, $H(Y|X)$, defined in (46), the equation (12) is valid and thus, similarly as in the previous subsection, we have

$$\frac{\Delta Q_W - \Delta Q_0 - \Delta Q_{0x}}{kT_W} - H(Y|X) = \frac{\Delta Q_W}{kT_W} - \frac{\Delta Q_0}{kT_W} \text{ and consequently}$$

$$H(Y|X) = -\frac{\Delta Q_{0x}}{kT_W} = -\frac{\Delta Q_{0x}}{kT_0}\cdot\beta < 0, \;\; \beta = \frac{T_0}{T_W}, \; T_W \geq T_0 > 0. \tag{47}$$

In agreement with the definitions (46), the relations (47) for the average noise information $H(Y|X)$ in the channel $\mathcal{K}(\cong \mathcal{L})$ are in accordance with the fact that noise heat $\Delta Q_{0x} > 0$ is inevitably drained off from the medium $\mathcal{L}$ (the sign $-$) at the temperature T_0. This reduces the mechanical work ΔA generated in the reversible part $\mathcal{O}$ of the cycle $\mathcal{O}'$ from the heat ΔQ_W when the temperature of $\mathcal{L}$ is at T_W. This fact is expressed by the denominator of the first ratio in (47). The heat ΔQ_{0x} represents the above-mentioned construction *influence* on the definition (expressed thermodynamically by $\mathcal{O}$) of the measuring process itself.

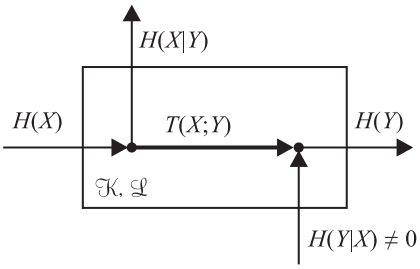


Fig. 7. The information transfer channel modelling an irreversible Carnot Cycle

It is obvious from (47), as well as from the equation (12), that the relations (46) describe a transfer channel $\mathcal{K}$ with negative values of noise added, and thus with the *subtractive* noise $H(Y|X) < 0$, which is caused by the draining off of the heat ΔQ_{0x} ($-\Delta Q_{0x} < 0$) from the medium $\mathcal{L}$. With regards to the information loss $H(X|Y)$ from the transferred information $H(X)$ in any input message $x \sim \Delta Q_W$ within the reversible part ($\mathcal{O}$) of the irreversible ($\mathcal{O}'$), the heat ΔQ_{0x} is that noise which causes *another* (average) information loss $H(Y|X)$ [the sign $-$ in (47)] in $\mathcal{O}''$.
For the transinformation value defined in (9), (10) and following the definitions (46) and (47), we have

$$T(X;Y) = \frac{\Delta Q_W}{kT_W} - \frac{\Delta Q_0}{kT_W} = \frac{\Delta Q_W}{kT_W} \cdot \eta_{max} = H(X) \cdot \eta_{max} \text{ but also} \tag{48}$$
$$T(Y;X) = \frac{\Delta Q_W - \Delta Q_0 - \Delta Q_{0x}}{kT_W} - \left(-\frac{\Delta Q_{0x}}{kT_W}\right) = H(X) \cdot \eta_{max}.$$

The left sides of the equations in (48) are equal, and the transinformation satisfies the requirement of symmetry (12). Also it is obvious that the relations (48) are the same as those stated in (39) for a noisless transfer.
The definitions (46), and the relations (48), together imply the following relations

$$\begin{aligned} &\Delta I' = T(X;Y) \quad \text{when} \quad \Delta Q_{0x} = 0, \quad \text{then} \quad \Delta I' = \Delta I, \\ &\Delta I' < T(X;Y) \quad \text{when} \quad \Delta Q_{0x} > 0, \quad \text{then} \quad \Delta I' < \Delta I. \end{aligned} \tag{49}$$

Satisfying its definition, and with the given meaning of the quantities $H(\cdot)$, $H(\cdot|\cdot)$, the transinformation $T(X;Y)$ is the *maximal* (average) amount of output information that can be gained within either a reversible or irreversible Carnot cycle. In accordance with the second part of Carnot's theorem this is also true for *any other* heat cycle in the medium $\mathcal{L}$ with the extreme working temperatures T_W and T_0, $T_W \geq T_0 > 0$. [These cycles are considered to be models of a transfer process with a given $H(X)$ in a channel $\mathcal{K}(\cong \mathcal{L})$.] The sharp inequality for the $\Delta I'$ in (49) is a consequence of the second part of Carnot's theorem. Thus, following from (46) and assuming that (12) is valid, the inequalities

$$\Delta I' \leq T(X;Y) < H(X) \tag{50}$$

can be considered to be the *information* formulation of Kelvin's and Thomson-Planck's theorem and thus Carnot's theorem [the second part and also the first part; compare with (40) and (49)]. Therefore they are the ***information*** formulation of the II. Principle of Thermodynamics (7).

Because the value η_{max} is the maximum of a set of efficiencies η, it is obvious that $T(X;Y) \stackrel{\triangle}{=} T_{max}(X;Y)$. And, because it satisfies the definition of information capacity as the maximum (supremum) of a set of transinformations, the transinformation $T(X;Y)$, stated in (48) and (39), is the capacity C_{T_W,T_0} for the given extreme temperatures T_W and T_0. Consequently on the set of all heat engines with these extreme temperatures we define

$$C_{T_W,T_0} \stackrel{\text{Def}}{=} T(X;Y), \;\; T(X;Y) = \frac{\Delta Q_W}{kT_W} \cdot \frac{T_W - T_0}{T_W}. \tag{51}$$

When the states of the medium $\mathcal{L}$ have gone through the irreversible cycle $\mathcal{O}'$ once, and in agreement with (34), we have the following for the change $\Delta S_{\mathcal{L}}$ of the heat entropy $S_{\mathcal{L}}$ of the whole irreversible Carnot engine:

$$\Delta S_{\mathcal{L}} = \oint_{\mathcal{O}''} \frac{\delta Q}{T} = \frac{\Delta Q_W}{T_W} - \frac{\Delta Q_0 + \Delta Q_{0x}}{T_0} = -\frac{\Delta Q_{0x}}{T_0} < 0 \tag{52}$$

and following from (47), we can write this informationally

$$\Delta S_{\mathcal{L}} = k{\cdot}H(Y|X) \cdot \beta^{-1}, \;\; \beta = \frac{T_0}{T_W}, \;\; T_W{\geq}T_0.$$

When the medium $\mathcal{L}$ has gone through the irreversible cycle $\mathcal{O}'$ once, and due to the fact that heat entropy is additive (considering substitute changes (19) are reversible), we have the following for the change $\Delta S_{\mathcal{AB}}$ of the heat entropy $S_{\mathcal{C}}$ in the system $(\mathcal{AB})$:

$$\Delta S_{\mathcal{AB}} = -\frac{\Delta Q_0}{T_W} + \frac{\Delta Q_0}{T_0} + \frac{\Delta Q_{0x}}{T_0} = \frac{\Delta Q_W}{T_W} \cdot \eta_{max} + \frac{\Delta Q_{0x}}{T_0} \tag{53}$$

and, following from (47) and (48), written informationally

$$\Delta S_{\mathcal{AB}} = k{\cdot}[T(X;Y) - H(Y|X)] \cdot \beta^{-1} = k{\cdot}T(X;Y) - \Delta S_{\mathcal{L}}, \;\; \beta = \frac{T_0}{T_W}, \;\; T_W{\geq}T_0.$$

The whole irreversible Carnot engine is an isolated system, in which the transformation of heat energy $\Delta Q_W \sim x$ to mechanical energy $\Delta A' \sim y$ is performed. As a consequence of the additivity of heat entropy when substitute changes are considered to be reversible, and when (52) and (53) are used, we have the following for the resulting change $\Delta S_{\mathcal{C}} = \Delta S_{\mathcal{L}} + \Delta S_{\mathcal{AB}}$ of the heat entropy $S_{\mathcal{C}}$ of the whole irreversible Carnot engine:

$$\Delta S_{\mathcal{C}} = -\frac{\Delta Q_{0x}}{T_0} + \left(\frac{\Delta Q_W}{T_W} \cdot \eta_{max} + \frac{\Delta Q_{0x}}{T_0} \right) = \frac{\Delta Q_W}{T_W} \cdot \eta_{max} \tag{54}$$

and following (48) we can write informationally

$$\frac{1}{k} \cdot \Delta S_{\mathcal{C}} = H(X) \cdot \eta_{max} = T(X;Y) = C_{T_W,T_0}.$$

The relation (54) for the result change $\Delta S_{\mathcal{C}}$ of the heat entropy $S_{\mathcal{C}}$ of the whole irreversible Carnot engine is the same as in the case of a noiseless transfer within a reversible Carnot engine (43). Following from (49) and (54) we can thus immediately derive

$$\Delta S_{\mathcal{C}} - k{\cdot}\Delta I' > 0 \;\text{ or also }\; \Delta(S_{\mathcal{C}} - k{\cdot}I') > 0, \tag{55}$$

[see (44)].

The inequality (55) reveals that the result change $\Delta S_{\mathcal{C}}$ of the heat entropy $S_{\mathcal{C}}$ of the whole

irreversible Carnot engine stated in (54), together with the output average information $\Delta I'$ defined in (46), satisfies Brillouin's (2) extended formulation of the II. Principle of Thermodynamics,

$$\mathrm{d}(S_{\mathcal{C}} - k{\cdot}I') \geq 0 \text{ or, stated informationally, } \mathrm{d}[T(X;Y) - H(Y)] \geq 0. \tag{56}$$

It is obvious from both (46) and (54) that for the difference in (55) we have

$$\frac{1}{k}\cdot\Delta S_{\mathcal{C}} - \Delta I' = H(X)\cdot\eta_{max} - \left(H(X)\cdot\eta_{max} - \frac{\Delta Q_{0x}}{kT_W}\right) = \frac{\Delta Q_{0x}}{kT_W} > 0 \tag{57}$$

and, using the result (52), we can write informationally,

$$\frac{1}{k}\cdot\Delta S_{\mathcal{C}} - \Delta I' = |H(Y|X)|\cdot\beta = |\Delta S_{\mathcal{L}}|\cdot\frac{1}{k}\cdot\beta,\ \ \beta = \frac{T_0}{T_W},\ \ T_W{\geq}T_0 > 0.$$

Thus the result change $\Delta S_{\mathcal{C}}$ of the heat entropy $S_{\mathcal{C}}$ of the whole irreversible Carnot engine is greater than the output information $\Delta I'$ transferred through it.
In this case the result of the information transfer is worse than the noiseless case by the difference (57). Following from (44) and (54) we can see that $\Delta S_{\mathcal{C}} = \Delta I = T(X;Y)$.
When we use the definitions (37) and (46) directly, and express η from (31) using (29) and (30), it becomes obvious that we can state

$$\Delta I - \Delta I' = H(X)\cdot\left[\eta_{max} - \left(\eta_{max} - \frac{\Delta Q_{0x}}{\Delta Q_W}\right)\right] = \frac{\Delta Q_W}{kT_W}\cdot\frac{\Delta Q_{0x}}{\Delta Q_W} = \frac{\Delta Q_{0x}}{kT_W} > 0. \tag{58}$$

The equality of the left sides in (57) and (58) is obvious. Following from the definition (46), and with $0 \leq \eta{<}\eta_{max}{<}1$, we can state

$$\Delta I' = \frac{\Delta Q_W}{kT_W}\cdot\eta_{max} - \frac{\Delta Q_{0x}}{kT_W} = \frac{\Delta Q_W}{kT_W}\cdot\eta_{max} - \frac{\Delta Q_{0x}}{kT_0}\cdot\beta \tag{59}$$

and following (52) we can write informationally

$$\Delta I' = H(X)\cdot\eta_{max} + \Delta S_{\mathcal{L}}\cdot\frac{1}{k}\cdot\beta = C_{T_W,T_0} + \Delta S_{\mathcal{L}}\cdot\frac{1}{k}\cdot\beta =$$

$$= T(X;Y) - |\Delta S_{\mathcal{L}}|\cdot\frac{1}{k}\cdot\beta = T(X;Y) + \Delta S_{\mathcal{L}}\cdot\frac{1}{k}\cdot\beta,\ \ \beta = \frac{T_0}{T_W},\ \ T_W{\geq}T_0 > 0.$$

The quantity $|\Delta S_{\mathcal{L}}|$ we term *production of heat entropy* in the medium $\mathcal{L}$ (within the irreversible cycle $\mathcal{O}'$). From the relations (59) it is clear that the change in the *structure* of the mechanical output (created by the positive addition of $\Delta A'$ to the output potential energy) is expressed by the value of $\Delta I'$. This is less than ΔI at the value $|\Delta S_{\mathcal{L}}|\cdot\beta$ in the reversible case.
Therefore *noise output* exists, draining off the transferred information from the maximum:

$$\Delta I = T(X;Y) = H(X)\cdot\eta_{max} \text{ to the value}$$

$$\Delta I' = H(X)\cdot\eta < T(X;Y),\ \text{Carnot's theorem (the second part)}.$$

In our case of carnotized information transfers running in what we can term *Carnot's transfer system* (i.e. Carnot's engine as a thermodynamic, average-value realization, or model, of a Shannon transfer chain), we have

$$\begin{aligned} \Delta I' < T(X;Y),\ \ \text{when}\ \ \eta < \eta_{max},\ \ \Delta I' = H(X)\cdot\eta, \\ \Delta I' = \Delta I = T(X;Y),\ \ \text{when}\ \ \eta = \eta_{max},\ \ \Delta I' = H(X)\cdot\eta_{max}. \end{aligned} \tag{60}$$

In both reversible and irreversible cases we can state

$$T(X;Y) = H(X) \cdot \eta_{max}, \text{ Carnot's theorem (the first part).} \tag{61}$$

Because the I. and the II. Principle of Thermodynamics hold, it is obvious that the *thorough* transfer of any input message $x \sim \Delta Q_W$ with the (average) information value $H(X)$, expressed by

$$T(X;Y) = H(X), \tag{62}$$

is only the limit ($\eta_{max} \to 1$), but not achievable in reality.
As in the reversible case, our heat transfer process completes with the addition of $\Delta S_{\mathcal{C}}$ to the whole thermodynamic entropy $S_{\mathcal{C}}$, and for the average information $\Delta I'$ gained from $H(X)$, we have

$$\Delta S_{\mathcal{C}} \geq \Delta I' = H(X) \cdot \eta \geq 0 \text{ where } \Delta S_{\mathcal{C}} = \Delta I = H(X) \cdot \eta_{max}. \tag{63}$$

The equality $\Delta I' = \Delta I$ is valid only in a reversible transfer system in which no heat dissipation, generated from its non-ideal properties, exists, i.e. where $\Delta Q_{0x} = 0$.
Our thermodynamic-information derivation based on a heat cycle demonstrates the fact that it is impossible, in the type of channel considered, for the *bound* (2; 15) information contained in an input message to be transferred without its (average) loss. Such information transfer can be worsened only by heat dissipation of energy, which means by *noise* heat ($\Delta Q_{0x} > 0$) generated by the irreversible processes in the channel [described by a transformer $\mathcal{L}$ of input heat, which has non-ideal properties (inner friction)]. Simultaneously the whole thermodynamic entropy of the extended isolated system in which this process is running increases, and maximum average value of the output transferred information diminishes (7).

Summarizing Note

Even in the case of a *noiseless* but repeatable transfer of an input message within our "heat" type channel it is *impossible* to transfer this message *without* the average loss of the information it contains. This loss is the consequence of both the I. and II. Principle of Thermodynamics which determine a cyclical energy transformation in an isolated system. The resulting entropy increase ΔS_C within this process is caused by the requirement that the transfer be repeatable, or cyclical. Consequently the reduction, denoted as $H(X|Y)$, of the input average amount of information $H(X)$, is *inseparable* from such a repeatable process, being its *necessary* condition;

repeatibility of our information transfer => (average) loss of information.

This is the condition for a cyclical transfer $\mathcal{O}$ [transforming input energy (coding x) to output energy (coding y)] to function, physically as expressed by (26) and informationally as expressed by the inequality in (40),

$$T(X;Y) < H(X).$$

This relation is the necessary condition (expressed informationally) for our channel $\mathcal{K}$ to function repeatedly. Therefore our conclusion

$$H(Y) < H(X)$$

information form of Thomson-Planck's formulation of the II. Principle of Thermodynamics, and our conclusion

$$T(X;Y) = H(Y)$$

is the information variant of the first part of Carnot's theorem.

Assuming that $H(Y|X) < 0$ [information variant of Kelvin's formulation of the II. Principle of Thermodynamics for irreversible cyclical transfer $\mathcal{O}'$ when the respective relation (34) is valid], we also have the information variant of the second part of Carnot's theorem

$$H(Y) < T(X;Y).$$

Following the last four relations we can derive the inequality

$$H(Y) \leq T(X;Y) < H(X),$$

which is the *complete information formulation* of the II. Principle of Thermodynamics in our model, introduced only in (7) and in this paper. Also it is the necessary condition of channel equality (12). This equality can be considered, as such, as a reasoning of the equivalence of various formulations of the II. Principle of Thermodynamics. Or, **we can see (12) as the most general formulation of the II. Principle of Thermodynamics** (7).

Consequently and, as the example, our model, in which the last formulation is valid, satisfies Shannon's coding theorem (3) for the case the channel capacity $C < H(X)$; thus transfer errors are the inevitable consequence of the fact the transfer exists and thus, it is the physical reasoning of the *DPE, Data Processing Enequality* (3), $H(Z) < H(Y) < H(X)$, when a transfer $(X \longrightarrow Y \longrightarrow Z)$ is considered be "carnotized".

5.3 Reverse reversible Carnot cycle and channel

Reverse reversible Carnot cycle works as a *heat pump* described in subsection 4.3. In this cycle, comprehended as a thermodynamic, average-value realization, or model of the transfer process in a channel $\mathcal{K} \cong \mathcal{L}$, which is transferring an (arbitrary) input message $x \in X$ containing the average information amount $H(X)$, we use these symbols and denotations:

ΔQ_0 the heat drained off from the cooler $\mathcal{B}$ within the isothermal expansion at T_0,

ΔA the mechanical work delivered to the cycle by the compression phase at T_W,

ΔQ_W the output heat delivered to the heater $\mathcal{A}$ by the isothermal phase at T_W.

Further, we define the values of changes of information entropies on the channel $\mathcal{K} \cong \mathcal{L}$ (with an information transfer process being realized by this cycle) by the changes of its physical entropies, for instance, in this way:

$$\begin{aligned} H(X) &\overset{\text{Def}}{=} \frac{\Delta A}{kT_W}, \text{ input entropy,} \\ &\quad \Delta A \cong x \text{ input message;} \\ H(Y) &\overset{\text{Def}}{=} \frac{\Delta Q_W}{kT_W} = \frac{\Delta Q_0 + \Delta A}{kT_W} \overset{\triangle}{=} \Delta I, \text{ output entropy,} \\ &\quad \Delta Q_W \cong y \text{ output message;} \\ H(Y|X) &\overset{\text{Def}}{=} \frac{\Delta Q_0}{kT_W} > 0, \text{ noise entropy,} \\ &\quad \Delta Q_0, \text{ noise "message".} \end{aligned} \tag{64}$$

So, we consider a channel with *additive* noise. It is clear that

$$H(Y|X) = \frac{\Delta Q_0}{kT_W} \cdot \frac{T_0}{T_0} = \frac{\Delta Q_0}{kT_0} \cdot \beta = \frac{\Delta Q_W}{kT_W} \cdot \beta = H(Y) \cdot \beta, \; \beta = \frac{T_0}{T_W} \tag{65}$$

can be validated.
The noise with information entropy $H(Y|X)$ is the integral part of the definition of the transfer information process. It is not generated by a positive production of the heat $\Delta Q_{0x} > 0$ in the working medium $\mathcal{L}$.[19]
We are supposing further that for changes of information at the values $H(X)$, $H(Y|X)$, $H(Y)$, $H(X|Y)$ defined by (64) the relations (12) are valid and, then,

$$\frac{\Delta A}{kT_W} - H(X|Y) \;=\; \frac{\Delta Q_0 + \Delta A}{kT_W} - \frac{\Delta Q_0}{kT_W} \tag{66}$$

and then

$$H(X|Y) \stackrel{\text{Def}}{=} 0.$$

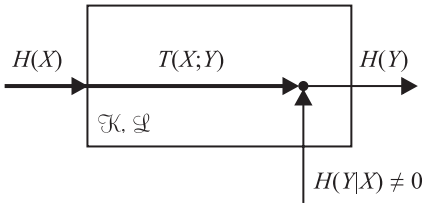


Fig. 8. The information transfer channel modelling a reverse and reversible Carnot Cycle

So, we have a channel *without losses*. For transinformation $T(X;Y)$, $T(Y;X)$, with respecting the definitinon (64) it is valid that

$$T(X;Y) = H(X) - H(X|Y) = \frac{\Delta A}{kT_W} - 0 = H(X) \tag{67}$$

and, also,

$$T(Y;X) = H(Y) - H(Y|X) = \frac{\Delta Q_0 + \Delta A}{kT_W} - \frac{\Delta Q_0}{kT_W} = \frac{\Delta A}{kT_W} = H(X). \tag{68}$$

The left sides of derivations (67) and (68) are then equal and they are valid for information amounts defined by (64) and for the result of derivation (66), the equution (12) for information conservation (both the average, information entropies and the "instantneous" ones) in the channel too. (Within one run of the system $\mathcal{L} \cong \mathcal{K}$ through the reverse Carnot cycle realizing a transfer process.)
It is evident that

$$\frac{H(X)}{H(Y)} = \frac{\dfrac{\Delta A}{kT_W}}{\dfrac{\Delta Q_W}{kT_W}} = \frac{\Delta A}{\Delta A + \Delta Q_0} = \eta_{\max} \tag{69}$$

and, consequently

$$H(X) = H(Y) \cdot \eta_{\max}, \tag{70}$$

[19] From the relations for η and $\eta_{\max}$ it follows that $\Delta Q_0 = f(T_0)$, where function $f(\cdot)$ is a not-negative function of the argument T_0, $f(T_0) \geq 0$, for which $\lim_{T_0 \to 0} f(T_0) = 0$ is valid.

where $\eta_{\max}$ is the efficiency of the relevant direct cycle. So it is valid, in accordance with relations (65) and (68), that

$$T(X;Y) = H(Y) \cdot \eta_{\max}. \tag{71}$$

Now, let us notice the changes of thermodynamic entropy in an isolated system in which the described process is running:

$$\begin{aligned} \Delta S_{\mathcal{AB}} &= \frac{-\Delta Q_0}{T_0} + \frac{\Delta Q_0}{T_W} = \frac{-\Delta Q_0}{T_0} \cdot \frac{T_W - T_0}{T_W} = \\ &= \frac{-\Delta Q_0}{T_0} \cdot \eta_{\max} = -H(Y) \cdot \eta_{\max} < 0. \end{aligned} \tag{72}$$

Thermodynamic (Clausius) entropy $S_{\mathcal{AB}}$ of the system $\mathcal{AB}$ is then lowering its value - the (thermodynamic, heat) distinguishability of the systems $\mathcal{A}$ and $\mathcal{B}$ is growing. Of course, it is at a detriment of the mechanical work ΔA delivered, or, of the entropy at the value $\frac{\Delta A}{kT_W}$. There is a need to gain this energy (entropy) and it is possible within such an isolated system by an *unnatural* process of transformation of heat to mechanical energy. But, this process is "running" at a background of a *natural* process of the heat transfer in accordance with the II. Principle of Thermodynamics.
We are considering such a reversible process, which is producing the mechanical work at a value $\Delta A^* \geq |\Delta A|$; we write it with regard to various directions of functioning of both these cycles;

$$\begin{aligned} \frac{\Delta A^*}{kT_W^*} &= H(X^*) \cdot {\eta_{\max}}^* = H(X^*) \cdot \frac{T_W^* - T_0^*}{T_W^*} = \Delta S_{\mathcal{A}^*\mathcal{B}^*}, \;\; T_W^* \geq T_0^* > 0, \\ \frac{|\Delta A|}{kT_W} &= H(Y) \cdot \eta_{\max} = H(Y) \cdot \frac{T_W - T_0}{T_W} = -\Delta S_{\mathcal{AB}}, \;\; T_W \geq T_0 > 0. \end{aligned} \tag{73}$$

For the whole change ΔS of the entropy S of the whole isolated system in which both these processes are running, following the II. Principle of Thermodynamics, it is valid that

$$\Delta S = \Delta S_{\mathcal{A}^*\mathcal{B}^*} + \Delta S_{\mathcal{AB}} \geq 0. \tag{74}$$

But, for $\Delta S_{\mathcal{AB}} \leq 0$, it must be valid that

$$\Delta S_{\mathcal{A}^*\mathcal{B}^*} \geq |\Delta S_{\mathcal{AB}}|. \tag{75}$$

This means that the greater decrease of entropy of the value $|\Delta S_{\mathcal{AB}}|$, the greater addition of $\Delta S_{\mathcal{A}^*\mathcal{B}^*}$ is to be generated, and, then, the whole entropy is growing at the value

$$\Delta S = \Delta S_{\mathcal{A}^*\mathcal{B}^*} - |\Delta S_{\mathcal{AB}}| \geq 0. \tag{76}$$

The equality occures when ${\eta_{\max}}^* = \eta_{\max}$. In another case ${\eta_{\max}}^* > \eta_{\max}$ which, e.g. for $T_0^* = T_0$, means that $\Delta Q_W^* > \Delta Q_W$ a $T_W^* > T_W$.

The environment of the entropy decrease is being exhausted, as for its structure which is conserving energy, at a greater degree (or at least at the same one) - its undistinguishability, disorder (chaos) is growing more (or at the same degree) than this decrease, the local growing of the order or organisation, is (9).

In the following parts of this text we will end our considerations about the analogy of heat transformation and information transfer. We will deal with direct processes with concentration of our attendance to their reversible type; the irreverse ones are expressible easily by those equivalent reverse ones.
However, we will return to the problem of a reverse cycle and its information-structural aspects in section 8 once more.

6. Gibbs paradox and information transfer

We observe an equilibrium system $\mathcal{A}$, $S* = S_{\text{Claus}} = S_{\text{Boltz}} = -kNB* = -kN \ln N$.
Let, in accordance with the *solution* of Gibbs paradox, the integration constant S_0 be the (change of) entropy ΔS, added to the entropy σ, *to figure out* the measured entropy S_{Claus} of the equilibrium state of the system $\mathcal{A}$ (the final state of Gay-Lussac experiment) at a temperature Θ. We have shown that without such correction, the lower entropy σ is evidenced, $\sigma = S_{\text{Claus}} - \Delta S,\ \Delta S = S_0$.
Following the previus definitions and results we have

$$\Delta S = \frac{\Delta Q_0}{\Theta} = -nR \ln \frac{n}{\gamma}, \tag{77}$$
$$\ln \gamma = \frac{\Delta S}{knN_A} + \ln n = \frac{\Delta S}{kN} + \ln N - \ln N_A,\ \gamma = N \ \Rightarrow\ \frac{\Delta S}{kN} = \ln N_A.$$

By the entropy ΔS the "lost" heat ΔQ_0 (at the temperature Θ) is defined.
Thus, our observation can be understood as an information transfer $\mathcal{T}$ in an information channel $\mathcal{K}$ with entropies (6)-(8), but now bound; we have these information entropies per one particle of $\mathcal{A}$:

$$\begin{aligned} &\textit{input}\ \ H(X) \overset{\text{Def}}{=} \frac{S*}{kN} = \ln \gamma = -B* = \ln N = -rB(r), \\ &\textit{output}\ \ H(Y) \overset{\text{Def}}{=} \frac{\sigma}{kN} = -B_{\text{Gibbs}} = -B_{\text{Boltz}} = -B(r), \\ &\textit{loss}\ \ H(X|Y) \overset{\text{Def}}{=} \frac{S_0}{kN}, \\ &\textit{noise}\ H(Y|X) = 0 \quad \text{by (12);} \\ &H(X|Y) = -rB(r) - [-B(r)] = B(r) \cdot (r-1) = (-B*) \cdot \frac{r-1}{r},\ r \geq 1. \end{aligned} \tag{78}$$

For a number m of cells of our railing in the volume V with $\mathcal{A}$, $m \leq N$, for a detail r of this our description of the "inner structure" of $\mathcal{A}$ (a thought structure of V with $\mathcal{A}$) and for the number q of diaphragmas creating our railing of cells and constructed in such a way that $q \in <1,\ m-1>$, we have $r = \dfrac{N-1}{q}$.
The *maximal* detail of our "description", the accuracy of our observation of the system $\mathcal{A}$, is achieved for $r = 1$. Then $B(r) = B*$ and, for the output, the input and the loss information entropies, it is valid that

$$H(Y) = H(X) = -B*,\ H(X|Y) = 0.$$

The *minimal* accuracy, detail of our description of the observed system should be for $r = \infty$. In this case we should place $q = 0$ diaphragmas, no railing is laid and $m = 1$. We are not

considering the "inner structure" of $\mathcal{A}$ in this case. Thus we define an output information source Y, bound, for which $H(Y) = -B_{\text{Gibbs}} = 0$. Then, the result of such "observation" is 0, and the loss information entropy is

$$H(X|Y) = \frac{S*}{kN} = \ln N = H(X).$$

Our observation of the equilibrium system $\mathcal{A}$, including the *mathematical correction* for the Gibbs paradox, is then describable by the Shannon *scheme* (12), where

$$H(X) = \frac{S_{\text{Claus}}}{kN},\ H(X|Y) = \frac{S_0}{kN},\ H(Y) = \frac{S_{\text{Claus}}}{kN},\ H(Y|X) = \frac{\Delta S}{kN}. \tag{79}$$

However, a real observation process described in (79), equivalent to that one with $r = 1$, is impossible (7).

We conclude this section by noting that the diminishing of the measured entropy value about ΔS against $S*$ awaited, evidenced by the **Gibbs paradox, *does not originate in an observed system itself***. Understood this way, **it is a contradiction of a *gnozeologic character* based on not respecting *real* properties of any observation** (7; 8). (The influence of our measuring, the properties of our measuring, are not included in our epistemology.)

6.1 Observation and II. thermodynamic principle

The heat ΔQ_0 is defined by, and *defines* the loss entropy $S_0 = \Delta S$ of our observation. We call it the *loss* heat. It expresses an energy needed for a realization of our observation method. Thus, it is natural to consider it be "paid" by the observed system $\mathcal{A}$ and its driving off from $\mathcal{A}$ (with temperature Θ) to $\mathcal{B}$ (with temperature $T_0 < \Theta$). Then,

$$\frac{\Delta Q_0}{T_0} = -\frac{Q - \Delta Q_0}{T_0} + \frac{Q}{T_0} = -\frac{Q - \Delta Q_0}{T_0} \cdot \frac{Q}{Q} \cdot \frac{\Theta}{\Theta} + \frac{Q}{T_0} \cdot \frac{\Theta}{\Theta} = \frac{Q}{\Theta}. \tag{80}$$

Our observation method, a bound information transfer $\mathcal{T}$, states the *identical* relation between two heats Q and ΔQ_0 at temperatures Θ and T_0 as a *reversible* (direct) Carnot cycle $\mathcal{O}$ with a *transformer* $\mathcal{L}$, the working temperature T_0 of its *cooler* $\mathcal{B}$ and Θ of its *heater* $\mathcal{A}$; it supplies a mechanical energy $\Delta A = Q - \Delta Q_0$. Thus, the relation (80) represents the integral formulation of the II. P.T., written generally as

$$\sum_{\mathcal{O}} \Delta S_{[\cdot]} \overset{\triangle}{=} \oint_{\mathcal{O}} \mathrm{d}S = \oint_{\mathcal{O}} \frac{\delta Q(\Theta)}{\Theta} = 0,\ \ \Theta \in (0, \infty),\ \ \mathcal{O} = \bigcup\nolimits_{\{\cdot\}} \mathcal{O}_{[\cdot]}. \tag{81}$$

The cycle $\mathcal{O}$ is a thermodynamic *realization* of our observation, transfer $\mathcal{T} \cong \mathcal{O}$ (7). The energy ΔA bears the whole (average) information $\sigma = S* - S_0 = kN \cdot H(Y)$, gained from $\mathcal{A}$ while it is observed by $\mathcal{T}$.

We consider a system of stochastic quantities, a transfer channel $\mathcal{K} \equiv (X, Y)$, with entropies from (6)-(8),

$$H(X),\ H(Y),\ H(X|Y),\ H(Y|X)$$

In the realization $\mathcal{O}$ of $\mathcal{T}$, $\mathcal{L} \cong \mathcal{K}$, the quantities X, Y, $X|Y$, $Y|X$ are bound, have a meaning of energies being input and being output from the transformer $\mathcal{L}$ undergoing $\mathcal{O}$[22]. "Bound" equality (12), proved in information theory (3; 20), written generally as

$$\sum_{\mathcal{T}} H_{[\cdot]} \overset{\triangle}{=} \oint_{\mathcal{T}} \mathrm{d}H = \oint_{\mathcal{O}} \mathrm{d}B = 0,\ \ \mathcal{T} = \bigcup\nolimits_{\{\cdot\}} \mathcal{T}_{[\cdot]}, \tag{82}$$

[22] Thus, our observing method includes, "in itself", its own repeatibility (7)

is valid for our $\mathcal{T} \cong \mathcal{O}$ or $\mathcal{L} \cong \mathcal{K}$ respectively, by the following assignment (7):

$$H(X) = \frac{Q}{kN\Theta} = \frac{\Delta Q_0}{kNT_0}, \quad H(Y) = \frac{\Delta A}{kN\Theta} = H(X) \cdot \eta_{max} \stackrel{\triangle}{=} \Delta I, \tag{83}$$
$$H(Y|X) = 0, \quad H(X|Y) = \frac{Q}{kN\Theta} \cdot \beta = \frac{\Delta Q_0}{kN\Theta},$$
$$H(X) - H(X|Y) = H(X) \cdot \eta_{max} = \Delta I < H(X),$$

where $\beta \stackrel{\triangle}{=} \frac{T_0}{\Theta} = \frac{\Delta Q_0}{Q}$.

For changes of bound information entropies of systems $\mathcal{L}$, $\mathcal{A}$, $\mathcal{B}$ (defined by a realization $\mathcal{O}$ of our observation $\mathcal{T}$), and for the whole isolated system $\mathcal{C}$ in which our observation is performed, it is valid that (7),

$$\Delta S_{\mathcal{L}} = \oint_{\mathcal{O}} \frac{\delta Q}{\Theta} = \frac{Q}{\Theta} - \frac{\Delta Q_0}{T_0} \stackrel{\triangle}{=} kN \cdot \Delta H_{\mathcal{L}} = 0, \tag{84}$$
$$\Delta S_{\mathcal{AB}} = -\frac{\Delta Q_0}{\Theta} + \frac{\Delta Q_0}{T_0} = \frac{\Delta Q_0}{T_0} \cdot \eta_{max} \stackrel{\triangle}{=} kN \cdot \Delta H_{\mathcal{AB}} = kN \cdot \Delta I \geq 0,$$
$$\Delta S_{\mathcal{C}} = \Delta S_{\mathcal{L}} + \Delta S_{\mathcal{AB}} = \frac{Q}{\Theta} \cdot \eta_{max} \stackrel{\triangle}{=} kN \cdot \Delta H_{\mathcal{C}} = kN \cdot \Delta I \geq 0.$$

The last inequality for $\Delta S_{\mathcal{C}}$ states that the *extended* II. P.T. for reversible processes (2; 7) is valid witin the system $\mathcal{C}$.[23] Written informationally,

$$\oint_{\mathcal{T}} \mathrm{d}H = \Delta(H_{\mathcal{C}} - I) = 0 \Rightarrow \Delta H_{\mathcal{C}} = \Delta I \geq 0. \tag{85}$$

For the detail $r \geq 1$, for the entropy growth $\Delta H_{\mathcal{C}}$ and for the *efficiency* η_{max} of the realized definition of our observation method (7) it is, by (12), (83)-(85), valid that

$$\Delta H_{\mathcal{C}} = -r \cdot B(r) - [-B(r) \cdot (r-1)] = \Delta I = -B_{\mathrm{Gibbs}}, \tag{86}$$
$$\Delta H_{\mathcal{C}} = \frac{-B*}{r} = \frac{-S*}{kNr} = \frac{Q}{kNr\Theta}, \text{ and then}$$
$$\Delta S_{\mathcal{C}} = \frac{Q}{r\Theta} \geq 0 \Rightarrow \eta_{max} = \frac{1}{r}, \ \beta = \frac{\Delta Q_0}{Q} = \frac{T_0}{\Theta} = \frac{r-1}{r}, \ r \geq 1.$$

The value η_{max}, set by the accuracy r of our observation, is the coefficient of the entropy growth for any *natural* process of heat transition between hot and cool environments. It is the efficiency of a cyclic reversible transformation using this transition. The value $1 - \eta_{max} = 1 - \frac{1}{r}$ sets the (average) information loss connected necessarily with an information transfer, observation realized this way (8).

[23] The same growth of entropy $S_{\mathcal{C}}$ could occur in an irrevirsible case too. Heat $\Delta Q_{0x} > 0$, from irreversibility of our (realized) observation, measuring, would diminish output $H(Y)$ from ΔI to $\Delta I - \frac{\Delta Q_{0x}}{kN\Theta}$; $\Delta H_{\mathcal{C}} = \Delta I \geq 0$ any way (7).

7. Equivalence principle of thermodynamics

The definition $\eta_{max} \overset{\text{Def}}{=} \frac{Q - \Delta Q_0}{Q} \in <0,1$ is a formulation of the *I. P.T.* too.
But, by derivations (83)-(86) it is clear that

$$\eta_{max} = \frac{\Delta H_{\mathcal{C}}}{H(X)}, \tag{87}$$
$$H(X) > 0 \Rightarrow \Delta H_{\mathcal{C}} \geq 0, \quad \Delta S_{\mathcal{C}} = kN \cdot \Delta H_{\mathcal{C}} \geq 0 \quad [II.\, P.T.].$$

Thus, in (87) the equivalent definition of η_{max} is given and it provides a formulation of the II. P.T. too. By (87) **the common formulation of the *Principle of Equivalence* of the I. and the II. P.T. is stated** (8).
Thus, by the derivation of (12) and by considering its equivalence with the information description of a reversible heat cycle (7), **the II. P.T. is prooved**.
Even more, the III. P.T. is deducible from (87) in that the interval of values of η_{max} must be, due to validity of the I. and the II. P.T., open on its left side $\Rightarrow T_0 > 0^\circ$ K.
At the end, let us formulate the theorem, the *Principle of Equivalence of the I., the II. and the III. P.T.*, the proof of which is provided by this paper.
Theorem (*Equivalence Principle of Thermodynamics*) (8).

Let (X, Y) be a system of (bound) stochastic quantities with information entropies $H(X)$, $H(Y)$, $H(X|Y)$, $H(Y|X)$ or, with relevant thermodynamic entropies $S_{[\cdot]} = (Q_{[\cdot]}) \cdot \Theta^{-1} = kN \cdot H(\cdot)$.
Let $H(X|Y) = \frac{\Delta Q_0}{kN\Theta}$, $\Delta Q_0 = Q \cdot \frac{T_0}{\Theta} = kN \cdot H(X) \cdot T_0$ where $\Theta \geq T_0 > 0$ and Θ^{-1} is Pfaffian integrating factor for $Q_{[\cdot]}$, N is a number of particles of $\mathcal{A} \cong \mathcal{X}$ (the selecting space of $X = [\mathcal{X},\ p(\cdot)]$).
Then, by (12) with $H(Y|X) = 0$ and for $S* \overset{\triangle}{=} kN \cdot H(X)$, $S_0 = (S*) \cdot \frac{T_0}{\Theta}$,

$$\frac{Q - \Delta Q_0}{Q} = \eta_{max} \in <0,1), \qquad [I.\, P.T.] \Leftrightarrow \tag{88}$$
$$\Leftrightarrow \frac{(S*) - S_0}{S*} = \frac{H(X) - H(X|Y)}{H(X)} = \eta_{max} \Rightarrow$$
$$\Rightarrow (S*) - S_0 = \Delta S_{\mathcal{C}} \geq 0 \text{ and}$$
$$H(X) - H(X|Y) = \Delta H_{\mathcal{C}} \geq 0, \qquad [II.\, P.T.] \Leftrightarrow$$
$$\Leftrightarrow \frac{\Theta - T_0}{\Theta} = \eta_{max} \Rightarrow \Theta \geq T_0 > 0, \quad [III.\, P.T.].$$

The II. P.T. is *derivable* logically by properties of a (bound) stochastic system *(X, Y)*; the others are derivable from this one; they all are equivalent.

Q.E.D.

8. Possible qualitative biology analogy

Let us think now, in a *free way* about possible applications of the rsults of this article in *biology* (10).[25] We can consider a principle loss of information (structure) within the process of proliferation of cells by their duplication (dividing). Within any dividing of a cell (*the predecessor*) its *follower (equivocant)* is generated and the a distortion of the *duplicated* (copied) structure (information) of the parent cell arises. This loss is measurable by the quantity of (average) information amount. It is a loss of part of message being copied (transferred), a loss of information within this process of such an information transfer. The whole structure of the cell is that message, including "a program" for its functionality in a texture.

8.1 Analogy I

The result of the *normal* generation of cells is that: the continuously decreasing precise of the structure of the followers is generated in a sequence of copying, dividing the cells, and, the whole biological organism, the collection of textures of cells, is aging by "the tooth of time" - by the loss of their *structure*, precision of the construction of the follower cells (for this, by the loss of both their inner and outer bounds). The notion of structure of an object (a message, a cell) is to be understood as follows: The object has more structure with the greater amount of information it contains, the more complicated it is - the greater number of parts bound mutually it contains, and, also, the less probability of its construction, of its structure is evidenced, or, the less stability as an isolated physical system it reveals. At the end, following a number of divisions, the incompatibility of the resulting structure of the last cell (the incompatibility of the information amount being represented by this cell) with such a minimal structure (information), which maintains its ability of both inner and outer communication in such a way that it is recognizible as "a certain cell" of the certain cell type - so as it is being alive, arises.

This mechanism can be described in a functional way by our *direct* 'carnotized' model of information transfer. Within any run of our model Carnot cycle the less output (average) information is gained in comparison with the input one (37),

$$\Delta I < \Delta H(X). \tag{89}$$

If this result information is used again the less information is generated etc. But, contemporarily, any run of the cycle generates the positive addition of entropy (thermodynamic) of a wider isolated system in which this transformation (transfer) runs (44),

$$\Delta S_{\mathcal{C}} > 0. \tag{90}$$

In this model it is the less and less distinguishability of parts $\mathcal{A}$ and $\mathcal{B}$ of the whole model heat engine, as for their heat contents - our thermodynamic model of that oldering by 'a tooth of time'. So, the sought 'gen of oldering' could be nothing else than a datum about the precision of that duplication; in our model its the efficicency of the transformation of the input energy,

$$\eta_{max} < 1. \tag{91}$$

So it is clear that the loss of information transferred this way is a suitable functional model for the case of the dividing of cells.

[25] The author is emphesizing strongly that the ideas expressed below are free *hypotheses* of a functional analogy type only.

As for cells, this principle reveals, signalizes itself by the phenomenon of the *shortening of telomeras*,[26]

$$l \cong f(S^+ - \Delta S_{\mathcal{AB}}) < S^+, \ \Delta S_{\mathcal{AB}} > 0. \tag{92}$$

where f is a growing function and S^+ is a starting value for the given generation run (10).

8.2 Analogy II

Within the *pathological* generation, proliferation of cells the opposite situation than it is in a normal case arises. In the pathological case we have the cells with a more precise structure (but, of an another type in comparison with the normal structure of its own original and 'normal' type is). But, this grow of structure in a certain locality in the whole organism is paid by a pumping off energy from an environment of this locality.

This environment is then in a lack of energy for its normal functionality; this energy is consumpted in the favour of this local grow of structure signalized by that phenomenon that the followers of the predecessors have the telomeras lengthened,[15]

$$l \cong f(S^+ - \Delta S_{\mathcal{AB}}) > S^+, \ \Delta S_{\mathcal{AB}} < 0. \tag{93}$$

where f is a growing function where f is a growing function and S^+ is a starting value for the given generation run.

This situation is describable again by our information-thermodynamic model, but by a *reverse* this time. In this model a *local* decreasing of entropy is evidenced (72),

$$\Delta S_{\mathcal{AB}} = -H(Y) \cdot \eta_{\max} < 0, \ H(Y) > 0. \tag{94}$$

as for the heat contents of the system $(\mathcal{AB})$, and, then, the grow of structure of such a locality (within a wider isolated system) appears.

However, for this decrease of entropy (for this grow of structure) the energy is needed (73),

$$\Delta A^* > 0. \tag{95}$$

but, delivered from the environment (75),

$$\Delta S_{\mathcal{A}*\mathcal{B}*} > 0. \tag{96}$$

of this locality which leads to a grow of entropy of this environment and, consequently, to a grow of entropy of the whole isolated wider system (organism) (74),

$$\Delta S > 0. \tag{97}$$

This is the price for the greater and greater structure of this its (open) part, the problematic locality. In the case of cells we can see the wasting away of the whole organism. Our reverse information-thermodynamic model authorizes us to an awaiting of a stable (moderate) higher body temperature of a patient and, also, a less temperature of the problematic texture (73),

$$T_W^* > T_W. \tag{98}$$

Also, as for therapy, we are authorized to consider an energy consumption from the organism, body as a whole, e.g. by its cooling (therapy by intensive freeze, swimming in a cold bath; especially in a not-operabile cases of such disease). This could be the way to limit, at least

[26] *Wen-Chi Hsueh, University of California, San Francisco; Proceedings of the National Academy of Sciences, 2007*

for a short time, the positive flow of heat supporting the cell metabolic processes in such problematic localities .This means the lowering of the temperature gradient between the sound and the ill textures (10).

For the pathology proliferating texture is consumpting the energy from the rest of organism it is logical to try to cut off this energy supply. From this point of view the therapy by the *devitalization* of a pathology tumour [*MUDr. Karel Fortýn, CSc., Ústav živočišné fyziologie a genetiky AV ČR v Liběchově, 1957, 1971-2001*] seems to be a right way.

9. Conclusion

Our paper is arranged as a sequence of successive steps, definitions and derivations (4)-(12) and interpretations (82)-(86) especially, stating gradually, the relation between combinatorial definition of Shannon (information) entropy and Boltzman and Clausius (thermodynamic) entropy, and, finally, resulting in (87) and (88). Although we combined the known facts about heat transformations and the Shannon's concept of an information transfer chain, *this combination presented has not been used yet by another else*, as far as the author of the paper is informed.

A next possible application of the opinion presented could be used for another transfer (transmission, transition) structures, e.g. electric circuits or computational processes and, also it could be used for stating of limits of possible behavior of anticipatory systems of various types, e.g. for dynamic systems with their bifurcations viewed as such a kind of information transfer process. The achieved results are planned to be used in studying relations of Thermodynamics and Logics and Thermodynamics of Computing.

Following our information-thermodynamic analyses, it seems to be right to await that this our attitude is capable, within the quantitave-functional-analogy point of view, to define properties, quality of very various phenomena.

10. References

[1] Bennett, Ch. H. The thermodynamics of computation – a Review. *Int. J. Theor. Phys.* 1982, *21* (12), 905–940.

[2] Brillouin, L. *Science and Information Theory*; Academia Press: New York, 1963.

[3] Cover, T. M.; Thomas, J. B. *Elements of Information Theory*; Wiley: New York, 1991.

[4] Gershenfeld, N. Signal entropy and the thermodynamics of computation. *IBM Systems Journal* 1996, *35* (3/4), 384.

[5] Hála, E. *Úvod do chemické termodynamiky*; Academia: Praha, 1975.

[6] Hašek, O.; Nožička, J. *Technická mechanika pro elektrotechnické obory II.*; SNTL: Praha, 1968.

[7] Hejna, B. Thermodynamic model of Noise Information Transfer. In *AIP Conference Proceedings*, Computing Anticipatory Systems: CASYS'07 – Eighth International Conference; Dubois, D., Ed.; American Institute of Physics: Melville, New York, 2008; pp 67–75.

[8] Hejna, B. Gibbs Paradox as Property of Observation, Proof of II. Principle of Thermodynamics. In *Computing Anticipatory Systems: CASYS'09 - Nineth International Conference*; Dubois, D. M., Ed.; 2009. Proceedings of CASYS'09 131-140, ISBN 978-0-7354-0858-6, ISSN 0094-243X, American Institute of Physics, Melville, New York, 2010

[9] Hejna B. *Informační Termodynamika I., Rovnovážná termodynamika přenosu informace*; ISBN 978-80-7080-747-7, VŠCHT Praha, Praha 2010.
[10] Hejna, B. Thermodynamics, Information and Telomera. In *Proceedings of International Conference Automatics and Informatics'10, I-59 –I-64*; Atanasoff, J., Ed.; Society of Automatics and Informatics: Sofia, Bulgaria, 2010.
[11] Horák, Z.; Krupka, F. *Technická fyzika*; SNTL/ALFA: Praha, 1976.
[12] Jaynes, E. T. Evolution of Carnot's Principle. Reprinted in Ericksen & Smith. 1988, *1*, 267–282.
[13] Jaynes, E. T.; Smith, C. R.; Ericksen, G. J.; Neudorfer, P. O. The Gibbs Paradox. *Kluwer Academic Publishers* 1992, 1–22.
[14] Košťál, K. *Statistická fyzika: Vybrané partie*; ČVUT: Praha, 1973.
[15] Kotek, Z.; Vysoký, P.; Zdráhal, Z. *Kybernetika*; SNTL: Praha, 1990.
[16] Kvasnica J. *Termodynamika*; SNTL: Praha, 1965.
[17] Landauer, M. Irreversibility and Heat Generation in the Computing Process.*IBM J. Res. Dev.* 2000, *44* (1/2).
[18] Maršák, Z. *Termodynamika a statistická fyzika*; ČVUT: Praha, 1995.
[19] Moore, W. J. *Fyzikální-chemie*; SNTL: Praha, 1981.
[20] Shannon, C. E. A mathematical theory of communication. *The Bell Systems Technical Journal* 1948, *27*, 379–423, 623–656.

5

Mesoscopic Thermodynamics in the Presence of Flow

I. Santamaría-Holek[1], R. Lugo-Frías[1], R. F. Rodríguez[2] and A. Gadomski[3]

[1]*UMJ-Facultad de Ciencias, Universidad Nacional Autónoma de México*
[2]*Instituto de Física, Universidad Nacional Autónoma de México*
[3]*Institute of Mathematics & Physics, University of Technology & Life Sciences*
[1,2]*México*
[3]*Poland*

1. Introduction

Thermodynamics is a powerful theory that has been succesfully applied to describe the properties and behavior of macroscopic systems under a very wide range of physical conditions (Callen, 1985). Thermal and caloric information is obtained by performing experiments that agree with the constraints imposed (Callen, 1985; Kondepudi & Prigogine, 1999; Ragone, 1995).

However, a considerable number of experimental, practical problems and systems of interest are often found in nonequilibrium or metastable states in which static thermodynamic relations are only valid locally, and frequently are insufficient to describe complicated time dependent situations (Demirel, 2007). Macroscopic systems in flow conditions are good examples of this peculiar behavior through non-Newtonian rheological effects, phase transitions and generalized statistics of turbulent motion. In fact, although there is a large amount of theoretical work proposing generalizations of the thermodynamic theory to nonequilibrium or quasiequilibrium situations (Beck & Cohen, 2003; Beris & Edwards, 1994; C. Beck & Swinney, 2005; de Groot & Mazur, 1984; Demirel, 2007; Kondepudi & Prigogine, 1999; Onuki, 2004; Rodríguez & Santamaría Holek, 2007), some of them very succesful (de Groot & Mazur, 1984; Demirel, 2007; Kondepudi & Prigogine, 1999), they are restricted due to the assumption and validity of local thermodynamic equilibrium. Recently, some interesting generalizations of thermodynamics to describe macroscopic systems in the presence of flow have been systematically developed in Refs. (Beris & Edwards, 1994; Onuki, 2004). Nevertheless, there are still a good deal of open questions, for instance those concerning the validity of the usual relations for the thermal and caloric equations of state, arising further research on this subject.

Related to these general questions, there are several particular manifestations of the effects of flow on the thermodynamic behavior of systems. For example, diffusion of suspended particles in a heat bath in equilibrium, may strongly differ from that when the particles diffuse in a heat bath under the presence of shear flow. These effects have been analyzed along the years in studies ranging from experiments (Breedveld, 1998; Guasto & Gollub, 2007; Pine, 2005; Taylor & Friedman, 1996) and computer simulations (Sarman, 1992) to kinetic

theory of gases (Rodríguez, 1983), projector operator techniques (Shea & Oppenheim, 1998), Langevin and Fokker-Planck dynamics (Drossinos & Reeks, 2005; Mauri & Leporini, 2006; Ryskin, 1988; Subramanian & Brady, 2004) and mesoscopic thermodynamics (Gadomski, 2008; Santamaría Holek, 2005; 2009; 2001).

The key point in these studies is that the presence of flow introduces non-thermal diffusion effects that can be related to the hydrodynamic viscosity and to hydrodynamic interactions between particles in the case of concentrated suspensions. These effects may modify the fluctuation-dissipation relations and, consequently, the corresponding expressions for the thermal energy and other state variables like pressure or chemical potential. This may lead to what is known as nonequilibrium state equations (Onuki, 2004).

In this chapter, we present a discussion of the general context involving thermodynamics in the presence of flows, emphasizing how the flow may modify the state equations as well as the transport coefficients. A selected presentation of some experimental results in which the effect of the flow clearly modifies the thermodynamics of the system is given and discussed in view of theoretical models attempting to describe them, mainly those coming from mesoscopic thermodynamics. How non-Newtonian effects (Santamaría Holek, 2005) arise from nonequilibrium state equations, how the transition to irreversibility in suspensions of non-Brownian particles ocurrs (Breedveld, 1998; Guasto & Gollub, 2007; Pine, 2005; Santamaría Holek, 2009; Taylor & Friedman, 1996) and how the nucleation in the presence of shearing flows is promoted or supressed (A. Penkova, 2006; Blaak, 2004), are the three main topics that we will analyze in more detail.

The original contribution of this chapter bears on the discussion of two recent studies on the problem of nucleation in the presence of flows (Blaak, 2004; Reguera & Rubi, 2003b) and the formulation of a new model that allows to explain experimental results (A. Penkova, 2006). These models are excellent illustrations of how the effects of flow affect the equations of state of fluid systems. Particularly, we will show how control volume thermodynamics in the presence of flow, a well established analysis of classical thermodynamics (Ragone, 1995), can be used to formulate a qualitative theoretical explanation of the experimental observations.

This chapter is organized as follows. Section 2 is devoted to analyze the effects of flow on diffusion from different points of view ranging from kinetic theory of gases, Langevin equation approach and mesoscopic nonequilibrium thermodynamics. In Section 3 we analyze how flow modifies the constitutive relations of materials via viscoelastic and non-Newtonian effects. Thereafter, the existence of nonquilibrium equations of state is discussed in Section 4 and then used to describe nucleation in the presence of flow in Section 5. Finally, we present our conclusions in Section 6.

2. Effects of flow on diffusion

The effects of flow on the diffusion of particles suspended in a simple liquid is an important problem which was analyzed many times along the last 30 years. Several approaches have been followed in order to understand and quantify the effects of the presence of a velocity gradient on the diffusion coefficient D of the particles. Usually, these are restricted to the case of shear flow because this case is more manageable from the mathematical point of view, and because there are several experimental systems that allow the evaluation and validity of the corresponding results. The approaches followed vary from kinetic theory and projector operator techniques, to Langevin and Fokker-Planck equations. Here, we summarize some recent contributions to this subject and their results.

2.1 Kinetic theory approach.

From a rigorous point of view, the problem of diffusion in the presence of flow should be analyzed by means of the kinetic theory gases. This has been done in a series of articles (Refs. (Ferziger & Kaper, 1972; Marchetti &Dufty, 1983; McLennan, 1989; Rodríguez, 1983)) and lead to a better understanding of the phenomenology presented by these systems. The rigorous derivations coming from kinetic theory serve to indicate and evaluate the range of validity of those coming from different approaches, some of them will be discussed later.

Essentially, the problem consists in the coupling of the dynamics of two subsystems, the host fluid and an ensemble of tagged particles that, in the case of massive and non-concentrated particles, constitute an ideal Brownian gas. Each susbsystem can be described by means of their corresponding phase space densities that evolve in time following a set of coupled Boltzmann equations.

In first approximation (Marchetti &Dufty, 1983), the system considered is a tagged particle in a low density gas of Maxwell molecules interacting through a r^{-5} force law. The fluid is assumed to be in a state of uniform shear flow with a large shear rate and the tagged particle is considered to be mechanically equivalent but with a larger mass than the fluid particles. The fluctuations of the tagged particle in the nonequilibrium gas can be described kinetically using the nonlinear Boltzmann and the Boltzmann-Lorentz equations.

$$\left(\frac{\partial}{\partial t}+\mathbf{v}\cdot\nabla\right)f=J\left[f,f\right] \tag{1}$$

$$\left(\frac{\partial}{\partial t}+\mathbf{v}\cdot\nabla\right)h=J\left[f,h\right] \tag{2}$$

where $f(\mathbf{x},t)$ and $h(\mathbf{x},t)$ are the phase space densities for the fluid and tagged particle, and $\mathbf{x}$ denotes the position and velocity of the later. $J\left[f,f\right]$ is the usual bilinear Boltzmann operator (Ferziger & Kaper, 1972; McLennan, 1989). The previous pair of equations is completed considering the evolution equation for the fluctuations of the tagged particle $C_{1,2}$

$$\left(\frac{\partial}{\partial t}+\mathbf{v}_1\cdot\nabla_1\right)C_{1,2}=J\left[f,C_{1,2}\right] \tag{3}$$

where

$$C_{1,2}(\mathbf{x}_1,t+\tau;\mathbf{x}_2,\tau)=\left\langle\delta\left[\mathbf{x}_1-\mathbf{x}_T(t+\tau)\right]\cdot\left\{\delta\left[\mathbf{x}_2-\mathbf{x}_T(\tau)\right]-\left\langle\delta\left[\mathbf{x}_2-\mathbf{x}_T(\tau)\right]\right\rangle\right\}\right\rangle \tag{4}$$

In the special case of Maxwell molecules with uniform shear flow one can obtain an exact set of kinetic equations for the average position and velocity, as well as their fluctuations, from Eqs. (1)-(3).

This kinetic description can be extended to the case where the tagged and fluid particles are assumed to interact via the same force law that holds for fluid particles. The mass m of the tagged particle is assumed to be large compared to that of a fluid particle (m_f). Therefore, the mass ratio $\epsilon = m_f/m$ is a small parameter in terms of which the Boltzmann- Lorentz collision operator may be expanded. If this expansion is carried out to the leading order, $\epsilon^{1/2}$, the Boltzmann-Lorentz operator reduces to a differential operator yielding a kinetic Fokker Planck equation for the tagged particle distribution F

$$\left(\frac{\partial}{\partial t}+\mathbf{v}\cdot\nabla\right)F=\frac{\partial}{\partial v_i}\left[A_i(\mathbf{r},\mathbf{v},t)+\frac{1}{2}\frac{\partial}{\partial v_j}D_{ij}(\mathbf{r},t)\right]F, \tag{5}$$

where the drift vector $A_i(\mathbf{r},\mathbf{v},t)$ and the diffusion tensor $D_{ij}(\mathbf{r},t)$ are given by

$$A_i(\mathbf{r},\mathbf{v},t) = \nu_1 \frac{n(\mathbf{r},t)}{n_0}\left[\mathbf{v}-\mathbf{v}_0(\mathbf{r},t)\right], \tag{6}$$

$$D_{ij}(\mathbf{r},t) = 2\rho_0^{-1}\left[\nu_1 p(\mathbf{r},t)\delta_{ij} + (\nu_1-\nu_2)\mathbb{P}^*_{ij}(\mathbf{r},t)\right]. \tag{7}$$

Here, ν_1 and ν_2 are constants proportional to $\epsilon^{1/2}$, and $n(\mathbf{r},t)$, $p(\mathbf{r},t)$ and $\mathbf{v}_0(\mathbf{r},t)$ are the local density, pressure and flow velocity of the fluid. The constant n_0 is the average number density and $\rho_0 = mn_0$. Finally, $\mathbb{P}^*_{ij}(\mathbf{r},t)$ is the traceless part of the fluid pressure tensor and represents the irreversible part of the momentum flux proportional to $\nabla\mathbf{v}_0$.
Although equation (7) is valid for an arbitrary nonequilibrium state of the fluid, for a fluid in uniform shear flow $v_{0,i}(\mathbf{r}) = \dot{\gamma}_{ij}r_j$, where $\dot{\gamma}_{ij}$ is the rate tensor, there are important simplifications. In this case the drift vector remains the same, but the diffusion tensor becomes independent of $\mathbf{r}$ and $\mathbf{v}$

$$D_{ij}(\mathbf{r},t) = 2\nu_1[k_BT(t)/m]\delta_{ij} + 2(\nu_1-\nu_2)\rho_0^{-1}\mathbb{P}^*_{ij}(t). \tag{8}$$

The results obtained from a kinetic description show that the drift vector and the diffusion tensor can be exactly calculated and that they depend on the nonequilibrium state of the gas only through the low order moments of the fluid distribution function. Even more, the fluid state in this case is specified by the nonequilibrium temperature $T(t)$ and the irreversible stress tensor $\mathbb{P}^*_{ij}(t)$ which is proportional to the velocity gradient $\nabla\mathbf{v}_0$.
These results imply that the extension of equilibrium theories to nonequilibrium states is not always valid in a straightforward way. Particularly, the diffusion tensor is proportional to the components of the pressure tensor or equivalently to the velocity gradient $\nabla\mathbf{v}_0$, which implies that the amplitude of the noise in the dynamics of the tagged particle is not simply thermal as in equilibrium since the diffusion tensor cannot be characterized entirely by the thermodynamic temperature. In similar manner, Eq. (5) does not depend on the irreversible heat flux. This is an anomaly of the Maxwell potential, for other potentials there will be an additional contribution to the drift vector that would depend on the any temperature gradient in the fluid.

2.2 Langevin equation approach

Another interesting approach to the effect of flow on diffusion is the one based on Langevin equations (Drossinos & Reeks, 2005; Swailes, 2009). Related to this approach is the interesting discussion on the existence and generality of effective nonequilibrium temperatures, which attracted much attention in recent years (Kurchan, 2005; Mauri & Leporini, 2006; Pérez-Madrid, 2005; Pérez-Madrid & Santamaría Holek, 2009; Popov & R., 2007). The possibility of using thermodynamic quantitites like the temperature or the chemical potential in systems out of equilibrium is attractive because it may strongly simplify the solutions to many practical problems (Demirel, 2007).
Specifically, the existence of a nonequilibrium temperature in sheared systems lead to a considerable theoretical (Criado-Sancho, 2005; Mauri & Leporini, 2006), numerical (Kurchan, 2005; O'Hern, 2004; Sarman, 1992) and experimental works with particular emphasis to the nonequilibrium properties of colloidal suspensions(Abou & Gallet, 2004; Bellon, 2001). These studies are relevant to us because kinetic nonequilibrium temperatures may not follow equipartition and are usually related to the breakdown of the fluctuation-dissipation theorem,

a central relation of Langevin descriptions (Criado-Sancho, 2005; Santamaría Holek, 2001; Subramanian & Brady, 2004).
Suspended objects diffuse in and are convected by the fluid; their motion may eventually be influenced by the presence of external fields. In Ref. (Drossinos & Reeks, 2005) the simultaneous diffusive and inertial motion of particles in simple shear flow is investigated to determine the effect of particle inertia on diffusive transport. First, a mesoscopic approach is used to describe the motion of a rigid spherical Brownian particle in a two dimensional simple shear flow with shear rate tensor $\dot{\gamma}_{ij}$, (i,j = x, y), with a Langevin equation in a Lagrangian reference frame. In this model the fluid velocity is along the x-direction, $v_i = \dot{\gamma}_{ij} r_j$, and the only nonzero element of $\dot{\gamma}_{ij}$ is $\dot{\gamma}_{xy} = \dot{\gamma}$. However, since the particle density (ρ_p) is much larger than the fluid density, (ρ_f), modifications to Stokes drag such as Basset history integral or Faxen's corrections, are neglected. Furthermore, only small-diameter, low-inertia particles ($a \geq 50\mu m$), are considered so that the gravitational settling and the Saffman lift force may be also neglected. The particle equation of motion is

$$\frac{d}{dt}\mathbf{u} = \beta\left(\underline{\dot{\gamma}} \cdot \mathbf{r} - \mathbf{u}\right) + \mathbf{f}(t), \tag{9}$$

where the random force per unit particle mass, $\mathbf{f}(t)$, is taken to be white in time,

$$\left\langle f_i(t) f_j(t') \right\rangle = q\delta_{ij}\delta\left(t - t'\right), \tag{10}$$

The friction coefficient β is the inverse particle's relaxation time, $\beta = 9\mu_f/(2\rho_p a^2)$, where μ_f is the fluid's dynamic viscosity. Since the Langevin equations are linear, particle velocity and position may be formally solved as functionals of the random force, and in the diffusive limit $t >> \beta^{-1}$, i. e., for times much larger than the particle relaxation time, they allow for the analytical evaluation of ensemble averaged products of particle position and velocity and two-point correlation functions, in terms of the random-force strength q. The authors carefully justify why they use the classical (equilibrium) form of the fluctuation-dissipation theorem (FDT): in a Langevin description the time scale of the white noise is considered to be much shorter than the time scale of the imposed flow. Thus, the non-equilibrium corrections would be of the order of the ratio of the fluid molecular relaxation time to the time scale of the imposed shear and may be neglected. In this case both the time scales are clearly separated and q may be determined solely from the classical form of the FDT,

$$\frac{q}{2\beta} = \frac{k_B T}{m}. \tag{11}$$

In the diffusion limit it is found that the combined effects of particle inertia and shear flow modify the amplitude and the time-dependence of the particle-velocity autocorrelation functions, a result which is expressed in terms of the Stokes number, $St = \dot{\gamma}/\beta$. The shear flow breaks macroscopic time reversibility and stationarity: the autocorrelation functions of the particle velocities are stationary and the velocity correlation along the shear is symmetric in the time difference τ, but the cross correlation is non-symmetric in τ function in the streamwise direction is non-stationary. The time decay of the velocity correlation along the flow is not a pure exponential and the underlying stochastic process is not an Ornstein-Uhlenbeck process.

Secondly, the authors give an Eulerian description of the motion of N independent and identical Brownian particles in terms of the phase space density $f(\mathbf{r},\mathbf{u};t)$ obeying the Fokker-Planck equation (FPE)

$$\frac{\partial f}{\partial t}+\nabla\cdot(\mathbf{u}f)=\beta\frac{\partial}{\partial \mathbf{u}}\cdot[(\mathbf{u}-\mathbf{v})\,f]+\frac{k_BT}{m}\beta\frac{\partial^2 f}{\partial \mathbf{u}^2}, \tag{12}$$

where $\nabla\equiv\partial/\partial\mathbf{r}$ and the classical FDT has also been used. The solution of this equation defines a Gaussian process which is used to evaluate density-weighted ensemble averages like particle concentrations, mean particle velocities and particle-velocity covariances. In particular, by taking velocity moments of the FPE the appropriate coupled mass and momentum equation for the particle phase are obtained for the case of a general flow and a white noise random force,

$$\frac{d\rho}{dt}=-\rho\nabla\cdot\langle\mathbf{v}\rangle, \tag{13}$$

$$\frac{d}{dt}\langle\mathbf{u}\rangle=\beta\left(\mathbf{v}-\langle\mathbf{u}\rangle\right)-\frac{1}{\rho}\nabla\cdot\left(\rho\left\langle\mathbf{u}'\mathbf{u}'\right\rangle\right). \tag{14}$$

Here $\mathbf{v}'=\mathbf{v}-\langle\mathbf{v}\rangle$ is the fluctuating component of the velocity and the overbar or angular brackets denote an average over a normalized velocity probability density. Thirdly, by using the analytical solution of the FPE, in conjuction with its first two-velocity moment equations given above, a convective-diffusion equation, or generalized Smoluchowski equation (GSE), may be derived without the use of non-perturbative approximations in the diffusive limit,

$$\frac{\partial\rho}{\partial t}+\nabla\cdot(\rho\mathbf{v})=\langle(\mathbf{u}-\mathbf{v})\rangle\,\mathbf{r}\cdot\nabla^2\rho\equiv\mathbb{D}\cdot\nabla^2\rho. \tag{15}$$

This equation incorporates inertial effects on diffusional transport for dilute suspensions and defines the diffusion tensor, $\mathbb{D}(St)$, dependent on both particle inertia and the shear rate through the Stokes number St.

The GSE is valid in the transition regime between the diffusion limit and the inertia-dominated limit and incorporates both particle transport mechanisms. The diffusion tensor is not symmetric reflecting the symmetry-breaking effect of the imposed shear. The total diffusion coefficients measuring the particle mean square displacement are always positive and depend on the shear rate and particle inertia. However, the stream-wise diffusion coefficient becomes negative with increasing Stokes number and one of the cross coefficients is always negative.

2.3 Mesoscopic nonequilibrium thermodynamics approach

A different formalism in which the diffusion of a Brownian gas in a fluid under stationary and non-stationary flow has been analyzed is mesoscopic nonequilibrium thermodynamics (MNET) (Pérez-Madrid, 1994; Rubi & Mazur, 1994; Rubi & Pérez-Madrid, 1999). This theory uses the nonequilibrium thermodynamics rules in the phase space of the system, and allows to derive Fokker-Planck equations that are coupled with the thermodynamic forces associated to the interaction between the system and the heat bath. The effects of this coupling on system's dynamics are not obvious. This is the case of Brownian motion in the presence of flow where, as we have discussed previously, both the diffusion coefficient and the chemical potential become modified by the presence of flow (Reguera & Rubi, 2003a;b; Santamaría Holek, 2005; 2009; 2001).

Using the principles of MNET, a thermodynamics for systems in a stationary state induced by an imposed flow was formulated based on the assumption of the existence of local equilibrium

in the phase space of the system (Santamaría Holek, 2005). This assumption enables one to interpret the probability density and its conjugated nonequilibrium chemical potential, as mesoscopic thermodynamic variables.

2.4 Mesoscopic nonequilibrium thermodynamics in the presence of flows

As in the Langevin description, the dynamic description of noninteracting Brownian particles moving in a fluid in stationary flow, demands a mesoscopic treatment in terms of the probability density $f(\mathbf{r},\mathbf{u},t)$. The evolution in time of this quantity is governed by the continuity equation

$$\frac{\partial f}{\partial t}+\nabla\cdot(\mathbf{u}f)=-\frac{\partial}{\partial \mathbf{u}}\cdot(f\mathbf{V}_{\mathbf{u}}), \tag{16}$$

where $f\mathbf{V}_{\mathbf{u}}$ is a probability current in $\mathbf{u}$-space. The analysis of the dynamics starts from the entropy production rate $\sigma(t)$ related to the probability diffusion process. This can be done by invoking the second principle of thermodynamics that the entropy production rate should be positive definite $\sigma(t)\geq 0$. To calculate $\sigma(t)$ it is convenient to use the relative canonical entropy, or generalized Gibbs entropy postulate (de Groot & Mazur, 1984; Pérez-Madrid, 1994; Rubi & Mazur, 1994),

$$\rho\Delta s(t)=-k_B\int f\ln\left|f/f_{leq}\right|d\mathbf{\Gamma}, \tag{17}$$

where k_B is Botzmann's constant, $\Delta s=s-s_{leq}$ is the difference of the specific nonequilibrium entropy $s(t)$ with respect to that characterizing the local equilibrium reference state s_{leq}, and $d\mathbf{\Gamma}=d\mathbf{r}d\mathbf{u}$. In addition, the mass density of the system of particles is defined by

$$\rho(\mathbf{r},t)=m\int f(\mathbf{r},\mathbf{u},t)d\mathbf{u}, \tag{18}$$

where m is the mass of a Brownian particle. Eq. (17) constitutes an irreversibility criterion assuring that the obtained equations satisfy an extremal principle in $\mathbf{\Gamma}$-space.
The local equilibrium distribution function $f_{leq}(\mathbf{r},\mathbf{u})$ can be described by the local Maxwellian with respect to the non-stationary convective flow $\mathbf{v}_0(\mathbf{r},t)$, that is

$$f_{leq}(\mathbf{r},\mathbf{u})=e^{-\frac{m}{k_BT}\left[\mu_B+\frac{1}{2}(\mathbf{u}-\mathbf{v}_0)^2\right]}, \tag{19}$$

where μ_B is the chemical potential at equilibrium. Using these ingredients, the balance equation for the internal energy and assuming isothermal conditions, the entropy production rate of the system was obtanied in the form (Santamaría Holek, 2005; 2001)

$$\sigma=-\frac{m}{T}\int f\mathbf{V}_{\mathbf{u}}\cdot\frac{\partial\mu}{\partial\mathbf{u}}d\mathbf{u}-\frac{m}{T}\int\mathbf{J}\cdot\nabla\left[\frac{1}{2}(\mathbf{u}-\mathbf{v}_0)^2\right]d\mathbf{u}-\frac{m}{T}\int\mathbf{J}_0\cdot\mathbf{F}d\mathbf{u}. \tag{20}$$

This entropy production rate contains three contributions. The first one is due to diffusion in $\mathbf{u}$-space and is proportional to the product of the mesoscopic probability current $f\mathbf{V}_{\mathbf{u}}$ and the derivative with respect to particle velocity $\mathbf{u}$ of the nonequilibrium chemical potential $\mu(\mathbf{r},\mathbf{u},t)$ per unit mass

$$\mu(\mathbf{r},\mathbf{u},t)=\frac{k_BT}{m}\ln|f(\mathbf{r},\mathbf{u},t)|+\frac{1}{2}(\mathbf{u}-\mathbf{v}_0)^2. \tag{21}$$

The second contribution originates from the product between the gradient of the local kinetic energy of the particles $\frac{1}{2}(\mathbf{u}-\mathbf{v}_0)^2$ (see also (de Groot & Mazur, 1984)), and the relative current $\mathbf{J}=(\mathbf{u}-\mathbf{v})f$, where $\mathbf{v}$ represents the average velocity field of the particles,

$$\rho\mathbf{v}(\mathbf{r},t)=m\int\mathbf{u}f(\mathbf{r},\mathbf{u},t)d\mathbf{u}. \tag{22}$$

Finally, the third contribution to the entropy production rate involves the product of the relative current $\mathbf{J}=(\mathbf{u}-\mathbf{v}_0)f$ times the force $\mathbf{F}=d\mathbf{v}_0/dt$ arising from the dependence on time of the external velocity field $\mathbf{v}_0(\mathbf{r},t)$.
A linear law for the mesoscopic current $f\mathbf{V}_\mathbf{u}$ in terms of the force was proposed, and when substituted into Eq. (16), it yields the Fokker-Planck equation

$$\frac{\partial f}{\partial t}+\nabla\cdot(\mathbf{u}f)=\frac{\partial}{\partial\mathbf{u}}\cdot\left\{\underline{\beta}\cdot(\mathbf{u}-\mathbf{v}_0)f-\underline{\zeta}\cdot\mathbf{F}f+\mathbb{D}_v\cdot\frac{\partial f}{\partial\mathbf{u}}\right\}. \tag{23}$$

The postulation of a linear law scheme necessarily introduces the Onsager coefficients $\underline{\beta}$, $\underline{\zeta}$ and $\underline{\epsilon}$. The diffusion coefficient in velocity space $\mathbb{D}_v$, is related to $\underline{\beta}$ and $\underline{\epsilon}$ through the expression

$$\mathbb{D}_v=\frac{k_BT}{m}\left[\underline{\beta}-\underline{\epsilon}\cdot\nabla\vec{v}_0\right]. \tag{24}$$

In Eq. (23), $\underline{\beta}$ is the friction coefficient that usually appears when describing Brownian motion in velocity space (Santamaría Holek, 2005; 2001), and corresponds to what is referred as a direct effect in nonequilibrium thermodynamics (de Groot & Mazur, 1984). The cross coefficient $\underline{\zeta}=\rho_p/\rho_f\underline{1}$ is related to inertial effects and ρ_p and ρ_f denote the particle and host fluid densities, and $\underline{1}$ is the unit tensor. The tensor $\underline{\epsilon}$ constitutes an important result of this analysis. It corresponds to a cross effect term proportional to the gradient of the imposed velocity flow: $\nabla\mathbf{v}_0$.
The expressions for $\underline{\epsilon}$ and ζ were obtained by using the generalized Faxén theorem giving the force experienced by a particle of arbitrary shape into a fluid under non-stationary flow conditions (Mazur & Bedeaux, 1974; Santamaría Holek, 2005). For spherical particles at low Reynolds numbers, $\underline{\epsilon}=\epsilon\underline{1}$, ϵ is given by

$$\epsilon=\frac{1}{6}\frac{m}{k_BT}a^2\beta_0^2k_\omega, \tag{25}$$

where $k_\omega=(1+2a\alpha_\omega)$, $\alpha_\omega=\sqrt{-i\omega/\nu}$ is the inverse viscous penetration length of the host fluid and ν its kinematic viscosity. Finally, $\beta_0=6\pi\eta a/m$ is the Stokes friction coefficient where η is the dynamic viscosity of the host fluid.
The main fact to stress here is that the fluctuation term in the Fokker-Planck equation (23) contains the diffusion coefficient in velocity space $\mathbb{D}_v$, Eq. (24), that is similar to the one obtained through kinetic theory, see Eq. (8). This coefficient contains the expected thermal contribution $k_BT\underline{\beta}/m$ and, more interesting, a non-thermal contribution coming from the cross effect coupling the diffusion of probability in configuration and velocity subspaces: $\frac{1}{6}a^2\beta_0^2\nabla\vec{v}_0$. This particular dependence of $\mathbb{D}_v$ leads to a modification of the spatial diffusion coefficient $\mathbb{D}(\mathbf{r},t)$ of the Brownian particles which is in agreement with experimental and simulation results (Pine, 2005; Sarman, 1992). However, before discussing this point, some comments on the diffusion or Smoluchowski equation are in turn.

2.5 Diffusion in the presence of shear flow

The Smoluchowski equation can be derived from Eq. (23) by restricting the description to the long-time diffusion regime, that is, when $t >> \beta_{ij}^{-1}$.

In references (Santamaría Holek, 2005; 2009; 2001), the Smoluchowski equation was obtained by calculating the evolution equations for the first moments of the distribution function. These equations constitute the hydrodynamic level of description and can be obtained through the Fokker-Planck equation. The time evolution of the moments include relaxation equations for the diffusion current and the pressure tensor, whose form permits to elucidate the existence of inertial (short-time) and diffusion (long-time) regimes. As already mentioned, in the diffusion regime the mesoscopic description is carried out by means of a Smoluchowski equation and the equations for the moments coincide with the differential equations of nonequilibrium thermodynamics.

The procedure is enterily similar to that followed in kinetic theory (McLennan, 1989). In particular, after assuming that stresses relax rapidly enough, the resulting Smoluchowski equation is

$$\frac{\partial \rho}{\partial t} = -\nabla \cdot \left(\rho \mathbf{v}_0 + \zeta \underline{\xi} \cdot \vec{F} \right) + \nabla \cdot \left(\mathbb{D} \cdot \nabla \rho \right), \tag{26}$$

where the *effective diffusion* coefficient is defined according to the relation

$$\mathbb{D} = \frac{k_B T}{m} \underline{\xi} \cdot (\underline{\mathbf{1}} - \underline{\xi} \cdot \nabla \mathbf{v}_0)^s - \frac{1}{6} a^2 \beta_0^2 k_\omega \underline{\xi} \cdot (\underline{\xi} \cdot \nabla \mathbf{v}_0)^s. \tag{27}$$

Here $\underline{\xi}$ is the mobility tensor satisfying $\underline{\xi} \cdot \underline{\beta} = \underline{\mathbf{1}}$. Again, this diffusion coefficient contains two contributions, one depending on the temperature T of the fluid and another one independent from it and proportional to the velocity gradient of the imposed external flow. Therefore, the first contribution is related to the usual thermal Brownian motion whereas the last one is related to non-thermal effects. Notice also that the Smoluchowski equation (26) also contains the usual convective term $\nabla \cdot (\rho \mathbf{v}_0)$.

Thus, the effect of the flow on diffusion is to introduce a new contribution term in the diffusion coefficient arising from spatial inhomogeneities of the velociy field. As a consequence, the flow also induces anisotropy in the Brownian motion. This can be seen from the fact that even if the host fluid is assumed isotropic and the particles spherical, implying $\underline{\beta} = \beta \underline{\mathbf{1}}$, the effective diffusivity is a tensor proportional to $\nabla \mathbf{v}_0$. It is convenient to mention that this non-thermal contribution, known as the shear-induced diffusion effect, is proportional to the size (surface) of the particle and breaks-down the fluctuation-dissipation relation of the Smoluchowski equation. The presence of a nonequilibrated bath therefore modifies the thermal energy available to the system in order to perform fluctuations. As shown before, similar results are obtained from kinetic theory (Rodríguez, 1983). These results suggest that the fluctuation dissipation relation used in Langevin-like approaches should be modified accordingly for appreciably large Reynolds numbers. Comparison of these results with other theories like, for instance, extended irreversible thermodynamics could be of interest.

2.6 Classical simulation results on diffusion in the presence of shear flow

It is convenient to emphasize that nonequilibrium molecular dynamics simulations with Lennard-Jones pure fluids and mixtures under strong shear flow conditions (Sarman, 1992), lead to a dependence of the self-diffusion tensor on the shear rate.

The results on diffusion reported show that the components of the diffusion tensor start increasing or decrasing linearly as a function of the shear rate $\dot{\gamma}$, (Sarman, 1992), in

agreement with the results obtained with kinetic theory and mesoscopic nonequilibrium thermodynamics. For larger values of $\dot{\gamma}$, a quadratic contribution seems to appear and modifies the initial tendency. In general, diagonal elements are positive whereas the non-diagonal elements are negative. However, despite the sign of each element of the diffusion tensor, the authors emphasize that the dependence on $\dot{\gamma}$ and the obtained anisotropy for the self-diffusion tensor evidences a break-down of the Curie principle and of the local equilibrium hypothesis of linear irreversible thermodynamics, thus implying an anisotropic kinetic energy distribution induced by the shear flow.

These observations pose an interesting question because the results reported in Refs. (Santamaría Holek, 2005; 2009; 2001) were obtained under the assumption of local equilibrium in phase space, that is, at the mesoscale. It seems that for systems far from equilibrium, as those reported in (Sarman, 1992), the validity of the fundamental hypothesis of linear nonequilibrium thermodynamics can be assumed at the mesoscale. After a reduction of the description to the physical space, this non-Newtonian dependedence of the transport coefficients on the shear rate appears. This point will be discussed more thoroughly in the following sections when analyzing the formulation of non-Newtonian constitutive equations.

2.7 The shear-induced diffusion effect and the transition to irreversibility

One of the most beautiful results manifesting the effect of flow on diffusion is the shear-induced diffusion effect, which is the quantitative manifestation of the transition to irreversibility in a set of particles obeying the classical reversible equations of motion.

Recently, this effect was studied in a series of experiments and numerical simulations were the transition from a dynamical reversible behavior to a dynamical irreversible chaotic behavior of a suspension of non-Brownian particles was directly observed (Drazer, 2002; Pine, 2005). The importance of these experiments lies in the fact that they shed light on the origin of the thermodynamic irreversibility and its relation to the chaotic dynamics of a system (Guasto & Gollub, 2007; Pine, 2005).

The experiments were performed with semi-diluted suspensions of spherical particles with diameters $d \simeq 0.2mm$. The host fluid had the same density as the particles and was highly viscous. The suspension was contained in a cylindrical Couette cell which was taken out from equilibrium by applying an oscillating shear flow $\dot{\gamma}\, cos(\omega t)$ with dimensionless amplitude γ and rate $\dot{\gamma} = \omega\gamma$, being ω the characteristic frequency of the oscillation. The Reynolds number of the flow was always much smaller than one, thus avoiding turbulent effects. For small oscillation amplitudes γ the motion of the particles was reversible, that is, after each period the particles came back to the original position. On the contrary, when increasing the oscillation amplitude γ the trajectories of the particles became chaotic implying that their reversible behavior was lost. From the thermodynamic point of view, it was interesting to know that the motion of the particles was characterized by means of the mean square displacement $\langle \Delta x^2 \rangle$, which scaled with time in the form: $\langle \Delta x^2 \rangle \sim 2d^2\dot{\gamma}t$, thus implying that the effective diffusivity scales with $D \sim d^2\dot{\gamma}$. Computer simulations using Stokesian dynamics (Bossis & Brady, 1989; Pine, 2005) modeled the experiment and determined a positive Lyapunov exponent of particle trajectories in phase space (Pine, 2005), thus implying that particle dynamics is chaotic. The transition was characterized by giving quantitative values for the critical strain amplitude at which it occurs.

Different theoretical descriptions were proposed to explain this effect. One of these models is based on a Smoluchowski equation in which the diffusivities have been constructed by analyzing the temporal behavior of the correlation of the position of the particles as

obtanied from simulations (Seriou & Brady, 2004). However, this description does not impose new thermodynamic postulates but it gives good quantitative results, though it lacks of a deeper fundamental basis. This fundamental analysis can be given in terms of mesocopic nonequilibrium thermodynamics (Santamaría Holek, 2009).

The explanation of the effect was formulated for a set of N particles described through a multivariate probability distribution. After deriving the corresponding Fokker-Planck equation along the lines indicated above, the description was reduced to the case of a single tagged particle in an effective medium composed by the host fluid and the remaining particles. The main result was a Smoluchowski equation for the single particle distribution function $\rho(\mathbf{r},t)$ entirely similar to Eq. (26), with an effective diffusion coefficient of the form (27). However, in the referred case the diffusion tensor depends on position and time, and on the fraction of volume $\phi = NV_p/V$ occupied by the particles in the sample. V_p and V are the volume of a particle and the total volume. These dependencies arise when reducing the N-particle description to the single-particle one, a process that involves averaging over positions and (Santamaría Holek, 2009) velocities of the remaining particles (Santamaría Holek, 2009).

For sufficiently large particles (with vanishing thermal contribution) it was possible to show that the mean square displacement is of the form

$$\langle r^2 \rangle \sim \frac{1}{3} g_{12}(\phi) d^2 \dot{\gamma} t, \tag{28}$$

expressed in terms of the number of cycles n hence $t = 2\pi n/\omega$; thus $\langle r^2 \rangle \sim \frac{2\pi}{3} g_{12}(\phi) d^2 \gamma_0 n$, which has the same scaling relation for the mean square displacement than that reported in the experiments in Ref. (Pine, 2005). Here, the dimensionless factor $g_{12}(\phi)$ depending on the volume fraction ϕ is a consequence of the existence of hydrodynamic interactions between particles.

2.8 Lattice Boltzmann simulations

Lattice Boltzmann simulations using the D2Q9 model (Aidun & Ding, 1998; Ladd, 1994) were also performed to study the transition to irreversibility in sheared suspensions of non-Brownian particles (Santamaría Holek, 2009). This method was used because it esentially controls the evolution in time of the host fluid without thermal fluctuations, thus giving in a natural way the hydrodynamic interactions that, for large shear rates and particle concentrations, dominate the dynamics of the system. Particle interactions were implemented with the method proposed in Ref. (Ladd, 1994). The numerical simulations were carried out in a cavity with the bounce-back boundary condition, which consists in reversing the incoming particle distribution function after the stream process.

From these simulations it was observed that the effective diffusion of the particles can be enhanced by increasing the particle concentration ϕ or by increasing the Reynolds number Re, even at small values of the Re. The transition to irreversibility was characterized through a power spectrum of particle trajectories, as shown in Fig. 1, where the trajectory of one of the sixteen particles is shown. In Fig. 1a, one can appreciate that the particle describes regular motion and even though in Fig. 1b we can notice some regularity in the motion. However, it is clear that hydrodynamic interactions induce correlations between all particles giving rise to new dissipation modes in the dynamics. By increasing the Reynolds number above its critical value (corresponding to the critical value of the strain amplitude) the motion becomes chaotic, as shown in Fig. 1e. Correspondingly, the spectrum shows (Fig. 1f) that almost all frequencies contribute with the same amplitude, thus suggesting the stochastization of the motion.

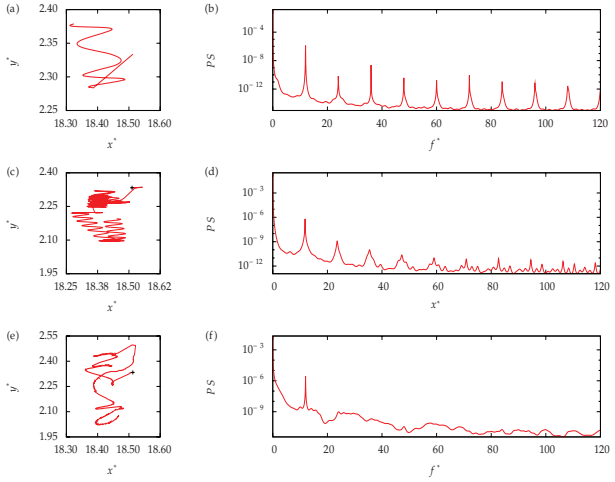


Fig. 1. Trajectories and power spectra of the same particle with equal initial conditions for (a-b) $Re = 0.01$ and (c-d) $Re = 0.07$ and (e-f) $Re = 0.08$ with a volume fraction of $\phi = 0.14$ and dimensionless frquency $f^* = 10.0$. The figure is a courtesy of Dr. G. Barrios.

To close this section, we want to emphasize that the explanation of the experiments and simulations on shear-induced diffusion shows the validity of the theoretical approach and its importance, because it was shown that the shear-induced diffusion and the transition to irreversibility are fully compatible with the thermodynamic irreversibility as obtained from the second principle and the generalized Gibbs entropy postulate.

3. Constitutive relations

The main question related to the thermodynamics in presence of flow is the formulation of constitutive relations between density currents and thermodynamic forces. In linear irreversible thermodynamics (de Groot & Mazur, 1984; Kondepudi & Prigogine, 1999) the constitutive relations are linear. Although this approximation is valid for a large number of systems, there are also many of them in which these linear relationships are not valid. A classical example of this are chemical reactions, where the currents are sometimes not linearly related to the forces, like in the mass action law (de Groot & Mazur, 1984; Kondepudi & Prigogine, 1999).

Flowing systems are also typical examples in which this linear behavior is broken. This is the case of rheology, the study of viscosity of complex fluids, that is, of fluids with internal structure that exhibit a combination of viscous and elastic behavior under strain (Beris & Edwards, 1994; Doi & Edwards, 1998). Examples of such fluids are polymer solutions and melts, oil and toothpaste, among many others.

There are several approaches leading to the formulation of these non-linear relationships. Unfortunately it is not possible to summarize them here because the field is very rich in the variety of systems and behaviors. Consequently, such a summary lies beyond the scope of this study. An excellent and detailed review is given in Ref. (Beris & Edwards, 1994). However, let us mention that many attempts have been made to bring together continuum mechanics theories and molecular models in order to formulate appropriate constitutive equations for

polymer solutions (Beris & Edwards, 1994; Doi & Edwards, 1998). One of the central aspects of this research is the incorporation of microscopic theories such as those by Zimm or Rouse (Doi & Edwards, 1998) in constitutive relations. However, these attempts are usually valid only in the dilute regime, although the proposed rheological equations are often evaluated quantitatively and qualitatively to describe concentrated colloidal suspensions or polymer solutions (James, 1972). From a phenomenological point of view, the first model attempting to describe the hydrodynamic behavior of complex fluids was proposed by Maxwell and describes what is now known as Maxwell materials (Meyers & Chawla, 2009). Other type of materials are Kelvin-Voigt and Oldroyd materials.
The main difference between constitutive relations for simple and complex fluids lies in the fact that simple fluids satisfy the linear relationships between the stress tensor $\underline{\sigma}(\mathbf{r},t)$ and the strain or deformation tensor $\underline{\theta}(\mathbf{r},t)$ or the velocity gradient $\nabla\mathbf{v}$. In contrast, complex fluids have to incorporate a relaxation term for the stresses and a non-linear dependence on the strain or velocity gradient tensors.

3.1 Simple viscoelastic materials

For simple viscoelastic materials, the main characteristic of the constitutive relations is the presence of a time derivative of the stress tensor, associated to a relaxation process. In the case of an incompressible Maxwell material, the constitutive relation is given through the evolution equation

$$\tau\frac{\partial}{\partial t}\underline{\sigma}+\tau\left(\mathbf{v}\cdot\nabla\underline{\sigma}\right)_{cont}=-\underline{\sigma}+\eta\left(\nabla\mathbf{v}+\nabla\mathbf{v}^{T}\right) \tag{29}$$

where τ is a relaxation time of the material and the contravariant convective term is given by the second order tensor $\{(\mathbf{v}\cdot\nabla\underline{\sigma})_{cont}\}_{ij}=v_k\nabla_k\sigma_{ij}-\sigma_{ik}\nabla_k v_j-\sigma_{kj}\nabla_k v_i$. This convective term has been proposed by Oldroyd in order to satisfy the material invariance condition (Beris & Edwards, 1994). As mentioned before, the last term of the equation is often expressed in terms of the rate of the deformation tensor $\underline{\theta}$. This model is related to microscopic approaches where the molecules are modeled as dumbells (Hinch, 1994; Rallison & Hinch, 1988). Clearly, this equation is coupled to the momentum equation and can be solved for many types of flows. Averaged over volume, Maxwell's model is frequently expressed in the more simple form

$$\frac{d}{dt}\hat{\underline{\sigma}}=-\frac{1}{\tau}\hat{\underline{\sigma}}+\frac{\eta}{\tau}\frac{d}{dt}\hat{\underline{\theta}}, \tag{30}$$

where $\hat{\sigma}$ and $\hat{\theta}$ are now time dependent quantities only. Assuming that the rate of deformation is constant $d\hat{\underline{\theta}}/dt=\dot{\underline{\theta}}$, one then obtains that the constitutive relation for a simple viscoelastic fluid is of the form

$$\hat{\underline{\sigma}}=\eta\dot{\underline{\theta}}\left(1-e^{-t/\tau}\right), \tag{31}$$

that is, after applied an external force inducing the deformation of the material, the stress tensor relaxes until reaching a stationary value. When an oscillatory shear is applied on the system then one obtains that the viscoelastic response of the fluid is characterized by a frequency dependent viscosity.

3.1.1 Non-Newtonian materials

Very frequently, complex fluid manifest non-linear constitutive relations between the stress tensor and the velocity gradient. These relations can be a consequence of the fact that the relaxation time dependens on a scalar function of the deformation tensor like in the FENE models (Hinch, 1994; Rallison & Hinch, 1988) or of the stress tensor, like in the case of the

so-called modified Maxwell models (Beris & Edwards, 1994), that is, $\tau = \tau(tr[\underline{\sigma}])$ where the dependence on the stress tensor is usually introduced via the invariants of $\underline{\sigma}$, for example, through $tr\underline{\sigma}$. As a consequence of this, the evolution equation for the fluid becomes of the form

$$\frac{\partial}{\partial t}\underline{\sigma} + (\mathbf{v}\cdot\nabla\underline{\sigma})_{cont} = -\frac{\underline{\sigma}}{\tau(tr[\underline{\sigma}])} + \frac{\eta}{\tau(tr[\underline{\sigma}])}\left(\nabla\mathbf{v} + \nabla\mathbf{v}^T\right). \tag{32}$$

Frequently, non-Newtonian materials show a power-law behavior of the viscosity as a function of the shear rate $\dot{\gamma}$ in which the exponent is a free parameter. An example of this is the following model for the relaxation time (Apelian, 1988)

$$\tau(tr[\underline{\sigma}]) = \frac{\tau_0}{1 + \{h(\tau_0)tr[\underline{\sigma}]\}^{\alpha-1}}, \tag{33}$$

where τ_0 is a charateristic relaxation time and h is a function that gives the correct dimensions to the formula. As a consequence of this type of dependencies, it is ussually obtained that the viscosity of the fluid and the normal stress difference Ψ depend on the shear rate in the form

$$\eta \sim \dot{\gamma}^{k(\alpha)} \quad \text{and} \quad \Psi \sim \dot{\gamma}^2, \tag{34}$$

where in general the exponent of the shear rate k depends on α.
These examples illustrate that viscoelastic and non-Newtonian materials do not respond instantaneously to the shear stresses applied on them and its response is not linear with respect to those stresses.

3.2 Mesoscopic constitutive relations for polymer solutions

Nonlinear constitutive relations for complex fliuds can also be obtained by using MNET (Málaga, 2006). As described in Section II, the description was given in terms of a probability density accounting for the state of the system through the position $\mathbf{r}$ and instantaneous velocity $\mathbf{u}$ of the molecules. The deformation of each molecule is modeled by introducing a local vector parameter $\mathbf{R}$.
Following the scheme of MNET, a Fokker-Planck equation was obtained from which a coarse-grained description in terms of the hydrodynamic equations was derived in turn. Molecular deformation and diffusion effects become coupled and a class of non-linear constitutive relations for the kinetic $\mathbb{P}^k$ and elastic parts $\mathbb{P}^E$ of the stress tensor are obtained. The expression for the stress tensor can be written in terms of dimensionless quantities like

$$\underline{\sigma} = \mathbb{P}^k + \mathbb{P}^E. \tag{35}$$

Each contribution can be obtained from its own evolution equation. The kinetic part of the stress tensor is governed by the equation

$$\frac{d\mathbb{P}^k}{dt} + 2\left[\mathbb{P}^k\cdot\left(\underline{\beta} + \nabla\mathbf{v}\right)\right]^s = 2\frac{k_B T}{m}\rho\left(\underline{\beta} - \mathbb{L}_{ur}\cdot\nabla\mathbf{v}_0\right)^s \tag{36}$$

where the superscript ^{s} stands for the symmetric part of a tensor, m is the mass of the molecule, $\underline{\beta}$ is the the friction tensor of the particles proportional to the corresponding Stokes coefficient (Santamaría Holek, 2001), and $\mathbb{L}_{ur}$ is an Onsager coefficient. Notice that in contrast with Eqs. (29) and (32), where the relaxation time is scalar, Eq. (36) contains a matrix of the relaxation times $\underline{\tau} \equiv \underline{\beta} + \nabla\mathbf{v}$. This means that in contrast with the usual phenomenological

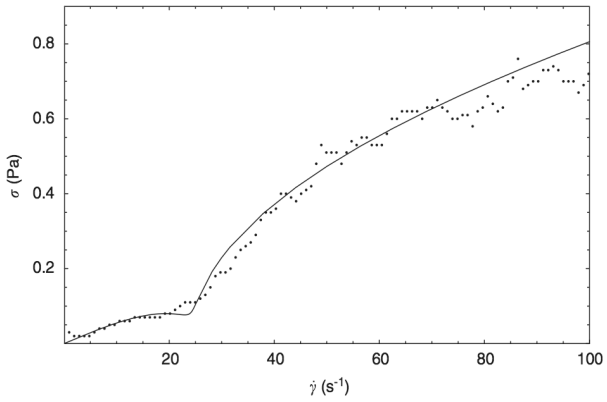


Fig. 2. Experimental data (points) for the non-Newtonian shear stress versus the shear rate of a dilute solution of wormlike micelles (Delgado & Castillo, 2005; 2007). Fit of the data using the isotropic friction model (line) in Eqs. (35), (36) and (38). Notice that the dependence of the effective viscosity as a function of the shear rate is not simply a power law. Figure taken from Ref. (Málaga, 2006).

relations such as (29) and (32), the relaxation time of (36) has an explicit relation in terms of the friction coefficients of the particles in the host fluid.

For a FENE-type model with nonlinear elastic force of the dumbell, the restituting elastic force of a molecule can be represented in the form: $\mathbf{F} = -\xi_0 F_0 \mathbf{R}$, where ξ_0 is the characteristic spring restitution coefficient per unit mass and $F_0 = L^2/\left(L^2 - tr\mathbb{A}\right)$ with L the maximum length of the polymer. The evolution equation for the elastic contribution is given in terms of the deformation or conformational tensor $\mathbb{A} = \langle \mathbf{RR} \rangle$, which is now related to the elastic part of the stress tensor by the formula $\mathbb{P}^E = -\xi_0 F_0 \mathbb{A}$. The equation is

$$\frac{d\mathbb{A}}{dt} - 2[\mathbb{A}\cdot\nabla\mathbf{v}]^s = 2\left\{\mathbb{L}_{RR}\cdot\left[\frac{kT}{m}\mathbb{I} - \langle \mathbf{FR}\rangle\right]\right\}^s + 2\left[\mathbb{L}_{Rr}\cdot(\nabla\mathbf{v}\cdot\nabla\mathbf{v}^T)\cdot\mathbb{A}\right]^s, \qquad (37)$$

where $\mathbb{L}_{Rr}$ and $\mathbb{L}_{RR}$ are Onsager coefficients characterizing the mobility and the coupling between the deformation of the molecule and drag forces, respectively. The mesoscopic formalism leads naturally to the Oldroyd type derivative that maintains material frame independence (Málaga, 2006). Eqs. (36) and (37) have similar features as those indicated by the Maxwell and the modified Maxwell models already discussed.

These equations were applied to describe different situations. For isotropic friction, the dynamics of dumbell-like molecule solutions under simple shear conditions was analyzed to obtain a correction to the usual FENE models that arises through the kinetic contribution to the stress tensor. A second application for non-constant friction coefficient β lead to the evolution equation

$$\frac{d\mathbb{A}}{dt} - \mathbb{A}\cdot\nabla\mathbf{v}_0 - \nabla\mathbf{v}_0^T\cdot\mathbb{A} = \frac{16}{9\sqrt{tr\mathbb{A}}}\frac{1}{De}\left(\mathbb{I} - \xi_0 F_0 \mathbb{A}\right) \qquad (38)$$

where $De = \beta\dot{\gamma}/\xi_0$ is the Deborah number. As a consequence of this equation, a non-linear relation between the elastic part of the stress tensor $\mathbb{P}^E$ and the shear rate $\dot{\gamma}$ appears. This relation was used to describe satisfactorily experiments with dilute miscellar solutions in water (Málaga, 2006), see Fig. 2.

4. Nonequilibrium equations of state

One interesting aspect of the previous discussion is that the expressions for the normal as well as for the diagonal elements of the stress tensor $\underline{\sigma}$ or, more generally, the pressure tensor $\mathbb{P}$ are functions of the applied velocity gradient or the strain rate. In this sense, these relations represent what is known as nonequilibrium equations of state. It was shown rigorously that at mesoscopic level, the chemical potential becomes modified by the external flow in a term proportional to its local kinetic energy, see Eq. (21).

In this section we want to address how the presence of flow may lead to the existence of these nonequilibrium equations of state at macroscopic level. For convenience we will only consider the case of particle suspensions (Santamaría Holek, 2005), that is, we will not consider the internal degrees of freedom of each molecule, as we did for the case of polymer solutions.

Let us recall that the evolution equation for the kinetic part of the pressure tensor $\mathbb{P}^k$ was given by equation (36). For times $t \gg \beta^{-1}$ the particles enter the diffusion regime governed by the corresponding Smoluchowski equation. In this regime, assuming a stationary process, the following constitutive relation for $\mathbb{P}^k$ was obtained

$$\mathbb{P}^k \simeq \frac{k_B T}{m}\rho\underline{1} - \left[\left(\underline{\eta}_B + \underline{\eta}_H\right)\cdot\nabla\vec{v}_0\right]^s, \tag{39}$$

where the Brownian ($\underline{\eta}_B$) and hydrodynamic ($\underline{\eta}_H$) viscosity tensors are defined by

$$\underline{\eta}_B \equiv \frac{k_B T}{m}\rho\underline{\beta}^{-1} \quad \text{and} \quad \underline{\eta}_H \equiv \frac{k_B T}{m}\rho\underline{\beta}^{-1}\cdot\underline{\epsilon}. \tag{40}$$

For spherical particles, the friction coefficient $\underline{\beta}$ can be identified with the Stokes friction coefficient $\underline{\beta} \simeq \beta\underline{1}$. Two cases are of interest.

4.1 Oscillatory flows in the creeping flow approximation

When the oscillatory flow is slow enough, by using Faxen's theorem it was shown that the friction coefficient depends on the frequency (ω) in the form (Mazur & Bedeaux, 1974)

$$\beta = \beta_0(1 + a\alpha_\omega), \tag{41}$$

where $\alpha_\omega = \sqrt{-i\omega/\nu(\omega)}$, is the inverse viscous penetration length of the host fluid with kinematic viscosity $\nu(\omega)$. The expression for ϵ is given through Eq. (25). Using these results it was found that the expressions for the Brownian $\underline{\eta}_B(\omega)$ and hydrodynamic $\underline{\eta}_H(\omega)$ viscosities are

$$\underline{\eta}_B \simeq \frac{k_B T}{m\beta_0}\rho\left[1 - \Delta_\omega(\tau\omega)^{\delta_\omega}\right] \quad \text{and} \quad \underline{\eta}_H \simeq \frac{1}{12}\beta_0 a^2\rho\left[1 + \Delta_\omega(\tau\omega)^{\delta_\omega}\right], \tag{42}$$

where Δ_ω and δ_ω are a scaling factor and an exponent that in general may depend on the particle volume fraction ϕ. For viscoelastic fluids the expressions corresponding to the normal stress difference and the non-diagonal elements of the pressure tensor are

$$\begin{aligned} P_{xx} - P_{yy} &\simeq \beta_0^{-1}\left[\underline{\eta}_B(\omega) + \underline{\eta}_H(\omega)\right]\dot{\gamma}^2, \\ P_{xy} &\simeq \beta_0^{-1}\left[\underline{\eta}_B(\omega) + \underline{\eta}_H(\omega)\right]\dot{\gamma}. \end{aligned} \tag{43}$$

This relation suggests that the normal stresses of simple viscoelastic fluids depend on the shear rate $\dot{\gamma}$, that is, are modified by the flow conditions $P_{ii} \propto \dot{\gamma}^2$, see Eq. (34).

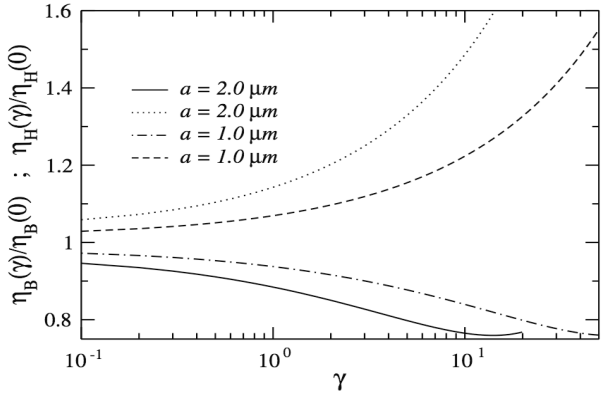


Fig. 3. Brownian (solid and dash-dotted lines) and hydrodynamic (dashed and dotted lines) viscosities in terms of the shear rate for different values of the radius of the Brownian particle. The figure shows the shear thinning and shear thickening effects associated to the Brownian and hydrodynamic contributions, respectively. The values of the parameters we have used are $\beta_0 = 10^6 s^{-1}$, $D_0 = 10^6 cm^2 s^{-1}$ and $\rho = 10^{-2} g cm^{-3}$. Figure taken from Ref. Santamaría Holek (2005).

4.2 Non-Newtonian behavior

When the applied shear is large enough, the friction coefficient becomes a function of the shear rate in the form (Santamaría Holek, 2005)

$$\underline{\beta}^{-1} = \beta_0^{-1} \left(\underline{1} - \mathbb{M} a \alpha_{\dot\gamma} \right) \tag{44}$$

where $\alpha_{\dot\gamma} = \sqrt{\dot\gamma / \nu(\omega)}$, is the inverse viscous penetration length due to the presence of a large shear flow and $\mathbb{M}$ is a geometrical factor depending on the boundary conditions over the surface of the particle. Using the corresponding expression for $\epsilon(\dot\gamma)$ it is found that in the concentrated case, the non-Newtonian expressions for the Brownian and hydrodynamic viscosities are

$$\underline{\eta}_B \simeq \frac{k_B T}{m \beta_0} \rho \left[1 - M_{xx} \Delta_\gamma (\tau \dot\gamma)^{\delta_\gamma} \right] \quad \text{and} \quad \underline{\eta}_H \simeq \frac{1}{12} \beta_0 a^2 \rho \left[1 + M_{xx} \Delta_\gamma (\tau \dot\gamma)^{\delta_\gamma} \right]. \tag{45}$$

where, as in the previous case, the scaling factor Δ_γ and the exponent δ_γ are ϕ dependent. For the present case, the expressions corresponding to the normal stress difference and the non-diagonal elements of the pressure tensor are

$$\begin{aligned} P_{xx} - P_{yy} &\simeq -\frac{M_{xy}}{3} \beta_0^{-1} a^2 \rho \Delta_\gamma \tau^{\delta_\gamma} \dot\gamma^{1+\delta_\gamma}, \\ P_{xy} &\simeq \beta_0^{-1} \left[\underline{\eta}_B(\dot\gamma) + \underline{\eta}_H(\dot\gamma) \right] \dot\gamma. \end{aligned} \tag{46}$$

These results, together with those for pure viscoelastic fluids clearly suggest the possibility of the existence of non-equilibrium equations of state, that is, equations of state that are modified by the presence of the flow. This possibility will be analyzed more thoroughly in the following sections.

5. Nucleation in the presence of flow

One of the interesting applications where the influence of flow on transport coefficients and on the constitutive and state equation is relevant, is the nucleation process. The effect of flow on crystallization is important because it often occurs in the presence of mechanical processing of the melt, such as extrusion, shearing, or injection, all these factors may affect the process drastically.

Despite the various existing approaches to the problem, we will restrict the presentation to that based on MNET (Gadomski, 2005; 2008; Reguera & Rubi, 2003b). In this approach the influence of a shear flow on the process of nucleation is analyzed through the derivation of a Fokker-Planck equation for the probability distribution that depends on size. This is done for a metastable phase subjected to a stationary flow. In addition to the correction of the diffusion coefficient, it was shown that to the flow also introduces modifications to the nucleation rate.

By considering a metastable phase in which the emerging clusters, liquid droplets or crystallites are embedded, the description is given in terms of the probability distribution $f(\mathbf{r}, \mathbf{u}, n, t)$, where now n denotes a cluster of n molecules. Following a scheme similar to that of Section 2, after proposing the Gibbs equation the authors assume that the local equilibrium distribution [see Eqs. (17) and (19)] has the form

$$f_{leq}(\mathbf{r}, \mathbf{u}) = exp\left\{ \left[\mu_B - \Delta G(n) + \frac{1}{2} m(n)(\mathbf{u} - \mathbf{v}_0)^2 \right] / k_B T \right\}, \tag{47}$$

The description given in Ref. (Reguera & Rubi, 2003b) is completely general since the particular for of the free energy difference is not specified. However, to fix ideas, one may assume that $\Delta G(n)$ is given by the expression of the classical nucleation theory

$$\Delta G(n) = -n\Delta\mu + v_a^* \tilde{\sigma}\, n^{2/3}, \tag{48}$$

where $\Delta\mu$ is the difference in chemical potential between the solid and the liquid phase, v_a^* is a constant given by $v_a^* = (3\sqrt{4\pi} v_a)^{2/3}$ where v_a is the molecular volume of the liquid and $\tilde{\sigma}$ is the interfacial free energy per unit area appearing because an interface between the solid nucleus and surrounding liquid is formed. In terms of the number of particles of the cluster with critical size n_{crit}, the free energy difference may be written in the form

$$\Delta G(n) = \Delta G_{crit} \left[3 \left(\frac{n}{n_{crit}} \right)^{2/3} - 2 \left(\frac{n}{n_{crit}} \right) \right], \tag{49}$$

where the height of the barrier is given by $\Delta G_{crit} = \frac{16\pi}{3} v_a^2 \tilde{\sigma}^3 / \Delta\mu^2$.

Using Eq. (47) and the rules of mesoscopic nonequilibrium thermodynamics, the resulting Fokker-Planck equation for the probability distribution is

$$\begin{aligned} \frac{\partial f}{\partial t} + \nabla \cdot (\mathbf{u} f) = \frac{\partial}{\partial n} \left(\frac{D_n}{k_B T} f \frac{\partial}{\partial n} \left[\Delta G(n) + \frac{m(n)}{2} (\mathbf{u} - \mathbf{v}_0)^2 \right] + D_n \frac{\partial f}{\partial n} \right) \\ + \frac{\partial}{\partial \mathbf{u}} \cdot \left\{ \underline{\beta} \cdot (\mathbf{u} - \mathbf{v}_0) f - \underline{\zeta} \cdot \mathbf{F} f + \mathbb{D}_v \cdot \frac{\partial f}{\partial \mathbf{u}} \right\}. \end{aligned} \tag{50}$$

Here D_n plays the role of a diffusion coefficient in cluster size space. Reducing the description to the diffusion regime ($t \gg \beta^{-1}$), an effective value for the nucleation barrier was obtained:

$\Delta\tilde{G} = \Delta G + (3/2)k_B T \ln|n|$. This result is interesting because it implies that the flow modifies the nucleation rate in the form

$$J \simeq k^+(n_{crit}) e^{\Delta\tilde{G}_{crit}/k_B T}, \tag{51}$$

where k^+ is the rate of attachment of particles to the cluster and $\Delta\tilde{G}_{crit}$ is the height of the nucleation barrier.

5.1 Extension of the classical nucleation theory

A different approach to the problem of nucleation in the presence of flow was proposed in Ref. (Blaak, 2004) by extending the classical nucleation theory in an intuitive and interesting form. Starting from Eq. (49), a macroscopic nonequilibrium equation of state was proposed for the Gibbs free energy difference by expanding in powers of the shear rate about the equilibrium values of both, the chemical potential difference and the surface energy

$$\Delta\mu = \Delta\mu^{eq}\left[1 + c_0\dot{\gamma}^2 + O(\dot{\gamma}^4)\right], \tag{52}$$

$$\tilde{\sigma} = \tilde{\sigma}^{eq}\left[1 + \kappa_0\dot{\gamma}^2 + O(\dot{\gamma}^4)\right],$$

where c_0 and κ_0 are, in principle, fitting constants and $\Delta\mu^{eq}$ and $\tilde{\sigma}^{eq}$ are the equilibrium values of the chemical potential difference and the surface energy. The second term at the right hand side of Eq. (52) is similar to the corrections previously obtained from MNET, see Eq. (21). The main difference is that in the latter case these corrections enter through the local kinetic energy of the system. Using these dependences, it was shown that the barrier height and the critical cluster size n_{crit} depend on the shear in the form

$$\Delta G_{crit} = \frac{16\pi\nu_s[\tilde{\sigma}^{eq}]^3}{3[\Delta\mu^{eq}]^2}\left[1 + (3\kappa_0 - 2c_0)\dot{\gamma}^2 + O(\dot{\gamma}^4)\right], \tag{53}$$

$$n_{crit} = \frac{32\pi\nu_s[\tilde{\sigma}^{eq}]^3}{3[\Delta\mu^{eq}]^3}\left[1 + (3\kappa_0 - 3c_0)\dot{\gamma}^2 + O(\dot{\gamma}^4)\right].$$

Brownian dynamic simulations performed with relatively small shear rates showed a shift of the maximum of the free energy difference as a function of the applied shear rate implying an increase in the nucleation barrier and the critical nucleation size (Blaak, 2004).
The authors conlcuded that their results indicate that crystal nucleation is suppressed by the presence of shear (Butler & Harrowell, 1995), however much larger shear rates were not accessed by simulations because the increase of the critical nucleus size would lead to simulation artefacts. However, this study seems to be necessary because experimental observations on lyzosyme, ferritin and apoferritin crystallization reported in Ref. (A. Penkova, 2006), showed that the presence of moderate flows may promote nucleation whereas large flows may suppress the nucleation. As a consequence, there is still an open question on the appropriate theoretical description of the effect of flow on nucleation.

5.2 Flow induced nucleation: A Control volume thermodynamics approach

Here, we will formulate a novel model of flow induced nucleation by using control volume thermodynamics. The general idea is very simple and goes through the lines indicated along this chapter, that is, by assuming that nonequilibrium equations of state can be formulated in a way consistent with the laws of thermodynamics when a flow is imposed on the system.

Consider that the system is an open control volume with size V that allows the transit of particles in solution through it and that, on average, remains with stationary pressure and temperature. The first law of thermodynamics for open systems states that the change of the total energy E of the system obeys the relation

$$\Delta E = Q + W\,, \tag{54}$$

where Q and W are the amount of heat and work exchanged between the system and the surroundings, and the change in the total energy is given by the relation

$$\Delta E = \Delta U + \Delta E_c + \Delta E_p. \tag{55}$$

Here, U is the internal energy of the system and E_c and E_p are their kinetic and potential energies. Assuming that the potential energy contribution is negligible in the case of our interest, then E_c is the kinetic energy of the particles in solution and we may write

$$E = U + \frac{1}{2}mN\overline{v}^2 + \frac{1}{2}\rho V\overline{v}^2\,,$$

where m is the mass of a particle, N is the total number of particles in the control volume and ρ is the density of the solvent, that we assume constant. Thus, the total mass of the solution mN crosses the container volume with average velocity $\overline{v}$.
In equilibrium, the Gibbs free energy G is defined in terms of the internal energy U, the entropy S and the total volume of the system V:

$$G(N,P,T) = U - TS + PV\,, \tag{56}$$

where P is the pressure. However, in order to be consistent with the first principle (55) for an open system, in formulating the definition of the free energy we have to take into account not only the internal energy U, but the total energy E of the system. Otherwise, it may occur that the free energy difference could not be equal to the work exchanged between the system and the surroundings. In view of these considerations, we propose the nonequilibrium free energy

$$G^{neq}(N,P,T,\overline{v}) = E - TS + PV\,. \tag{57}$$

Now, by using Euler's thermodynamic equation, $U = TS - PV + \mu N$, one can write the nonequilibrium Gibbs free energy in terms of the chemical potential μ and the kinetic energy of particles in solution in the form

$$G^{neq} = \mu N + \frac{1}{2}mN\overline{v}^2 + \frac{1}{2}\rho V\overline{v}^2\,. \tag{58}$$

Using Eq. (58), we may define the effective chemical potential of the particles in solution μ^{neq}

$$\mu^{neq} = \mu + \frac{1}{2}m\overline{v}^2\,. \tag{59}$$

Notice that Eq. (59) is fully compatible with Eq. (52) and with (21). However, it is important to stress that Eq. (59) is exact and a natural consequence of the first law of thermodynamics for open systems.

5.3 The kinetic energy contribution to nucleation

Consider now that nucleation of crystals/aggregates may occur in the system. Following the lines of classical nucleation theory, the description of the nucleation process can be given by means of the free energy difference between the solution G_s^{neq} and the solution/cluster compound G_c^{neq}. Thus, assuming that the particles in solution move at a constant average velocity $\overline{v}$, that is, local diffusion is negligible, the Gibbs free energy of the particles in solution G_s^{neq} takes the form

$$G_s^{neq} = \mu_s^{neq} n_s + \frac{1}{2}\rho V \overline{v}^2 , \tag{60}$$

where μ_s and n_s are the chemical potential and the number of particles in solution, respectively. When the cluster of particles starts to form, the Gibbs free energy of the system changes and therefore, the solution/cluster compound G_c becomes

$$G_c^{neq} = (n_s - n_c)\mu_s^{neq} + \mu_c n_c + \tilde{\sigma} A + \frac{1}{2}\rho V \overline{v}^2 , \tag{61}$$

where μ_c and n_c are the chemical potential and the number of particles of the cluster with surface energy $\tilde{\sigma}$ and surface A. Using Eqs. (60) and (61), one obtains the following expression for the total change in the Gibbs free energy of the system $\Delta G^{neq} = G_c^{neq} - G_s^{neq}$

$$\Delta G^{neq} = -\Delta\mu^{neq}\frac{4\pi R^3}{3v_s} + 4\pi\gamma R^2 \tag{62}$$

where $\Delta\mu^{neq} = \mu_s - \mu_c + \frac{1}{2}m\overline{v}^2 = \Delta\mu + \frac{1}{2}m\overline{v}^2$ and we have assumed that the cluster has spherical shape with radius R. Eq. (62) shows that for a constant value of $\Delta\mu$, the Gibbs free energy free energy barrier decreases for increasing values of the flow velocity $\overline{v}$. This means that the flow gives energy to the particles in order to overcome the barrier. This result agrees partially with those reported in (Blaak, 2004) and, more importantly, with the experimental observation (A. Penkova, 2006) that for low and moderate flow velocities nucleation is enhanced. However, it disagrees with the experimental fact that for large flow velocities nucleation is inhibited or supressed. To explain this second effect, it is necessary to consider the deformation of the nuclei. For another discussion on nucleation in soft-matter colloid-type systems, see (Gadomski, 2005; 2006).

5.4 The elastic contribution

The stresses applied by the solvent on the nucleus' surface will induce its deformation and, as a consequence, an elastic contribution to the free energy G_{el} will emerge. This contribution is proportional to the restitution force that tends to maintain the spherical shape of the nucleus of volume V_c and, consequently, it is proportional to the elastic stress per unit volume $P^E = tr[\mathbb{P}^E]$, thus leading to the expression: $G_{el} \propto P^E V_c$.

The average elastic stress per unit volume of the solution/cluster compound is given through the relation: $\mathbb{P}^E = -\xi_0 F_0 \mathbb{A}$ with $\mathbb{A}$ the conformation tensor accounting for the average deformation of the nuclei. In the present case, the characteristic restitution coefficient per unit mass ξ_0 will be related to the Lame coefficients of the material and the dimensionless factor F_0 will be held constant for simplicity. Thus, from Eq. (37), the average deformation of the nuclei in terms of the flow velocity is

$$\mathbb{A} = \frac{kT}{F_0}\mathbb{I}\cdot\left(\mathbb{I} - \frac{L_{Rr}}{L_{RR}F_0}(\nabla\mathbf{v}\cdot\nabla\mathbf{v}^T)\right)^{-1} , \tag{63}$$

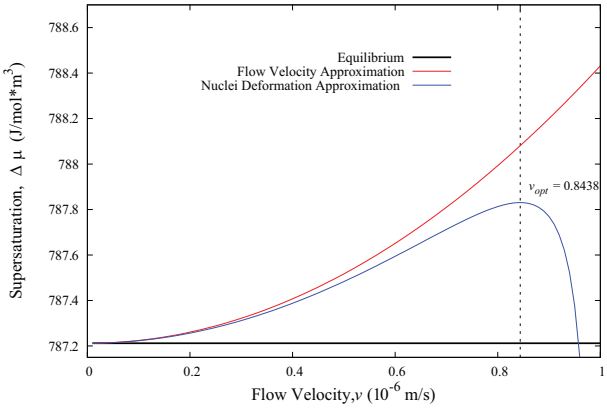


Fig. 4. Supersaturation as a function of the averge flow velocity for a lysozime solution as given by Eq. (67). The existence of an optimal flow velocity $\overline{v}_{opt}$ gives rise to two nulceation regimes. For moderate flow velocities $\overline{v} < \overline{v}_{opt}$ nucleation is enhanced whereas for large flow velocities $\overline{v} > \overline{v}_{opt}$ becomes inhibited and even supressed.

where we have assumed stationary conditions: $d\mathbb{A}/dt - 2[\mathbb{A} \cdot \nabla \mathbf{v}]^s = 0$ and used the relation $\langle \vec{F}\vec{R} \rangle = F_0 \mathbf{A}$. Now, assuming that the average flow is in the x-direction and calculating the trace of $\mathbb{A}$ we may finally write

$$P^E \approx -\xi_0 kT \left(\frac{\tilde{a}}{\tilde{a} - b\overline{v}^2} \right), \tag{64}$$

where $\tilde{a} = L_{RR} F_0$ and $b = L_{Rr} R_{crit}^{-2}$. Here we have assumed that $|\nabla \mathbf{v}| \sim \overline{v}/R_{crit}$ with the characteristic length given by the radius of the critical nucleus R_{crit}. Eq. (64) shows that the elastic contribution to the stresses in the nuclei is a function of the temperature T and the average flow velocity $\overline{v}$. This effect modifies the total free energy of the particles in the solution/cluster compound G_c by adding the term $G_{el} = \Delta P^E V_c$, with $\Delta P^E = P^E - P_0^E$ with $P_0^E = -\xi_0 kT$ the value of P^E at zero flow. Therefore, the expression for the Gibbs free energy taking into account the elastic contribution is

$$G_c = (n_s - n_c)\mu_s^{neq} + \mu_c n_c + \gamma A + \Delta P^E V_c \,. \tag{65}$$

Now, using Eqs. (60) and (65), the total change in the Gibbs free energy in terms of the radius of the nucleus is given by

$$\Delta G^{neq} = -\Delta\tilde{\mu}^{neq} \frac{4\pi R^3}{3v} + 4\pi\gamma R^2 \,, \tag{66}$$

where

$$\Delta\tilde{\mu}^{neq} = \Delta\mu + \frac{1}{2} m\overline{v}^2 + \xi_0 v_s kT \left(1 - \frac{\tilde{a}}{\tilde{a} - b\overline{v}^2} \right) . \tag{67}$$

This result implies that the steady state nucleation rate J becomes a function of the average flow velocity

$$J(\overline{v}) \simeq \exp\left[-\Delta G_{crit}^{neq}(\overline{v})/kT \right] \tag{68}$$

because the height of the free energy barrier becomes a function of $\overline{v}$: $\Delta G_{crit}^{neq} \sim \tilde{\sigma}^3/[\Delta\tilde{\mu}^{neq}]^2$, see Eq. (53). The main consequence arising from the last term of (67) is that an optimal average flow velocity $\overline{v}_{opt}$ appears. This means that there are two regimes for nucleation in the presence of flow. For moderate flow velocities, $\overline{v} < \overline{v}_{opt}$, nucleation is enhanced whereas for large flow velocities $\overline{v} > \overline{v}_{opt}$ nucleation is inhibited or even supressed. The existence of an optimal flow velocity for nucleation was established experimentally in Ref. (A. Penkova, 2006) for lyzosyme, ferritin and apoferritin solutions.

6. Conclusions

In this chapter, we have discussed three important aspects of thermodynamics in the presence of flow. By considering different points of view ranging from kinetic and stochastic theories to thermodynamic theories at mesoscopic and macroscopic levels, we addressed the effects of flow on transport coefficients like diffusion and viscosity, constitutive relations and equations of state. In particular, we focused on how some of these effects may be derived from the framework of mesoscopic nonequilibrium thermodynamics.

Our analysis started by considering how the flow introduces corrections on transport coefficients and illustrated how this may be relevant for giving a sound theoretical description of experiments. The first case considered was the contribution on the diffusion coefficient due to a shear flow. This contribution, proportional to the velocity gradient or shear rate, was first obtained by means of kinetic theory for Lorentz fluids and after that corroborated in a more general fashion by MNET and Langevin calculations. Computer simulations and experiments in which this correction was observed were also discussed. Comparison with experiments illustrating how these contributions are relevant for giving a sound theoretical explanation of the shear induced effect and the transition to the irreversibility was also presented.

Afterwards, we studied the existence of non-linear constitutive relations on several systems like, for example, suspensions of Brownian and non-Brownian particles and polymer solutions. These non-linear relations are understood in the sense that the currents, like for instance, the diffusion current or the momentum current are not simply proportional to their conjugated thermodynamic forces. In particular, the relation between the stress tensor of a fluid and the velocity gradient, which is linear for simple Newtonian fluids, becomes non-linear for many soft-matter systems like polymer solutions, polymer melts, glasses, etc. This fact makes these materials non-Newtonian and viscoelastic. We illustrated the phenomenological approach to this problem by mentioning Maxwell and extended Maxwell fluid models, and then showed how similar non-linear constitutive relations emerge from the mesoscopic thermodynamics analysis by considering Faxén's theorem and its generalizations to the case of time dependent flows.

These analyses allowed us to present a discussion on how the equilibrium equations of state become modified when the system is subject to sufficiently strong flows, because not only the non-diagonal components of the stress tensor become modified, but also the normal components. In addition, through the discussion we have shown that the chemical potential becomes corrected by a term proportional to the local kinetic energy of diffusion, and therefore, is a flow dependent quantity. This correction was analyzed in the context of nucleation, were it was found that both the nucleation barrier and consequently the nucleation rate can be written in terms of the kinetic energy of the flow.

Using control volume thermodynamics, we showed that the nonequilibrium equations of state heuristically proposed by Frenkel and co-workers is in fact consistent with the general expression of the first law of thermodynamics, in which the heat and the work exchanged

between the system and its surroundings are responsible for the change of the internal, kinetic and potential energies of the system. Our proof uses arguments similar to those previously used by Oono's when discussing why nonequilibrium thermodynamics may treat both thermodynamic and hydrodynamic fluctuations. Finally, we have proposed a model taking into account internal restitutive stresses of the nucleus that allows to describe in a qualitative form the experiments on lysozyme, ferritin and apoferritin nucleation in the presence of flow. The main result is that the supersaturation parameter and consequently the nucleation rate exhibit two regimes. One in which nucleation is enhanced by low and moderate flow velocities and other one (large flow velocities) in which nucleation is inhibited or even supressed by the flow. These results lead to the existence of an optimal nucleation flow velocity, a quantity that was reported in experiments.

7. Acknowledgments

We thank Dr. G. Barrios for kindly sharing Fig. 1 with us. This work was supported by Grant DGAPA-UNAM IN-102609. Rl and RLF also acknowledge financial support by CONACyT.

8. References

A. Penkova, e. a. (2006). Nucleation of protein crystals under the influence of solution shear flow., *Annals of the New York Academy of Sciences* 1077: 214–31.

Abou, B. & Gallet, F. (2004). Probing a nonequilibrium einstein relation in an aging colloidal glass, *Phys. Rev. Lett.* 93: 160603.

Aidun, C. K. & Ding, E. J. (1998). Direct analysis of particulate suspensions with inertia using the discrete Boltzmann equation, *J. Fluid Mech.* 373: 287.

Apelian, M. R., e. a. (1988). Impact of the constitutive equation and singularity on the calculation of stick-slip flow: The modified upper-convected maxwell model, *J. Non-Newtonian Fluid Mech.* 27: 299–321.

Beck, C. & Cohen, E. (2003). Superstatistics, *Physica A* 322: 267–275.

Bellon, L. e. a. (2001). Violation of the fluctuation-dissipation relation during the formation of a colloidal glass, *Europhys. Lett.* 53: 511.

Beris, A. N. & Edwards, B. J. (1994). *Thermodynamics of flowing systems*, Oxford, New York.

Blaak, R., e. a. (2004). Homogeneous nucleation of colloidal melts under the influence of shearing fields, *J. of Phys.: Cond. Matt.* 16: S3873–S3884.

Bossis, G. & Brady, J. F. (1989). The rheology of brownian suspensions, *J. Chem. Phys* 91: 1866.

Breedveld, V., e. a. (1998). The measurement of the shear-induced particle and fluid tracer diffusivities in concentrated suspensions by a novel method, *J. Fluid Mech.* 375: 297–318.

Butler, S. & Harrowell, P. (1995). Kinetics of crystallization in a shearing colloidal suspension, *Phys. Rev. E* 52: 6424–6430.

C. Beck, E. C. & Swinney, H. (2005). Fluctuations of entropy and log-normal superstatistics, *Phys. Rev. E* 72: 056133.

Callen, H. B. (1985). *Thermodynamics and an introduction to thermostatistics*, John Wiley and Sons, New York, Second Edition.

Criado-Sancho, M. e. a. (2005). Nonequilibrium kinetic temperatures in flowing gases, *Phys. Lett. A* 350: 339.

de Groot, S. R. & Mazur, P. (1984). *Non-equilibrium Thermodynamics*, Dover, New York.

Delgado, J. & Castillo, R. (2005). *Unpublished data* .

Delgado, J. & Castillo, R. (2007). Shear-induced structures formed during thixotropic loops in dilute worm-micelle solutions, *J. Coll. Int. Sci.* 312: 481.

Demirel, Y. (2007). *Non-equilibrium Thermodynamics*, Elsevier, Amsterdam.

Doi, M. & Edwards, S. F. (1998). *The Theory of Polymer Dynamics*, Oxford University Press, New York.

Drazer, D., e. a. (2002). Deterministic and stochastic behaviour of non-brownian spheres in sheared suspensions, *J. Fluid Mech.* 460: 307–335.

Drossinos, Y. & Reeks, M. W. (2005). Brownian motion of finite-inertia particles in a simple shear flow, *Phys. Rev. E* 71: 031113.

Ferziger, J. H. & Kaper, H. G. (1972). *Mathematical Theory of Non-Uniform Gases*, North Holland, Amsterdam.

Gadomski, A. e. a. (2005). On temperature and space-dimension dependent matter agglomerations in a mature growing stage, *Chemical Physics* 310: 153–161.

Gadomski, A. e. a. (2006). Formation of soft material in terms of a coupled matter migration-and-relaxation picture: a synchronous mode, *Materials Science-Poland* 24(No. 4): 935–949.

Gadomski, A. e. a. (2008). *Statistical Mechanics Research*, Nova Science Publishers, Inc. New York, chap. 1, pp. 13-98.

Guasto, J. S. & Gollub, J. P. (2007). Hydrodynamic irreversibility in particle suspensions with nonuniform strain, *Phys. Rev. E* 81: 061401.

Hinch, E. J. (1994). Uncoiling a polymer molecule in a strong extensional flow, *J. Non-Newtonian Fluid Mech.* 54: 209–230.

James, D. F. (1972). Constitutive Equations for Dilute Polymer Solutions, *J. Rheol.* 16: 175.

Kondepudi, D. K. & Prigogine, I. (1999). *Modern Thermodynamics: From Heat Engines to Dissipative Structures*, John Wiley and Sons, Chichester.

Kurchan, J. (2005). In and out of equilibrium, *Nature* 433: 222–225.

Ladd, A. J. C. (1994). Numerical simulations of particulate suspensions via a discretized Boltzmann equation. Part 1 and Part 2., *J. Fluid Mech.* 271: 285–339.

Málaga, C., e. a. (2006). Mesoscopic constitutive relations for dilute polymer solutions, *Physica A* 369: 291–300.

Marchetti, M. C. & Dufty, J. W. (1983). Tagged particle fluctuations in uniform shear flow, *J. Stat. Phys.* 32: 255–277.

Mauri, R. & Leporini, D. (2006). Violation of the fluctuation-dissipation theorem in confined driven colloids, *Europhys. Lett.* 76: 10221028.

Mazur, P. & Bedeaux, D. (1974). A generalization of faxén's theorem to nonsteady motion of a sphere through an incompressible fluid in arbitrary flow, *Physica A* 76: 235–246.

McLennan, J. A. (1989). *Introduction to Non-equilibrium Statistical Mechanics*, Prentice-Hall, Englewood Cliffs, New Jersey.

Meyers, M. A. & Chawla, K. K. (2009). *Mechanical Behavior of Materials*, Cambridge University Press, Cambridge.

O'Hern, C. S. e. a. (2004). Effective temperatures in driven systems: Static versus time-dependent relations, *Phys. Rev. Lett.* 93: 165702.

Onuki, A. (2004). *Phase Transition Dynamics*, Cambridge University Press, New York.

Pérez-Madrid, A. (2005). A model for nonexponential relaxation and aging in dissipative systems, *J. Chem. Phys.* 122: 214914.

Pérez-Madrid, A. & Santamaría Holek, I. (2009). Fluctuation theorems for systems under Fokker-Planck dynamics, *Phys. Rev. E* 79: 011101.

Pérez-Madrid, e. a. (1994). Brownian motion in the presence of a temperature gradient, *Physica A* 212: 231–238.

Pine, D. J., e. a. (2005). Chaos and threshold for irreversibility in sheared suspensions, *Nature* 438: 997–1000.

Popov, A. V. & R., H. (2007). Ontology of temperature in nonequilibrium systems, *J. Chem. Phys.* 126: 244506.

Ragone, D. V. (1995). *Thermodynamics of materials*, John Wiley and Sons, New York.

Rallison, J. M. & Hinch, E. J. (1988). Do we understand the physics in the constitutive equation?, *J. Non-Newtonian Fluid Mech.* 29: 37–55.

Reguera, D. & Rubi, J. (2003a). Homogeneous nucleation in inhomogeneous media. i. nucleation in a temperature gradient, *J. Chem. Phys.* 119: 9877.

Reguera, D. & Rubi, J. (2003b). Homogeneous nucleation in inhomogeneous media. ii. nucleation in a shear flow, *J. Chem. Phys.* 119: 9888.

Rodríguez, R. F., e. a. (1983). Fokker-Planck and Langevin descriptions of fluctuations in uniform shear flow, *J. Stat. Phys.* 32: 279–298.

Rodríguez, R. F. & Santamaría Holek, I. (2007). Superstatistics of brownian motion: A comparative study, *Physica A* 385: 456–464.

Rubi, J. M. & Mazur, P. (1994). Simultaneous brownian motion of n particles in a temperature gradient, *Physica A* 250: 253–264.

Rubi, J. M. & Pérez-Madrid, A. (1999). Inertial effects in non-equilibrium thermodynamics, *Physica A* 264: 492.

Ryskin, G. (1988). Brownian motion in a rotating fluid: Diffusivity is a function of the rotation rate *Phys. Rev. Lett.* 61: 01442.

Santamaría Holek, I., e. a. (2005). Mesoscopic thermodynamics of stationary non-equilibrium states, *New J. Phys.* 7: 35.

Santamaría Holek, I., e. a. (2009). Transition to irreversibility in sheared suspensions: An analysis based on a mesoscopic entropy production, *Phys. Rev. E* 79: 031201.

Santamaría Holek, I., e. a. (2001). Diffusion in stationary flow from mesoscopic nonequilibrium thermodynamics, *Phy. Rev. E* 63: 051106.

Sarman, S., e. a. (1992). Mutual and self-diffusion in fluids undergoing strong shear, *Phys. Rev A* 46: 893–902.

Seriou, A. & Brady, J. F. (2004). Shear-induced self-diffusion in non-colloidal suspensions, *J. Fluid Mech.* 506: 285–314.

Shea, J. E. & Oppenheim, I. (1998). Fokker-Planck and non-linear hydrodynamic equations of an inelastic system of several brownian particles in a non-equilibrium bath, *Physica A* 250: 265.

Subramanian, G. & Brady, J. F. (2004). Multiple scales analysis of the Fokker-Planck equation for simple shear flow, *Physica A* 334: 343–384.

Swailes, D. C., e. a. (2009). Stochastic transport of particles in straining flows, *Phys. Rev. E* 79: 036305.

Taylor, G. I. & Friedman, J. (1996). *Low Reynolds Number Flows*, National Committe on Fluid Mechanics Films, Encyclopedia Britannica Educational Corp., United States.

6

Non-Instantaneous Adiabats in Finite Time

Delfino Ladino-Luna and Ricardo T. Páez-Hernández
Universidad Autónoma Metropolitana-A
México

1. Introduction

Since the pioneer paper of Curzon and Ahlborn (1975), the so called finite time thermodynamics has been in development. Curzon and Ahlborn proposed a model of thermal engine called endoreversible cycle or Curzon and Ahlborn cycle, shown in Figure 1, with the so called Curzon-Ahlborn-Novikov efficiency,

$$\eta_{CAN} = 1 - \sqrt{T_C / T_H}\,, \tag{1}$$

where T_C is the cold reservoir temperature and T_H is the hot reservoir temperature. This endoreversible cycle is an engine in which the endoreversibility condition, $Q_H / T_{HW} = Q_C / T_{CW}$, is fulfilled and the entropy production during the exchange of heat, Q_H and Q_C, between the system and its reservoirs of heat is only taken into account. The temperatures of working substance are T_{HW} and T_{CW}. The relation between these temperatures in the cycle is $T_C < T_{CW} < T_{HW} < T_H$.

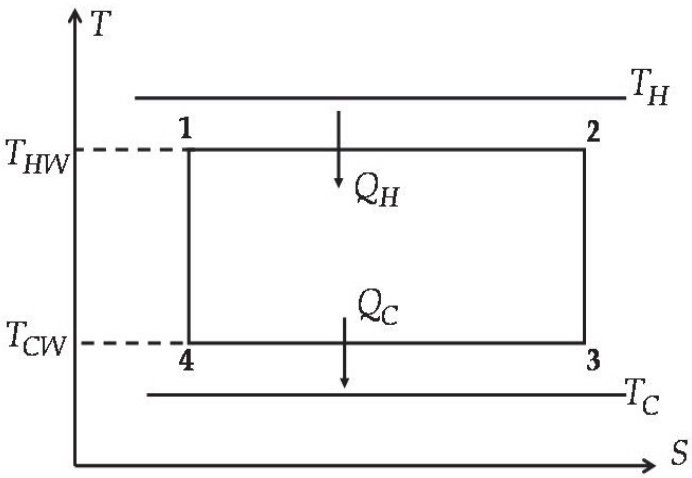


Fig. 1. Curzon and Ahlborn cycle in the entropy S vs temperature T plane.

As we can see, Carnot efficiency, η_C, is obtained when the temperatures of reservoirs are the same as engine temperatures, which means $T_{HW} = T_H$ and $T_C = T_{CW}$ in Figure 1, namely,

$$\eta_C = 1 - \frac{T_C}{T_H} = 1 - \frac{T_{CW}}{T_{CH}} \tag{2}$$

Equation (1) was previously advanced by Chambadal (1957) and Novikov (1958), among others, and has been recovered by some procedures (Salamon, et al., 1976; Rubin, 1979a, 1979b, 1980; Bejan, 1996; Gutkowicz-Krusin et al., 1978 among others). Particularly, the optimal configuration of heat engines was studied (Rubin, 1979a), and it was introduced a procedure in which the power output of cycle is taken as a function of the compression ratio by using the parameter $\lambda \sim \left[\ln V_{\max} - \ln V_{\min}\right]^{-1}$, where $V_{\max}$ and $V_{\min}$ are the maximum and the minimum volumes spanned in the cycle, respectively (Gutkowicz-Krusin et al., 1978). More recently, this subject has been also studied by other authors (Badescu, 2004; Amelkin, et al., 2004, 2005; Song et al., 2006, 2007). Even more Angulo-Brown (1991) introduced an optimization criterion of merit for the Curzon and Ahlborn cycle taking into account entropy production, the ecological criterion, through the function,

$$E = P - T_C \sigma , \tag{3}$$

where P is the power output, T_C is the temperature of cold reservoir an σ is the total entropy production. The function in (3) is known as ecological function, and at maximum of this function the efficiency of Curzon and Ahlborn cycle can be written as,

$$\eta_E = 1 - \sqrt{(\varepsilon^2 + \varepsilon)/2} . \tag{4}$$

A comparison of values obtained with the previous expressions of efficiency, for some plants reported in the literature of finite time thermodynamics, is shown in Table 1.

Notice that the ecological criterion proposed by Angulo-Brown for finite-time Carnot heat engines, Equation (3), represents a compromise between the high power output P and a loss power output, $T_C \sigma$. However Yan (1993) showed that it might be more reasonable to use $E_0 = P - T_0 \sigma$ if the cold reservoir temperature T_C is not equal to the environments temperature T_0 because in the definition of E two different quantities, exergy output, P, and a non-exergy $T_C \sigma$, were compared together. The criterion with function E_0 is more reasonable than that presented by Angulo-Brown. Nevertheless, since $E_0 \to E$ when $T_0 \to T_C$ it can be used as the optimization of E without loss of generality.

Recently, following the procedure of Gutkowicz-Krusin et al (1978) the form of the ecological function and its efficiency was found using the Newton heat transfer law and ideal gas as working substance (Ladino-Luna & de la Selva, 2000), and using Dulong-Petit heat transfer law for ideal gas as working substance (Ladino-Luna,2003).

It is important to remark that Curzon and Ahlborn efficiency is an adequate approximation for conventional power plants, and ecological efficiency is the adequate approximation for modern power plants (nuclear and others), as it is shown in Table 1.

On other hand, in nature there are no endoreversible engines. Thus, some authors have analyzed the non-endoreversible Curzon and Ahlborn cycle. Ibrahim et al. (1991), and Wu and Kiang (1992) proposal include a non-endoreversibility parameter to take into account

Plant	T_C / T_H	η_C	η_{CAN}	η_E	η_{obs}
West Thurrock (coal fired steam plant), U K	298/838	0.64439	0.40367	0.50905	0.360
Lardarello (geothermal steam plant), Italy	353/523	0.32505	0.17845	0.24818	0.160
Central steam power station, U K	298/698	0.57307	0.3466	0.44809	0.280
Steam power plant, U S A	298/923	0.67714	0.43179	0.55447	0.400
Combined-cycle (steam and mercury), U S A	298/783	0.61941	0.38308	0.48744	0.340
Doel 4 (nuclear pressurized water reactor), Bélgium	283/566	0.50000	0.29289	0.38763	0.350
Almaraz II (nuclear pressurized water reactor), Spain	290/600	0.51667	0.30478	0.40127	0.345
Sizewell B (nuclear pressurized water reactor), U K	288/581	0.50430	0.29594	0.39114	0.363
Cofrentes (nuclear boiling water reactor), Spain	289/562	0.48577	0.28290	0.37603	0.340
Heysham (nuclear advanced gas cooled reactor), U K	288/727	0.60385	0.37060	0.47413	0.400

Table 1. Values of different efficiency expressions for the cycle in Figure 1. *T* is in Kelvin scale.

internal irreversibilities in the cycle. Later, Chen (1994, 1996) analyzed the effect of thermal resistances, heat leakage and internal irreversibility with these parameter definite as,

$$I_S \equiv \Delta S_C / \Delta S_H , \tag{5}$$

where ΔS_C is the entropy change during heat exchange from the engine to the cold reservoir, and ΔS_H is the entropy change during heat exchange from the hot reservoir to the engine. Chen et. al. (2004, 2006) carried out the ecological optimization for generalized irreversible Carnot engine with heat resistance, heat leakage and internal irreversibility for newtonian heat transfer law. Zhu et. al (2003) used a generalized convective heat transfer law $Q \propto (\Delta T)^n$, and generalized radiative heat transfer law $Q \propto \Delta(T^n)$. More recently the ecological optimization for generalized irreversible universal heat engine, including Diesel, Otto, Bryton Atkinson, Dual and Miller cycles, with heat resistance, heat leakage and internal irreversibility was carried out for newtonian heat transfer law (Chen et al., 2007). The non-endoreversible Curzon and Ahlborn cycle model is shown in Figure 2. The efficiency of Curzon and Ahlborn cycle using the parameter in Equation (5), at maximum power output was found as (Chen, 1994, 1996),

$$\eta_m = 1 - \sqrt{I_S \varepsilon}, \quad I_S > 1 . \tag{6}$$

On other hand, Angulo-Brown et al (1999) showed that a general property of endoreversible Curzon and Ahlborn cycle demostrated previously (Árias-Hernández & Angulo-Brown, 1997) can be extended for a non-endoreversible Curzon and Ahlborn cycle. Besides, Velasco et. al. (2000) follow the idea in Chen (1994, 1996), and they found expressions to measure possible reductions of non-desired effects in heat engines operation. They pointed out that I_S is not depending of ε and re-wrote Equation (6) as,

$$\eta_m = 1 - \sqrt{\varepsilon / I}, \quad I \equiv 1 / I_S, \quad 0 < I < 1 . \tag{7}$$

Even more, Angulo-Brown et. al. (2002) applied variational calculus to show that both the saving function (Velasco et. al., 2000) and a modified ecological criterion are equivalent.
These previous results have been found assuming an ideal gas as working fluid. However, in a real context, a thermal engine works with a non-ideal gas. The performannce of a finite time cycle with a van der Waals gas as working fluid was analyzed among others by Agrawal & Menon (1990), and more recently by Ladino-Luna (2005).

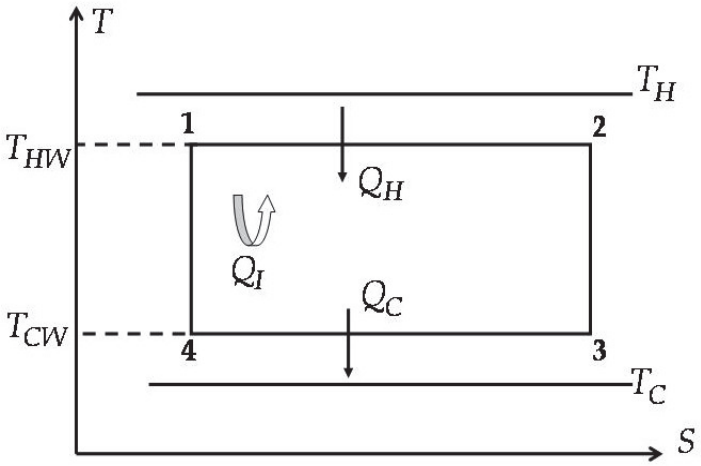


Fig. 2. Non-endoreversible Curzon and Ahlborn cycle in the *S*-*T* plane. Q_I is the generated heat by internal processes.

In the present work it is shown that some of internal irreversibilities can be taken into account for a more general expressions of both power output and ecological function, with a non-linear heat transfer law like $\frac{dQ}{dt} \propto (\Delta T)^k$, assuming the Curzon and Ahlborn cycle with non-instantaneous adiabats. Approximate efficiencies curves at maximum power output and at maximum ecological function are shown for $k = \frac{5}{4}$, that is in case of the Dulong and Petit heat transfer law. Comparative tables of values of efficiencies are shown for certain power plants reported in some papers, (Árias-Hernández & Angulo-Brown, 1997; Velasco et al., 2000 and more recently Ladino-Luna, 2008). The cycle analysis shall be treated assuming both ideal gas and van der Waals gas as working fluids. Also, we show that power output and ecological function have a similar property when the compression ratio is taking into account, e. g. the efficiency obtained at maximum of each of objective function with the definition $\varepsilon = T_C / T_H$ is a bound for efficiencies, when an engine is modeled as a Curzon and Ahlborn cycle working at maximum of each objetive function and the time of all the processes in the cycle is took into account (Ladino-Luna & de la Selva, 2000; Ladino Luna, 2003). All of quantities have been taken in the International Measurement System.

2. Non-instananeous adiabats with Newton´s heat transfer law

From classical thermodynamics, the efficiency of a reversible thermal engine working between two temperatures $T_{HW} > T_{CW}$ is known when heat exchanged is also known. In this description, the temperatures of working gas in the isothermal processes, T_{HW} and T_{CW}, are assumed to be the same as that of the corresponding reservoirs. As a consequence the process associated with the heat transfer between the engine and the reservoirs is ignored. The upper limit of the efficiency of any heat engine corresponds to the Carnot cycle, η_C, in which the temperatures of the reservoirs are the same as the temperatures of the heat engine in Figure 1, as it was shown in Equation (2). Thus, the definition of efficiency of an engine working in cycles leads to the Carnot efficiency, fulfilling,

$$\frac{Q_H - Q_C}{Q_H} \leq 1 - \frac{T_{CW}}{T_{HW}} \tag{8}$$

It is important to note that the following expressions, with the change $\varepsilon \equiv T_C / T_H$, Carnot efficiency $\eta_C \equiv 1 - \varepsilon$, Curzon and Ahlborn-Novikov efficiency $\eta_{CAN} = 1 - \sqrt{\varepsilon}$ and ecological efficiency $\eta_E = 1 - \sqrt{(\varepsilon^2 + \varepsilon)/2}$, and using the Newton heat transfer law, appear as a function of ε, so they can be written in general as,

$$\eta = 1 - z(\varepsilon)\,, \tag{9}$$

Thus, the problem of finding the efficiency of a heat engine modeled as a Curzon and Ahlborn cycle, and using either of two alternatives, maximizing power output or maximizing ecological function, becomes the problem of finding a function $z = z(\varepsilon)$, as complicated or simple as to allow the heat transfer law being used and the chosen procedure. So, substituting $z = z(\varepsilon)$ in expression (9) the efficiency is obtained as

$$\eta = \eta(\varepsilon)\,. \tag{10}$$

So we can use the Newton's heat transfe law, and also we can assume the same thermal conductance α in two isothermal processes of Curzon and Ahlborn cycle. The heat exchanged between the engine and its surroundings can be expressed as,

$$\frac{dQ_H}{dt} = \alpha(T_H - T_{HW}) \qquad \text{and} \quad \frac{dQ_H}{dt} = \alpha(T_C - T_{CW})\,. \tag{11}$$

2.1 The Gutkowics-Krusin, Procaccia and Ross model

With the previous ideas, to make the present paper self-contained we include in this section a brief explanation and some results of model used by Gutkowicz-Krusin et al (1978), and others that we need for our present purposes. In their model Gutkowicz-Kru et al. consider a working substance inside of a cylinder with a movable piston as engine, and also they considered an ideal gas as working fluid, contained in the cylinder and the mass of piston not negligible. The inertia of the movable piston does not affect the endoreversible character of Curzon and Ahlborn cycle to consider the expansion of gas, and because the volume occupied by the gas in the expansion and compression can be written as

$$V = l \cdot A \,, \tag{12}$$

where V is the volume occupied by the gas, A is the cross section area (constant) of the cylinder and l is the distance traveled by the piston in the expansion or compression of gas. The acceleration of the piston during the processes is

$$\frac{d^2 l}{dt^2} = \frac{1}{A} \cdot \frac{d^2 V}{dt^2} \,, \tag{13}$$

so that, from the pressure definition,

$$pressure = \frac{force}{area} \,, \tag{14}$$

and with the Newton's second law, namely $force = (mass) \cdot (acceleration)$, we can write,

$$\frac{1}{A}\left(m \frac{d^2 l}{dt^2} \right) = \frac{force}{area} = pressure \,. \tag{15}$$

For the *gas+piston* system, the difference in internal and external pressures is expressed by

$$p - p_{ext} = \frac{m}{A} \frac{d^2 V}{dt^2} \,. \tag{16}$$

On other hand, conservation of energy law of the system can be written as

$$\frac{dU}{dt} = \frac{dQ}{dt} - p_{ext} \frac{dV}{dt} - \frac{m}{A} \frac{d^2 V}{dt^2} \frac{dV}{dt} \,, \tag{17}$$

where the last term represents the power output during the movement of piston to take volumen V. Substituting (16) in (17) it is obtain

$$\frac{dU}{dt} = \frac{dQ}{dt} - p_{ext} \frac{dV}{dt} - (p - p_{ext}) \frac{dV}{dt} \quad \text{ó} \quad \frac{dU}{dt} = \frac{dQ}{dt} - p \frac{dV}{dt} \,, \tag{18}$$

which means that the system is in mechanical equilibrium with its surroundings. Now, we can find the form of time for each process using the Newton's heat transfer law, Equation (11). For isothermal processes, using an ideal gas we have $U = U(T) = constant$, indicating that (18) is reduced to the expression,

$$\frac{dQ}{dt} = p \frac{dV}{dt} \,. \tag{19}$$

Due to the equation of state for ideal gas, (19) can be written as

$$\frac{dQ}{dt} = \frac{RT}{V} \frac{dV}{dt} = RT \frac{d}{dt} (\ln V) \,; \tag{20}$$

The power, defined by the quotient of the total work output W and the total time t_{tot} is as,

$$P=\frac{W}{t_{tot}}=\frac{\alpha(T_{HW}-T_{CW})\left(\ln\frac{V_3}{V_1}+\frac{1}{\gamma-1}\ln\frac{T_{CW}}{T_{HW}}\right)}{\left(\frac{T_{HW}}{T_H-T_{HW}}+\frac{T_{CW}}{T_{CW}-T_C}\right)\ln\frac{V_3}{V_1}}. \tag{21}$$

α is the thermal conductance, $\gamma\equiv C_P/C_V$; t_{tot} is the cycle period and the adiabatic processes are not instantaneous. In fact,

$$t_{TOT}=t_1+t_2+t_3+t_4 \tag{22}$$

where the times for the isothermal processes have been found to be,

$$t_1=\frac{RT_{HW}}{\alpha(T_H-T_{HW})}\ln\frac{V_2}{V_1}\ \text{ and }\ t_3=\frac{RT_{CW}}{\alpha(T_{CW}-T_C)}\ln\frac{V_4}{V_3} \tag{23}$$

and the times for the adiabatic processes have been assumed to be:

$$t_2=f_1\ln\frac{V_3}{V_2}\ \text{ and }\ t_4=f_2\ln\frac{V_4}{V_1}, \tag{24}$$

with

$$f_1\equiv\frac{RT_{HW}}{\alpha(T_H-T_{HW})}\ \text{ and }\ f_2\equiv\frac{RT_{CW}}{\alpha(T_{CWiw}-T_C)}, \tag{25}$$

where R is the general constant of gases. The heat flows, Q_H and Q_C are assumed to be given by Newton's heat transfer law, as (11). The power output is written in terms of the variables $u=T_{HW}/T_H$ and $z=T_{CW}/T_{HW}$ from which we obtain $P=P(u,z)$ as,

$$P=\frac{\alpha T_H(1-z)\left[1+\lambda\ln z\right]}{\frac{1}{1-u}+\frac{z}{uz-\varepsilon}}, \tag{26}$$

and its maximization conditions $\partial P/\partial u=0$ and $\partial P/\partial z=0$ allow to obtain

$$u=\frac{z+\varepsilon}{2z}, \tag{27}$$

and

$$(z^2-\varepsilon)(1+\lambda\ln z)=\lambda(z-\varepsilon)(1-z); \tag{28}$$

where λ represents the external parameter,

$$\lambda=\frac{1}{(\gamma-1)\ln(V_3/V_1)} \tag{29}$$

meaning that

$$P_{\max}=P_{\max}(u(z),z), \tag{30}$$

that is $P_{\max}$ is a projection on the (z,P) plane. It is also found that at the maximum power condition z is given by a power series in λ:

$$z_P = \sqrt{\varepsilon} + \tfrac{1}{2}(1-\sqrt{\varepsilon})^2\lambda + \tfrac{1}{4}(1-\sqrt{\varepsilon})^2\left[(1-\sqrt{\varepsilon})^2 / 2\sqrt{\varepsilon} - \ln\varepsilon\right]\lambda^2 + O(\lambda^3) \tag{31}$$

Upon substiting Equation (31) in Equation (9) and because the terms in the series (31) are positive, an upper bound for the efficiency is obtained when $\lambda = 0$, i.e. when the engine size goes towards infinity, it is the following one:

$$\eta_{\max} = 1 - z_P(\lambda = 0) = \eta_{CAN} \tag{32}$$

In the next section we construct the equation analogous to (31) for the ecological function following the Gutkowicz-Krusin, Procaccia and Ross model outlined here.

2.2 The ecological function

In the ecological function, Equation (3), we take P from Equation (26) and the entropy production term σ as $\sigma = \Delta S / t_{tot}$, where ΔS represents the entropy change caused at the isothermal processes because of the heat transfers Equation (11),

$$\sigma = \frac{1}{t_{tot}}\left(\frac{Q_C}{T_C} - \frac{Q_H}{T_H}\right). \tag{33}$$

t_{tot} is given by Equations (22) to (25), and in terms of the variables (u, z, ε), σ becomes,

$$\sigma = \alpha\frac{T_1}{T_2}\frac{\left(\ln\frac{V_3}{V_1} + \frac{1}{\gamma-1}\ln z\right)(z-\varepsilon)}{\ln\frac{V_3}{V_1}\left[\frac{1}{1-u} + \frac{z}{zu-\varepsilon}\right]}, \tag{34}$$

where, thanks to the endoreversibility condition, we have used the thermostatic results $V_2 / V_1 = V_3 / V_4$ and $V_2 = V_3\left(T_{CW} / T_{HW}\right)^{\frac{1}{\gamma-1}}$, where λ is given by Equation (29).
With Equations(26) and (34) the expression for the ecological function becomes

$$E = \alpha T_1\frac{(1+\varepsilon-2z)(1+\lambda\ln z)}{\frac{1}{1-u} + \frac{z}{zu-\varepsilon}} \tag{35}$$

Figure 3 shows the behavior of $P/\alpha T_H$, $\sigma/\alpha T_H$ and $E/\alpha T_H$ in the u constant plane, at $\lambda = 0$ and ε a given constant value. It is apparent that the maximum power output is achieved with high production of entropy, it is also apparent that zero entropy production is achieved with zero power output, while the function E represents the maximum possible power output with the minimum possible entropy production.
Upon maximizing the two variables function $E = E(u,z)$ (ε defined positive and λ defined semipositive, being external parameters), we obtain for $\partial E/\partial u = 0$ and $\partial E/\partial z = 0$, at first $u = u(z)$, as in case of maximizing power output, and later the following relation between the variables z and u,

$$\left[2(1+\lambda\ln z)z - \lambda(1+\varepsilon-2z)\right](z-\varepsilon)(zu-\varepsilon) = (1+\varepsilon-2z)(1+\lambda\ln z)(1-u)\varepsilon z\,. \tag{36}$$

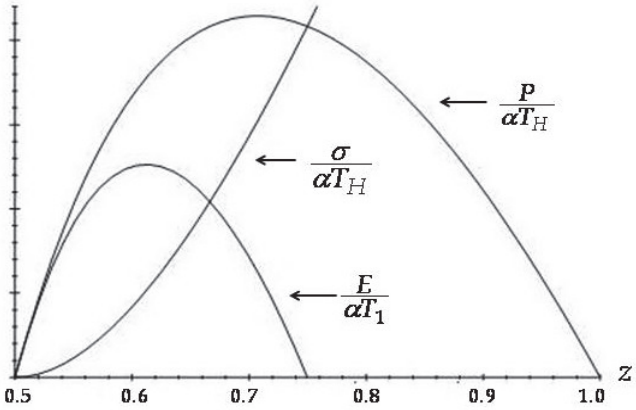


Fig. 3. Behavior of power output, entropy production and ecological efficiency. Here, $T_H = 400\,K$, $\varepsilon = 0.75$, $\lambda = 0$.

Substituting u from (27) in Equation (36) it is obtain the equation that z obeys at the maximum of the ecological function, namely,

$$\left[2(1+\lambda \ln z)z - \lambda(1+\varepsilon-2z)\right](z-\varepsilon) = (1+\varepsilon-2z)(1+\lambda \ln z)\varepsilon\,. \tag{37}$$

If we suppose $z \equiv z(\lambda)$ given by the power expansion,

$$z_P \equiv b_0 + b_1\lambda + b_2\lambda^2 + b_3\lambda^3 + \ldots\,, \tag{38}$$

we find, upon taking the implicit successive derivatives of z_P with respect to λ in Equation (37) and equating them with the coefficients b_i in Equation (38),

$$\begin{aligned} z_E = \sqrt{\tfrac{1}{2}(\varepsilon+\varepsilon^2)}\left\{1+\left[\tfrac{1}{4}(1+3\varepsilon)\sqrt{\tfrac{2}{\varepsilon+\varepsilon^2}}-1\right]\lambda+\left(\frac{\frac{1}{16}(1+3\varepsilon)}{\varepsilon+\varepsilon^2}-\frac{1}{2}\sqrt{\frac{2}{\varepsilon+\varepsilon^2}}\ln\sqrt{\tfrac{1}{2}(\varepsilon+\varepsilon^2)}\right)\times\right. \\ \left.\left(1+3\varepsilon-4\sqrt{\tfrac{1}{2}(\varepsilon+\varepsilon^2)}\right)\lambda^2+O(\lambda^3)\right\} \end{aligned} \tag{39}$$

Furthermore, using (39), we can write the efficiency as a power series in λ,

$$\eta_E \equiv 1 - z_E(\varepsilon,\lambda) \tag{40}$$

In the particular case when $\lambda = 0$ we find the value

$$z_{EO}(\varepsilon,\lambda=0) = \sqrt{\frac{(\varepsilon+\varepsilon^2)}{2}}\,, \tag{41}$$

and the corresponding value for the ecological efficiency with instantaneous adiabats is as:

$$\eta_{EO} = 1 - z_{EO}(\varepsilon,\lambda=0) = 1 - \sqrt{\tfrac{1}{2}(\varepsilon+\varepsilon^2)}\,, \tag{42}$$

which is the maximum possible one, since all the terms in Eq.(39) are positive.

2.3 The linear approximation

As we can see in Equations (31) and (39), it can be taken a linear approximation for the efficiency η in terms of compression ratio, namely $V_{\max}/V_{\min}$, and of the ratio T_C/T_H, obtaining an expression like $F(\eta, V_{\max}/V_{\min}, T_C/T_H)=0$, with the same form regardless it was obtained by maximization of power output or maximization of ecological function. It permits analyze the behavior of compression ratio in respect to T_C/T_H. It can verified that $r_C \to \infty$ and $\lambda \to 0$ lead to the Curzon-Ahlborn-Novikov efficiency, now written as $\eta_{CAN} \equiv \eta_P(\lambda=0) = \eta_{PO}$. From (31) the linear approximation can be obtained,

$$\eta_{PL}(\lambda) = 1-\sqrt{\varepsilon} - \tfrac{1}{2}(1-\sqrt{\varepsilon})^2\lambda\,, \tag{43}$$

and the corresponding linear approximation of ecological efficiency is as,

$$\eta_{EL}(\lambda) = 1-\sqrt{\tfrac{1}{2}(\varepsilon^2+\varepsilon)} - \left[\tfrac{1}{4}(1+3\varepsilon) - \sqrt{\tfrac{1}{2}(\varepsilon^2+\varepsilon)}\right]\lambda\,. \tag{44}$$

As can be seen, the linear approximation of efficiency, maximizing power output or ecological function, has the form,

$$\eta_{JL}(\varepsilon,\lambda) = \eta_{JO} - b_J(\varepsilon)\lambda = \eta_{JO} - \frac{b_J(\varepsilon)}{(\gamma-1)\ln r_C}, \tag{45}$$

where b_J is de coefficient of linear term in λ, being $\lambda = \left[(\gamma-1)\ln r_C\right]^{-1}$, and the subscript J is substituting by *P* or *E*, for each of cases: maximization of power output or maximization of ecological function. That is, for maximum power output we have η_{PL}, η_P and b_P; and for maximum ecological function we have η_{EL}, η_E and b_E. So, for a particular value of efficiency we have $r_C = r_C(\varepsilon)$. The general expression of $r_C(\varepsilon)$, from (45), is obtain as,

$$r_C = \exp\left\{\frac{b_J}{(\gamma-1)(\eta_{JO}-\eta_{JL})}\right\}. \tag{46}$$

Taking η_{JO} as Curzon and Ahlborn-Novikov efficiency or ecological efficiency, it is true,

$$0 < \eta_{JL} < \eta_{JO}\,. \tag{47}$$

A particular value of efficiency η_{JL} permits find the interval $0<\varepsilon<1$ in which r_C satisfies,

$$r_C > 1\,, \tag{48}$$

and Equation (48) permits find, from (46),

$$\frac{b_J}{(\gamma-1)(\eta_{JO}-\eta_{JL})} > 0\,, \tag{49}$$

which leads to inequality

$$` \eta_{JL} < \eta_{IO} , \tag{50}$$

as a necessary condition because η_{JO} must be an upper bound for η_{JL}. For monoatomic gases $\gamma = 1.67$, with η_{JL} as a variable parameter, the variation of r_C can be obtained, and we can see that $r_C \to \infty$ when $\eta_{JL} \to \eta_{JO}$, as it should be. For each temperatures in Table 1 the variation of r_C is obtained from (46). By example in the West Turrock plant, $T_H = 838\ K$ and $T_C = 298\ K$, with $\eta_{PO} = \eta_{CA} = 0.403367$. Figure 4 shows the behavior of r_C respect to η_{PL}. Using the ecological function for the same plant, $\eta_{EO} = 0.50905$, and Figure 5 shows the behavior of r_C respect to η_{EL}. There is a minimum value of compression ratio greater than 1. On other hand for a particular value of r_C and for values of the used parameters in Figures 4 and 5, the behavior of η_{JL} can be considered as function of $\varepsilon \equiv T_C / T_H$, where we can be see the correctness of (50), so $\eta_{JL} \to \eta_{JO}$ only when $\varepsilon \to 1$, as it is shown in Figure 6, for the ecological function with $r_C = 10$, closer to compression values found in thermodynamics textbooks (among others Burghardt, 1982). In addition, the values of efficiency obtained naturally with the linear approximation are closer to real values than the corresponding values of η_{CA}, and η_E. The physically possible values of r_C take places when the values of λ that comply $0 < \lambda < 1$.

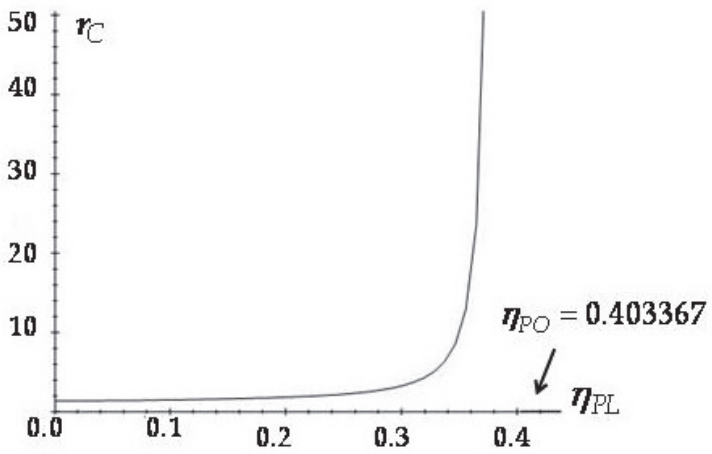


Fig. 4. Behavior of r_C in respect to variation of η_{PL} in the interval $[0, 0.403367)$.

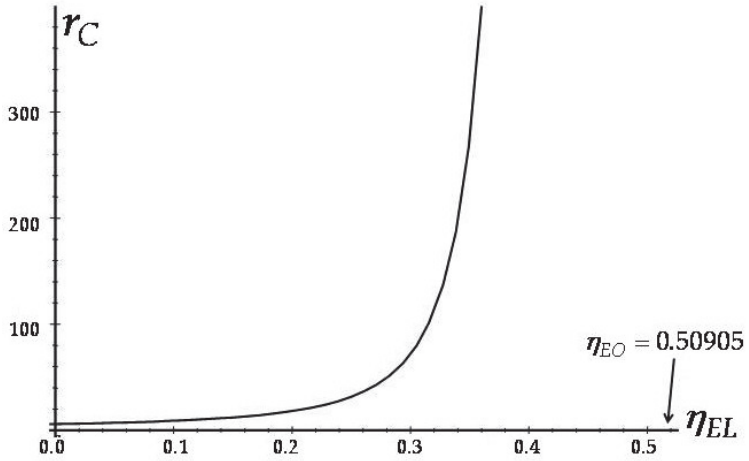


Fig. 5. Behavior of r_C in respect to variation of η_{EL} in the interval $[0, 0.50905)$.

In real compressors, named alternative compressors with dead space, percent of volume in the total displacement of a piston into a cylinder is named dead space ratio, defined as $c = (volume\ of\ dead\ space) / (volume\ of\ displacement)$, (Burghardt, 1982). In case of a Curzon and Ahlborn cycle $c = (\min imum\ volume) / (\max imum\ volume)$ is the reciprocal of r_C. Experimentally it is found that $3\% \le c \le 10\%$, so $100/3 \ge r_C \ge 100/10$, or $33 > r_C \ge 10$. Compression ratio is a useful parameter to model the behavior of a thermal engines, but it is not easy to include this parameter in design of power plants, would be interesting find a model in which r_C could be explicitly incorporated in design power plants.

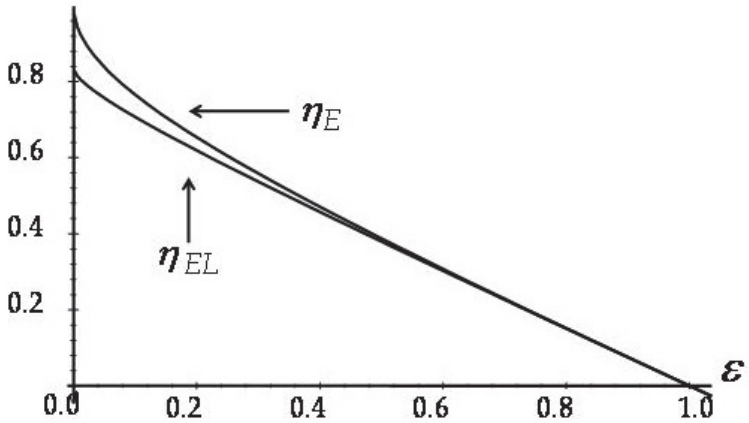


Fig. 6. Comparison between ecological efficiency at zero order and at linear order.

Supposing these plants are working as a Curzon and Ahlborn cycle, we found that linear approximation of efficiency, Equation (45), permits us to find intervals of efficiency values near to experimental values of efficiency than others. Table 2 shows a comparison between real values and linear approximation values, assuming ideal gas as working fluid ($\gamma = 1.67$), making clear the need for a closer approximation, nevertheless table shows the closeness of the linear approximation.

Nuclear power plant	$T_C\ (K)$	$T_H\ (K)$	η_{obs}	η_{EL} , $10 \le r_C < 33$
Doel 4 (Belgium),	283	566	0.35000	0.37944 to 0.38224
Almaraz II (Spain)	290	600	0.34500	0.39234 to 0.39539
Sizewell B, (U K)	288	581	0.36300	0.38277 to 0.38563
Cofrentes (Spain)	289	562	0.34000	0.36844 to 0.37103
Heysham (U K)	288	727	0.40000	0.46036 to 0.46506

Table 2. Comparison of values of experimental efficiencies and values of linear ecological approximation.

3. Non-instantaneous adiabats with Dulong-Petit's heat transfer law

The ecological efficiency has also been calculated using Dulong and Petit's heat transfer law (Angulo-Brown & Páez-Hernández, 1993; Arias-Hernández & Angulo-Brown, 1994), maximizing ecological function. Their numerical results have shown that the efficiency value changes with the heat transfer law one assumes. Velasco et al. (2000) studied both the power and the ecological function optimizations, using Newton's heat transfer law. It is worthwhile to point out that in all of the above quoted calculations the time for the adiabatic processes is not taken into account explicitly.

In the present section the power output P, and ecological function E are chosen to be maximized. Use is made of the more general Dulong and Petit heat transfer law, and the time for all the processes of the Curzon and Ahlborn cycle is explicitly taken into account, to see if the construction of a function $\eta_{PDP} = \eta_{PDP}(\lambda, \varepsilon)$ and $\eta_{EDP} = \eta_{EDP}(\lambda, \varepsilon)$ is possible. That the time for the adiabats can in principle to be an arbitrarily chosen function of the time of the isothems, here however it is chosen in the same way as in previous section for the purposes of comparison. The Dulong and Petit law has been chosen because the main ocurring heat transfers are conduction through the wall separating the working fluid from the thermal bath, and convection takes place within the working fluid. Radiative heat transfer is of smaller magnitude (O'Sullivan, 1990). With the optimization of the power

output of a Curzon and Ahlborn engine, it is shown an approximate expression for efficiency by means of also the Dulong and Petit's heat transfer law, and the corresponding zero order term in a power series of the parameter λ above cited. We follow the procedure employed in the previous section.

3.1 The power output efficiency

Let us assume a gas in a cylinder with a piston as a working fluid that exchanges heat with the reservoirs like in previous section, and let us use a heat transfer law of the form:

$$\frac{dQ}{dt} = \alpha (T_f - T_i)^k \tag{51}$$

where $k > 1$, α is the thermal conductance which is assumed the same for both reservoirs, dQ/dt is the rate of heat Q exchange and T_i and T_f are the temperatures for the heat exchange process considered. From the first law of thermodynamics applied to gas under mechanical equilibrium condition, i.e., $p = p_{ext}$, we obtain

$$\frac{dU}{dt} = \frac{dQ}{dt} - p\frac{dV}{dt} \ , \tag{52}$$

and assuming an ideal gas as working substance $U = U(T)$. One has in case of isothermal processes $\frac{dU}{dt} = \frac{dU}{dV}\frac{dV}{dt} = 0$. Using Eq. (51) we obtain, for the isothermal processes that

$$\frac{dQ}{dt} = p\frac{dV}{dt} \quad \text{or} \quad \alpha (T_f - T_i)^k = \frac{RT_i}{V}\frac{dV}{dt} \ . \tag{53}$$

Equation (53) implies that the time of the process along the first isothermal process is

$$t_1 = \frac{RT_{HW}}{\alpha (T_H - T_{HW})^k} \ln \frac{V_2}{V_1} \tag{54}$$

and analogously, the time along the second isothermal process is

$$t_3 = \frac{RT_{CW}}{\alpha (T_{CW} - T_C)^k} \ln \frac{V_3}{V_4} \ , \tag{55}$$

The corresponding heat exchanged Q_H and Q_C become, respectively,

$$Q_H = RT_{HW} \ln \frac{V_2}{V_1} \ , \quad Q_C = RT_{CW} \ln \frac{V_4}{V_3} \ , \tag{56}$$

where, R is the general gas constant and V_1, V_2, V_3, V_4, are the corresponding volumes for the states 1,2,3,4 in Figure 1.

While it is true that the speed for the adiabatic branches is independent from the speed of the isothermal ones in the cycle, but with a non null value, in order to obtain a more realistic result it will be assumed that their speed follows a similar law to the isothermal ones.

The previous assumption means that the rate of change of volume in the first adiabat is the same that in the first isotherm. Under this assumption, the time along the adiabatic processes is respectively,

$$t_2 = \frac{RT_{HW}}{\alpha(T_H - T_{HW})^k(\gamma-1)}\ln\frac{T_{HW}}{T_{CW}},\ \text{and}\ t_4 = \frac{-RT_{CW}}{\alpha(T_{CW} - T_C)^k(\gamma-1)}\ln\frac{T_{CW}}{T_{HW}}. \tag{57}$$

where $\gamma \equiv C_P / C_V$ has been used. With these results we can now compute the form for the power output, given by

$$P = \frac{W}{t_{tot}} = \frac{Q_1 + Q_2}{t_{tot}}, \tag{58}$$

where $t_{tot} = t_1 + t_2 + t_3 + t_4$. Power output is written as

$$t_{TOT} = \frac{R}{\alpha}\left[\frac{T_{HW}}{(T_H - T_{HW})^k} + \frac{T_{CW}}{(T_{CW} - T_C)^k}\right]\ln\frac{V_3}{V_1}, \tag{59}$$

by using $V_2 / V_1 = V_3 / V_4$ and $V_2 = V_3\left(T_{CW} / T_{HW}\right)^{\frac{1}{\gamma-1}}$; so that after making the exchange of variables as in Equation (26), P becomes,

$$P = \frac{T_1^k\alpha(1-z)(1+\lambda\ln z)}{\frac{1}{(1-u)^k} + \frac{z}{(zu-\varepsilon)^k}}, \tag{60}$$

with the same parameters as in previous section. By means of $\partial P / \partial u = 0$ and $\partial P / \partial z = 0$ we obtain,

$$u = \frac{z^{\frac{2}{k+1}} + \varepsilon}{z + z^{\frac{2}{k+1}}} \tag{61}$$

and,

$$\left[-z(1+\lambda\ln z)(zu-\varepsilon) + \lambda(1-z)(zu-\varepsilon) + zku(1-z)(1+\lambda\ln z)\right]\cdot\left((zu-\varepsilon)^k + z(1-u)^k\right)$$

$$-z(1-z)(1+\lambda\ln z)(zu-\varepsilon)\cdot\left(uk(zu-\varepsilon)^{k-1} + (1-u)^k\right) = 0 \tag{62}$$

Substituting the variable u in Equation (62) with the help of Eq. (61), the resulting expression is the following one, which shows the implicit function $z = z(\lambda, \varepsilon)$, for a given k,

$$\left[z^{\frac{2}{k+1}}(z-\varepsilon)\left(\lambda(1-z) - z(1+\lambda\ln z)\right) + zk(z^{\frac{2}{k+1}} + \varepsilon)(1-z)(1+\lambda\ln z)\right](z^{\frac{2k}{k+1}} + z)$$

$$-z(1-z)(1+\lambda\ln z)\left[z^2 + \varepsilon z^{\frac{2k}{k+1}} + z^{\frac{2}{k+1}}(z-\varepsilon)\right] = 0\,. \tag{63}$$

Because the solution of Eq. (63) is not analytically feasible when k is not an integer, the case discussed here is $k = 5/4$, the Dulong and Petit's heat transfer law. So one can take the reasonable approximations only for the exponents in Equation (63),

$$\frac{2}{k+1} \approx 1\ \text{and}\ \frac{2k}{k+1} \approx 1 \tag{64}$$

obtaining

$$(1+\lambda)((k\varepsilon+zk)(1-z)-z(z-\varepsilon))+\lambda(1-z)(1-\varepsilon)-(1+\lambda\ln z)(1-z)z=0\,. \tag{65}$$

Equation (65) allows to obtain the explicit expression for the function $z=z(\varepsilon,k)$ for $\lambda=0$,

$$z_{OP}(\varepsilon,k)=\frac{(k-1)(1-\varepsilon)\pm\sqrt{(\varepsilon-1)^2(1-k)^2+4k^2\varepsilon}}{2k}\,. \tag{66}$$

Taking now $k=5/4$ in Equation (66) we obtain the following value for the physically acceptable solution of (63), namely,

$$z_{OPDP}=\frac{1-\varepsilon+\sqrt{\varepsilon^2+98\varepsilon+1}}{10}\,. \tag{67}$$

The numerical results for $\eta_{OPDP}=1-z_{OPDP}$ are shown in Table 3, compared with η_{CAN} and the observed efficiency η_{obs}, where can be seen that are in good agreement with the reported values. Figure 7 shows the comparison between η_{OPDP} and η_{CAN} with the temperatures of the reservoirs in real plants (Angulo-Brown & Páez-Hernández, 1993; Velasco et al., 2000).

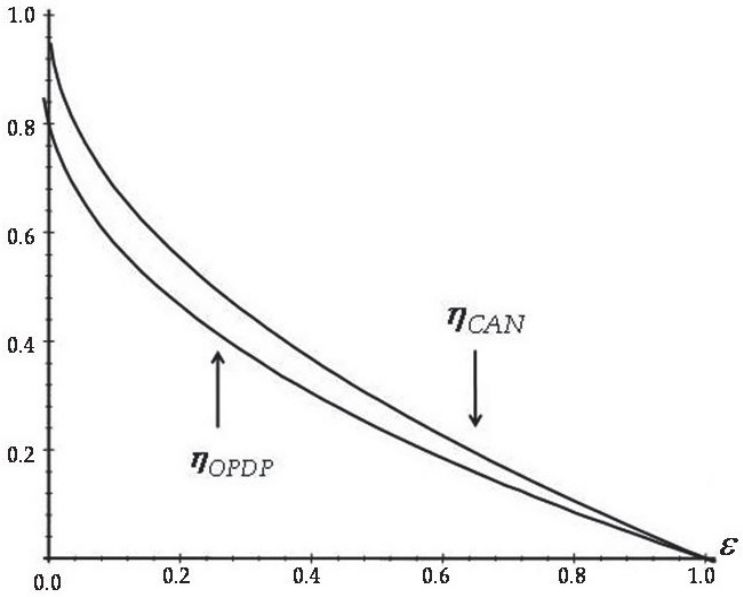


Fig. 7. Comparison between η_{OPDP} obtained here and η_{CAN}, in real plants.

Power plant	T_C	T_H	η_{CAN}	η_{OPDP}	η_{OBS}
Steam power plant, West Thurrock , U K	298	838	0.40367	0.33577	0.360
Geothermal steam plant, Lardarello, Italy	353	523	0.17845	0.1453	0.160
Steam power plant, U S A	298	923	0.43179	0.36006	0.400
Combined cycle plant (steam-mercury), U S A	298	783	0.38308	0.31804	0.340

Table 3. Comparison of Curzon and Ahlborn and observed efficiencies with the here approximated obtained efficiency.

Now assuming that z obtained from equation (65) can be expressed as a power series in the parameter λ, we have the following expression for η_{PDP},

$$\eta_{PDP} = 1 - z_{PDP}(\lambda, \varepsilon) = 1 - z_{OPDP}[1 + B_1(\varepsilon)\lambda + B_2(\varepsilon)\lambda^2 + O(\lambda^3)]. \tag{68}$$

We can find the coefficients B_j, $j = 1, 2, ...etc.,$ through successive derivatives respect to λ. The two first ones coefficients are:

$$B_1(\varepsilon) = \frac{16(1 - z_{OPDP})(\varepsilon - z_{OPDP})}{z_{OPDP}(5 - 4\varepsilon - 40 z_{OPDP})} \tag{69}$$

and

$$B_2 = \frac{4(z_{OPDP} - 1)(z_{OPDP} - \varepsilon)}{(1 + 9\varepsilon - 10 z_{OPDP})^2} \left\{ \frac{[(1 - \varepsilon + 10 z_{OPDP}) \ln z_{OPDP} + 8 z_{OPDP} - 4\varepsilon - 4](\varepsilon + 1 - 10 z_{OPDP})}{1 + 9\varepsilon - 10 z_{OPDP}} \times \right.$$

$$\left. \frac{40(z_{OPDP} - 1)(z_{OPDP} - \varepsilon)}{1 + 9\varepsilon - 10 z_{OPDP}} - \left[(9\varepsilon - 1 - 10 z_{OPDP}) \ln z_{OPDP} + 4 + 4\varepsilon - 8 z_{OPDP}\right] \right\} \tag{70}$$

which are positive for ε values in the interval $0 < \varepsilon < 1$, as we can see in Figure 8.

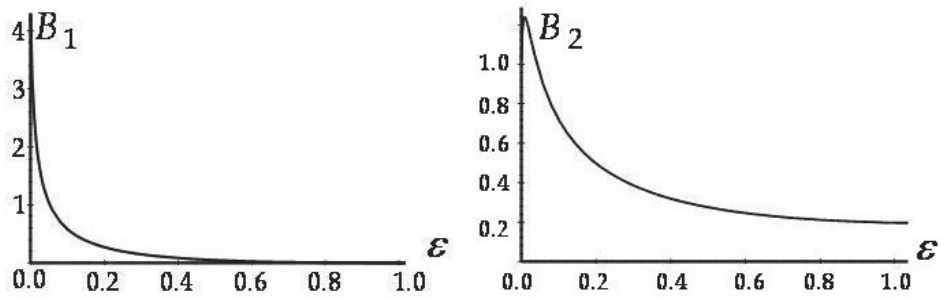


Fig. 8. First and second order coefficients, $B_1 = B_1(\varepsilon)$, $B_2 = B_2(\varepsilon)$, of (2.21) for $0 < \varepsilon < 1$.

3.2 Ecological efficiency

Now we consider the entropy production given by

$$\sigma \equiv \frac{\Delta S}{t_{tot}} = \frac{1}{t_{tot}}\left(-\frac{Q_H}{T_H} + \frac{Q_C}{T_C}\right) \tag{71}$$

which becomes,

$$\sigma = \frac{\alpha T_H^k (1+\lambda \ln z)(z-\varepsilon)}{T_C\left(\frac{1}{(1-u)^k} + \frac{z}{(zu-\varepsilon)^k}\right)}; \tag{72}$$

so that the ecological function for Curzon and Ahlborn engine takes the form,

$$E(x,z) = \frac{T_H^k \alpha (1+\lambda \ln z)(1+\varepsilon-2z)}{\frac{1}{(1-u)^k} + \frac{z}{(zu-\varepsilon)^k}}. \tag{73}$$

As in the previous sections we find the function $z(\varepsilon)$ that follows from the maximization of function $E(u,z)$, which permits obtain the corresponding efficiency for the value $k = 5/4$, namely, the Dulong and Petit heat transfer law, previously defined. Upon setting $\partial E / \partial u = 0$ and $\partial E / \partial z = 0$, we obtain from the first condition that

$$u = \frac{z^{\frac{2}{k+1}} + \varepsilon}{z + z^{\frac{2}{k+1}}}, \tag{74}$$

and from the second one,

$$\frac{((1+\varepsilon-2z)\lambda - 2z(1+\lambda \ln z))(zu-\varepsilon)}{(1+\lambda \ln z)(1+\varepsilon-2z)(zu-\varepsilon-kuz)z} - \frac{(1-u)^k}{(zu-\varepsilon)^k + z(1-u)^k} = 0. \tag{75}$$

Substituting now Equation (74) for u in Equation (75) we are led to the following expression,

$$(-2(1+\lambda \ln z)z + (1+\varepsilon-2z)\lambda)(z^2 + z^{\frac{k+3}{k+1}})(z-\varepsilon) = z(1+\lambda \ln z)(1+\varepsilon-2z)(z^{\frac{k+3}{k+1}} - \varepsilon z^{\frac{2}{k+1}} - (z^{\frac{k+3}{k+1}} + z\varepsilon)k). \tag{76}$$

The analytical solution of Eq. (76) is not feasible when the exponents of z are not integers, which is the present case, because with $k = 5/4$, Dulong and Petit's heat transfer law, one has $(k+3)/(k+1) = 17/9$ and $2/(k+1) = 8/9$.

The numerical solution of Eq. (76) shows that any solution falls into the region bounded by solutions for $\lambda = 0$ and $\lambda = 1$, (Ladino-Luna, 2008). It can be appreciated that within the interval $0 \le \varepsilon \le 1$, which is the only one physically relevant, the curve (76) can be fitted with a parabolic curve. The simplest approximation that allows for a parabolic fit for $0 \le \lambda \le 1$ is the following modification of the exponents:

$$\frac{k+3}{k+1} \sim 2, \quad \frac{2}{k+1} \sim 1. \tag{77}$$

These approximations allow the following approximate analytical expression for $z(\varepsilon,\lambda)$

$$2(-2(1+\lambda\ln z)z+(1+\varepsilon-2z)\lambda)(z-\varepsilon)-(1+\lambda\ln z)(1+\varepsilon-2z)((z-\varepsilon)-(z+\varepsilon)k)=0. \quad (78)$$

For the case $\lambda=0$, that corresponds to instantaneous adiabats, and taking $k=5/4$ in Equation (78), the value of the positive root $z_{OEDP}(\varepsilon)$ is obtained,

$$z_{OEDP}=\frac{1-\varepsilon+\sqrt{649\varepsilon^2+646\varepsilon+1}}{36}. \quad (79)$$

The negative root has no physical meaning because efficiencies must always be positive. Figure 9 shows a comparison between fitted numerical values of η_{MEDP} (Angulo-Brown & Páez-Hernández, 1993; Árias-Hernández & Angulo-brown, 1994) and $\eta_{OEDP}=1-z_{ODP}$.

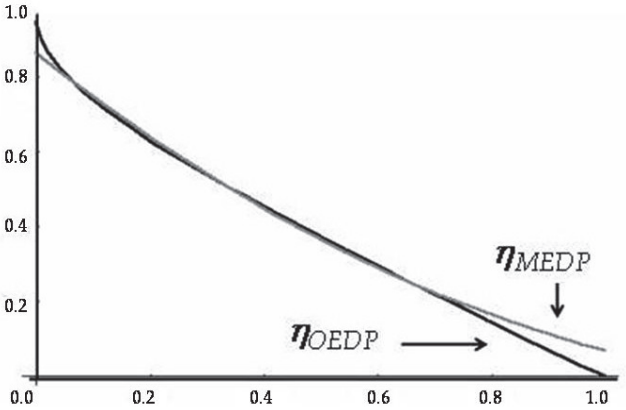


Fig. 9. Approximated Ecological efficiency η_{OEDP}, compared with a fitted of η_{MEDP}.

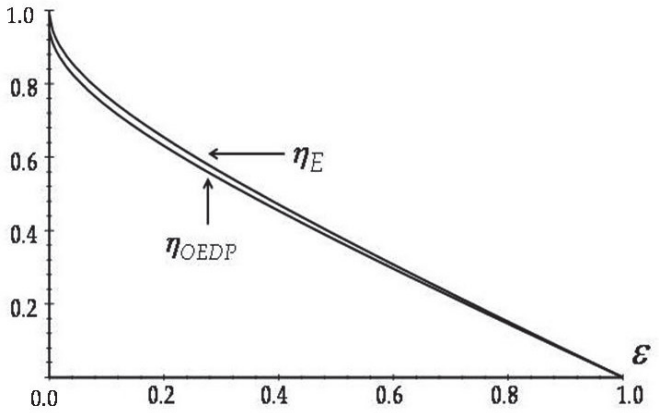


Fig. 10. Ecological efficiency for Newton's heat transfer and Dulong-Petit's heat transfer.

Notice that η_{OEDP} is a better result from a theoretical point of view, because it goes to zero as $\varepsilon \to 1$ as it should be. Figure 10 shows the comparison between η_E and η_{OEDP} where $\eta_{OEDP} < \eta_E$.

Let it be assumed now that z given by equation (78) is a power series in the parameter λ, i.e.,

$$\eta_{EDP} = 1 - z_{EDP}(\lambda, \varepsilon) = 1 - z_{EDP}(1 + b_1(\varepsilon)\lambda + b_2(\varepsilon)\lambda^2 + O(\lambda^3)), \tag{80}$$

and let us proceed to the calculation of the coefficients of the powers in λ. To this end one takes $z_0 = z_{EDP}(\varepsilon, \lambda = 0)$ and from (78) the coefficients are calculated by successively taking the derivative with respect to λ and evaluating at $\lambda = 0$. The first two are:

$$b_1(\varepsilon) = \frac{-2z_0 + 2\varepsilon - 6z_0\varepsilon + 2\varepsilon^2 + 4z_0^2}{z_0(-9z_0 - \frac{1}{4}\varepsilon + \frac{1}{4})} \tag{81}$$

and

$$b_2(\varepsilon) = \frac{1}{2z_0}(A_1(\varepsilon) + A_2(\varepsilon)), \tag{82}$$

where

$$A_1(\varepsilon) = \frac{4b_1}{(1 - \varepsilon - 20z_0)^2}\{-160(z_0 - \varepsilon)(1 + \varepsilon - 2z_0) + [(-36z_0 + 9\varepsilon + 1)\ln z_0 - 18z_0$$

$$+9\varepsilon + 1 + \frac{9(\varepsilon^2 + \varepsilon)}{z_0} + 8(1 + \varepsilon - 2z_0) - 16(z_0 - \varepsilon)](-1 + \eta + 20z_0)\}, \tag{83}$$

and

$$A_2(\varepsilon) = -\frac{8(z_0 - \varepsilon)(1 + \varepsilon - 2z_0)}{(1 - \varepsilon - 20z_0)^2}\left[(36z_0 + \varepsilon - 1)\ln z_0 + 48z_0 - 40\varepsilon - 8 + \frac{1 + \varepsilon - 2z_0}{z_0}(z_0 + 9\varepsilon)\right] \tag{84}$$

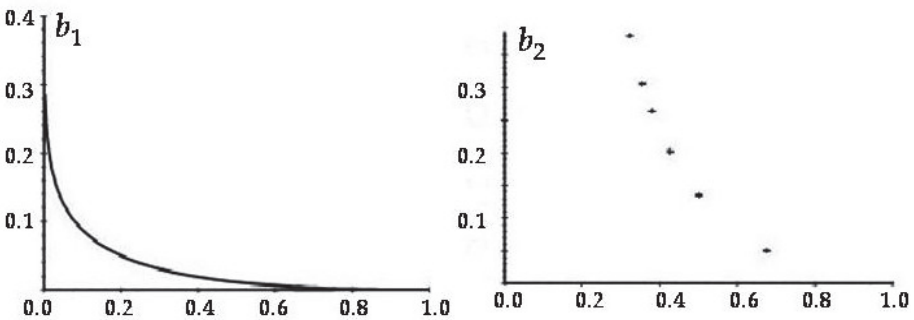


Fig. 11. First and second order coefficients $b_1 = b_1(\varepsilon)$ and $b_2 = b_2(\varepsilon)$ in (80)

To assume that (80) is valid it requires that $\eta_{EDP} \leq 1$. For this to be so, b_1 and b_2 must to be positive when $0 < \varepsilon < 1$ and when $\lambda < 1$, i.e. there must exist an interval for ε into which the coefficients have positive values, near to zero. Figure 11 shows that in fact this is correct. This guarantees that Eq. (80) is valid and that $1 - z_0$ is an upper bound of η_{DP}, but not the upper bound η_C.

4. The van der Waals gas

The internal energy in the case of a van der Waals gas for n moles, with a change of temperature $\Delta T = T - T_0$, at volume V, and with the characteristic constant a of the system, and the constant heat capacity C can be written as,

$$U = nC(T - T_0) - \frac{an^2}{V}. \tag{85}$$

So that taking the temporary derivative for an adiabatic process,

$$\frac{dU}{dt} = \frac{an^2}{V^2} \cdot \frac{dV}{dt}, \tag{86}$$

the first law of thermodynamics leads to

$$\frac{dU}{dt} = \frac{dQ}{dt} - p_{ext}\frac{dV}{dt} - (p - p_{ext})\frac{dV}{dt}, \tag{87}$$

taking p as the internal pressure and p_{ext} as the pressure of surroundings. Combining Equations (86) and Eq. (87), in mechanical equilibrium, we obtain

$$\frac{an^2}{V^2}\frac{dV}{dt} = \frac{dQ}{dt} - p\frac{dV}{dt}, \tag{88}$$

so that, for a non linear heat transfer law, more general than Dulong and Petit heat transfer law, as

$$\frac{dQ}{dt} = \alpha(T - T_0)^k, \tag{89}$$

with the constant thermal conductance α, and the constant exponent k, $k > 1$, from Equation. (88), in an isothermal process,

$$\left(\frac{an^2}{V^2} + p\right)\frac{dV}{dt} = \alpha(T - T_0)^k. \tag{90}$$

On other hand, the state equation for a van der Waals gas, with a constant b characteristic of the system, which is a more realistic model for a real gas, takes the following expression, with constant parameters *a* and *b*,

$$p = \frac{nRT}{V - nb} - \frac{an^2}{V^2}, \tag{91}$$

whose derivative respect T at $p = cons\tan t$ leads to

$$0 = \frac{nR}{V - nb} + nRT\left(-\frac{1}{(V - nb)^2}\frac{\partial V}{\partial T} + \frac{2a}{V^2}\frac{\partial V}{\partial T} \right). \tag{92}$$

By taking $n = 1$, Equation (91) into Equation (90) leads to

$$\frac{RT_0}{V - b}\frac{dV}{dt} = \alpha (T - T_0)^k . \tag{93}$$

In the case of a Curzon and Ahlborn cycle (Figure 1), for the heat exchange between the engine and the reservoirs, Equation (93) leads to the time of the isothermal processes by taking its integration. Moreover in the case of adiabatic processes $dQ / dt = 0$, so that Equation (87) reduces to

$$\frac{dU}{dt} = -p\frac{dV}{dt}, \tag{94}$$

and one can obtain

$$C_V \ln T = -R\ln(V - b), \tag{95}$$

or as it is usually written,

$$T(V - b)^{\frac{R}{C_V}} = constant . \tag{96}$$

Also, the duration time of the adiabatic processes can be obtained by integration of (93). Therefore the duration time of all processes in the cycle can be obtained from Equation (94), and Equation (96) leads to the relation between temperatures of the engine and the changes of volume in the adiabatic transformation.

4.1 Power output and ecological function

Taking into account the difference of temperatures between the engine and its reservoirs (Figure 1), it can be written the time for all of the processes in the cycle from Equation (93) For the isothermal processes, Eq. (93) can be written as

$$\frac{RT_{HW}}{V - b}\cdot\frac{dV}{dt} = \alpha (T_H - T_{HW})^k , \quad \text{and} \quad \frac{RT_{CW}}{V - b}\cdot\frac{dV}{dt} = \alpha (T_{CW} - T_C)^k , \tag{97}$$

and by direct integration of Equationss (97) we obtain

$$t_1 = \frac{RT_{HW}}{\alpha (T_H - T_{HW})^k}\ln\frac{V_2 - b}{V_1 - b}, \quad \text{and} \quad t_3 = \frac{RT_{CW}}{\alpha (T_{CW} - T_C)^k}\ln\frac{V_4 - b}{V_3 - b}, \tag{98}$$

Analogously, the time for the adiabatic processes can be obtained as

$$t_2 = \frac{RT_{HW}}{\alpha (T_H - T_{HW})^k}\ln\frac{V_3 - b}{V_2 - b}, \quad \text{and} \quad t_4 = \frac{RT_{CW}}{\alpha (T_{CW} - T_C)^k}\ln\frac{V_1 - b}{V_4 - b}, \tag{99}$$

Now, taking into account Equation (96), for the first adiabatic process,

$$T_{HW}(V_2-b)^{\frac{R}{C_V}}=T_{CW}(V_3-b)^{\frac{R}{C_V}},\quad \text{or,}\quad \ln\frac{V_3-b}{V_2-b}=\frac{C_V}{R}\ln\frac{T_{HW}}{T_{CW}}, \tag{100}$$

and for the second adiabatic process,

$$T_{CW}(V_4-b)^{\frac{R}{C_V}}=T_{HW}(V_1-b)^{\frac{R}{C_V}},\quad \text{or,}\quad \ln\frac{V_1-b}{V_4-b}=\frac{C_V}{R}\ln\frac{T_{CW}}{T_{HW}}, \tag{101}$$

and the combination of Eqs. (100) and (101) allows to obtain the relation

$$\frac{V_3-b}{V_2-b}=\frac{V_4-b}{V_1-b},\quad \text{or,}\quad \frac{V_3-b}{V_4-b}=\frac{V_2-b}{V_1-b} \tag{102}$$

The power output P can be written simplyfing with the same used parameters as,

$$P=\frac{\alpha T_H^k(1-z)\left[\lambda_{VW}\ln z+1\right]}{\frac{1}{(1-u)^k}+\frac{z}{(zu-\varepsilon)^k}}, \tag{103}$$

where $\lambda_{VW}=\left[(\gamma-1)\ln\frac{V_3-b}{V_1-b}\right]^{-1}$. One can see that $b\to 0$ leads to $\lambda_{VW}\to\lambda$ in Equation (29), and one can see that $b\to 0$ and $k\to 1$ reduce (103) to expression of P, such as it was found previously (Ladino-Luna, 2002, 2005). An expression of power series in λ_{VW} leads to the efficiency that can be obtained following the procedure in those references.

In the case of ecological function it is necessary to build the entropy production σ, $\sigma=\frac{\Delta S}{t_{tot}}$, so that (33) can be written since

$$\Delta S=\Delta S_{1\to 2}+\Delta S_{3\to 4}, \tag{104}$$

where $\Delta S_{1\to 2}$ is the change of entropy in the first isothermal branch and $\Delta S_{3\to 4}$ is the change of entropy at the second isothermal branch. For heat reservoirs, $\Delta S=\frac{Q}{T}$, assumed as it is only in the transfer processes between the reservoirs and the engine,

$$\Delta S_{1\to 2}=\frac{Q_{1\to 2}}{T_H}=R\frac{T_{HW}}{T_H}\ln\frac{V_2-b}{V_1-b}\quad\text{and}\quad \Delta S_{3\to 4}=\frac{Q_{3\to 4}}{T_C}=R\frac{T_{CW}}{T_C}\ln\frac{V_4-b}{V_3-b}; \tag{105}$$

so that Eq. (104) can be written as

$$\Delta S=R\left[\frac{T_{HW}}{T_H}\ln\frac{V_2-b}{V_1-b}-\frac{T_{CW}}{T_C}\ln\frac{V_3-b}{V_4-b}\right], \tag{106}$$

and by using (102) and (96) one can obtain the entropy production as,

$$\sigma=\frac{\alpha T_H^k(\varepsilon-z)\left[\lambda_{VW}\ln z+1\right]}{\frac{1}{(1-u)^k}+\frac{z}{(zu-\varepsilon)^k}}, \tag{107}$$

then, by using (3), (98) and (99) the ecological function can be written as

$$E = \frac{\alpha T_1^k (1 - 2z + \varepsilon)\left[\lambda_{VW} \ln z + 1\right]}{\frac{1}{(1-u)^k} + \frac{z}{(zu-\varepsilon)^k}} . \tag{108}$$

One can see that the structure of Eq. (108) leads to the case with Newton's heat transfer law when the limit $k \to 1$ is. It is also obtained the case of Newton heat transfer with an ideal gas as the working substance when $k = 1$ and $b = 0$. A general form of ecological function and power output function can be obtained by replacing λ_{VW} instead of λ, and with approximations for the cases when $k > 1$. z_{EDP} and η_{EDP} (Ladino-Luna, 2008) are modified with the substitution $V - b$ instead of V.

The corresponding maximization of ecological function taking Dulong-Petit's heat transfer and a van der Waals gas as the working substance can be found with the substitution λ_{VW} instead of λ in all of the process to build the ecological efficiency. In the case of power output with the same substitution, we obtain the approximate formula for the efficiency when λ_{VW} goes to zero, and a similar power series of the efficiency as a function of λ_{VW},

$$\eta_{PDPVW} = 1 - z_{PDP}(\lambda_{VW}, \varepsilon) = 1 - z_{OPDP}(1 + b_1(\varepsilon)\lambda_{VW} + b_2(\varepsilon)\lambda_{VW}^2 + O(\lambda_{VW}^3)) , \tag{109}$$

wher z_{OPDP} is the same approximate efficiency previously found in section 3, following the procedure by Ladino-Luna (2003). At the limit $\lambda_{VW} \to 0$ we obtain $\eta_{PDPVW}(\lambda_{VW} = 0) = \eta_{OPDP}$ where η_{OPDP} is the same approximate efficiency found in Section 3. As one can see, $\eta_{PDPVW}(\lambda_{VW} = 0) < \eta_{CAN}$, so η_{CAN} can be consider as an upper bound for the efficiencies that taking into account the time of the adiabatic processes in the Curzon and Ahlborn cycle.

5. Conclusions

A first result is the fact that the efficiency for a Carnot type engine depends on the size of the engine, the compression ratio, as represented by the parameter $\lambda \sim \left[\ln(V_3 / V_1)\right]^{-1}$ or $\lambda_{VW} = \left[(\gamma - 1)\ln\frac{V_3 - b}{V_1 - b}\right]^{-1}$. Leading term in power series corresponds to the exact value numerically calculated without explicitly taking into account the dependence on λ, and is an upper bound for the value of the efficiency η_{DP}; in fact the larger the ratio V_3 / V_1 (or $(V_3 - b)/(V_1 - b)$), the larger the efficiency becomes. The comparison between the upper bound of the efficiency calculated with the proposed approximations and a fitted curve obtained of the numerical values from cited references shows the goodness of the made approximations in case of $k = 5/4$. It is worthwhile mentioning that exist an interval for ε, ε ~0.5, were the approximation employed is acceptable within 5% of the true value of $z(\varepsilon)$ for $0 \le \lambda \le 1$ as shown. A last result is shown in Figures where one can appreciate that the difference between using Newton's or Dulong-Petit's heat transfer laws does not lead to an important difference in the value of the ecological efficiency. It has also been shown that for the Dulong-Petit heat transfer law and the ideal gas law, the limit $\lambda \to 0$ reduces to the reported result. Also, the results suggest that can be extended a new interpretation as the way to real performance of the thermal plants. It shows a mixture between Newton and Dulong-Petit heat transfer laws. Non-endoreversible cycles could be analyzed using non-instantaneous adiabats together with non-linear heat transfer .

6. Acknowledgements

The authors would like to thank for the support of CONACYT (México) by the SNI program.

7. References

Agrawal D.C. & Menon, V.J. (1990). The Carnot cycle with the van der Waals equation of state, *Eur. J. Phys.*, Vol. 11, pp. 88-90

Agrawal, D.C.; Gordon, J.M. & Huleihil, M. (August 1994). Endoreversible engines with finite-time adiabats, *Indian J. Eng. Mat.Sc.*, Vol. 1, pp. 195-198

Amelkin, S.A.; Andresen, B.; Burzler, J.M.; Hoffman, K.H. & Tsirlin, A.M. (2004). Maximum power processes for multi-source endoreversible heat engines; *J. Phys. D: Appl. Phys.* Vol. 37, pp. 1400-1404

Amelkin, S.A.; Andresen, B.; Burzler, J.M.; Hoffman, K.H. & Tsirlin, A.M. (2005). Thermomechanical systems with several heat reservoirs: Maximum power processes; *J. Non-Equil. Thermod.* Vol. 30, pp. 67-80

Angulo-Brown, F. (june 1991). An ecological optimization criterion for finite-time heat engines, *J. Appl. Phys.*, Vol. 69, pp. 7465-7469

Angulo-Brown, F.; Ares de Parga, G. & Árias-Hernández, L.A. (2002). A Variational Approach to Ecological-type Optimization Criteria for Finite-time Thermal Engine Models. *J. Phys. D: Appl.Phys.*, Vol. 35, pp. 1089-1093

Angulo-Brown, F.; Árias-Hernández, L.A. & Páez-Hernández, R.T. (1999). A General Property of Non-endoreversible Thermal Cycles. *J. Phys. D: Appl. Phys.*, Vol. 32, pp. 1415-1420

Angulo-Brown, F. & Páez-Hernández, R. (1993). Endoreversible thermal cycle with a non linear heat transfer law, J. Appl. Phys., Vol. 74, pp. 2216-2219

Arias-Hernández, L. A. & Angulo-Brown, F., (1997). Thermodynamic optimization of endoreversible engines, *Rev. Mex. Fís.*., Vol. 40, pp. 866-877

Badescu, V. (March 2004). Optimal paths for minimizing lost available work during usual finite-time heat transfer processes, *J. Non-Equil. Thermod.* Vol. 29, pp. 53-73

Bejan, A. (1996). Entropy Generation Minimization: The New Thermodynamics of Finite-size Devices and Finite-time processes. *J. Appl. Phys.*., Vol. 79, pp. 1191-1218

Burghardt, M.D. (1982). *Engineering thermodynamics,* section 5.2, Harper and Row (Ed.) New York, Unites States of America

Curzon, F. L. & Ahlborn, B., (1975). Efficiency of a Carnot engine at maximum power output, *Am. J. Phys.*, Vol. 43, pp. 22-42.

Chambadal, P. (1957). *Récupération de chaleur 'a la sortie d'un reactor*, chapter 3, pages 39-58, Armand Colin (Ed.) Paris, France

Chen, J. (1994). The Maximum Power Output and Maximum Efficiency of an Irreversible Carnot Heat Engine, *J. Phys. D: Appl. Phys.*, Vol. 27, pp. 1144-1149

Chen, J. (1996). The Efficiency of an Irreversible Combined Cycle at Maximum Specific Power Output, *J. Phys. D: Appl. Phys.* Vol. 29, pp. 2818-2822

Chen, L.;. Zhang, W. & Sun, F. (2007). Power, efficiency, entropy generation rate and ecological optimization for a class of generalized universal heat engine cycles, *Applied Energy*, Vol. 84, pp. 512-525

Chen, L.; Zhou, J.; Sun, F. & Wu, C. (2004). Ecological optimization of generalized irreversible Carnot engines. *Applied Energy*, Vol. 77, pp. 327-338

Chen, L.; Zhu, X.; Sun, F. & Wu, C. (2006). Exergy-based ecological optimization of linear phenomenological heat transfer law irreversible Carnot engines, *Applied Energy*, Vol. 83, pp. 573-582

Gutkowics-Krusin, D., Procaccia, I. & Ross, J. (1978) On the efficiency of rate processes. Power and efficiency of heat engines, *J. Chem. Phys.,* Vol. 69, pp. 3898-3906.

Ibrahim, O.M.; Klein, S.A. & Mitchel, J.W. (1991). Optimum heat power cycles for specified boundary conditions, *Transactions of the ASME*., Vol. 113, pp. 514-521

Ladino-Luna, D. (2003). Efficiency of a Curzon and Ahlborn engine with Dulong-Petit heat transfer law, *Rev. Mex. Fís.*, Vol. 49, pp. 97-91

Ladino-Luna, D. (2005). Van der Waals gas as working substance in a Curzon and Ahlborn-Novikov engine, *Entropy*, Vol. 7, pp. 108-121

Ladino-Luna, D. (2008). Approximate ecological efficiency with a non-linear heat transfer. *J. of Energy Institute,* Vol. 81, pp. 114-117

Ladino-Luna, D. & de la Selva, S. M. T., (2000). The ecological efficiency of a thermal finite time engine, *Rev. Mex. Fís.*, Vol. 46, pp. 52-56

Novikov, I.I. (1958). The efficiency of atomic power stations. *Journal of Nuclear Energy II*, Vol. 7, pp. 125-128

O' Sullivan, C. T. (1990). Newton's law of a cooling-a critical assessment, *Am. J. Phys.*, Vol. 58, pp. 956-960

Rubin, M. (1979a). Optimal Configuration of a Class of Irreversible Heat Engines. I. *Phys. Rev. A.*, 19, pp. 1272-1276

Rubin, M. (1979b). Optimal Configuration of a Class of Irreversible Heat Engines. II. *Phys. Rev. A*. Vol. 19, pp. 1277-1288

Rubin, M. (1980). Optimal Configuration of an Irreversible Heat Engine with Fixed Compression Ratio, *Phys. Rev. A.* Vol. 22, pp. 1741-1752

Salamon, P.; Andresen, B. & Berry, R.S. (1976). Thermodynamics in Finite Time. I. Potencials for Finite-time processes. *Phys. Rev. A,* Vol. 15, pp. 2094-2101

Song, H.; Chen, L.; Sun, F. & Wu, C. (2006). Optimal configuration of a class of endoreversible heat engines with linear phenomenological heat transfer law [$q\Delta(T^{-1})$] *J. Appl. Phys*., Vol. 100, 124907; doi:10.1063/1.2400512 (13 pages)

Song, H.; Chen, L. & Sun, F. (2007). Endoreversible heat-engines for maximum power-output with fixed duration and radiative heat-transfer law, *Applied Energy*, Vol. 84, pp. 374-388

Velasco, S., Roco, J. M. M., Medina, A. & White, J. A. (2000). Optimization of heat engines including the saving of natural resources and the reduction of thermal pollution, J. Phys. D, Vol. 33 , pp. 355-359

Wu C. & Kiang, R.L. (1992). Finite-time thermodynamic analisys of a Carnot engine with internal irreversibility. *Energy, The Int. J.*, Vol. 17, pp. 1173-1178

Yan, Z. (1993). Comment on "ecological optimization criterion for finite time engines". *J, Appl. Phys.,* Vol. 73, pp. 3583

Zhu, X ¸Chen, L.; Sun, F. & Wu, C. (2003). The ecological optimization of generalized irreversible Carnot engine with generalized heat transfer law. *Int. J. Ambient Energy,* Vol. 24, pp. 189-194

7

Heterogeneous Melting in Low-Dimensional Systems and Accompanying Surface Effects

Dmitry G. Gromov and Sergey A. Gavrilov
Moscow Institute of Electronic Technology
Russia

1. Introduction

It is well known that the smaller the size of low-dimensional objects the lower melting temperatures they have. This effect is described in its various manifestations in numerous both old enough and comparatively recent publications.
However, for quite objective reasons, the mechanisms of the observed phenomena in low-dimensional systems are still discussed, as it remains unclear whether these phenomena can be called melting or not.
The evolution of different material thin films during heating has been repeatedly studied. Heating is known to result in thin-film disintegration into droplets (or, in other terms, beads), if a film has been deposited on an inert surface. However, there exists no unified understanding of this phenomenon; hence, there is no generally accepted terminology. It should be noted that in various research this phenomenon goes under different terms, e.g. breaking up, rupture, reflow, agglomeration, disintegration, dewetting, decay or beading of a thin film, melting, surface premelting, surface-induced melting.
The aim of this chapter is to analyze the specific features of heterogeneous melting and to show, on the basis of experimental research, that heterogeneous melting mechanism starts to play an important role as the system size decreases and initiates a number of processes that are frequently encountered in micro- and nanotechnologies. It should be noted that the reason for different points of view on this phenomenon is hidden in the peculiarities of the melting process.
Classical equilibrium thermodynamics strictly determines the temperature of the phase transition solid-liquid and vice versa. According to it, it is the temperature at which one can observe the equality of the Gibbs potential of solid and liquid phases.
In accordance with equilibrium thermodynamics, at constant pressure in a one-component system phase transition occurs at a specific temperature and it should be accompanied by a sudden change in heat release or absorption, i.e. phase transition has no extension in time or hysteresis. It should be noted, that although interphase border is an integral part of any system where phase transition takes place, classical equilibrium thermodynamics does not pay any attention to the possible contribution of this border to phase transitions.
While observing the phenomenon researchers often see a paradox. The melting process, accompanied by heat absorption and not a spontaneous one, and therefore requiring more energy for the system, occurs at the reference temperature. This temperature is defined as equilibrium melting temperature, and the melting process here does not require

overheating. At the same time the crystallization process, which is accompanied by heat release and is spontaneous and therefore does not require more energy for the system, occurs at the temperature lower than the reference one, i.e. this process needs overcooling. Thus, phase transition is asymmetric.

The cause of this paradox becomes clear only after taking into account the contribution of interphase border to phase transition. Here we can distinguish two extreme cases: homogeneous and heterogeneous. In a homogeneous case, phase transition is the transition from a monophase state, in which there is no surface, to a two-phase state, in which phases are separated from each other by the interphase surface. So, a homogeneous case requires energy to create phase interface. Heterogeneous case occurs when two-phase state exists meaning that phase interface exists as well, i.e. there is no need for applying energy.

In the real material world, any system is finite or at least semi-infinite, i.e. regardless of its size somewhere it comes to an end with a surface. As a surface is the main defect of a three-dimensional crystal lattice, its vibration spectrum differs from the bulk spectrum. Vibration amplitude of atoms located on the surface is always higher than that in the bulk. It is an established fact that vibration amplitude of atoms located close to the surface is 40-100% higher than atoms in the bulk. Consequently the Debye temperature of a surface layer is approximately half the Debye temperature characteristic of the bulk phase. This means that the crystal surface can melt at a temperature of half the melting temperature of the bulk material.

That is why melting always begins from the surface, the melt front moves inward the crystal. Thus, the melting process is heterogeneous, it does not require energy to create interface surface of solid and liquid phases, and the material completely melts at the equilibrium melting temperature indicated in reference books.

For the same reason of higher amplitude of near-surface atomic vibrations, crystallization cannot start from the liquid-phase surface. Because of this, unlike the melting process, the crystallization process is homogeneous and needs overcooling of the system by dozens or even hundreds degrees relative to the equilibrium temperature of the phase transition.

At the same time, as a solid-phase surface is artificially introduced, the heterogeneous crystallization, which almost does not require any overcooling of the liquid phase, also easily proceeds. The example of such process can be crystal growth as in the case of pulling out single crystals from a melt. Here to fulfill the phase transition overheating or overcooling should be only a few degrees; there is practically no hysteresis and the process is close to equilibrium.

One should not forget that along with the process of phase transition – melting-crystallization, the process of solidification-dilution also takes place and this process is not phase transition. The essence of this process consists in changing viscosity (or fluidity) of liquid phase depending on temperature. Being independent, this process has a great influence on the process of crystallization. The greatest influence of this process can be observed in creating solid amorphous state when because of higher viscosity with lower temperature, crystalline phase cannot occur.

2. Heterogeneous melting in semi-infinite systems

As it is stated earlier equilibrium temperature of phase transition from solid to liquid state T_∞ in an equilibrium thermodynamic system is the temperature when one can observe the equality of the Gibbs potential of solid and liquid phases (Fig. 1):

$$G_S = G_L \tag{1}$$

i.e.

$$H_S - T_\infty S_S = H_L - T_\infty S_L \tag{2}$$

From equation (2) the entropy change in transition from solid to liquid state, we have

$$\Delta S(T_\infty) = \frac{\Delta H(T_\infty)}{T_\infty} \tag{3}$$

where $\Delta H(T_\infty)$ is the change in the enthalpy of the system upon transition from a solid state to a liquid state at the corresponding temperature T_∞.

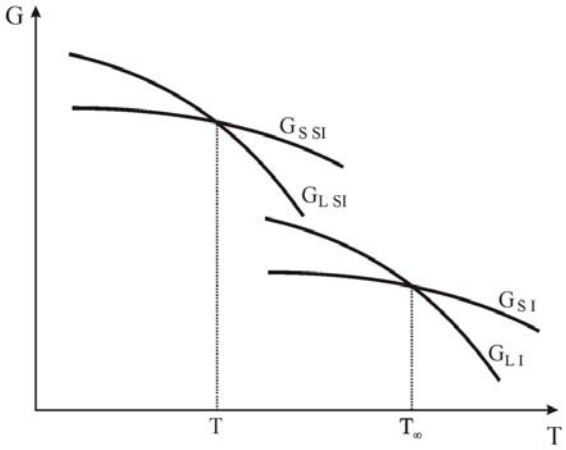


Fig. 1. The dependence of Gibbs energy from the temperature of solid and liquid phases in a single-component infinite (G_I) and semi-infinite (G_{SI}) system.

Equation (2) is correct for infinite systems where there is no contribution of energy of the surface to the total energy of the system. For the semi-infinite system having a limiting surface, under assumption that the volume remains unchanged after the phase transition, equality (2) transforms into the equation:

$$H_S - TS_S + \sigma_S \frac{A_S}{V} = H_L - TS_L + \sigma_L \frac{A_L}{V} \tag{4}$$

where T is temperature of the phase transition in semi-infinite system, σ is surface energy, A is surface area, V is volume. One should bear in mind that H and S must be on volume unit (J/m^3).

Additional components in equation (4) in comparison with equation (2) lead to the higher Gibbs potential and curve shift of free enthalpy changes of solid and liquid phases as in Fig.1. The result is that equilibrium point of phase shifts from T_∞ to lower temperatures T.

From transposition of the components in equation (4) we see

$$T\Delta S(T) = \Delta H(T) + \frac{\sigma_L A_L - \sigma_S A_S}{V} \tag{5}$$

Entropy is a thermodynamic function defining the degree of an order in system. The degree of an order changes in transition from ordered crystalline state to disordered liquid state. This leads us to the fact that although entropy depends on the system temperature, the change of entropy in the transition from crystalline state to liquid state does not depend much on temperature at which it occurs, i.e.

$$\Delta S(T) \approx \Delta S(T_\infty) \tag{6}$$

To obtain the dependence between melting temperatures T_∞ in an infinite system and T in a semi-infinite system, we substitute (6) for (5) considering (3)

$$T = T_\infty \left(\frac{\Delta H(T)}{\Delta H(T_\infty)} + \frac{\sigma_L A_L - \sigma_S A_S}{V \Delta H(T_\infty)} \right) \tag{7}$$

Equation (7) depicts the change of melting temperature due to contribution of the surface under constant pressure.
However, there may be an assumption that melting heat does not depend on the temperature at which the phase transition occurs:

$$\Delta H(T) \approx \Delta H(T_\infty) \tag{8}$$

Taking into account that under phase transition the area surface does not change A_L=A_S, assumption (8) and that V=Ah, we obtain a well-known equation that describes the decreasing temperature of thin film depending on its thickness h

$$T = T_\infty \left(1 + \frac{\Delta\sigma}{\Delta H(T_\infty)} \frac{1}{h} \right) \tag{9}$$

where $\Delta\sigma$ is the change in surface energy as a result of transition from the crystalline state to the liquid state.
From equation (9) we can derive the expression for the approximate evaluation of the equilibrium thickness of a molten surface layer as a function of the temperature

$$h = \frac{T_\infty \Delta\sigma}{(T - T_\infty)\Delta H(T_\infty)} \tag{10}$$

We calculated the dependence of the thickness of the molten layer on temperature for lead using the following initial data. For the free upper surface of a lead film we take the energies of the lead surface adjacent to vacuum σ_S = 0.56 J/m^2 and σ_L = 0.44 J/m^2. We assume that the surface layer undergoing the melting process is initially a thin lead film laying on the lead surface. Our comparison of the calculated and experimental data show that the boundary energy between solid phases or liquid and solid phases is best represented by the arithmetic mean of the surface energies of the corresponding materials. Because of this, we use σ_S = 0.56 J/m^2 also for the lead–lead solid phase boundary and σ_L = 0.5 J/m^2 for the liquid lead–solid lead boundary. The melting heat of lead (per unit volume) was $\Delta H(T_\infty) = 2.7\times10^8$ J/m^3.
Fig. 2 illustrates the estimated dependency of the thickness of molten surface layer on the temperature of lead. The kind of dependency in Fig. 2 coincides with experimental curve, which is the result of the medium-energy ion dissipation method that allows us to determine the depth of unregulated layer on the surface. As it is shown in Fig. 2, the essence

of this phenomenon is that at temperatures much lower than the temperature of melting, an unregulated liquid layer appears on the surface of semi-infinite crystalline body. The thickness of the layer rises fluently at first, and then it rises very abruptly as the temperature rises, leading to complete melting at the equilibrium temperature T_{∞}. B.Pluis and his colleagues named this phenomenon "surface-induced melting". As we see it, this term reflects the essence of the phenomenon in the best way. Another term for this phenomenon is "heterogeneous melting".

Three distinctive areas can be distinguished in the phenomenon of heterogeneous melting (Fig. 2). Area III, which is situated higher than the equilibrium temperature of melting T_{∞} is the area where the crystalline phase is completely molten and the liquid phase alone exists. Area I and Area II both are characterized by coexistence of crystalline and liquid phases of the matter. However, in Area I, where the thickness of the molten layer is lesser than 20 nm, the quantity of the matter in liquid state does not correspond to physicochemical concept of "phase". In this area, the liquid layer can be regarded rather as an adsorption layer that lowers the surface energy of semi-infinite crystalline phase. In this case, the liquid layer and the semi-infinite crystalline phase are a single system. In Area II the liquid layer on the surface becomes a phase and can be regarded as an independent system that coexists in contact with another system, i.e., the crystalline phase. Properly, it is Area II that can be rightfully regarded as the area of the surface-induced melting process.

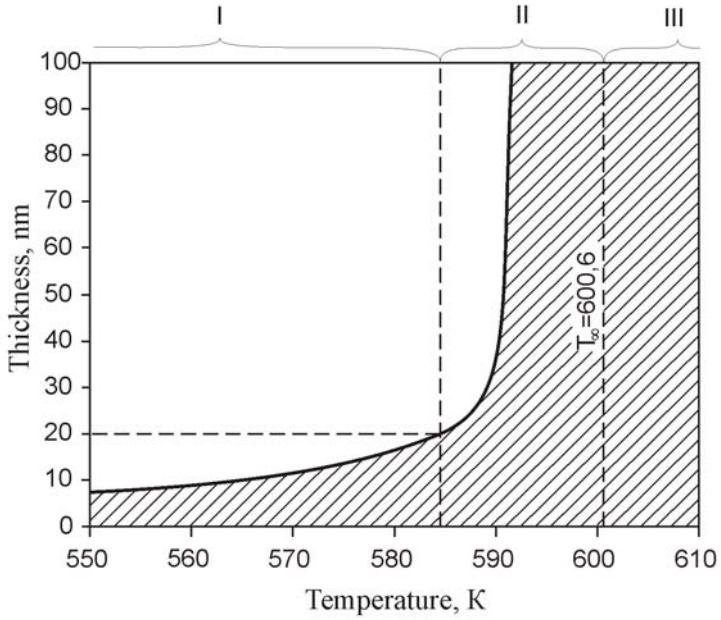


Fig. 2. Heterogeneous melting of lead: the estimated dependency of the thickness of the molten surface layer on temperature.

Thereby, the shape of the curve in Fig. 2 leads to two important conclusions:

- The existence of a surface in a semi-infinite system, and, as a consequence, the appearance of the phenomenon of melting which is caused by this surface, lead to the

fact that the complete melting of the semi-infinite phase occurs at the temperature of equilibrium phase conversion of an infinite system of the same kind of matter virtually in all cases. In other words, the absence of necessity to input energy to create a surface in a semi-infinite system always leads to heterogeneous melting, which requires no overheating comparative to equilibrium melting temperature of an infinite system;

- At the temperature that is lower than the reference temperature, on the surface of the solid phase there exists a layer of liquid phase of definite thickness, which is in equilibrium with the solid phase. The lower the temperatures, the thinner the layer of the liquid phase on the surface.

3. The dependence of melting temperature of nano-scale systems on size

The appearance of a liquid layer on the surface of crystalline semi-infinite objects or macro-objects at temperatures substantially lower than the reference temperature of melting has virtually no effect on the behavior of these systems. But considerable temperature range of the heterogeneous melting process leads to the fact that equilibrium temperature of low-scale systems can alter substantially as the size decreases.

For systems having the size (diameter of nano-crystal, thickness of thin film) of several dozens nanometers the equilibrium temperature of melting in accordance with (7) and (10) can fall by hundreds of degrees lower than the equilibrium melting temperature of volumetric matter T_∞.

Analyzing expression (7) one can come to conclusion that the value of melting temperature decline is defined by three causes. First and foremost of these causes is reduction of specific surface energy σ as a result of phase transition from solid to liquid state. The second cause is the possibility of the nano-object A surface area contraction as a result of phase transition from solid to liquid state. This cause is dependent on the first one, since the alteration of form in crystalline state is virtually impossible. But this possibility appears in liquid state, when such property as fluidity emerges. As a result, the main mechanism of mass transfer is no longer the diffusion resulting from heat-driven movement of atoms, but the movement of atoms by means of the force of surface-tension. The third cause that becomes noticeable if the first two causes are present is the reduction of the quantity of heat required for melting relating to the phase transition $\Delta H(T)$, if the melting process begins at a lower temperature due to the first two causes. In other words, at temperatures lower than T_∞, the system requires a lesser quantity of heat to change from solid state to liquid one.

Taking into account the aforesaid, all nano-scaled systems can be conventionally divided into two groups. To the first group belong systems that do not possess the possibility of changing their form. Consequently, the reduction of the temperature of melting for these systems is due to the first cause alone. To the second group belong systems that can change their forms. As a result, for these systems temperature falls due to the combined effect of all three causes.

To the first group belong systems in which the reduction of surface energy is possible only due to the transition from solid state to liquid state. This group includes nano-clusters and nano-crystals in free state that have the optimal form considering the ratio of surface area to volume. Furthermore, to this group also belong threadlike nano-crystals (nanowires) inside another solid phase and thin films, confined from above and below by thick layers of other solid materials. These systems do not have optimal forms and are not in free state, thereby they cannot change their form during melting.

To the second group belong systems that can change their form in the process of melting. An example of such process is the process of melting-dispersion of a film into drops. The energy expediency of this process is shown in Fig. 3, which illustrates the result of the calculations regarding the change of the surface area of a system consisting of a set of spherical objects depending on their radius. It is assumed that the volume of the system is constant, so if the radius of the spherical objects are changed, that leads to a corresponding change in their number. The character of dependence points to the fact that the increase of the size of the spherules leads to significant decrease of the surface area of the system. That way, if a system which contains $1{,}9\times10^9$ spherical objects with 10 nm radius each changes to the state of $1{,}5\times10^4$ objects with 500 nm radius each (given that V=const), the surface area, and, consequently, the absolute surface energy change 50-fold. For comparison, in the first group of systems the characteristic values of alteration of surface energy during the transition from solid to liquid state can amount to from 1.3 to 2-fold.

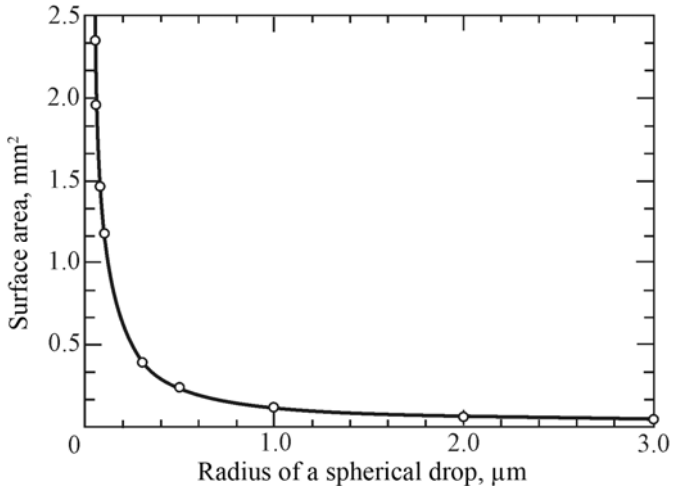


Fig. 3. Calculations regarding the change of the surface area of an array of spherical drops depending on their radius.

Along with the surface area alteration possibility, a significant effect can also be brought by noticeable differences in temperature dependencies of thermal capacities of solid and liquid phases. As it is illustrated in Fig. 4, where nickel is used as an example, ΔCp of phase transition is less if it occurs at some temperature T_i lower than T_∞. The consequence is temperature dependence ΔCp and $\Delta H(T)$ of phase transition.
Consequently, for systems belonging to the second group an additional noticeable reduction of melting temperature can be expected as compared to the systems of the first group.

3.1 The dependence of melting temperature of nano-scale systems on their size without the change of form (nano-crystals, nano-clusters in free state)

If nano-clusters and nano-crystals are in free state and have an optimal form considering the ratio between surface area and volume, then they virtually don't change their form during melting. If we assume that the form of nano-crystals is close to being spherical, it is easy to see that for this kind of systems the expression (7) is transformed into

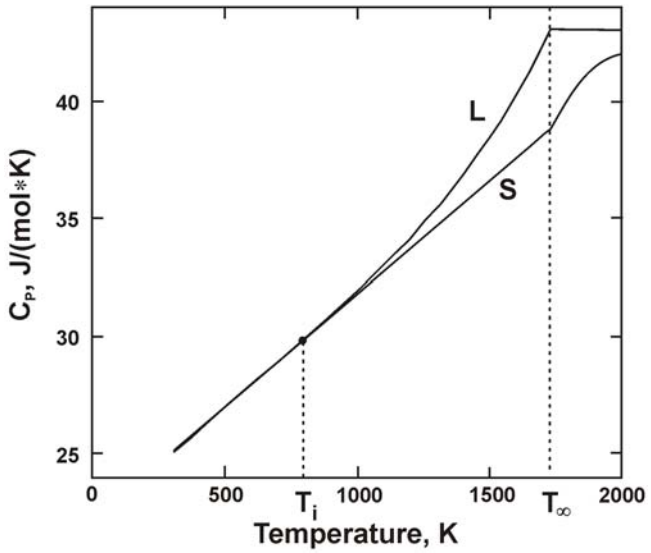


Fig. 4. Temperature dependence of specific thermal capacity of nickel in solid and liquid state.

$$T = T_\infty \left(1 + \frac{3\Delta\sigma}{r\Delta H(T_\infty)}\right) \tag{11}$$

where r is the radius of the nano-crystal.

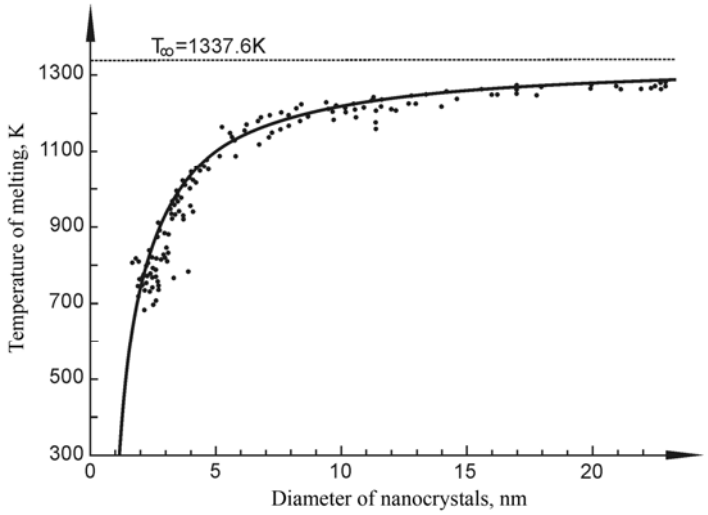


Fig. 5. Experimental values and calculated curve of dependence of gold nano-crystals melting temperature on their size according to (11).

As an example we shall calculate, using (11), the dependence of equilibrium melting temperature of gold nano-crystals on their size and compare the results of the calculations with available experimental data. The following surface energy values of gold bordering on a vacuum were taken: σ_S=1.325 J/m^2 and σ_L=1.125 J/m^2. The molar melting heat of gold is $\Delta H(T_\infty)$=12.68 kJ/mol, consequently the melting heat of a unit of volume of gold which is required for substitution into expression (11) amounts to 1.243×10^9 J/m^3. Fig.5 illustrates the array of experimental values of melting temperature of gold nano-crystals depending on their size obtained by Ph.Buffat and J.-P.Borel, as well as the calculated dependence of melting temperature according to expression (11). As can be seen, the calculated curve coincides well with the experimental data. Therefore, expression (11) reflects the experimental observations quite well.

3.2 Dependence of nano-scale system melting temperature in free state on their dimensions with changes of their state (thin films)

The peculiarity of the second group, which is represented by thin films in free state is that due to large surface area they are not optimized as far as minimum energy is concerned. From the analysis given in the Section 2 follows that the equilibrium temperature of thin films melting on inert surface is the temperature, at which thickness of the molten layer is equal to that of the thin film, and according to observations, in reality, it instantly disperses into drops. Due to this process surface energy sharply declines because of shrinking surface area – this implies optimization of shape. Experimental investigation implies that due to this factor temperature of thin films melting considerably differs from that of extensional objects, if compared to nano-crystals in free state. Temperature of Cu film melting (thickness-100 nm) on SiO_2 is below the reference temperature of melting of extensional Cu for 340 K, and the temperature of Ni film melting (thickness-100nm) on SiO_2 differs from the reference temperature of extensional Ni melting for 625 K.

The difficulty of melting temperature calculation for the second group according to formula (7) is that it includes the heat of melting $\Delta H(T)$ which itself depends on the melting temperature.

Excess absolute surface energy of a nano-scale object and possibility to lower it as a result of transition from solid to liquid state and decrease in the surface area is a kind of "trigger" for the process of melting at lower temperature. Due to this it is possible to find the temperature of a certain nano-object melting by means of iterations – in successive and multiply repeated calculations of melting temperature value T and corresponding heat of melting $\Delta H(T)$.

The first step is calculations of T according to (7). Next $\Delta H(T)$ is calculated at temperature T obtained at the first step according to the Kirchhoff equation

$$\Delta H(T) = \Delta H(T_\infty) + \int_{T_0}^{T} \Delta C_p dT \ , \tag{12}$$

where ΔC_p is the change in the heat capacity caused by the transition from solid to liquid state. The $\Delta H(T)$ value obtained is substituted into (7), and the melting temperature T is again calculated. Calculations are repeated until the melting temperature T ceases to change. This value is taken to be the melting temperature of the thin film of the selected thickness. In order to obtain the melting temperature of a thin film of another thickness, volume V substituted into (7) is corrected according to thickness changes and the calculations are repeated.

Theoretical curves of Cu thin films melting temperature are calculated according to their thickness by means formula (9) and by means of described iteration method using formulae

(7) and (12) as well. The object of modeling is a copper film area with surface 1000 μm in diameter deposited on the surface of tungsten. The following initial data are used. For the free upper copper film surface, we use σ_S = 1.72 J/m², and for the lower film surface bordering on tungsten, σ_S = 2.83 J/m². This value occurs theoretically only, when σ_S equals the surface energy of tungsten. It is assumed that in calculations according to (7) and (12) the film disperses into spherical drops 3 μm in diameter as a result of melting; each drop touches the surface of tungsten at a single point, and σ_L = 1.265 J/m² therefore corresponds to the surface energy of liquid copper. We assume that the volume of copper does not change as a result of melting and dispersion. The molar heat of copper melting is $\Delta H(T)$ = 13 kJ/mol, hence heat of melting per volume unit necessary for (7) or (12) is 1.826×10^9 J/m³.
The calculated curves are compared with the experimental melting temperatures of different thin copper films shown in Fig. 6.
It is noteworthy that experimental melting temperature of a rather thick film of 100 nm is considerably lower than that of bulk Cu (~160 ^{0}C). If we go back to Fig.5, we notice that the melting temperature of gold nano-crystals (the diameter of gold nano-crystals is significantly smaller – 20 nm) differs from bulk gold melting temperature by 80 °C).
Curve 1 demonstrates calculations according to (9) which does not take into consideration possibility of form alteration and heat of melting. We can see that this curve lies well above the experimental values – i.e. in case the surface area remains unchanged during the process of melting, the dependence of the thin film melting temperature will be closer to that of bulk melting temperature. Thus, formula (9) cannot describe temperature changes of melting accompanied by the dispersion of a film into drops.
Curve 2 obtained iteratively with the use of (7) and (12) shows the inclusion of the temperature dependence of the heat of melting and a change in the surface area of the system caused by its dispersion into drops decreases the melting temperature. Curve 2 is much closer to the experimental values.

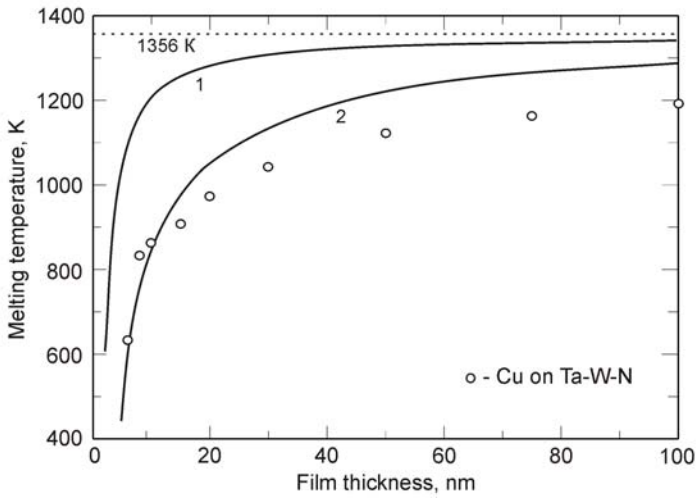


Fig. 6. Experimental Cu thin film melting temperatures on the surface of amorphous alloy Ta-W-N and calculated Cu thin film melting temperature dependencies on their thickness on the surface of tungsten. Curve 1- calculation according to (9), curve 2 – calculations by the iteration method according to (7) and (12).

As with copper, we calculate the film thickness dependence of theoretical melting temperatures of nickel thin films by the same iterative procedure with the use of (7) and (12). The calculations are made for nickel films on Al_2O_3 surface. We take σ_S = 1.86 J/m² for the free upper surface of nickel films and σ_S =2.6 J/m² for the lower surface bordering on Al_2O_3. As with copper, it is assumed that melting causes film dispersion into spherical drops 3 μm in diameter, which touch the surface at a single point each; we therefore use σ_L=1.62 J/m², is liquid nickel surface energy. The molar melting heat of nickel is $\Delta H(T)$=17.6 kJ/mol, its molar volume is 6.6 cm², and the heat of nickel melting is therefore $\Delta H(T) = 2.67\times10^9$ J/m³. We use the temperature dependence of the heat capacity of solid to liquid transition for bulk nickel

$$\Delta C_p = 4 - 0,003T + 5,1\times10^{-6}T^2 - 1,5\times10^5 T^{-2} \tag{13}$$

obtained from the data represented in fig.4.

Fig. 7 shows that the dependence calculated for nickel films on Al_2O_3 (curve 1) lies slightly above the experimental melting temperatures, as in the case with copper.

Further we analyze possible reasons for the observed deviations of the theoretical film thickness dependences of the temperatures of fusion obtained using the procedure suggested above from the experimental dependences.

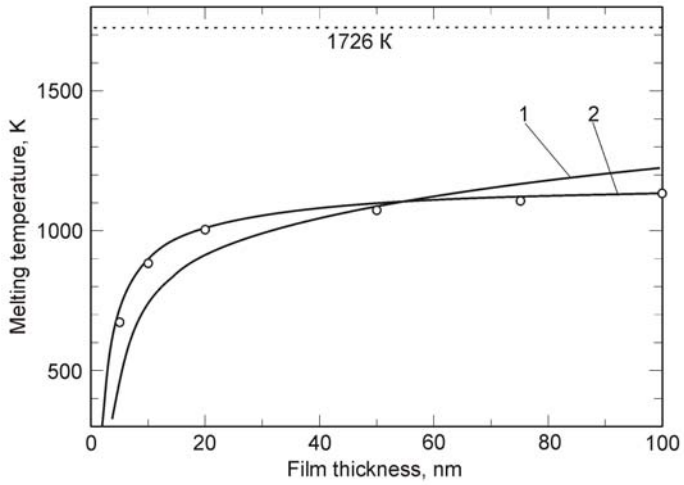


Fig. 7. Experimental and theoretical film thickness dependencies of melting temperature of thin nickel films calculated by the method of iterations according to (7) and (12). Curves (1) and (2) were calculated for Ni on Al_2O_3 from the data using the $\Delta C_p(T)$ dependence for volume nickel (Eq. (13)) and dependence (14), respectively.

The conclusions that can be drawn from our study of the contributions of the terms present in (7) to changes in the temperature of fusion of a system whose absolute surface energy can change as a result of shape changes are the following:

- contribution of changes in the surface energy to a decrease in the temperature of fusion is determined in (7) by the difference $T_\infty(\sigma_L A_L - \sigma_S A_S)/(\Delta H(T_\infty)V)$. Since $\sigma_S A_S T_\infty/(\Delta H(T_\infty)V) >> \sigma_L A_L T_\infty/(\Delta H(T_\infty)V)$, $\sigma_L A_L T_\infty/(\Delta H(T_\infty)V)$ does not make a significant

contribution. It follows that the size of drops and their shape influence changes in melting temperature insignificantly; i.e. the dispersion of a thin film into drops itself is of importance, but it does not matter whether these drops are ideally or non-ideally spherical and 1 or 5 μm in diameter.

- At the same time, $\sigma_S A_S T_\infty/(\Delta H(T_\infty)V)$ makes a substantial contribution. It follows that the initial energies $\sigma_S A_S$ of the upper and lower (at the boundary with a different material) thin film surfaces make a substantial contribution to melting temperature decrease.
- term $\sigma_S A_S T_\infty/(\Delta H(T_\infty)V)$ makes the most significant contribution because of different temperature dependences of the heat capacities C_p of the liquid and solid phases. The temperature dependences of the heat capacities of the liquid and solid nickel phases are shown in Fig. 4. The larger this difference, the larger the contribution of this term to a decrease in the temperature of fusion.

It follows from these conclusions that there can be several reasons for deviations of the theoretical curves calculated in this work from the experimental data.

The first reason for these deviations is the use of the temperature dependences of the heat capacities of volume liquid and solid phases in integration according to (12).

It follows from the introduction that the amplitude of oscillations of crystal atoms situated close to the surface is larger by than in the bulk. As a consequence, the Debye temperature of surface layers is lower by than that of the bulk material. Since the heat capacity is related to the values specified in accordance with the Debye quantum theory of heat capacity, the heat capacity of surface layers should be higher than the heat capacity of bulk materials. The amplitudes of atomic oscillations are higher in the liquid than in the solid phase, and the temperature dependence of the heat capacity of the liquid phase is in the majority of cases steeper compared with the solid phase. Extending this conclusion to the surface layer, we are led to suggest that the difference between the temperature dependences of the heat capacities of the surface layer in the liquid and solid state (that is, $\Delta Cp(T)$) increases compared with the bulk material.

It should in addition be noted that it follows from the same reasoning concerning the difference in the phonon spectra between the surface layer and volume phase that the heat capacity of the surface layer depends on film thickness as long as the film thickness is comparable with the surface layer thicknesses at which the amplitudes of atomic oscillations differ from those in the bulk.

Based on these considerations, we corrected the coefficients of the temperature dependence Cp(T) to fit the calculated dependence of melting temperature to the dependence observed experimentally for the thin nickel film on Al_2O_3The results are shown in Fig. 7 (curve 2). The corrected dependence has the form

$$\Delta C_p = 4 - 0,0128T + 13,7 \times 10^{-6} T^2 - 1,5 \times 10^5 T^{-2} \quad (14)$$

The second reason for the deviations can be the accuracy of the surface energy σ_S values that we use. The reliability of surface energy values leaves much to be desired because of the problems related to their measurement. The question of what energy value should be used for the interphase boundary between two materials remains open. What is more, we use a constant surface energy value, whereas, like ΔH, it depends on temperature. The discrepancies between the calculated curve (Fig. 6, curve 2) and experimental data on copper can therefore be caused by the use of the σ_S value for tungsten in our calculations, whereas the experimental data on the copper film were obtained on a Ta–W–N amorphous alloy whose surface energy is unknown but is likely higher than that of tungsten.

There is one more reason, if the first one is not responsible for the deviations. This is the accuracy of the temperature dependences of the heat capacity C_p of the solid and liquid phases. Calculations show that, for instance, a 10% error in the temperature dependence for liquid copper toward an increase in its heat capacity results in a noticeable (by about 50 K) shift of the calculated film thickness dependence of melting temperature toward lower temperatures.

3.3 The dependence of melting temperature of nano-scale systems on their size without form alteration (nanowires within other solid material)

As one more example of systems, relating to the first group, one can exemplify the melting of In, Cd and Zn nanowires within porous Al_2O_3.

Porous anodic alumina (PAA) is a closely packed structure, consisting of many hexagonal cells, in the center of which is a vertical pore. PAA attractiveness is in its low range of pore sizes in diameter and this size is well controlled in the course of its formation. Deposition of In, Cd and Zn in these pores is implemented by electrochemical method using pulsed deposition, which ensures a void-free PAA pore filling by metal.

Fig. 8 illustrates scanning electronic microscopy (REM) photography of some porous element, where nanowires of In are deposited in its pores. As it follows from the results of microscopy, metal homogeneously fills these pores and metallic wires have no voids. That is why the diameter of nanowires can be considered equal to that of pores in anodic oxide.

Melting temperature of In, Cd and Zn metallic nanowires is defined by means of differential scanning calorimetry (DSC) and with the help of incipient thermal absorption, which is expressed in DSC graphics as deviation from the linear law. Phase transitions aren't discovered in initial structures of PAA without a deposited metal at temperatures from 323K to 773K, therefore observed thermal absorption in objects with a metal deposition may have been caused only by the process of melting.

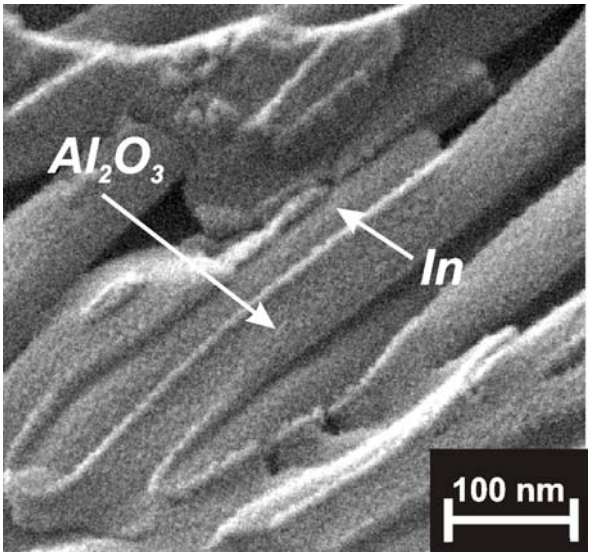


Fig. 8. SEM photographs of In nanowires in porous alumina host.

Fig. 9 illustrates an observed dependence of melting temperature of In nanowires in pores of different diameter. As we can see, along with a downscale of nanocrystals to 30 nm melting temperature reduces. Nevertheless, with the reduction of size to 20 nm, the increase of melting temperature takes place. Fig.9 draws attention to the fact that the reduction of melting temperature of nanowires, enclosed in PAA host, is inconsiderable (3K), unlike in the previous cases. This dependence is not only the feature of In nanowires. This type of so-called non-monotonous dependence is also observed for Cd and Zn nanowires.

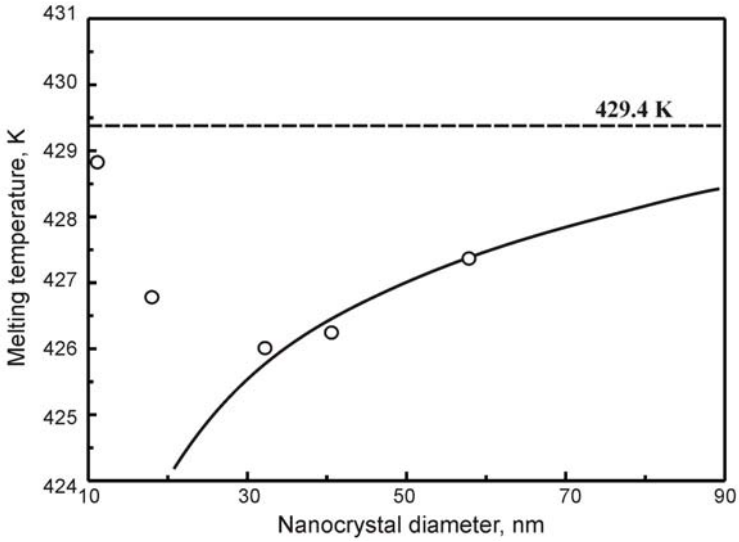


Fig. 9. Experimental melting points of In nanowires in PAA and estimated dependence of melting temperature using the equation (15).

Formula (7) enables us to calculate temperature dependences of melting for In, Cd, Zn nanowires. The length of cylinder is expected to remain unchangeable, and its diameter ranges from 8 to 20 nm. As the diameter of nanowires is much less than its height, it means that the calculation of surface area is added up to determining of the cylindrical surface. Taking into account the impossibility of nanowire shape alteration in PAA, (7) transforms into

$$T = T_{\infty}\left(1 + \frac{2\Delta\sigma}{r\Delta H(T_{\infty})}\right) \tag{15}$$

Initial data required for calculations are represented in Table 1. Since temperature dependences C_{pL} and C_{pS} of In coincide, ΔC_p equals zero in this case.

As it is shown in Fig.9, the calculated curve coincides with experimental values at reducing In nanowire diameter to 30nm. The same coincidence of calculated and experimental data in the same range are obtained for Cd and Zn.

These derived temperature dependences prove that when there is no possibility to alter surface area, the reduction of melting temperature (systems of the first group) with a size changing is very little. The crucial moment is this: temperature dependence $\Delta H(T)$ of systems relating to the first group doesn't make any impact on the reduction of melting temperature.

Material	σ_S, mN/m	σ_L, mN/m	T_∞, K	$\Delta H(T_\infty)$, $J/m^3 \times 10^9$	ΔC_P
In	633	556	429,4	0,42	0
Cd	606	560	593,9	0,47	-8,99-0,0258T-1,79^{-5}T^2+2,63^{5}T^{-2}
Zn	830	767	692	0,78	21,2-0,13T+4,9^{-5}T^2-8.55^{5}T^{-2}
AL_2O_3	2600				

Table 1. Initial data for the calculation of melting temperature of In, Cd and Zn nanowires.

From our point of view, the reason for the increase of melting temperature in nanowires with diameter<30 nm is not grounded in an amount of substance, that in this case doesn't correspond a physicochemical notion "phase", as it was remarked in Section 2. This range of sizes is influenced by the neighboring material. The analysis of causes of melting temperature increase with size reduction of 20 nm is examined below.

4. The role of heterogeneous melting in various physical and chemical processes

The analysis, presented in the previous sections, shows that if there is a surface, there are reasons of thermodynamic and energy character, which determine the reduction of melting temperature in low-sized objects along with the decrease of their dimensions. However, a considerable temperature of heterogeneous melting results in becoming of this kind of melting as an initiator or participant of other processes in low-scale systems and thus it makes a significant influence on their behavior.

4.1 The behavior of thin films on inert substrates

As it is stated above, thermal treatment of thin films from a few to hundreds of nanometers in thickness, deposited on an inert substrate (i.e. having no chemical reaction with it) leads to its micron-sized droplets decay.

The study of various material thin films (in particular Cu, Au, Ag, Ni, Fe) of different thickness has demonstrated that this process has no strictly determined temperature. It can occur in some temperature range, but its temperature is substantially lower than that specified in a reference book of a three-dimensional material.

For instance, copper film 20 nm in thickness starts to disintegrate into drops at 883 K after 300 s, while at 743 K this process appears after ~13000 s. The film 100 nm in thickness begins to dissociate at 1013 K after 420 s, while it also dissociates at 913 K, but after ~9000 s. The disintegration of copper thin film occurs locally at any place of the sample and then spreads further frontally (Fig.10). As a result, it is possible to distinguish three sample areas: the area where the film remains continuous (Fig.10 (a)); the area where the film is totally dissociated into droplets (Fig.10 (e)), and the third one - narrow area or a front, where dispersion takes place (Fig.10 (b-d).

Fig.11 represents dependences of time to melting on annealing temperature for copper films of different thickness. These experimental points are perfectly approximated (with correlation ratio 0, 98-0, 99) to the following equation

$$t = \frac{A}{\exp\left(\frac{E_{act}}{RT}\right)} \tag{16}$$

which follows the Arrhenius equation.

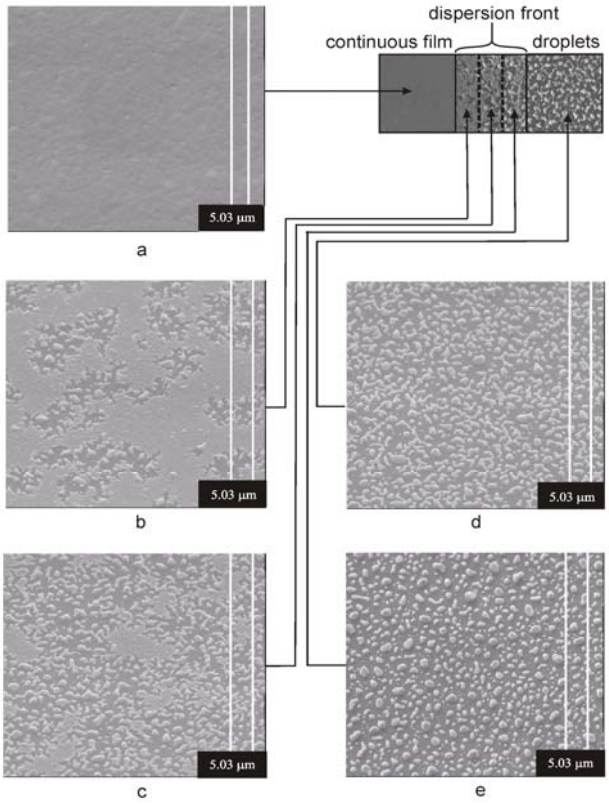


Fig. 10. SEM photographs of the copper thin film at different stages of melting: a) region of continuous film; b) – d) region, where melting of film occurs; e) region, where this film dissociated into droplets.

This approximation of experimental data enables to calculate activation energy of melting-dispersion for films of different thickness. In accordance with these calculations, E_{act} varies from 120 to 260 kJ/mole depending upon copper film thickness. For example, dependence t=f(T) for copper film 100 nm in thickness is approximated by means of this equation:

$$t = 1 \cdot 10^{-11} \cdot \exp\left(-\frac{260 \cdot 10^3}{8,31 \cdot T} \right) \tag{17}$$

Since this type of dependence is also common for diffusion, one can presume that the observed phenomenon is of diffusion nature. In this case diffusion coefficient D_0, calculated from (17), is equal to ~10000 cm²/s, that is two orders of magnitude more than specified in reference books copper self-diffusion coefficient, which is 70 cm²/s.

Values E_{act} and D_0 obtained are similar to the ones experimentally defined by A.E. Dolbak and his colleagues, who studied copper diffusion from its thin strips ~12 nm in thickness

along pure silicon surface. This high diffusion coefficient D_0 is explained in terms of surface diffusion that occurs via the solid - state spreading or "unrolling carpet" mechanism.

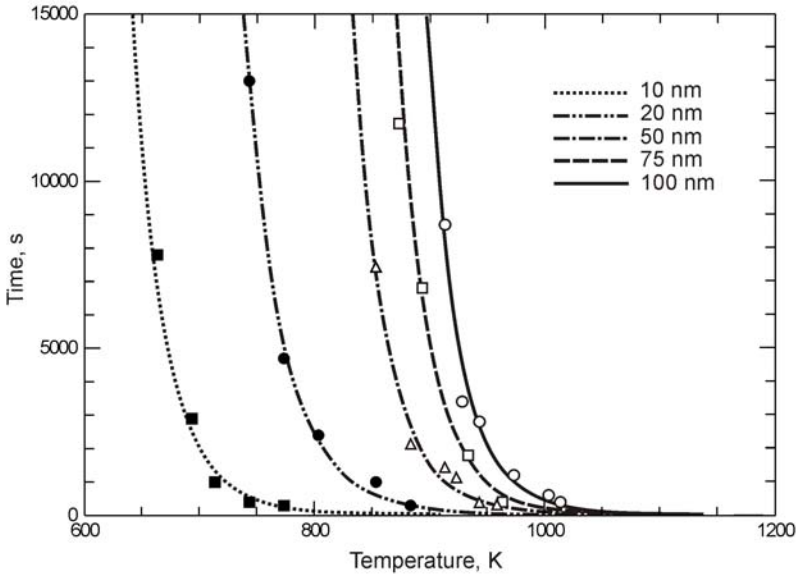


Fig. 11. Dependences of melting time in copper thin films of various thickness on heating temperature in a vacuum.

That's why it seems quite sensible to explain the dissociation of thin film into droplets by surface diffusion. But it does not seem to be correct. Diffusion is defined to be a mass transfer as the result of thermal movement of substance particles in the direction of their concentration decrease. Volume diffusion results in their uniform and even distribution of particles throughout the volume and surface diffusion – over the surface. In this case we deal with quite an opposite effect: the process goes from the uniform surface distribution of particles to their collecting into droplets, i.e. in the direction of concentration increase.

At the same time, we should draw attention to the following issue. Fig.12 demonstrates the area, defined experimentally depending upon the thickness where dispersion of nickel thin film can be observed. The upper bound of shaded area is the dependence of dispersion temperature, where it begins after 5s. The lower one shows when the same process begins after 5 h. Curve 1 illustrated on Fig.12, is an estimated dependence, attained by iteration method, using equations (7) and (12), which is shown on Fig.7 (Section 3.2). As we can see, the estimated dependence is almost wholly situated in the shaded area of experimental observations.

Furthermore, if we try to employ a theoretical dependence in the area of film heavy gages, it seems rather unusual. In compliance with this dependence, a rather thick film 0,5 – 1,0 μm in thickness is to melt at 300 – 400 K below than that specified in a reference book. Fig.13 represents photomicrographies of iron film surface at 1323 and 1273 K, which are similar to nickel regarding its behavior. Comparison of Fig.13 (a) and (b) enables us to assume that film 0, 5 μm in thickness after its annealing at 1323 K had such a morphology that could be formed only with a liquid phase, thus it confirms calculations result.

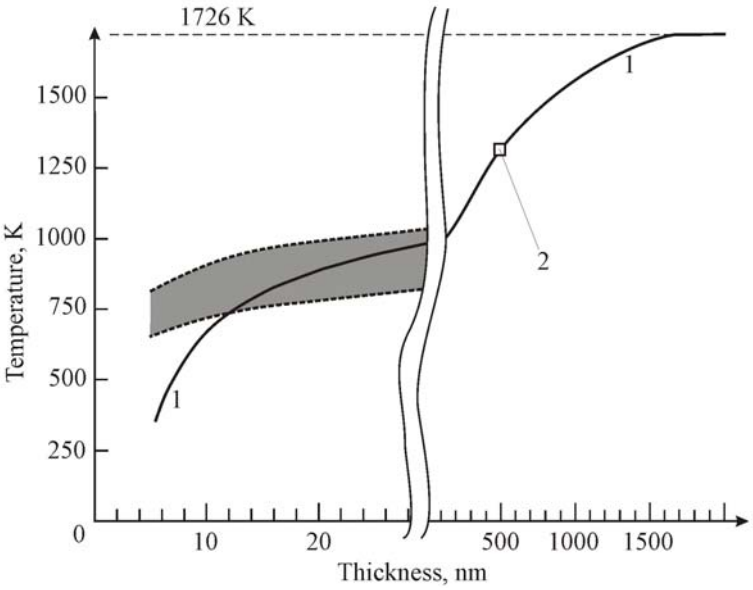


Fig. 12. Temperature range, where dispersion of nickel thin film in droplets is observed, and calculated curve of melting temperature dependence of nickel thin films on its thickness. Curve 1 – calculation by means of iteration method using formulas (7) and (12); 2 – control data point.

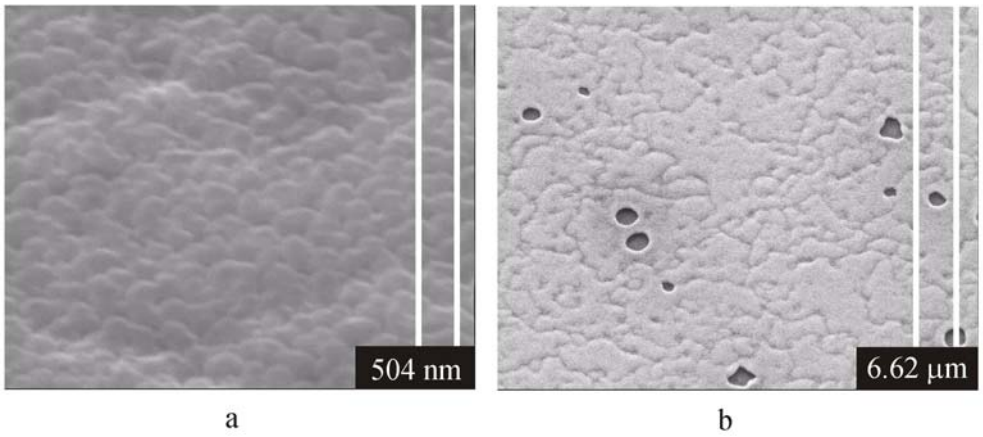


Fig. 13. SEM photographs of the iron film 0, 5 μm in thickness: a) – after annealing at 1273 K; b) – after annealing at 1323 K

There is an opinion that these facts are just coincidence, as the process of dispersion is a purely kinetic one and its usage for description of equilibrium thermodynamics relationships is not quite justifiable. Moreover, this process is an irreversible that gives evidence of its non-equilibrium state.

However, it is necessary to remember (Fig. 2) that in a semi-infinite system, where the heterogeneous melting occurs, the equilibrium melting temperature is a temperature at which the crystalline phase disappears completely. But the surface causes the emergence of the liquid phase long before this temperature is achieved. Similarly the equilibrium melting temperature of a thin film described by expression (7) is a temperature at which the whole film turns liquid but virtually the liquid phase on the film surfaces arises even earlier. Thus, theoretical curve 1 in Fig. 12, derived from expression 7, indicates that there are objective energy (i.e. thermodynamic) reasons producing a considerable displacement of equilibrium between liquid and solid phases towards a lower temperature as a thin film thickness is decreasing. In fact, owing to the heterogeneous melting the liquid phase and the thin film dispersion into droplets, caused by this phase emergence, arise even earlier at lower temperatures. Curve 1 in Fig. 12 should have coincided with the upper boundary of the shaded area.

If this process depended on kinetic laws only, we should have expected the film structure to affect the temperature and dynamics of this process significantly. However, the study of the thin copper films 50 nm thick on the Al_2O_3 surface showed the following: the temperature range, within which films with an average grain size of 90 nm disperse, is 10-20 K higher than that for the films with a grain 25 nm thick. Such an insignificant difference also shows that despite the evident importance of kinetics (exponential dependence of process rate), the dispersion process results from thermodynamics to a greater extent.

Kinetic peculiarities of the thin film dispersion into droplets may be conditioned by such liquid properties as viscosity or fluidity and by evaporation, all of them highly dependent on temperature.

Thin silver film behaviour during the annealing was studied. It was found (Fig. 14) that the material mass decreases with time, diminishing most intensively during the first 15-20 minutes of high-temperature exposure.

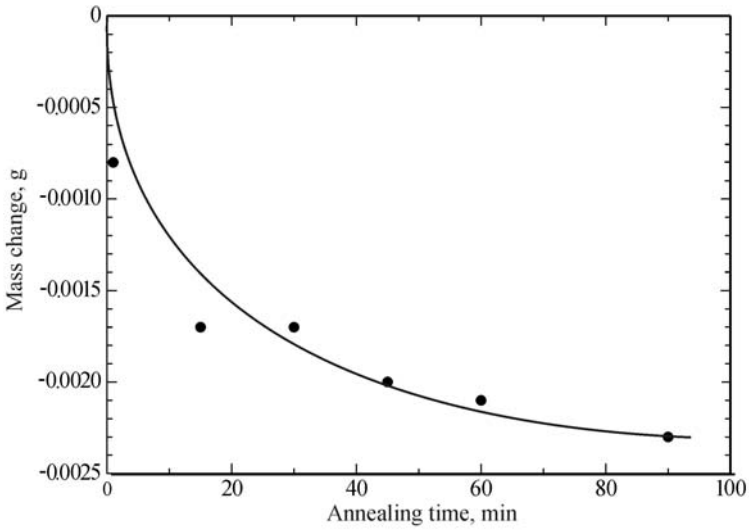


Fig. 14. The dependence of thin silver film mass variation (630 nm thick) versus the annealing time at temperature 820 °C.

The results of this experiment indicate that during the melting-dispersion process of a thin film into droplets a part of a substance evaporates. Moreover, we were able to observe a mass loss even at the temperature lower than the equilibrium melting temperature but at which the process of dispersion into droplets occurs.
The information described allows us to assume the following mechanism of the thin film dispersion into droplets. This mechanism accounts for the absence of a particular temperature of this process. Thin film excess surface energy causes this process. To reduce surface energy, we can decrease specific surface energy σ and thin film surface area A. But surface area shrinkage in a crystalline state is extremely complicated. Because of this, during heating the heterogeneous melting occurs faster.
On the upper and lower thin film surfaces there arises an inherent liquid phase that wets well and partially (not always completely) dissolves crystal grains of the thin film on the boundaries (Fig. 15 (a)). According to the experimental results obtained by AFM a thin film, e.g. 50 nm thick, consists of no more than two layers of crystalline grains. Liquid layer thickness depends on temperature. For example, dispersion of a gold film 50 nm thick was observed within the temperature range 903-1003 K. Formula (10) indicates that both at 903 K and at 1003 K on the upper and lower surfaces of a thin gold film there has to exist a liquid layer of more or less uniform thickness: approximately 12 and 16 nm, respectively.

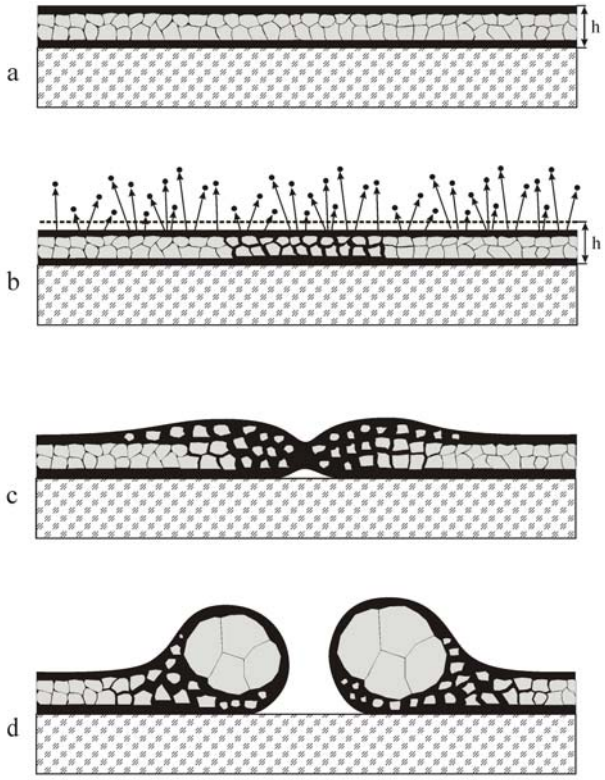


Fig. 15. Schematic diagram illustrating the melting-dispersion of a thin film on the inert surface.

As a liquid layer appears, there starts a sufficiently intensive process of evaporation. This reduces a film thickness, decreases its equilibrium melting temperature and diminishes the quantity of a solid crystalline phase (Fig. 15 (b)).

During the heating time grains with boundaries are sufficiently wetted, and in a local region the liquid phase penetrates into the entire film thickness (Fig. 15 (b)). Liquid evaporation and propagation resulting from the wetting determines the kinetic component of thin film dispersion into droplets: at higher temperatures fluid viscosity is lower and dispersion starts quickly; at lower temperatures viscosity is higher, wetting occurs more slowly, and dispersion process starting only in a while is slower.

Owing to the fact that the liquid phase is fluid, sufficiently wets its inherent crystalline solid phase and poorly wets the substrate inert surface, the surface tension force reduces the thin film surface area, spending the excess surface energy of the thin film. As a result, the liquid phase leaks on its inherent solid phase entraining crystallites it contains and causing a thin film rupture (Fig. 15 (c)). Surface energy lowering shifts the equilibrium between the liquid and solid phases towards higher temperatures, which makes impossible the existence of this amount of the liquid phase at the given temperature. Because of this, the process of liquid crystallization starts easily at the same temperature (as there is no need to spend energy to form an interface). Liquid crystallization is accompanied by a local heat release, which causes the effect of "partial melting" of the polycrystalline film remained in the solid state (Fig. 15 (d)). As a result, there arises a moving front where the melting-dispersion process occurs.

Thus, it is evident that the process of thin film dispersion is not just a melting process. However, it is quite obvious that the heterogeneous melting is the main component of this process and determines its temperature.

The studies of kinetics of thin film dispersion into droplets revealed an anomaly in the behavior of thin films less than 30 nm thick – nonmonotonic melting temperature dependence. This anomaly overlaps with the results obtained for filamentary nanocrystals (see part 3.3).

Fig. 16 (a) and (b) represents dependences of the equilibrium melting temperature and activation energy of dispersion of thin Cu, Ag and Au films into droplets. These metals belong to the same group of the periodic table. These diagrams show certain regularity. According to Fig. 16 (a), melting temperature dependences of thin films on their thickness are monotonic for Cu and Ag; but the thinner thickness, the higher melting temperature thin Au films (less than 20 nm thick) have, instead of expected temperature lowering. Besides, the diagram illustrating the activation energy dependence for Au also has a minimum on the same thickness, but activation energy falls steadily with reducing thickness for copper. Ag has an intermediate position in the group between Cu and Au in the periodic table and, thus, behaves accordingly: given the stable monotonic melting temperature dependence, the activation energy dependence obtains a minimum. I.e., Ag films 10 nm thick have higher energy of dispersion into droplets than a film 40 nm thick has. Thus, there is an evident regularity: as an atomic number of an element in a group increases, thin film dispersion into droplets is more difficult.

In section 2 we have already mentioned that there exists a problem of a physicochemical concept of "phase" for a thin film with thickness less than a certain thickness. Phase is a certain amount of substance which has an interface, and at any point of which thermodynamic functions are constants. However, as the thin film thickness reduces, both of its surfaces approach each other gradually, and at the certain moment this principle is violated and a thin film no longer corresponds to the "phase" concept.

There are three evident cases here.
The first extreme case is a characteristic feature of thick films several micrometers thick. The substrate effect is unnoticeable in this case. The film has the properties of an actually solid material and behaves like a classical thermodynamic phase, which borders upon the other phase – the substrate. The surface contribution to the combined energy of such a system is hard to notice. The film and the substrate can be viewed as two independent systems.
The second extreme case is a feature of the thinnest films and corresponds to a high degree of substrate effect. The peculiar feature of this case is that the thin film is, in essence, an adsorbed layer reducing the substrate surface energy. The thin film is not a phase in a classical physicochemical understanding. The thin film and the substrate are a single system in this case. Thin film material atoms tend to spread uniformly on the substrate surface area. According to the definition, the surface diffusion phenomenon relates to this very case.

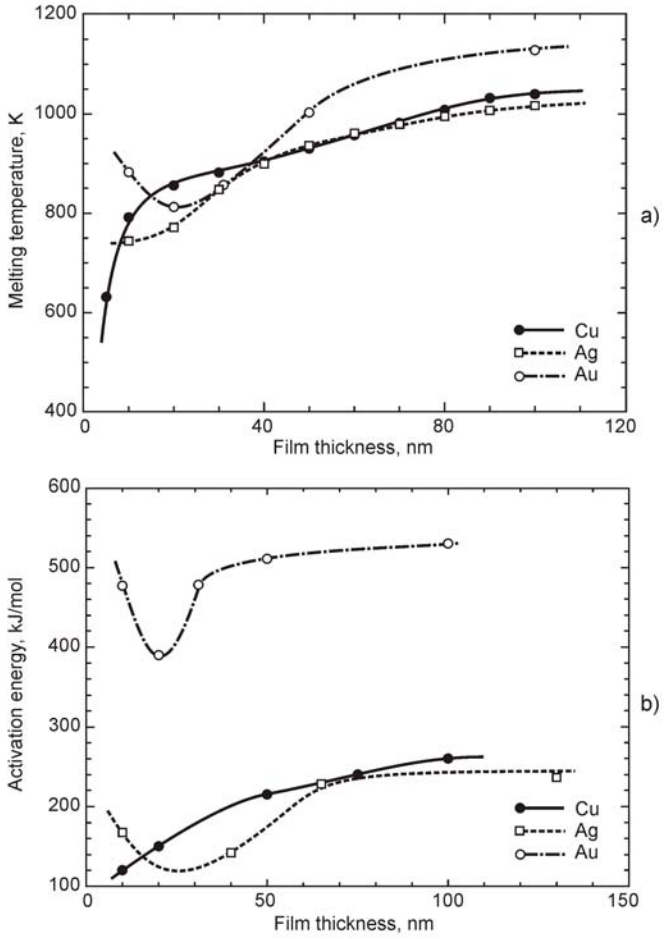


Fig. 16. The dependences of melting temperature and activation energy of thin films of Cu, Ag and Au on their thickness.

The third case is intermediate between the first one and the second one. When the film thickness exceeds the thickness of the second case, the thin film starts displaying properties of an individual independent object with its own volume and constrained by the surfaces above and below. In this case the substrate affects the film, but the film behaves like an independent thermodynamic system in non-equilibrium due to a substantial surface energy contribution. This system tends to transform into such a condition so that a volume energy distribution prevails over the surface one. As a result, the film disperses into droplets, each becoming an independent object. The fact that the thin film disperses into droplets proves it is not a phase. However, the film and the substrate can be viewed as two different systems.
The transitional nature of the third case lies in the fact that the substrate effect gradually weakens and the thin film individuality increases as the thin film thickness grows.
Thus, the main reason of non-monotonic dependence of melting temperature of a thin Au film and filamentary In, Cd and Zn nanocrystals on their size (thickness, diameter) is increasing effect of the material, which a low-sized system adjoins to. This is a transitional area between 2-D film shrinking into a 3-D droplet and surface diffusion and adsorption process.

4.2 Filling narrow trenches with copper

The phenomenon of heterogeneous melting of a thin film can be used for filling with copper of narrow trenches and contact windows with a high aspect ratio.
A thermal treatment at 800 ^{o}C of thin copper layer, which is conformally deposited on the structure of trenches on the surface of an amorphous Ta-W-N allow, which does not interact with it, leads to the phenomenon of dispersion of a thin film on spherical droplets accumulating mainly in the crest of the structure (Fig. 17 (a, b)). To suppress this phenomenon, the films of the Ta-W-N allow and copper are separated by a wetting layer of titanium. It is found that the titanium layer thickness up to 10 ÷ 15 nm does not change the nature of wetting and copper at the same temperatures continue to disperse into droplets, as in Fig. 17 (b). (It should be noted here that the film thickness corresponds to the thickness of the film on a flat surface. In view of the differences of deposition rate a film thickness on the sides were somewhat lower.)
However the situation changes dramatically for thicknesses of titanium layer over 15÷20 nm: wetting appears, and copper start to flow into the trenches (Fig. 17 (c)).
This result is a good indirect evidence that copper thin film at 800 ^{0}C had properties of fluid and was in liquid state: material is selected into spherical droplets, if surface wetting is absence, and fills the cavities, if there is good surface wetting.
In addition, in example of titanium layer we see again a criticality of thin film thickness up to 20 nm from point of view a physicochemical notion “phase”.

4.3 Degradation of aluminum contacts to the semiconductor in silicon integrated circuits

When metallization of silicon integrated circuits was single-level and single-layer, and its material was aluminum, failures began to appear with gradual decreasing the element sizes. It was established that the reason of the appeared frequent failures of integrated circuits with the one layer aluminum metallization was the shorting of *p*–*n* junctions due to the heat treatment. A selective removal of the aluminum film from the substrate surface near the contact window after heat treatment at 470°C for 30 min indicates catastrophic degradation of the morphologic perfection of the contact interface in the form of deep holes in silicon, that shows SEM-micrography (Fig. 18).

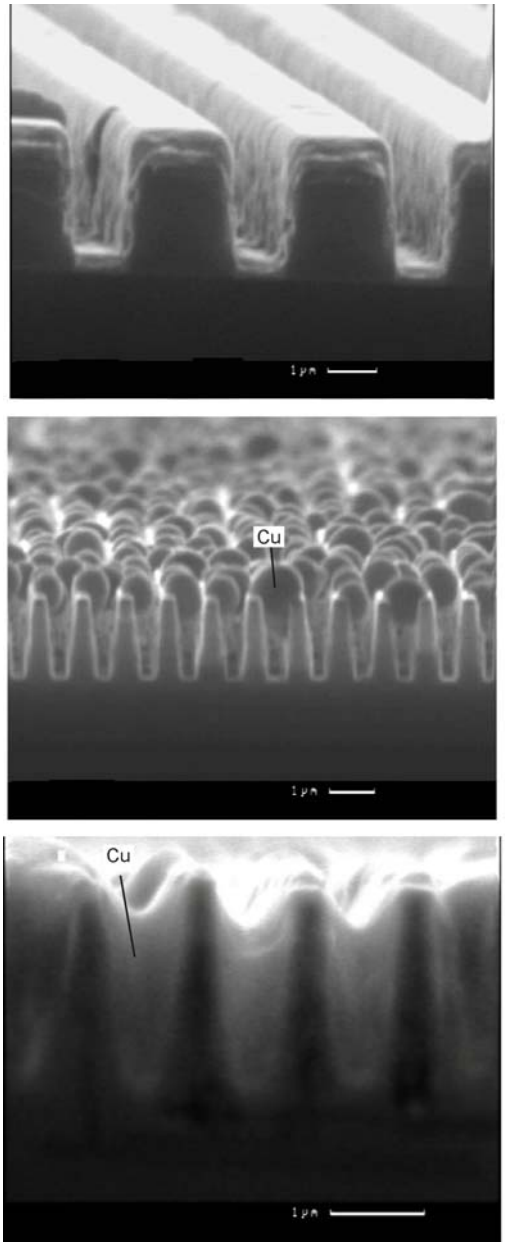


Fig. 17. Heterogeneous melting of the copper thin film deposited on the structure with trenches in SiO_2 coated by the layer of the amorphous Ta–W–N alloy: (a) dispersion into droplets as a result of heat treatment at 1123 K and (b) flow in the trenches as a result of heat treatment at 1123 K owing to the introduction of the wetting 20 nm thick titanium layer between the Ta–W–N layer and the copper film.

The Al–Si system has a phase diagram of the eutectic type. Aluminum almost does not dissolve in Si, while up to 1.65 at % Si can dissolve in aluminum. The eutectic melts at a temperature of 577°C. A result of such character of the interaction in the system is the dissolution of silicon from the substrate in aluminum at relatively low temperatures; i.e., aluminum acts as a solvent of silicon.

It was experimentally established that the appearance of punctures of *p*–*n* junctions is associated with the above vacancy holes formed as a result of inhomogeneous dissolution of silicon in aluminum (Fig18). The holes are filled with aluminum, which contains dissolved silicon and can cause the shorting of *p*–*n* junctions. The depth of penetration of aluminum into silicon estimated from the hole depth in local regions near the contact is almost unchanged in the range 300–500°C and is 40–90 nm. In the temperature range 500–577°C, the penetration depth sharply increases to 1.25–1.45 μm.

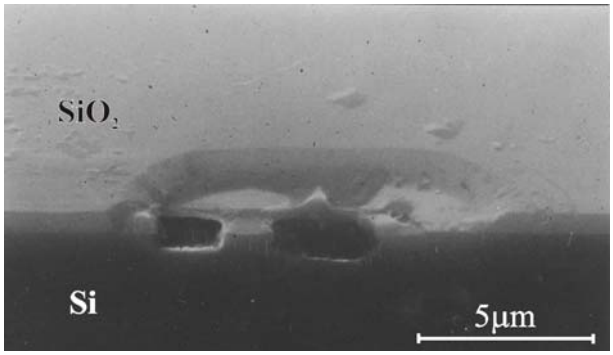


Fig. 18. SEM micrography of the surface morphology of the contact window after heat treatment at 470 °C for 30 minutes and remove the aluminum film.

There was no question that this inhomogeneous dissolution of silicon, filling of the formed holes with aluminum, and the above sharp increase in the penetration depth of aluminum into silicon in the range 500–577°C are due to the appearance of the liquid phase that "melts" silicon, although the heat treatment temperature is lower than the eutectic melting temperature. This circumstance is also confirmed by the fact that the amount of silicon dissolved in aluminum significantly exceeds the amount that should dissolve in aluminum according to the phase diagram at the aforementioned temperature of heat treatment.

It was believed that the reason for the appearance of the liquid phase at a lowered temperature can be local sites with a temperature above 577°C, which can be formed as a result of a strong exothermic reaction between aluminum and the residual natural silicon oxide

$$3SiO_2 + 4Al \rightarrow 3Si + 2Al_2O_3$$

which results in an energy released of 220 kJ per mole of the reacting oxide. However, the natural oxide is removed before the deposition of the aluminum film. Although the quasi-oxide SiO_x can remain on the silicon surface, its thickness is ~0.4–0.5 nm (i.e., 1–2 monolayers), which is unlikely sufficient to heat the film to the eutectic melting temperature due to the heat released as a result of the quasi-oxide reduction.

Moreover, the occurrence of meltings was detected at a lower temperature. Heat treatment of integral circuits structures having contact to the silicon layer of titanium, two levels of current-carrying interconnects Al-Si (1%)-Ti (0.5%), separated by a dielectric layer, and

diffusion-barrier layer of TiN, separating the contact layer of titanium and the first level of current-carrying interconnects, at 450 °C resulted in the formation of voids filled with aluminum in silicon, as can be seen in Fig. 19. Analysis of the distribution share of degraded contacts area of the plates showed that this distribution is uneven and random. It should be noted that most nucleation of voids starts at the edges of contact windows (Fig. 19).

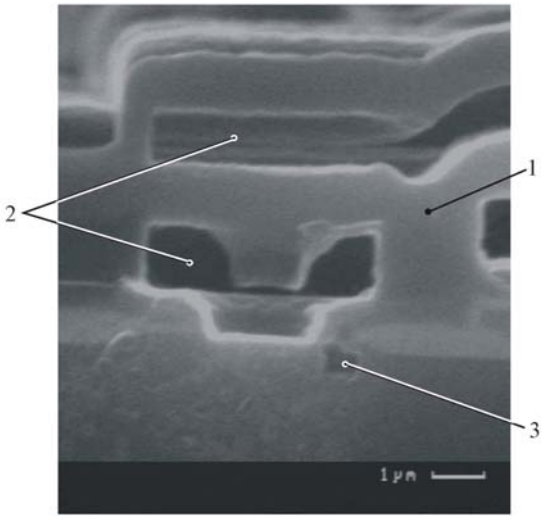


Fig. 19. SEM micrography of the integrated circuit structure cross-section with a two-level metallization after heat treatment at 450 °C for 30 min, and the selective removal of Al: 1 - interlevel SiO_2; 2 - the aluminum interconnects, 3 - penetration of aluminum into silicon in the contact window area.

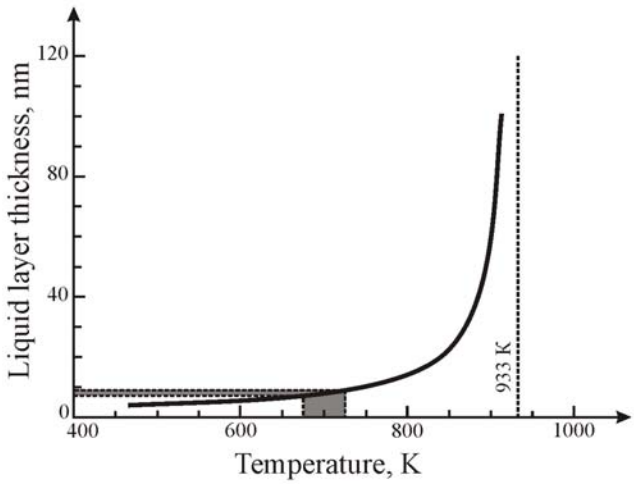


Fig. 20. Temperature dependence of liquid layer thickness on the surface of aluminium, calculated using relationship (10).

For the liquid phase to appear, the formation of the local sites with a temperature higher than the eutectic temperature is not necessary. The formation of such local sites is reasonably explained from the standpoint of the heterogeneous melting under consideration. The calculations (Fig. 20) performed using relationship show that even at temperatures of 400–450°C, which are substantially lower than the eutectic temperature (577°C), a 8–10 nm thick liquid layer should exist on the aluminum surface.

4.4 layer-by-layer growth of crystals and epitaxy

In our opinion, the phenomenon of the heterogeneous melting allows one to understand the vapor–liquid–solid mechanism of epitaxy, which occurs at a temperature below the reference melting temperature, and the mechanism of layer-by-layer growth of crystals.

The mechanism of layer-by-layer growth is observed in the growth from gaseous phase, melts, and solutions. Crystals frequently grow by the crystallization immediately of polymolecular layers, whose thickness is 10–1000 nm, rather than by monomolecular layers as follows from the Gibbs–Kossel–Stranski model. The essence of this mechanism is as follows. Near the crystal surface, there arises a layer that has a polymolecular thickness and completely consists of the crystal-forming material supersaturated with respect to the surrounding mother medium. The layer almost instantaneously crystallizes on a crystal face as a whole plate. The release of the heat of crystallization produces the crystallization pressure, which, in turn, causes outflow of the crystal-forming material from the newly formed crystal surface. Thereafter, the process of formation of the supersaturated polymolecular layer again starts near the crystal surface. Thus, the crystal growth takes a periodicity and becomes pulsating owing to the existence of the incubation period, which is necessary for the formation of each new layer. This layer, which they called the transition layer, "represents as if an independent second medium, micromedium, thinnest shell surrounding the growing crystal. In the medium, a new force field acts and the distribution of particles follows the laws of this field, namely, the potential field of forces."

However, a substantial gap in this mechanism is a total absence of a clarity regarding the reasons for the appearance of the layer near the crystal surface and its crystallization at some thickness.

Fig. 21. Layer-by-layer grown crystal.

This gap can be filled by the approach based on the mechanism of heterogeneous melting. Recall that the heterogeneous melting is the equilibrium coexistence of a liquid layer with some thickness quite specific at a given temperature and the bulk crystalline phase, e.g., a

crystal. In our opinion, the layer of the liquid phase on the crystal surface formed as a result of the heterogeneous melting is the transition layer proposed above. Thus, it is more correct to consider that this is not a layer near the crystal surface but it is a layer on the crystal surface, which certainly exhibits a particular ordering dictated by the crystalline phase.
Taking into account the heterogeneous melting, the mechanism of layer-by-layer growth should be complemented as follows. At the initial instant of time, at a temperature of the crystal growth, owing to the heterogeneous melting phenomenon, there exists (on the crystal surface) a fairly ordered layer of the own liquid phase with an equilibrium thickness, whose atoms occupy positions dictated by the potential field of the crystal (Fig.22(a)).

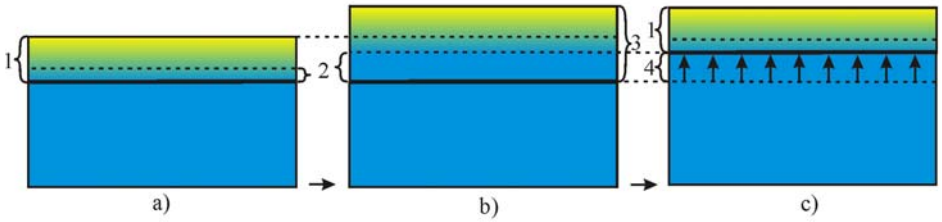


Fig. 22. Schematic diagram of layer-by-layer growth of crystals and epitaxy: 1 – liquid layer with an equilibrium thickness; 2 – ordered part of the liquid layer; 3 – increase of liquid layer thickness beyond equilibrium one; 4 – crystallized layer.

Under conditions of overcooling or supersaturation, under which the crystal grows, a new material continuously transfers to the liquid phase from the mother phase (gaseous phase, melt, or solution that differ from the liquid phase on the crystal surface in the composition, degree of ordering, or concentration) and is continuously ordered inside this liquid phase according to the distribution of the potential field of the crystal. As a result, the thickness of the liquid layer on the crystal surface increases and becomes larger than the equilibrium thickness (Fig. 22 (b)).
The disturbance of the equilibrium state leads to the fact that the layer cannot exist at the given temperature, which inevitably causes the crystallization and a decrease in the thickness of the liquid layer to an equilibrium value (the degree of excess of this thickness over the equilibrium value is determined by the work on displacing the interface between the liquid and crystalline phases) (Fig. 22 (c)). An important peculiarity of the layer-by-layer growth mechanism is that, even if the released heat of crystallization does not displace the liquid-phase atoms (which has not had time to occupy ordered positions) to the mother phase, this heat is responsible for the strong disordering of the remaining liquid-phase layer, so that, at this instant of time, the liquid—crystal interface appears to most abrupt.
Then, the incubation period of ordering of the liquid phase under the action of the potential field of the crystal starts again with a continuous increase in its thickness due to the transfer of the material from the mother phase.
In our opinion, this mechanism of layer-by-layer growth also takes place in the epitaxy, in particular, the vapor-phase epitaxy, which proceeds by so-called vapor—liquid—solid mechanism.
For example, in the chloride method of silicon vapor-phase homoepitaxy, the temperature at which a perfect epitaxial layer grows is 1250°C (the so-called diffusion region where the limiting stage of the process is the transfer of reagents to the solid-phase surface), which is

almost 200°C lower than the reference melting temperature of silicon. In this case, the liquid layer was experimentally observed on the silicon substrate surface. The calculations performed using relationship (10) show (Fig. 23) that the heterogeneous melting should really cause, at a temperature of 1250 ^{0}C, the appearance of the 4-nm-thick liquid layer on the surface. However, at a temperature of 900°C, which corresponds to the kinetic region, where the limiting stage of the process is considered to be the chemical reaction itself, the equilibrium thickness of the liquid layer according to the calculations should be only 0.5-1.0 nm (i.e., 2-3 mono-atomic layers). This corresponds to the vapor-solid mechanism observed for the kinetic region.

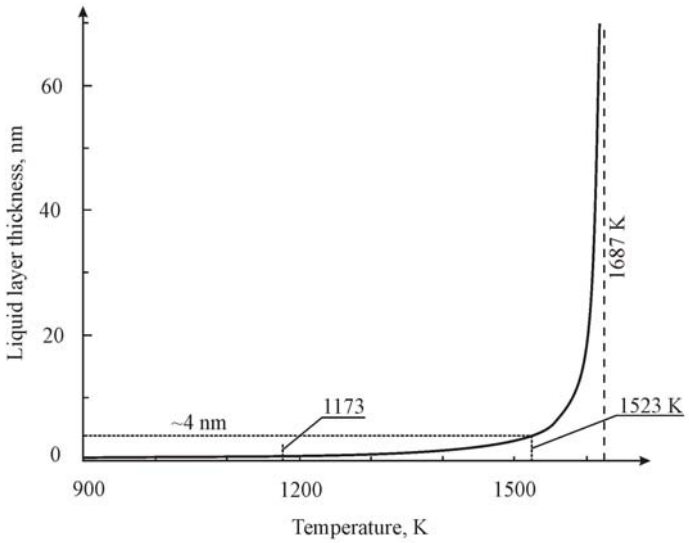


Fig. 23. Temperature dependence of liquid layer thickness on the surface of silicon, calculated using relationship (10).

Taking into account heterogeneous melting the mechanism of layer-by-layer growth allows one to understand how it is possible to decrease the temperature of the silicon homoepitaxy process, which is necessary to suppress the process of autodoping without substantial damage to the epitaxial layer quality. On the one hand, a decrease in the temperature of the process leads to a decrease in the equilibrium thickness of the liquid layer and, on the other hand, to an increase in the viscosity of the liquid and, hence, to a longer ordering of the atoms in the liquid layer. In this connection, the vapor-gas mixture should be added portion-wise. The time interval between the portions is determined by three subsequent stages: 1st stage - ordering of the liquid layer atoms, 2d stage - a decrease in the process temperature in order to induce the crystallization of the ordered liquid layer; and 3d stage - the return of the temperature to the previous level.

4.5 Growth of carbon nanotubes

Growth of carbon nanotubes is also based on the mechanism similar to the mechanism of layer-by-layer crystal growth with the participation of the liquid phase formed as a result of the heterogeneous melting at a temperature lower than the reference melting temperature.

A necessary participant of the process is a set of metal-catalyst droplets. Among such metals are Fe, Co, Ni, Pd, and others, which at high metal concentration have a phase equilibrium diagram with carbon, shown in Fig. 24 for the nickel-carbon system as an example. A specific feature of the diagram is a higher solubility of carbon in liquid nickel as compared to solid nickel.

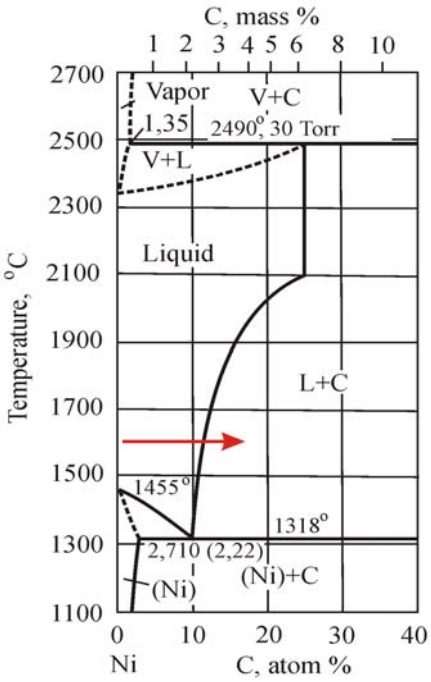


Fig. 24. Phase diagram of the nickel-carbon system.

As stated above, as the thin-film thickness decreases, the phase equilibrium between the liquid and solid phases for the film is shifted towards lower temperatures. For example, for the 50-nm-thick nickel film, this leads to the dispersion of the film into droplets with the participation of the liquid phase at a temperature that is 700° C lower than the reference melting temperature of nickel. Moreover, it was shown that a decrease in the sizes is accompanied by changing all the other equilibrium curves, so that the entire phase diagram is modified and shifted toward lower temperatures.

As a result, the mechanism of growth of carbon nanotubes occurs as follows. At a growth temperature (600—800°C), a nickel droplet 10—50 nm in size represents a crystalline core surrounded by a liquid shell, with an equilibrium thickness, formed due to the heterogeneous melting (Fig. 25 (a)).

When the vapor-gas mixture is fed to the liquid shell surface, there arises a catalytic reaction of decomposition of a carbon-containing organic compound with dissolution of carbon in the nickel liquid phase. A saturation of the nickel liquid layer with carbon increases its thickness and decreases the surface curvature. This disturbs the system from equilibrium (Fig. 25 (b)): as the liquid-layer thickness increases, the lines of equilibrium between the liquid layer and the crystalline core shift towards higher temperatures (see relationship (10)).

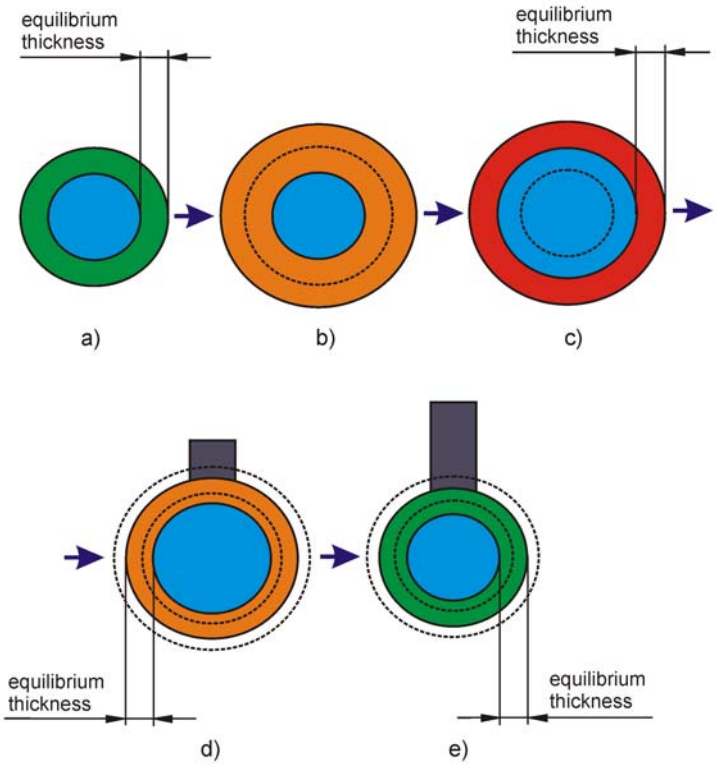


Fig. 25. Scheme of carbon nano-tubes growth.

This results in the initiation of the crystallization, which increases the volume of the crystalline core of the nickel droplet at the expense of nickel from the liquid shell and as a consequence, decreases the volume of the liquid shell (Fig. 25 (c)). In this case, since the solubility of carbon in the nickel liquid phase is much higher than that in the solid phase, carbon remains in the liquid phase and its concentration in the phase sharply increases. Eventually, this results in a supersaturation of the liquid with carbon and the transition to the two-phase region L+C of the coexistence of crystalline carbon and liquid; i.e., carbon precipitates as an individual crystalline phase in the form of nanotubes (Fig. 25(d)).
The released heat of crystallization leads to a partial melting of the crystalline core of the nickel droplet and the system reverts to its original state (Fig. 25(e)). Then, the entire process is repeated. Thus, the process of growth of carbon nanotubes is also periodic and pulsating.

5. Conclusions

Melting of real objects is always a process of heterogeneous melting, since the real objects always possess a surface. It follows from the classical thermodynamics that a liquid layer appears on a solid-phase surface at temperatures that are lower than the reference equilibrium melting temperature of the bulk material. Depending on temperature, certain thickness of the liquid layer on the surface is in equilibrium with the other crystalline phase.

Complete melting of a semi-infinite system occurs near the reference equilibrium melting temperature and does not need a substantial superheating.
The phenomenon of heterogeneous melting is particularly marked in low-dimensional systems, when an equilibrium thickness of the liquid layer becomes comparable to the sizes of objects. In this case, as the sizes decrease, the equilibrium melting temperature significantly decreases to the level lower than the reference temperature.
The process of heterogeneous melting initiates a number of other processes and mechanisms, such as the dispersion of a thin film into droplets, filling of narrow holes with a material of the conformally deposited thin film, the mechanism of vapor-liquid-solid epitaxy, the mechanism of layer-by-layer crystal growth, and the mechanism of growth of carbon nanotubes.

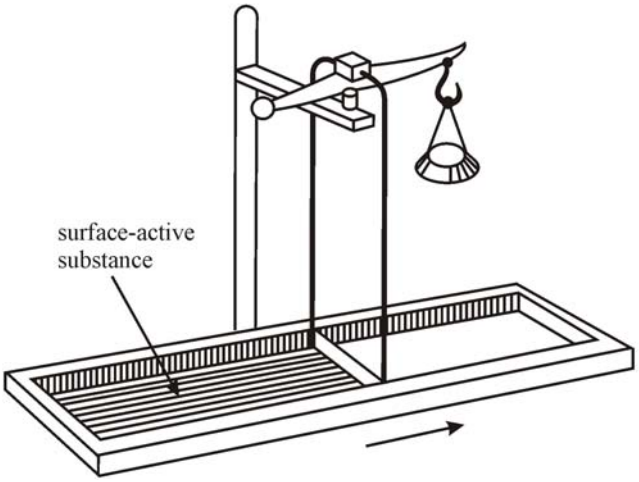


Fig. 26. Langmuir-Adam's Scales scheme.

Understanding of the heterogeneous nature of melting, perhaps, allows us to take another look at the processes of surface and grain boundary diffusion. Scales of Langmuir-Adams, which are designed to measure the surface tension forces, can exemplify this (fig.26). Scales are the movable frame, which touches the surface of the liquid and can easily move. If we add a surfactant (a surface active substance) onto one side of the frame, the frame will move in the direction indicated by the arrow. This process can be considered as a proof that the surface tension force does work on the movement of the frame, applying more energy surface on the other side of the frame. But at the same time, this process can turn out to be a rapid spread of a surface active substance on the surface (which determines the frame speed limit) i.e. the process of surface diffusion. Thus, taking into consideration both processes leads us to the conclusion that high surface diffusion coefficients can be explained by the work done by the surface tension force.
Probably heterogeneous melting process also affects the process of interaction of metal layers with the silicon substrate, which, while observing, first demonstrates the formation of an amorphous (i.e. liquid-like) layer at the interface metal-silicon, and only after that the formation of crystalline silicide compounds.
Diversity of phenomena, which can be explained by the heterogeneous melting participating in them, in our opinion, indicates the reliability of the approach proposed in this work.

6. References

Andrievskii R. A. & Ragulya A. V. (2005) *Nanostructured Materials*, Akademiya, Moscow, [in Russian]

Belousov, S. S., Gavrilov, S. A., Gromov, D. G., Redichev, E. N. & Chulkov I. S. (2007). Investigation and Modeling of Low-Dimensional System Melting Temperature. *Izv. Vyssh. Uchebn. Zaved., Elektron.*, No. 1, pp. 15-21, ISSN 1561-5405 [in Russian]

Buffat, Ph. & Borel, J.-P. (1976). Size effect on the melting temperature of gold particles. *Physical Review A*, Vol. 13, No 6, pp. 2287-2298, ISSN 1050-2947

Buzdugan, A. A., Gavrilov, S. A., Gromov, D. G., Redichev, E. N. and Chulkov, I. S. (2008). Phenomenological Description of Dispersion of 8–60-nm-Thick Silicon Thin Films into Drops on Al_2O_3 Inert Surface. *Semiconductors*, Vol. 42, No. 13, pp. 1487–1491, ISSN 1063-7826

Cahn, J. W. (1977). Critical Point Wetting. *Journal of Chemical Physics*, Vol. 66, p. 3667-3772, ISSN 0021-9606

Cahn, J. W. (1982). Transitions and Phase Equilibria Among Grain Boundary Structures. *Proceedings of Conference on the Structure of Grain Boundaries, Caen, France. Journal de Physique*, Vol. 43, pp. C6 199-213, ISSN 0449-1947

Chernov, A. A. & Mikheev, L. V. (1988). Wetting of Solid Surfaces by a Structured Simple Liquid: Effects of Fluctuations. *Physical Review Letters*, Vol. 60, No 24, p. 2488-2491, ISSN 0031-9007

Chernov, A. A. & Mikheev, L. V. (1989). Wetting and Surface Melting: Capillary Fluctuations vs. Layerwise Short-Range Order. *Physica A*, Vol. 157, pp. 1042-1058, ISSN 0378-4371

Chernov, A. A. & Yakovlev, V. A. (1987). Thin Boundary Layers of the Melt of a Biphenyl Single Crystal and Its Premelting. *Langmuir*, Vol. 3, pp. 635-640, ISSN 0743-7463

Chistyakov, Yu. D. & Rainova, Yu. P. (1979). *Physicochemical Foundations of Microelectronic Technology.* Metallurgiya, Moscow [in Russian].

Dostanko, A. P. , Baranov, V. V. & Shatalov V. V. 1989. *Current Carrying Film Systems for VLSI Application,* Vysheishaya Shkola, Minsk [in Russian].

Emsley J. (1991). *The Elements,* Oxford University Press, Oxford

Geguzin, Ya. E. (1984). *Physics of Sintering,* Nauka, Moscow, [in Russian]

Goodman, R. M., Farrell, H. H. & Somorjai, G. A. (1968). Mean Displacement of Surface Atoms in Palladium and Lead Single Crystals. *Journal of Chemical Physics*, Vol. 48, pp.1046-1051,

Gromov, D. G., Gavrilov, S. A. & Redichev, E. N. (2005). Influence of the Thickness of Copper Films in Cu/W-Ta-N, Cu/C, and C/Cu/C Layered Structures on the Temperature of the Melting-Dispersion Process. *Russian Journal of Physical Chemistry*, Vol. 79, No 9, pp.1394-1400, ISSN 0036-0244

Gromov, D.G. & Gavrilov S.A. (2009). Manifestation of the heterogeneous mechanism upon melting of low-dimensional systems. *Physics of the Solid State*, Vol. 51, No. 10, pp. 2135–2144, ISSN 1063-7834

Gromov, D.G., Gavrilov, S.A., Redichev, E.N. & Ammosov R.M. (2007). Kinetics of the melting-dispersion process in copper thin films. *Physics of the Solid State*, Vol. 49, No 1, pp.178-184, ISSN 1063-7834

Gromov, D.G., Gavrilov, S.A., Redichev, E.N., Chulkov, I.S., Anisimov M. Y., Dubkov, S.V. & Chulkov, S.I. (2010). Non-monotonic dependence of temperature of Au nanometer films dissociation into droplets on their thickness on Al_2O_3 surface. *Applied Physics A*, Vol. 99, pp. 67-71, ISSN 0947-8396

Gromov, D.G., Gavrilov, S.A., Redichev, E.N., Klimovitskaya, A.V. & Ammosov R.M. (2006). The factors that determine the temperature of fusion of Cu and Ni thin films on inert surfaces. *Russian Journal of Physical Chemistry*, Vol. 80, No 10. pp.1650-1655, ISSN 0036-0244

Gromov, D.G., Mochalov, A.I., Klimovitskiy, A.G., Sulimin, A.D. & Redichev, E.N. (2005). Approaches to diffusion barrier creation and trench filling for copper interconnection formation. *Applied Physics A*, Vol. 81, № 7, pp.1337-1343, ISSN 0947-8396

Gusev, A. I. & Rempel A. A. (2004). *Nanocrystalline Materials*, Fizmatlit, Moscow, Cambridge International Science, Cambridge

Gusev, A. I. (2005). *Nanomaterials, Nanostructures, and Nanotechnologies*, Fizmatlit, Moscow, [in Russian]

Handbook of Physical Quantities. (1996). I. S. Grigoriev & E. Z. Meilikhov (Eds.), Energoatomizdat, Moscow, CRC Press, Boca-Raton, FL, United States

Hansen, M. & Anderko, K. (1958). *Constitution of Binary Alloys,* McGraw-Hill Book Company, New York

Kiselev, V. F., Kozlov, S. N. & Zoteev, A. V. (1999). *Fundamentals of the Solid Surface Physics,* Moscow State University, Moscow [in Russian]

Komnik, Yu. F. (1979). *Physics of Metal Films: Dimensional and Structural Effects*, Atomizdat, Moscow, [in Russian]

Müller, P. & Kern R. (2003). Surface melting of nanoscopic epitaxial films. *Surface Science,* Vol. 529, pp. 59-94, ISSN 00396028

Pluis, B., Denier van der Gon, A. W. and van der Veen J. F. (1990). Surface-induced melting and freezing. I. Medium-energy ion scattering investigation of the melting of Pb{*hkl*} crystal faces. *Surface Science*, Vol. 239, p.265-281, ISSN 00396028

Processes of the Real Crystal Formation. (1977). N. V. Belov (Ed.), Nauka, Moscow [in Russian].

Redichev, E.N., Gromov, D.G., Gavrilov, S.A., Mochalov, A.I. & Ammosov, R.M. (2006). Combined method of copper electroplating deposition and low temperature melting for damascene technology. *Proceedings of SPIE*, Vol. 6260, SPIE, Bellingham, WA, pp. 62601H1-62601H1.

Roldugin, V. I. (2008). *Physicochemistry of the Surface,* Intellekt, Dolgoprudnyi, Moscow region, Russia, [in Russian]

Shigeta, Y. & Fukaya, Y. (2004). Structural phase transition and thermal vibration of surface atoms studied by reflection high-energy electron diffraction. *Applied Surface Science.* Vol. 237, p. 21-28, ISSN 0169-4332

Suzdalev, I. P. (2006). *Nanotechnology: Physical Chemistry of Nanoclusters, Nanostructures, and Nanomaterials*, KomKniga, Moscow, [in Russian]

Tartaglino, U., Zykova-Timan, T., Ercolessi F. & Tosatti, E. (2005). Melting and nonmelting of solid surfaces and nanosystems. *Physics Reports*, Vol. 411, No 5, pp. 291-321, ISSN 0370-1573.

Wautelet, M., Dauchot, J. P. & Hecq, M. (2000). Phase diagrams of small particles of binary systems: a theoretical approach. *Nanotechnology*, Vol. 11, pp. 6-9, ISSN: 1550-7033

8

Pressure Effects on Thermodynamics of Polymer Containing Systems

Shichun Jiang[1] and Hongfei Li[2]
[1]*School of Materials Science and Engineering, Tianjin University, Tianjin*
[2]*State Key Laboratory of Polymer Physics and Chemistry,*
Changchun Institute of Applied Chemistry,
Chinese Academy of Sciences, Changchun,
P. R. China

1. Introduction

The role of pressure on the phase diagram of polymer liquids and also polymer mixtures has been intensely studied in the past decades, and there has been increased interest in the effects of pressure on the miscibility of polymers(An et al. 1997; An & wolf, 1998; Blaum & Wolf, 1976; Geerissen et al. 1985; Hammouda & Bauer, 1995; Hosokawa et al. 1993; Lefebvre et al. 2000; Maderek et al. 1983; Rabeony et al. 1998; Wolf & Blaum, 1976, 1977; Wolf & Jend, 1977,1978). One reason is the need for such data to more fully understand polymer miscibility in relation to the various proposed theories and equations of state. Another is the realization that such pressure effects could be important in many situations where such blends are used, e.g., when mixing a blend in an extruder or in forming articles from a blend by injection molding. These needs have led to the development of pressure cells that can be used with both light and neutron scattering such that the phase behavior and interaction strengths of blends can be measured. In past work, a wide range of phase behavior at chosen composition or near critical point of polymer solutions and polymer blends was found(Beiner et al. 1998, 2002; Blaum & Wolf, 1976; Hammouda et al. 1997; Janssen et al. 1993; Lefebvre etal 1999; Schwahn et al. 2001; Wolf & Blaum, 1977; Wolf & Jend, 1977; Zeman &Patterson, 1972; Zeman et al. 1972). There are also many works on the theories about the pressure effects on the thermodynamics of polymer liquid and blends(An et al. 1997; An & wolf, 1998; Dudowicz & Freed, 1995, 2006; Kumar, 2000; Patterson & Robard, 1978; Walsh & Rostami, 1985). As several outstanding problems remain unexplained in these blends, we decided to investigate the dependence on pressure, an independent thermodynamic variable.

The phase behavior of polymer liquids is commonly described in the terms of the lattice model of Flory and Huggins (FH), and the thermodynamics of typical polymer containing systems are understood in the framework of the incompressible random phase approximation. According to original FH theory, the rigorous incompressible system should be unaffected by pressure. In contrast to rigid lattice theories, equation-of-state (EOS) theories are capable of predicting the thermodynamics of polymer containing systems.

Alternatively, the influence of pressure also can be deduced by calculating the pressure coefficient dT_{cr}/dP of the critical solution temperature, i.e., the slope of upper critical solution (UCS) in *P*- *T* space is described in terms of the excess functions using equations 1 and 2(Imre et al. 2001) :

$$\left(\frac{dT}{dP}\right)_c = \lim_{\substack{t\to T_C \\ x\to x_C}} \frac{\partial^2 v_M / \partial x^2}{\partial^2 S_M / \partial x^2} = \lim_{\substack{T\to T_C \\ x\to x_C}} T\frac{\partial^2 v_M / \partial x^2}{\partial^2 h_M / \partial x^2} \tag{1}$$

Here x is composition (mole fraction), v_M, S_M and h_M are solution molar volume, entropy and enthalpy, and v^E and h^E (in equation 2) are the excess volume and enthalpy, respectively. Subscript c denote the critical quantities. The strict equalities in Eq. 1 simplify should v^E and h^E have identical functional forms describing their T and x dependences.

$$\left(\frac{dT}{dP}\right)_c \approx \frac{T_c v_c^E}{h_c^E} \tag{2}$$

Eq. 2 is valid only at the critical composition. If, as is commonly so, $h^E>0$, the sign of $(dT/dP)_c$ is governed by the sign of v_c^E, but one must keep in mind that Eq. 2 is only approximate, and it is formulated in detail in reference (Imre et al. 2001).

The Clausius-Clapeyron equation(Ma, 1982) describes the coexistence line of two phases in a *T*-*P* plane according to $(dT/dP)_\phi = T\Delta V_M / \Delta H_M$, which is similar to Eq.2. Since ΔH_M, the enthalpy change on mixing, is positive at the critical point, the sign of this derivative is controlled by ΔV_M, the volume change on mixing. In most polymer blends investigated(Beiner et al. 1998) to date $(dT/dP)_\phi > 0$ implying that $\Delta V_M > 0$.

In this study we carried out cloud-point measurements for binary system trans-decahydronaphthalene(TD)/ polystyrene (PS) in a pressure range up to 800bar to determine coexistence curves and critical lines. The purpose of this paper is to test whether the Sanchez-Lacombe lattice fluid theory (SLLFT) can describe the equilibrium behavior and pressure effects of polymer containing systems. The spinodals, the binodals, the FH interaction parameter, the enthalpy of the mixing, and volume change of the mixing for TD/PS system were calculated as a function of pressure, temperature and composition on the basis of the SLLFT.

2. Theoretical background

In the lattice fluid theory, as formulated by Sanchez and Lacombe(Lacombe & Sanchez, 1976; Sanchez & Lacombe, 1976), the energy of mixing for binary polymer containing systems is related to the Gibbs energy per mer (indicated by the double bar) of the mixture (index M) and that of the pure components (index1or2) by

$$\Delta\overline{\overline{G}}_M = \Delta G_M / rN = \overline{\overline{G}}_M - (\phi_1\overline{\overline{G}}_1 + \phi_2\overline{\overline{G}}_2) \tag{3}$$

According to SLLFT(Lacombe & Sanchez, 1976; Sanchez & Lacombe, 1976, 1978), the thermal equation of state and Gibbs energy per mer for the pure components and the binary mixtures are expressed in eqs 2-5, respectively(An et al. 1997; An & wolf, 1998),

$$\tilde{\rho}_i^2 + \tilde{P}_i + \tilde{T}_i\left[\ln(1-\tilde{\rho}_i) + (1-\frac{1}{r_i^0})\tilde{\rho}_i\right] = 0 \tag{4}$$

$$\overline{\overline{G}}_i = \frac{G_i}{r_i^0 N_i} = \varepsilon_{ii}^*\left\{-\tilde{\rho}_i + \tilde{P}_i\tilde{\upsilon}_i + \tilde{T}_i\left[(\tilde{\upsilon}_i - 1)\ln(1-\tilde{\rho}_i) + \frac{1}{r_i^0}\ln(\frac{\tilde{\rho}_i}{\omega_i^0})\right]\right\} \tag{5}$$

$$\tilde{\rho}_i^2 + \tilde{P}_i + \tilde{T}_i\left[\ln(1-\tilde{\rho}_i) + (1-\frac{1}{r_i^0})\tilde{\rho}_i\right] = 0 \tag{6}$$

$$\overline{\overline{G}}_M = \frac{G_M}{rN} = \varepsilon^*\left\{-\tilde{\rho} + \tilde{P}\tilde{\upsilon} + \tilde{T}\left[(\tilde{\upsilon} - 1)\ln(1-\tilde{\rho}) + \frac{1}{r}\ln\tilde{\rho} + \frac{\phi_1}{r_1}\ln(\frac{\phi_1}{\omega_1}) + \frac{\phi_2}{r_2}\ln(\frac{\phi_2}{\omega_2})\right]\right\} \tag{7}$$

where r represents the average number of lattice sites occupied by a molecule in the mixtures, rN is the total number of lattice sites occupied in the mixtures, and ϕ_i stands for close-packed volume fractions of component i.

In eqs 4-7, the different symbols have the following meaning: $\tilde{\rho}_i$, $\tilde{P}_i$, $\tilde{\upsilon}_i$, $\tilde{T}_i$ and $\tilde{\rho}$, $\tilde{P}$, $\tilde{\upsilon}$, $\tilde{T}$ are the reduced variables of densities, pressures, volumes, and temperatures for the pure components and for the binary mixtures, respectively, which have the following definitions:

$$\tilde{x} = x/x^* \qquad x = P,\ T,\ P_i,\ T_i \tag{8}$$

$$\tilde{\upsilon}_i = V_i/V_i^* \qquad \tilde{\upsilon} = V/V^* \tag{9}$$

$$V_i^* = r_i^0 N_i \upsilon_i^* \qquad V^* = (r_1 N_1 + r_2 N_2)\upsilon^* \tag{10}$$

$$\tilde{\rho}_i = 1/\tilde{\upsilon}_i \qquad \tilde{\rho} = 1/\tilde{\upsilon} \tag{11}$$

According to the combining rules(Lacombe & Sanchez, 1976; Rodgers et al. 1991), the close-packed volume of a mer, υ^*, and the average number of lattice sites occupied, r, in the mixture are written by:

$$\frac{1}{\upsilon^*} = \frac{\phi_1}{\upsilon_1^*} + \frac{\phi_2}{\upsilon_2^*} \tag{12}$$

$$\frac{1}{r} = \frac{\phi_1}{r_1} + \frac{\phi_2}{r_2} \tag{13}$$

Therefore, total volume of binary mixtures becomes

$$V = rN\upsilon^*\tilde{\upsilon} \tag{14}$$

where the symbols ρ_i^*, P_i^*, υ_i^*, T_i^* and ρ^*, P^*, υ^*, T^* are the close-packed density, scaling pressure, volume per mer and scaling temperature for the pure component and for the binary mixtures, respectively, which can be expressed in terms of the SLLFT (Lacombe & Sanchez, 1976; Sanchez & Lacombe, 1976):

$$P_i^* = \varepsilon_{ii}^* / \upsilon_i^* \qquad P^* = \varepsilon^* / \upsilon^* \tag{15}$$

$$T_i^* = \varepsilon_{ii}^* / k \qquad T^* = \varepsilon^* / k \tag{16}$$

k is the Boltzmann constant. The parameters ω_i and ω_i^0 are the number of configurations available to a r_i-mer and r_i^0-mer in the close-packed state. In this calculation, they are given as their maxima and

$$\omega_i = \delta_{i,\max} r_i / \sigma_i e^{r_i - 1} \qquad \omega_i^0 = \delta_{i,\max}^0 r_i^0 \Big/ \sigma_i^0 e^{r_i^0 - 1} \tag{17}$$

where $\delta_{i,\max}$ and $\delta_{i,\max}^0$, the maxima of flexibility parameters of an r_i-mer and of an r_i^0-mer, should be(Sanchez & Lacombe, 1976):

$$\delta_{i,\max} = z(z-1)^{r_i - 1} \qquad \delta_{i,\max}^0 = z(z-1)^{r_i^0 - 1} \tag{18}$$

σ_i and $\sigma^0{}_i$ represent the symmetry numbers of the r_i-mer and the r_i^0-mer and $\sigma_i = \sigma^0{}_i = 1$. z is the coordination number of the lattice which equals 12 in the present work.

For the pure component i, the total interaction energy per mer is ε_{ii}^*. In a binary mixture, there are three kinds of interaction energies called ε_{11}^*, ε_{12}^* and ε_{22}^* at normal pressure when there are only non-specific interactions between two components, where ε_{12}^* is defined as the interaction energy of a mer belonging to component 1 when it is surrounded by z mers belonging to component 2. Therefore the total interaction energy per mer in a binary mixture can be written as(Lacombe & Sanchez, 1976)

$$\varepsilon^* = \phi_1^2 \varepsilon_{11}^* + 2\phi_1 \phi_2 \varepsilon_{12}^* + \phi_2^2 \varepsilon_{22}^* \tag{19}$$

If there are specific interactions as described by Sanchez and Balazs(Sanchez & Balazs, 1989), the total interaction energy per mer becomes a function of temperature and is derived as following:

$$\varepsilon_T^* = \phi_1^2 \varepsilon_{11}^* + 2\phi_1 \phi_2 f_{12}^* + \phi_2^2 \varepsilon_{22}^* \tag{20}$$

where f_{12}^* represents the total interaction Helmholtz energy and

$$f_{12}^* = \varepsilon_{12}^* + \delta\varepsilon^* - kT \ln\left[\frac{1+q}{1+q\exp(-2\delta\varepsilon^*/zkT)}\right] \tag{21}$$

$\delta\varepsilon^*$ corresponds to the increment of the total 1-2 interaction energy with specific interactions and q is the ratio of the statistical degeneracies of the nonspecific and specific interaction states.

According to our previous calculations(An et al. 1997; An & wolf, 1998), the total interaction energy per mer should be linearly dependent on pressure besides as a function of temperature, *i.e.*,

$$\varepsilon_{T,P}^* = \phi_1^2 \varepsilon_{11}^* + 2\phi_1 \phi_2 g_{12}^* + \phi_2^2 \varepsilon_{22}^* \tag{22}$$

where g_{12}^* describes the total interaction Gibbs energy between two components and

$$g_{12}^* = f_{12}^* + (P - P_0)\upsilon_{12}^* \tag{23}$$

where P_0 is the reference pressure (normally P_0 is chosen as 1 bar or atmosphere) and υ_{12}^* represents the change of interaction volume due to increasing of pressure. When the specific interactions are absent at normal pressure, *i.e.*, $\delta\varepsilon^* = 0$, $P = P_0$, then g_{12}^* will be reduced into ε_{12}^*. In this paper, the spinodals and binodals for the systems TD/PS are calculated by means of a new method without the derivatives of Gibbs energy(Horst, 1995; Horst & Wolf, 1992).

On the basis of the FH lattice model(Flory, 1953), the FH interaction parameter, γ, can be expressed as

$$\gamma = \frac{\dfrac{\Delta G_M^{FH}}{kTV} - \left(\dfrac{\phi_1}{V_1}\ln\phi_1 + \dfrac{\phi_2}{V_2}\ln\phi_2\right)}{\phi_1\phi_2} \tag{24}$$

where V is the total volume of a binary polymer system; V_i and ϕ_i represent the molecular volume and the volume fraction of component i, respectively; and ΔG_M^{FH} is the Gibbs energy of mixing in the FH theory.

According to our previous assumptions(An et al. 1997; Sun et al. 1999), eq. 24 can be rewritten by means of the Sanchez-Lacombe (SL)theory,

$$\gamma = \frac{1}{(\upsilon^*\tilde{\upsilon})\phi_1\phi_2}\left\{\begin{aligned}&\frac{-\tilde{\rho}+\tilde{P}\tilde{\upsilon}}{\tilde{T}} - \left[\phi_1\frac{-\tilde{\rho}_1+\tilde{P}_1\tilde{\upsilon}_1}{\tilde{T}_1} + \phi_2\frac{-\tilde{\rho}_2+\tilde{P}_2\tilde{\upsilon}_2}{\tilde{T}_2}\right] + (\tilde{\upsilon}-1)\ln(1-\tilde{\rho}) + \frac{1}{r}\ln\tilde{\rho}\\ &-\left[\phi_1\left((\tilde{\upsilon}_1-1)\ln(1-\tilde{\rho}_1) + \frac{1}{r_1^0}\ln\tilde{\rho}_1\right) + \phi_2\left((\tilde{\upsilon}_2-1)\ln(1-\tilde{\rho}_2) + \frac{1}{r_2^0}\ln\tilde{\rho}_2\right)\right]\\ &+\phi_1\left(\frac{1}{r_1^0}\ln\omega_1^0 - \frac{1}{r_1}\ln\omega_1\right) + \phi_2\left(\frac{1}{r_2^0}\ln\omega_2^0 - \frac{1}{r_2}\ln\omega_2\right)\end{aligned}\right\} \tag{25}$$

3. Pressure induced phase separation of TD/PS polymer solutions

The cloud-point temperatures (T_{cl}'s) of trans-decahydronaphthalene (TD)/ polystyrene (PS, M_w=270 000) solutions were determined by light scattering measurements over a range of temperatures(1- 16 °C), pressures(100 to 900 bar), and compositions (4.2-21.6 v% polymer). The system phase separates upon cooling and T_{cl} was found to increase with rising pressure for constant composition. In the absence of special effects this finding indicates positive excess volume for the mixing. Special attention was paid to the demixing temperatures as function of pressure for different polymer solutions and the plots in *T-*ϕ plane (where ϕ signifies volume fractions). The cloud-point curves of polymer solution under pressures were observed for different compositions, which demonstrates that pressure has more effect on the TD/PS solutions that far from the critical point than that near the critical point. Figure 1 shows isopleths (i.e., lines of constant composition of the solution) cloud-point curves observed for TD/PS polymer solutions on a temperature-pressure plane. The open circles indicate experimental points and the solid curves describe the behavior of the data points. Curves were measured at ϕ = 4.2, 8.4, 12.8, 17.2, and 21.6(v% at P=1bar), respectively.

The clout point curves at each composition behave similarly to one another. The region above the curve is the one-phase homogeneous region, while the region below represents the two-phase region. As shown, the phase separation pressures increase with raising temperature.

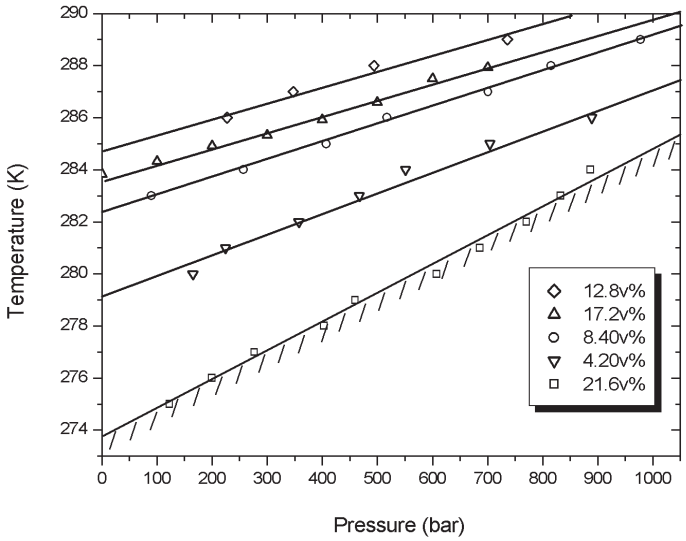


Fig. 1. Pressure dependence of phase transition temperature for TD/PS polymer solution at the indicated compositions (In SI units: 1bar = 10^5 N m^{-2}).

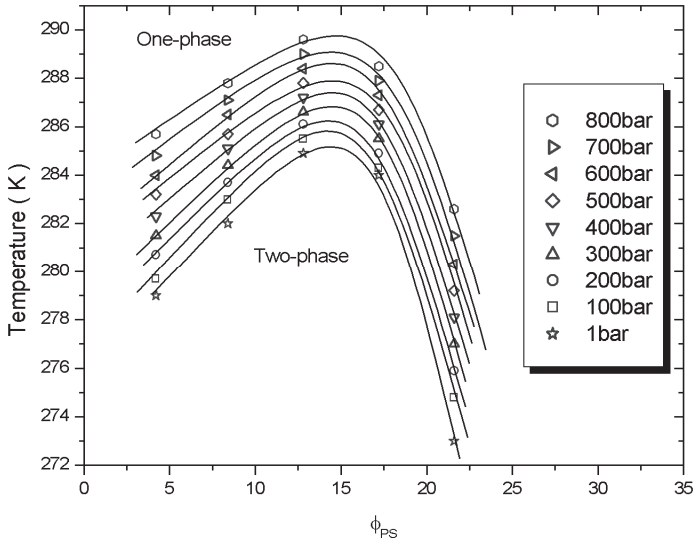


Fig. 2. Coexistence curves of TD/PS system at various pressures on a T-ϕ plane.

With the pre-selected polymer concentrations, isobaric critical lines can be constructed from the isopleths in Figure 1. The thus obtained critical lines from 100 bar to 800 bar, respectively, show in Figure 2 on a T-ϕ plane. In Figure 2 the coexistence boundaries are described on a T-ϕ plane at the indicated pressures (bar)(the boundary under 1 bar was extrapolated from Figure 1). The shape of the coexistence curve depends only slightly on pressure, which can be recognized by comparing the curve at 100bar with the curve at 800 bar, and this system shows that it is an upper critical solution temperature (UCST) behavior system.

PVT Diagram

The PVT property of PS was determined with a PVT-100 dilatometer. The experimental data was determined as a series of isotherms. Figure 3 presents the selected isobars with the interval of pressure 200bar and shows the specific volume, V_{sp}, of PS as a function of temperature and pressure. The open circles represent the experimental data, the solid lines describe the behavior of the data points in the equilibrium melt state and extrapolated to the glassy state of PS, and the dot lines AB, CD, EF, and GH separate the diagram into five parts; with the increase of temperature, the five parts includes three parts that are glassy (the range between AB and CD), exceptional (the range between CD and EF), and melt states (the range between EF and GH), separately. In the exceptional area, a glass is reformed by pressurization from the melt during isothermal compression (Schmidt & Maurer, 1998; Tait 1888). From Figure 3, it can be seen that the change of the PS specific volume in the glassy state is much smaller than that in the melt state with the changes of temperature and pressure. The PS specific volume decreases rapidly with the increase of pressure in the melt state. From the experimental results of the phase behavior of TD/PS polymer solution under pressure we knew that during the experiment for the measurements of TD/PS polymer solution phase behavior under pressure, PS is in the equilibrium state and at relative low temperature that is the glassy temperature for bulk PS. We deduced that the scaling parameters of PS fitted from low temperature and high temperature must be different. In order to get the data in equilibrium state at low temperature, we extrapolated the lines from the melt state to the glassy state. Both of the data from the equilibrium melt state and the extrapolated were used to obtain two series scaling parameters of PS by fitting the PVT data to the SL equations of state as described above and to estimate which series scaling parameters could be used to describe the thermodynamics of TD/PS polymer solution under pressure better.

Scaling Parameters

Fitting the PVT data that in the range between AB and CD (extrapolated to the glassy state temperature in Figure 3) and in the range between EF and GH (equilibrium melt state in Figure 3) to the SL equations of state, *i.e.*, eqs 4 and 6, we obtained two series scaling parameters, each includes three scaling parameters P^*, ρ^*, and T^* (or ε^*, υ^*, and r) of PS. The scaling parameters of PS and TD are listed in Table 1. In order to evaluate the scaling parameters that obtained from different state of the sample, we used them for the calculations in SLLFT based on the data in Figure 2 in atmospheric pressure. The calculated spinodals are shown in Figure 4. From Figure 4, it can be seen that the critical temperature (T_c) calculated with the scaling parameters obtained from the extrapolated data of PS is close to the experimental results (Figure 4b). We can know that the scaling parameters

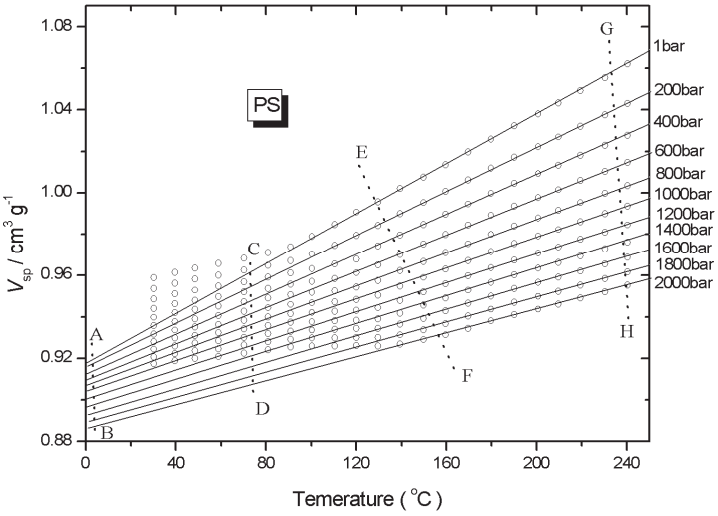


Fig. 3. Specific volume, V_{sp}, of PS as a function of temperature and pressure. The open circles represent the experimental data, respectively, and the solid lines AB, CD, EF and GH separate the diagram into five parts, which including glassy, exceptional and melt states.

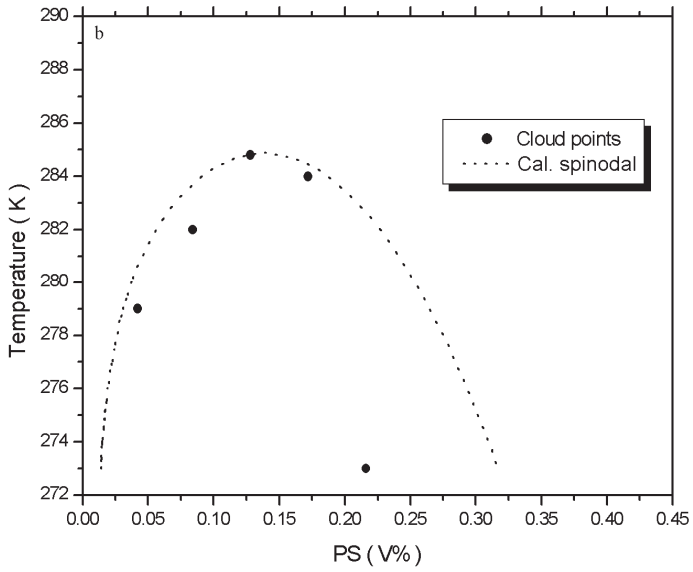


Fig. 4. The calculated spinodals of TD/PS polymer solutions by SL theory with two series of PS scaling parameters and compared with the experimental result at atmosphere (a: the calculation according to the scaling parameter of PS* in Table 1; b: the calculation according to the scaling parameters of PS in Table 1).

obtained from the glassy state temperature (the temperature range for the measurements of TD/PS phase behavior under pressure) of PS fit the experimental results better than that obtained from the equilibrium melt state (Figure 4a). The reason is that the phase separation behavior of the polymer solutions was measured in the temperature range that bulk PS is at glassy state temperature but in equilibrium state. The further calculations were based on the series scaling parameters obtained from the data of PS that extrapolated to the glassy temperature in Figure 3.

Pressure dependence of phase separation behaviors

From eqs 20, 21 and 23, the interaction Gibbs energy without specific interactions can be re-expressed into(An & Wolf, 1998)

$$g_{12}^{*} = \varepsilon_{12}^{*}\left(P = P_0\right) - \left(P - P_0\right)\upsilon_{12}^{*} \tag{26}$$

where $\varepsilon_{12}^{*}\left(P = P_0\right)$ for infinite molecular weight results to be –12.7K and for the present system the corresponding evaluation yields $\upsilon_{12}^{*}N_A = 0.033\,\mathrm{cm^3/mol}$, where N_A is the Avogadro number.

	T_i^*/K	P_i^*/bar	ρ_i^* / g cm^{-3}
TD	621	3110	0.935
PS	714	4432	1.068
PS*	725	4041	1.025

Table 1. Scaling parameters of TD and PS.The scaling parameters for TD were taken from the literature (Inre et al. 2001); those of PS stem from the evaluation of PVT data in the melt state and PS* stem from the evaluation of PVT data extrapolated from the measurements at higher temperature into the *T*-region of experiment for TD/PS polymer solution.

Spinodals and binodals under pressure

Like the spinodals and binodals of ternary and quaternary polymer blends were calculated with the method(Horst 1995; Horst & Wolf, 1992) (the knowledge of the first and second derivatives of ΔG with respect to the composition variables is not required), the spinodals and the binodals of the TD/PS system were calculated with the SL theory under different pressures.

Figure 5 shows the calculated results of the spinodals and the binodals compared with the experimental data as showed in Figure 2. The dashed and the solid lines represent the spinodals and the binodals calculated with the SL theory under indicated pressures, respectively; the solid circles represent the experimental data, respectively, which were obtained from Figure 2. In Figure 5 the qualitative agreement between the spinodals and the binodals calculated and the experimental cloud points is acceptable at different pressure. From Figure 5, it can be seen that the critical temperatures (T_c) calculated by means of FL theory under different pressure increase with pressure.

Interaction energy parameters under pressure

In previous paper (An et al, 1997; An & Wolf, 1998), the only one adjustable interaction energy parameter, g_{12}^{*}/k, in the SLLFT was evaluated by comparing the theoretical and the experimental phase diagrams at atmospheric pressure. For the high pressure, g_{12}^{*}/k was

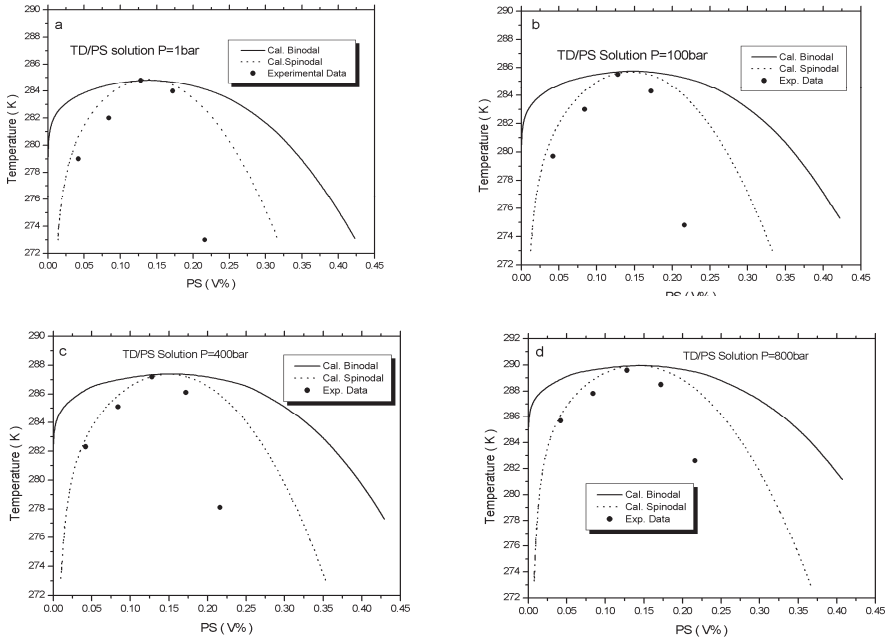


Fig. 5. Spinodals and binodals of TD/PS polymer solutions calculated with SL theory under the indicated pressures (the experiment data obtained from Figure 3).

calculated by eq. 25, and then used for the phase diagram calculation under pressure. In this work, however, g_{12}^{*}/k can be fitted from experimental cloud points, and the fitted g_{12}^{*}/k with pressure was shown in Figure 6. The solid circles are the fitted results from experimental cloud points and the dot curve represents the values of fitted data. From Figure 6 we can see that the interaction energy parameter deduces with pressure, which is different from the result that calculated from eq. 26. In that equation, the interaction energy parameter will be linear reduced with pressure, but for the result from Figure 6 is not the case.

FH interaction parameter

In the Flory-Huggins theory model (Koningsveld & Staverman, 1968), the interaction parameter (γ) was considered as a function of composition and temperature. The Flory-Huggins interaction parameter was dependent on the concentration of the components, temperature and pressure in the mean-field lattice-gas (MFLG) model and semi-phenomenological treatment(van Opstal & Koningsveld, 1992). According to the results of our calculation, the Flory-Huggins interaction parameter should be as a function of *P*, *T*, and ϕ for the system of TD/PS.

The FH interaction parameter, γ, was calculated with eq. 24 under different pressures. Figure 7 shows the temperature and blend composition dependences of γ for the TD/PS polymer solutions under the indicated pressures.

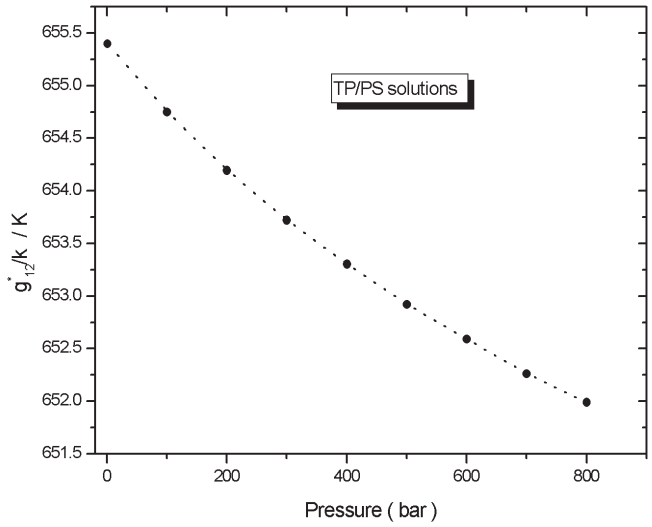


Fig. 6. Fitted interaction energy parameter under pressure with experimental results by the means of SL theory.

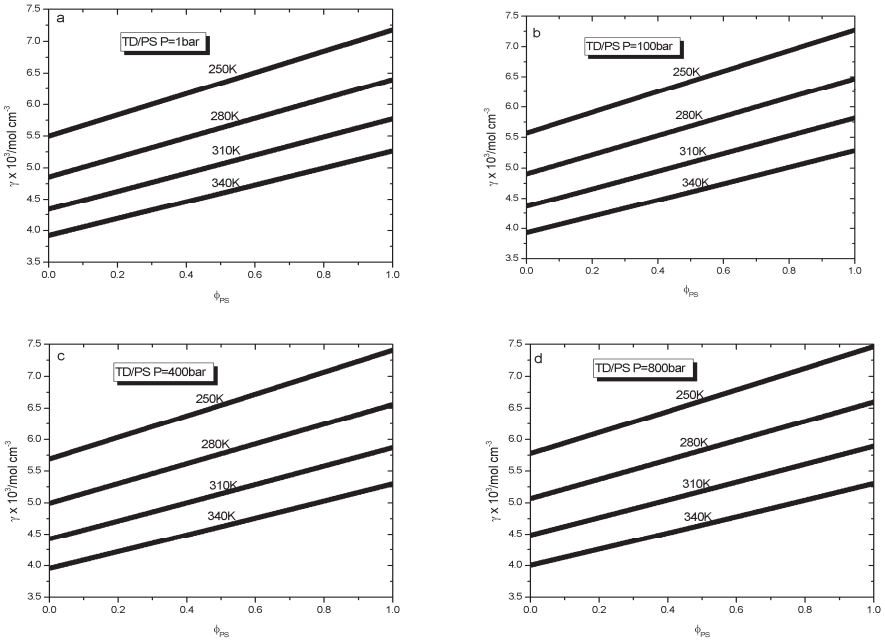


Fig. 7. Temperature and polymer solution composition dependences of the FH interaction parameter, γ , for the TD/PS polymer solutions under the indicated pressures.

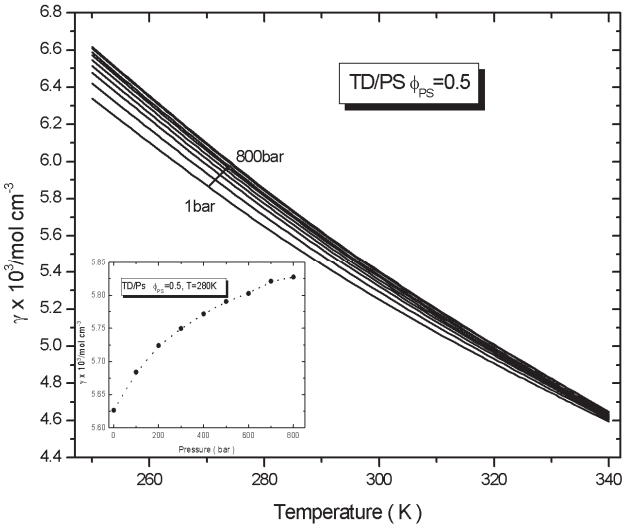


Fig. 8. Temperature dependences of the FH interaction parameter, , for the TD/PS polymer solutions at the indicated composition (the curves are the isobaric curves under 1, 100, 200, 300, 400, 500, 600, 700 and 800bar. The inserted figure shows pressure dependences of the FH interaction parameter, , for the TD/PS polymer solutions with fixed composition and temperature).

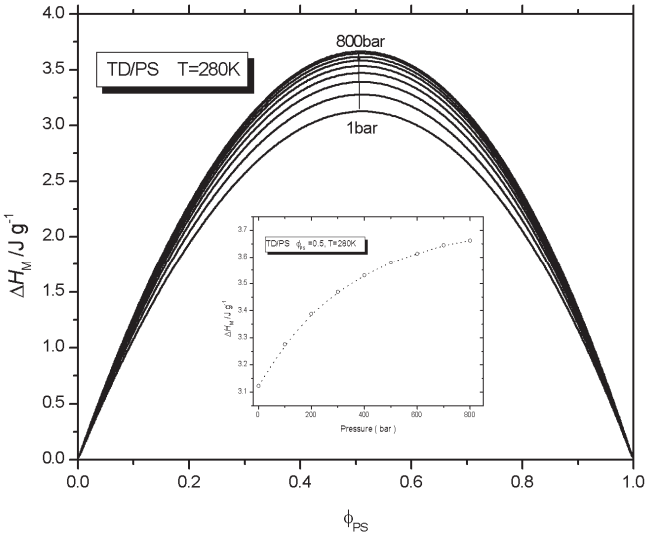


Fig. 9. Pressure and polymer solution composition dependences of the enthalpy of the mixing, ΔH_M , for the TD/PS polymer solutions at the indicated temperature (the inserted figure shows pressure dependences of the enthalpy of the mixing, ΔH_M , for the TD/PS polymer solutions with fixed composition and temperature).

Figure 7 shows the dependence of Flory-Huggins interaction parameter on pressure and temperature against concentration. We can see that the Flory-Huggins interaction parameter is linear increasing with the composition of PS increases. It also can be observed that the FH interaction parameter reduces with the increasing of temperature at the fixed composition and pressure, which agrees with the experiment result that this system exhibits UCST behavior. The effects of pressure and temperature on γ are identical with that derived by de Loos *et al*. for the systems polyethylene/ethylene(de Loos et al. 1983).

Figure 8 shows the plots of the Flory-Huggins interaction parameters against temperature at fixed PS concentration ($\phi_{PS} = 0.5$) under different pressures. It is shown that γ is almost linear with the temperature and decreases with the increasing of temperature, and the inserted figure shows that γ increases with pressure at fixed temperature and composition, which is consistent with the experimental measured results that TD/PS system shows low critical solution pressure (LCSP) behavior.

From these results calculated, the relations of temperature, pressure, and concentration of TD/PS polymer solutions are presented. The influence of pressure on calculated Flory-Huggins interaction parameter of TD/PS polymer solution shows that Flory-Huggins interaction parameter increases on raising the pressure, i.e., miscibility is reduced.

Enthalpy of mixing

It is known that the Gibbs energy can be split into its enthalpic and entropic parts in the SL theory(Lacombe & Sanchez, 1976; Sanchez & Lacombe, 1978). The enthalpy of mixing per mer, $\Delta\overline{\overline{H}}_M$, is given by

$$\Delta\overline{\overline{H}}_M = \frac{\Delta H_M}{rN} = \overline{\overline{H}}_M - \varphi_1\overline{\overline{H}}_1 - \varphi_2\overline{\overline{H}}_2 = \varepsilon^*\left(-\tilde{\rho} + \tilde{P}\tilde{\upsilon}\right) - \phi_1\varepsilon_{11}^*\left(-\tilde{\rho}_1 + \tilde{P}_1\tilde{\upsilon}_1\right) - \phi_2\varepsilon_{22}^*\left(-\tilde{\rho}_2 + \tilde{P}_2\tilde{\upsilon}_2\right) \quad (27)$$

where $\overline{\overline{H}}_M$ and $\overline{\overline{H}}_i$ are the enthalpies per mer of the blend and component i , respectively. In order to investigate the pressure effects on ΔH_M of TD/PS polymer solution, we calculated ΔH_M under different pressures with eq. 27 at fixed temperature, and the results are shown in Figure 9. The insert Figure shows the pressure effects on ΔH_M in TD/PS polymer solution, when the composition ϕ_{PS} =0.5 and the temperature is 280K. From Figure 9, we can see that the calculated ΔH_M increases with pressure, which means that increasing pressure on TD/PS polymer solution is not favorable for TD/PS system to be homogeneous polymer solution.

Volume change of mixing

As the Clausius-Clapeyron indicates that the phase behavior of a mixture under pressure is governed by the sign of the excess volume of mixing; for polymer solution is the same. Therefore we will analyse the reduced excess volume or the volume change of mixing. In the SL theory (Lacombe & Sanchez, 1976; Sanchez & Lacombe, 1978), volume changes upon mixing are calculable. The volume change of mixing, ΔV_M , is respectively given by

$$\Delta V_M = V - V_1 - V_2 = V^*\tilde{\upsilon} - V_1^*\tilde{\upsilon}_1 - V_2^*\tilde{\upsilon}_2 \quad (28)$$

where V_i and V are the volumes of pure component i and the solution, respectively.

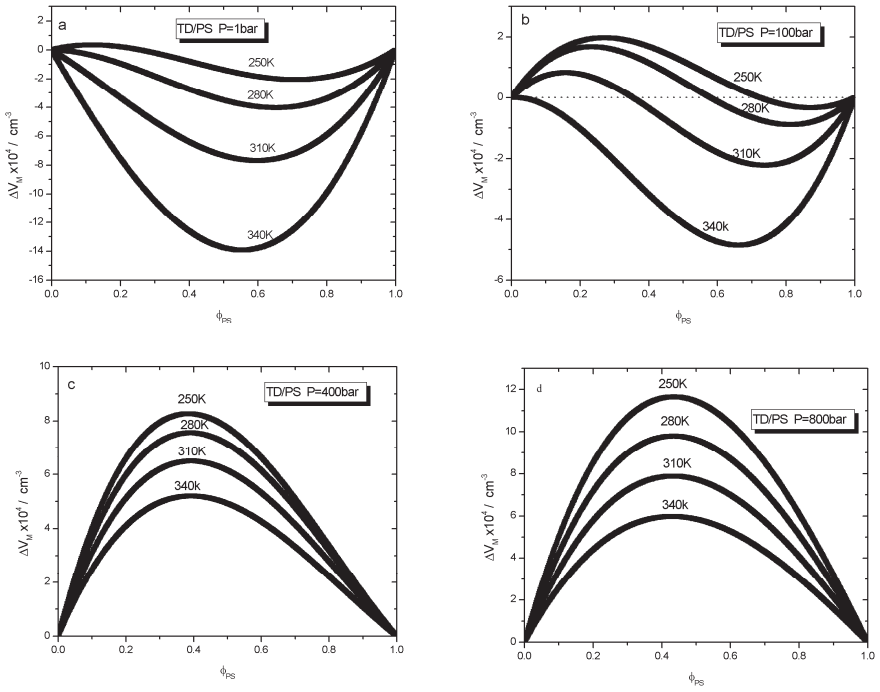


Fig. 10. Temperature and polymer solution composition dependences of the volume change of the mixing, ΔV_M, for the TD/PS polymer solutions under the indicated pressures.

Figure 10 shows the temperature and polymer solution composition dependences of ΔV_M calculated with eq. 28 for the TD/PS solutions under the indicated pressures. For fixed pressure and composition, the volume change of mixing reduces with temperature increasing. Within the calculated range, for low pressure and high temperature it is negative but becomes positive on raising the pressure or reducing the temperature.

In order to observe the pressure effects on the volume change of mixing for TD/PS polymer solution, we calculated the excess volume of TD/PS polymer solutions at 280K for various pressures. Figure 11 shows the volume change of TD/PS polymer solutions vs composition calculated with eq. 28 at 280K for the indicated pressures. We can know more clearly from Figure 11 that for low pressure the volume change is negative but becomes positive on raising the pressure. The change of sign occurs at about 100bar. At this pressure ΔV_M is still positive for low concentrations of PS, while for high PS concentrations it is already negative. This means that in this pressure range the miscibility on the PS-rich side is improved by raising pressure and is lowered on the TD-rich side, i.e., the phase diagram becomes distorted. Therefore, the shift of the temperatures of mixing and demixing on varying the pressure is small, so that the shape of the phase diagram hardly changes.

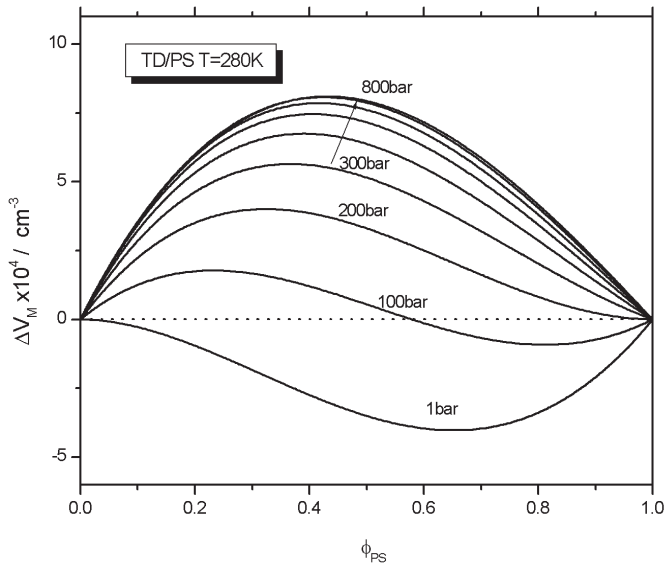


Fig. 11. Pressure and polymer solution composition dependences of the volume change of the mixing, ΔV_M, for the TD/PS polymer solutions at the indicated temperature.

4. Pressure induced miscibility in PEO/(P (EO-b-DMS) mixtures

The cloud-point temperatures (T_{cl}'s) of poly (ethylene oxide) (PEO) and poly (ethylene oxide-b-dimethylsiloxane) (P (EO-b-DMS)) homopolymer and block-oligomer mixtures were determined by turbidity measurements over a range of temperatures (105 to 130 °C), pressures (1 to 800 bar), and compositions (10-40 wt% PEO). The system phase separates upon cooling and T_{cl} was found to decrease with rising pressure for constant composition. In the absence of special effects this finding indicates negative excess volumes. Special attention was paid to the demixing temperatures as function of pressure for different polymer mixtures and the plots in T-Φ plane (where Φ signifies volume fractions). The cloud-point curves of polymer mixture under pressures were observed for different compositions. Figure 12 shows isopleths (i.e., curves of constant composition of the mixtures) cloud-point curves observed for polymer mixtures on a temperature-pressure plane. The open circles indicate experimental points and the solid curves are just a guide for the eye. The curves were measured at Φ = 9.7, 19.4, 29.2, and 39.1(PEO v% at P=1bar), respectively. The region above the curves is the one-phase homogeneous region, while the region below represents the two-phase region. As shown, the demixing pressures increase with increasing temperature.

This binary system has been studied at four compositions under pressure. With the reservations concerning polymer concentrations, isobaric critical lines can be constructed from the isopleths of Figure 12. The thus obtained critical lines from 1 bar to 600 bar, respectively, show in Figure 13 on a T-ϕ plane(the could points at P=1bar were measured with laser apparatus at atmosphere). In Figure 13 the coexistence boundaries are described on a T-ϕ plane at the indicated pressures (bar). The shape of the coexistence curve depends

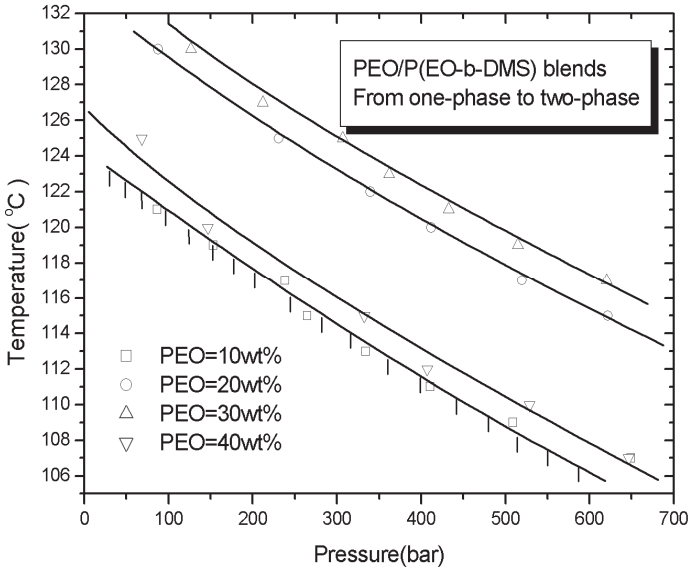


Fig. 12. Pressure dependence of T_{cp} for the system PEO/P(EO-b-DMS) (In SI units: 1bar = 10^5 N m^{-2}).

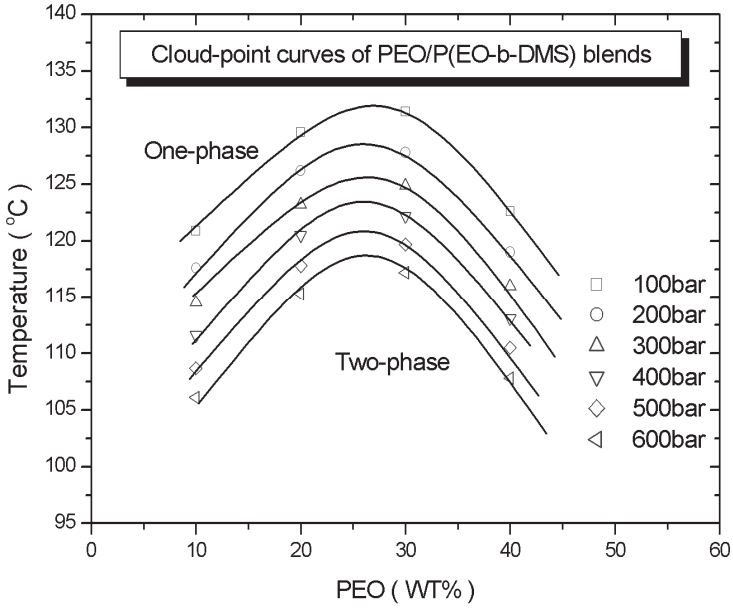


Fig. 13. Coexistence curves of PEO/P(EO-b-DMS) system at various pressures on a T-ϕ plane.

only slightly on pressure, which can be recognized by comparing the curve at 1bar with the curve at 600 bar, and this system shows that it is a UCST behavior system.

Scaling Parameters.

Within the framework of the Sanchez-Lacombe theory, we set up the following combining rules for the molecular weight and the scaling parameters for P(EO-b-DMS) block copolymer:

$$M_{COP}^{SL} = M_{P(EO-b-DMS)} + M_f \tag{29}$$

$$P_{COP}^{*SL} = \phi_{PEO} P_{PEO}^{*} + \phi_{PDMS} P_{PDMS}^{*} + \phi_{PEO}\phi_{PEMS} P_f^{*} \tag{30}$$

$$T_{COP}^{*SL} = \phi_{PEO} T_{PEO}^{*} + \phi_{PDMS} T_{PDMS}^{*} + \phi_{PEO}\phi_{PDMS} T_f^{*} \tag{31}$$

$$\rho_{COP}^{*SL} = \phi_{PEO} \rho_{PEO}^{*} + \phi_{PDMS} \rho_{PDMS}^{*} + \phi_{PEO}\phi_{PDMS} \rho_f^{*} \tag{32}$$

where M_{COP}^{SL}, P_{COP}^{*SL}, T_{COP}^{*SL} and ρ_{COP}^{*SL} are the modified molecular weight and reduced pressure, temperature, density of the block copolymer according to the new combining rules for Sanchez-Lacombe theory; $M_{P(EO-b-DMS)}$, P_{PEO}^{*}, P_{PDMS}^{*}, T_{PEO}^{*}, T_{PDMS}^{*}, ρ_{PEO}^{*}, and ρ_{PDMS}^{*} are the real molecular weight for the block copolymer, the reduced pressure, temperature and density of PEO and PDMS; M_f, P_f^{*}, T_f^{*}, and ρ_f^{*} are the adjusting parameters for the molecular, reduced pressure, temperature, and the density of the block copolymer; ϕ_{PEO} and ϕ_{PDMS} are the volume fraction of PEO and PDMS in the block copolymer. According to the new combining rules, we can get the characteristics of the homopolymer and the block copolymer, and they are collected in Table 2.

The scaling parameters for PEO were taken from the literature(Konowalow, 1903), those of P(EO-b-DMS) were calculated with the chemical structure of P(EO-b-DMS) and P(EO-b-DMS)* were calculated with the new combine rules for the diblock copolymer parameters. In this paper, the parameters of P(EO-b-DMS)* were used for the calculation as the parameters of block copolymer component.

In order to evaluate the new combining rules and the scaling parameters that obtained in Table 2, we used them for the calculations in SLLFT based on the experimental data in Figure 14 which obtained at atmospheric pressure. The calculated spinodals with the different scaling parameters and the measured data are shown in Figure 14. From Figure 14, it can be seen that the critical temperature (T_c) calculated with the scaling parameters obtained from the new combining rules for P(EO-b-DMS) is close to the experimental results (Figure 14b). We can know that the scaling parameters obtained from the new combining rules for the block copolymer fit the experimental results better than that obtained from the molecular structure (Figure 14a). The reason is that PEO/P(EO-b-DMS) polymer mixture is a special polymer blend system, the structures of the two components i.e., PEO and P(EO-b-DMS) are special, and the molecular weight of PEO is much larger than that of the block copolymer; at the same time, maybe the two components constitute the cluster in the mixtures, and the structures of the components are not the original structures of PEO and P(EO-b-DMS). So the molecular weight and the scaling parameters obtained with the new combining rules for the block copolymer fit the experimental data better. The further calculations were based on the series scaling parameters obtained from the new combining rules for the block copolymer.

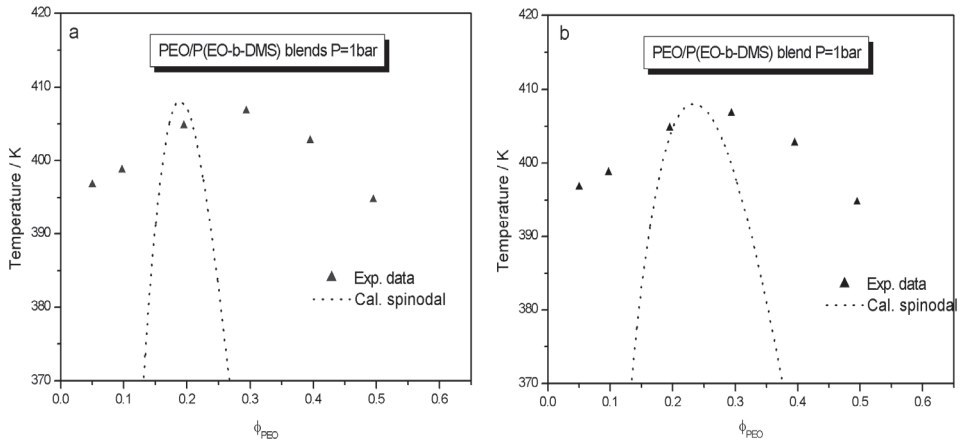


Fig. 14. The calculated spinodals of PEO/P(EO-b-DMS) polymer mixtures by SL theory compared with the experimental result (a: the calculation according to the scaling parameter of P(EO-b-DMS) in Table 1; b: the calculation according to the scaling parameters of P(EO-b-DMS) * in Table1).

Polymer	M_w/kg mol^{-1}	T_i*/K	P_i*/bar	ρ_i^* / g cm^{-3}
PEO	35.0	656	4922	1.178
P(EO-b-DMS)	1.8	609	4413	1.160
P(EO-b-DMS)*	3.0	622	4555	1.165

Table 2. Sample characteristics of PEO and P(EO-b-DMS).

Spinodals and binodals under pressure

Like the spinodals and binodals of ternary and quaternary polymer blends were calculated with the method(Horst 1995; Horst & Wolf, 1992) (the knowledge of the first and second derivatives of ΔG with respect to the composition variables is not required), the spinodals and the binodals of the PEO/P(EO-b-DMS) system were calculated with the SL theory under different pressures. Figure 15 shows the calculated results of the spinodals and the binodals compared with the experimental data as showed in Figure 14. The dashed and the solid lines represent the spinodals and the binodals calculated with the SL theory under indicated pressures, respectively; the open circles represent the experimental data, respectively, which were obtained from Figure 14. In Figure 4 the qualitative agreement between the spinodals and the binodals calculated and the experimental cloud points is acceptable at different pressure. From Figure 15, it can be seen that the critical temperatures (T_c) calculated by means of FL theory under different pressure increase with pressure.

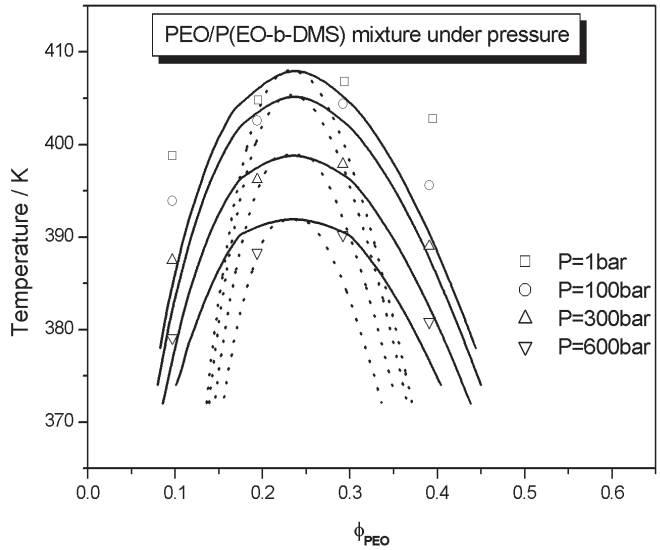


Fig. 15. Spinodals and binodals of P(EO-b-DMS) polymer mixtures calculated with SL theory (the experiment data obtained from Figure 13).

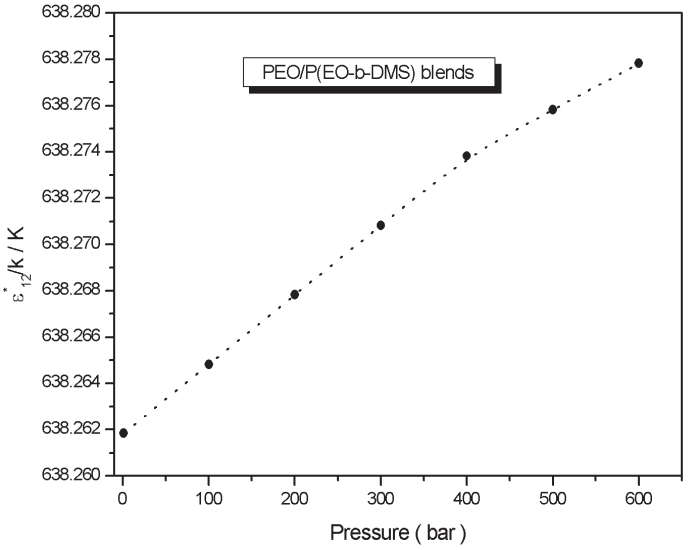


Fig. 16. Fitted interaction energy parameter under pressure with experimental results by the means of SL theory.

Interaction energy parameters under pressure

In the reports(An et al. 1997; An & Wolf, 1998), the only one adjustable interaction energy parameter, g_{12}^{*}/k, in the SLLFT was evaluated by comparing the theoretical and the

experimental phase diagrams at atmospheric pressure. For the high pressure, g_{12}^{*}/k was calculated by eq. 26, and then used for the phase diagram calculation under pressure. In this work, however, g_{12}^{*}/k can be fitted from experimental cloud points, and the fitted g_{12}^{*}/k with pressure was shown in Figure 16. The solid circles are the fitted results from experimental cloud points and the dot curve represents the values of fitted data. From Figure 16 we can see that the interaction energy parameter increases with pressure, which is different from the result that calculated from eq. 26. In that equation, the interaction energy parameter will be linear coherent with pressure, but for the result from Figure 16 is not the case.

FH interaction parameter

In the Flory-Huggins theory modified(Koningsveld & Staverman, 1968), the interaction parameter (γ) was considered as a function of composition and temperature. The Flory-Huggins interaction parameter was dependent on the concentration of the components, temperature and pressure in the MFLG model and semi-phenomenological treatment(van Opstal & Koningsveld, 1992). According to the results of our calculation, the Flory-Huggins interaction parameter should be as a function of P, T, and ϕ for the system of PEO/P(EO-b-DMS).

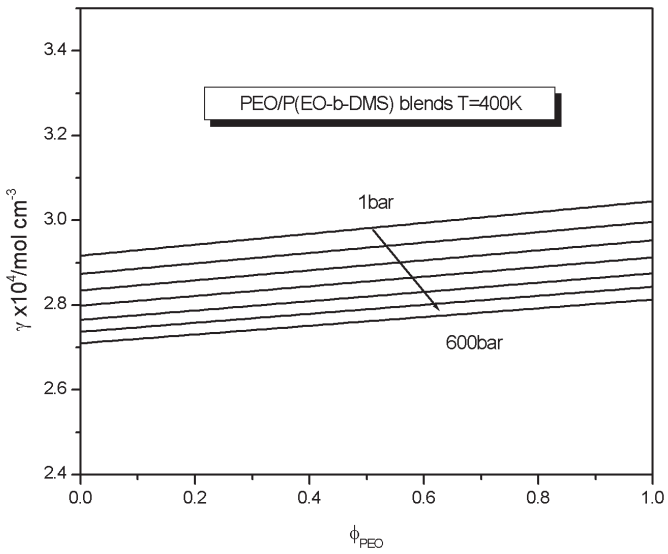


Fig. 17. Pressure and polymer mixture composition dependences of the FH interaction parameter, γ, for the P(EO-b-DMS) polymer mixturess under the indicated temperature.

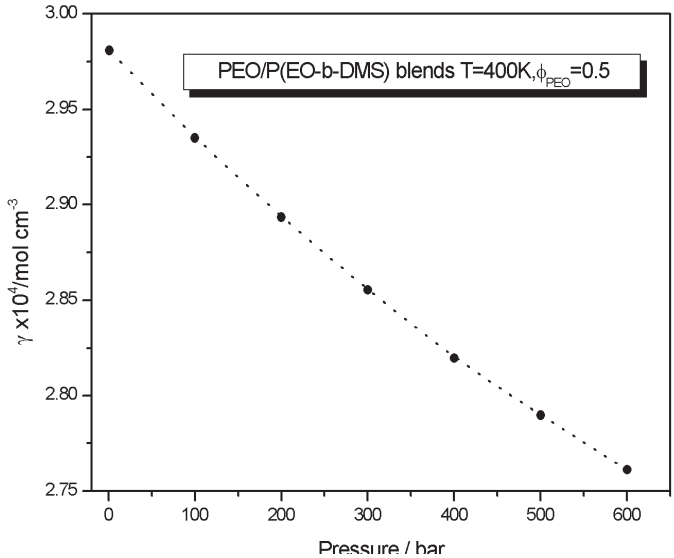


Fig. 18. Pressure dependences of the FH interaction parameter, γ, for the P(EO-b-DMS) polymer mixtures at the indicated composition and temperature.

The FH interaction parameter, γ, was calculated with eq. 25 under different pressures. Figure 17 shows the temperature and blend composition dependences of γ for the PEO/P(EO-b-DMS) mixtures under the indicated pressures.

Figures 17 shows the dependence of Flory-Huggins interaction parameter on pressure and temperature against concentration. We can see that the Flory-Huggins interaction parameter is almost linear increasing the composition of PEO increases at fixed temperature and pressure. It also can be observed that the FH interaction parameter reduces with the increasing of temperature at the fixed composition and pressure, which agrees with the experiment result that this system exhibits UCST behavior. The effects of pressure and temperature on γ are identical with that derived by de Loos *et al.* for the systems polyethylene/ethylene(de Loos et al. 1983).

Figure 18 shows the plots of the Flory-Huggins interaction parameters against PEO concentration at fixed temperature (T=400K) under different pressures. It is shown that γ reduces with pressure at the fixed PEO concentration in the mixture, which is consistent with the experimental measured result that PEO/P(EO-b-DMS) system shows pressure-induced miscibility behavior.

From these results calculated, the relations of temperature, pressure, and concentration of PEO/P(EO-b-DMS) mixtures are presented. The influence of pressure on calculated Flory-Huggins interaction parameter of PEO/P(EO-b-DMS) mixtures shows that Flory-Huggins interaction parameter deduces on raising the pressure, i.e., miscibility is enhanced.

Enthalpy of mixing

In order to investigate the pressure effects on ΔH_M of PEO/P(EO-b-DMS) mixture, we calculated ΔH_M under different pressures with eq. 30 at 400K, and the results are shown in Figure 19. Figure 20 shows the pressure dependences of calculated ΔH_M on the fixed

temperature and composition. From Figure 19 and 20, we can see that the calculated ΔH_M is not affected so much by the pressure and temperature. But in all calculated case, $\Delta H_M > 0$, which means that increasing pressure on PEO/P(EO-b-DMS) mixture is favorable for PEO/P(EO-b-DMS) system to be homogeneous polymer blend.

Volume Change of Mixing.

The volume change of mixing, ΔV_M, and the fractional volume change of mixing, $\Delta V_M / V$, is respectively given by eq. 28 and

$$\frac{\Delta V_M}{V} = 1 - \frac{1}{\tilde{\upsilon}}(\phi_1 \tilde{\upsilon}_1 + \phi_2 \tilde{\upsilon}_2) \tag{33}$$

Figure 21 shows the temperature and polymer mixture composition dependences of $\Delta V_M / V$ calculated with eq. 33 for the PEO/P(EO-b-DMS) mixtures under the indicated pressures. For fixed pressure and composition, the volume change of mixing increases with temperature increasing. Within the calculated range $\Delta V_M / V$ is negative, which means this system exhibits abnormal phase behavior under pressure.

In order to observe the pressure effects on the volume change of mixing for PEO/P(EO-b-DMS) mixture, we calculated the excess volume of PEO/P(EO-b-DMS) mixtures at 400K for various pressures. Figure 22 shows $\Delta V_M / V$ of PEO/P(EO-b-DMS) mixtures vs composition calculated with eq. 33 at 400K for the indicated pressures. We can know more clearly from Figure 22 that for low pressure the volume change is more negative and $\Delta V_M / V$ decreases on raising the pressure.

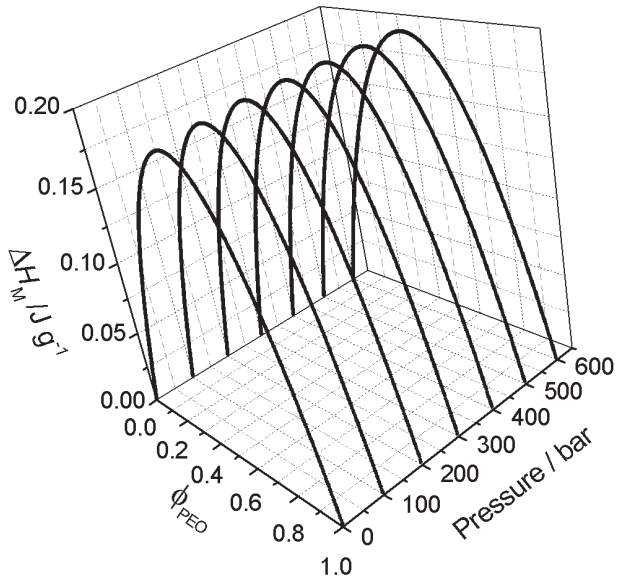


Fig. 19. Pressure and composition dependences of the calculated enthalpy of the mixing, ΔH_M, for the P(EO-b-DMS) polymer mixtures at the indicated temperature (T=400K).

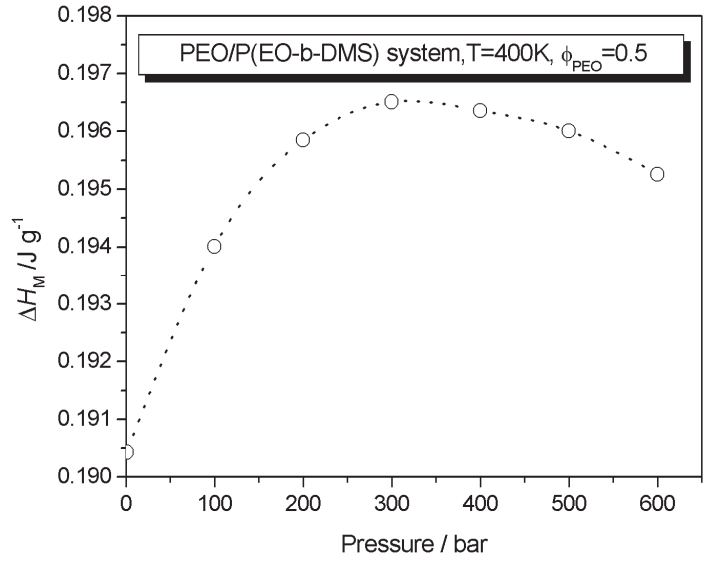


Fig. 20. Pressure dependences of the enthalpy of the mixing, ΔH_M, for the P(EO-b-DMS) polymer mixtures with fixed composition and temperature).

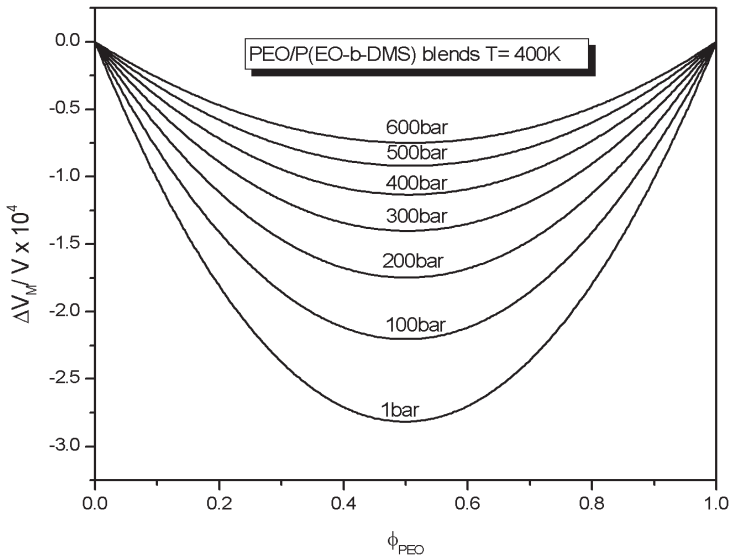


Fig. 21. Pressure and polymer mixture composition dependences of the volume change of the mixing, $\Delta V_M / V$, for the P(EO-b-DMS) polymer mixtures under the indicated temperature.

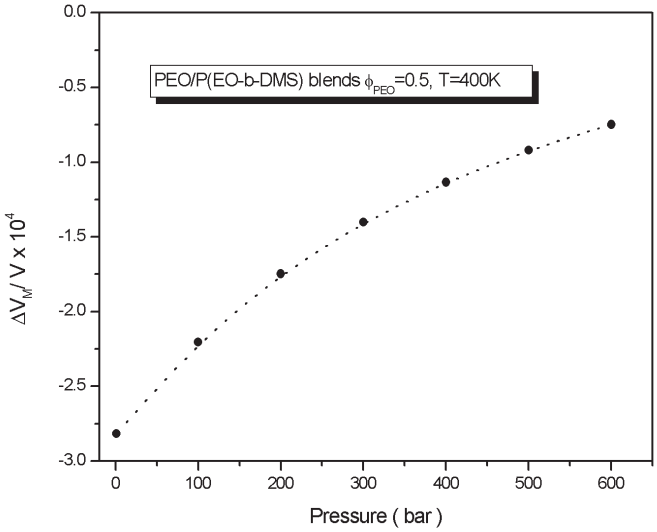


Fig. 22. Pressure dependences of the volume change of the mixing, $\Delta V_M / V$, for the P(EO-b-DMS) polymer mixtures at the indicated temperature and composition.

After calculating the ΔH_M and ΔV_M for PEO/P(EO-b-DMS) system, we know that just as the Clausius-Clapeyron equation(Ma, 1982) predicted the coexistence line of two phases in a T-P plane of PEO/P(EO-b-DMS) system according to $(dT / dP)_\varphi = T\Delta V_M / \Delta H_M$. Since ΔH_M , as calculated according to the Sanches-Lacombe theory the enthalpy change on mixing, is positive at the critical point, the sign of this derivative is controlled by ΔV_M , the volume change on mixing. In the calculated temperature and pressure range, ΔV_M is negative in this system, which indicates that $(dT / dP)_\varphi < 0$ for PEO/P(EO-b-DMS) mixtures. In most polymer blends investigated (Geerissen et al, 1985) to date $(dT / dP)_\varphi > 0$ implying that $\Delta V_M > 0$, which means that PEO/P(EO-b-DMS) mixture display the abnormal pressure effects on the phase separation and the root is $\Delta V_M < 0$.

5. Abnormal pressure dependence of the phase boundaries in TL/ PEO/ P(EO-b-DMS) ternary mixtures

The phase separation conditions for polymer-containing systems can change markedly as the systems pressured and the pressure effect on the phase behavior of mixtures consisting of polymers is of great technical and fundamental interest, and there are general observations on how the critical temperatures of polymer containing systems vary with pressure. The thermodynamic behavior of polymer blends is well understood in terms of the mean field Flory-Huggins theory and their deviations near the critical point when thermal composition fluctuations become dominant (An & Wolf, 1998; Hammouda & Bauer, 1995). For those systems that display UCST behavior (that is, phase separation upon cooling), it is nearly always found that the critical temperature, T_c, increases with pressure, i.e., the effect of increasing pressure is to decrease the range of miscibility. Similarly, in mixtures that display LCST behavior (phase separation upon heating), T_c also increases with pressure;

thus, increasing pressure almost always reduces the miscibility range in UCST systems and increases it in LCST systems. Most studies of polymer blends so far show an increase of the phase boundary with pressure which is quite obvious as the free volume decreases with pressure and thereby diminish the entropy of mixing(Janssen et al. 1995; Rudolf & Cantow, 1995; Schwahn et al. 1995).

The phase separation of polymer containing system is primarily driven by the reduced entropy of mixing as compared to small molecule analogs. This basic fact is captured by incompressible Flory-Huggins (FH) theory. Note that the FH theory, being incompressible, would suggest that pressure is an irrelevant variable. In contrast to FH theory, as showed above, the experimental results proved that pressure can play an important role in the phase separation of polymer containing systems(Hajduk et al., 1995, 1996; Jiang et al. 2002).

The behavior of ternary polymer mixtures containing a diblock copolymer with homopolymer and toluene as a function of mixture composition and temperature were investigated to obtain experimental phase diagram for solvent/copolymer/ homopolymer mixture. In order to avoid the complications associated with the microphase separation of block copolymers, the molar mass of block copolymer was kept low in our experiment (Madbouly & Wolf, 2002).

The presently available experimental information concerning pressure effects on polymer containing ternary system reveals an abnormal phenomena: under demixing condition, (dT/dP) changed from positive to negative with different concentration of toluene (TL) in the ternary system. The purpose of this study is to determine the coexistence surface of the ternary system consisting of polymers with different pressures. The phase boundaries at normal pressure, chosen temperatures and the chosen compositions, various pressures are measured. Further, the pressure effects on (dT/dP) of ternary are compared with that of binary system without solvent.

The measured cloud point curves of polymer containing binary mixtures and ternary systems at normal pressure are shown in figure 23. Figure 23a shows the measured transition temperatures of PEO/P(EO-b-DMS). The open circles indicate experimental points that got from the measurements of the transition temperature during the system changed from one-phase to two-phase, i.e. the heating process and the solid curve describes the behavior of the data points; the open squares indicate experimental points that got from the measurements of the transition temperature during the system changed from one-phase to two-phase, i.e. the cooling process and the dot curve describes the behavior of the data points. The lines are polynomial fits and serve as a guide for the eye (for other figures are the same). Figure 23b shows the cloud points of TL/PEO/P(EO-b-DMS) ternary system at 35, 45, 55, and 60 ^{o}C under atmosphere pressure. The solid squares on the curves at 35 and 45 ^{o}C represent the ternary system composition that were chosen for the pressure investigation, and the solid squares on the curves at 55 and 60 ^{o}C are the measured critical points at those temperatures and they were also investigated under pressure. It was found that toluene has much effect on the phase transition temperature of this ternary system, on the contrary, the shape of the coexistence curve depends only slightly on the ratio of PEO:P(EO-b-DMS) in the ternary system, which can be recognized by comparing the curve at 35 ^{o}C with the curve 60 ^{o}C, so the position of cloud-point curves depend on the concentration of TL in the ternary system.

The determination of the exact position of the critical point on the cloud-point curve is a problem not only with multicomponent systems. In the case of binary mixtures the coexistence curves are often very flat in the neighborhood of the critical concentration is

difficult to determine. It is established practice then to use the rectilinear diameter and assume that upon linear extrapolation the locus of the mean of the tie lines passes through the critical point. This implies that, if the volume fraction is used as the concentration variable, the volume ratio of the vanishing phases at the critical point should equal unity (Konowalow, 1903).

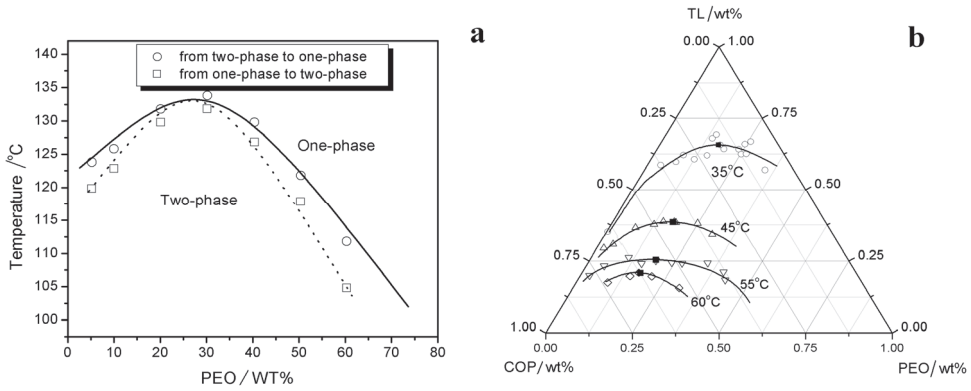


Fig. 23. Phase diagrams under normal pressure (a: phase diagram of PEO/P(EO-b-DMS) binary mixtures under normal pressure; b: phase diagrams of TL/PEO/P(EO-b-DMS) ternary solutions under normal pressure at indicated temperatures).

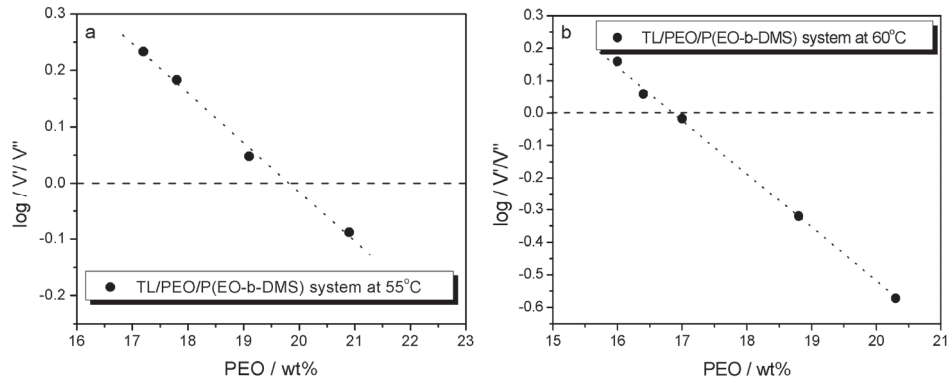


Fig. 24. Phase-volume ratio measurement for TL/PEO/P(EO-b-DMS) mixtures (a) at 50 °C and (b) at 60 °C.

This does not mean that the diameter passing through the critical point is linear. For single-component systems Bruhat has shown that the gas-liquid diameter may be curved close to the critical point (Bruhat, 1930).

This argument can be extended to cover systems consisting of more than two components. Therefore, a necessary and sufficient criterion for determining whether a composition is the critical concentration appears to be the equality of the phase volumes as measured at a temperature very close to the cloud-point of that concentration. As shown in figure 24, we just obtained the critical point for the mixtures at 55 °C and 60 °C which represented in

figure 23b, however, at the lower temperatures, it is very difficult for us to observe the macrophase separation in this system even the experiment lasted for more than one month. Figure 25 shows isopleths (i.e., lines of constant composition of the solution) cloud-point curves observed for ternary polymer containing solutions on a temperature-pressure plane, and the pressure effect on the binary polymer mixtures has been mentioned in section 4. The open circles indicate experimental points and the solid curves describe the behavior of the data points. The region above the curve is the one-phase homogeneous region, while the region below represents the two-phase region. As shown in reference (Jiang et al, 2002), the demixing pressures decrease with increasing temperature for the binary polymer mixtures, and (dT/dP) of the binary system is negative, and this means that the pressure decreases the phase separation temperature of PEO/P (EO-b-DMS) blends. Figure 25 shows the experimental measured phase transition temperatures under pressure for the compositions that showed in figure 23b as solid squares. As shown in Figure 25, the abnormal pressure effects on the ternary system was found, i.e., at low TL concentration of this system, pressure induces the compatibility, but for high TL concentration, pressure decreases the compatibility of this system.

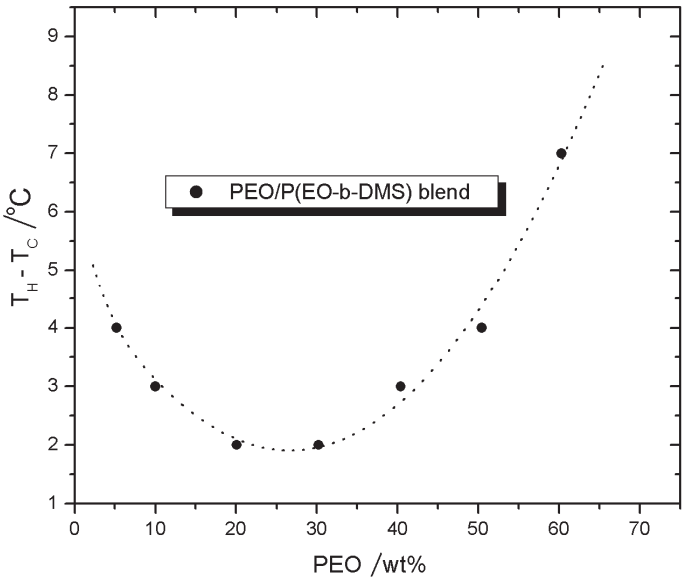


Fig. 25. Pressure dependence of T on T-P plane (the transition temperatures for TL/PEO/P(EO-b-DMS) ternary solutions at indicated temperatures(60 °C(○), 55 °C (△), 45°C(□), and 35 °C(▽)) and the compositions corresponding to Figure 1b indicated.

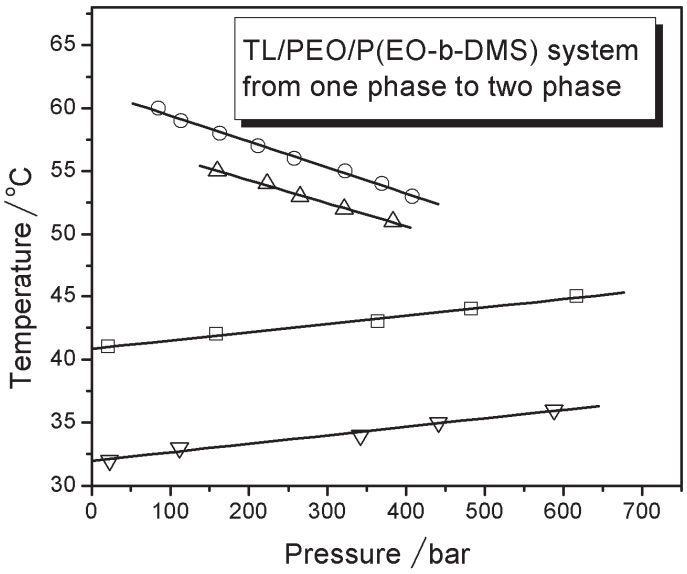


Fig. 26. T_H-T_C vs PEO wt% in binary system(T_H and T_C are the phase transition temperatures of heating and cooling process).

With the reservations concerning the critical polymer concentration, indicated in the figure 23a, figure 26 was constructed from the phase diagram of PEO/P(EO-b-DMS) binary mixtures as shown in figure 23a by plotting the (T_H-T_C) vs. PEO WT% (T_h corresponding to the temperature of heating process, i.e., the spinodal temperature; T_c corresponding to the temperature of cooling process, i.e., the binodal temperature). The transition value of the (T_H-T_C) was considered as the critical composition of this binary system.

The pressure effects on the demixing temperatures of polymer containing binary system and ternary system are shown in figure 27. Figure 27a shows the slopes of isopleths (dT/dP) in reference (Jiang et al, 2002) vs. PEO composition for PEO/P(EO-b-DMS) binary mixtures. From this figure, we not only can recognize the degree of the pressure effect on the phase separation temperature of PEO/P(EO-b-DMS) mixtures, i.e. $dT/dP \cong 28$ °C/kbar, but also can find that the pressure affected phase separation temperature of this binary system is nearly independent of the composition. Figure 27b shows the slopes of isopleths (dT/dP) in figure 25 vs. TL concentration for TL/PEO/P(EO-b-DMS) ternary solutions. The pressure effect on the ternary system where TL=0 is according to the results from figure 26 and figure 27a, which is the pressure effect on the critical composition of PEO/P(EO-b-DMS) binary mixtures. Figure 27b shows a transition for (dT/dP) with different concentration of TL within the measured composition in the ternary system. The slope of UCS in ($P < T$) space is described in terms of the excess functions using eqs 1 and 2 (Imre et al, 2001).

The influence of pressure can be deduced by calculating the pressure coefficien (dT_{cr}/dP) of the critical solution temperature, which given by (Wolf & Blaum, 1977):

$$dT_{cr} / dP \cong T\Delta V_m / \Delta H_m \tag{34}$$

where ΔV_m and ΔH_m are the excess volume and the enthalpy of mixing. Since ΔH_m is positive at the critical point, the sign of this derivative is controlled by ΔV_m. The Clausius-Clapeyron equation (Ma, 1982) describes the coexistence line of two phases in a *T*-*P* plane according to $(dT/dP)_\phi = T\Delta V_m / \Delta H_m$, which is similar to eq. 2 and 34. Since ΔH_m, the enthalpy change on mixing, is positive at the critical point, the sign of this derivative is controlled by ΔV_m, the volume change on mixing. In most polymer blends investigated to date (dT/dP) > 0 implying that $\Delta V_m > 0$. A simple equation of state, such as the lattice fluid model (which extends FH theory through the addition of free volume), shows that to leading order for binary mixtures (Sanchez & Lacombe, 1976):

$$\Delta V_m = \phi(1-\phi)\{4\chi - [(\varepsilon_{11} - \varepsilon_{22}) / RT]^2\} \tag{35}$$

ϕ is the volume fraction of component-1, χ is the interaction parameter, and R is the gas constant. For many common blends, such as those studied in past work, the χ parameter is larger, and therefore the lattice model predicts $\Delta V_m > 0$. In the contrast, for a carefully selected system with chemically similar monomers, where χ is positive but small, there exists a possibility for $\Delta V_m < 0$. These predictions are consistent with the work of Foreman and Freed using the lattice cluster model (Horst, 1995; Horst & Wolf, 1992). According to figure 27b and equation 35, we can qualitatively predict the pressure effect on this ternary system, i.e., the pressure effects on ΔV_m in this system; at the same time, we also get the information on the pressure effects on χ parameter for this system, i.e., with the increasing of TL in this system, χ parameter increases from small to large.

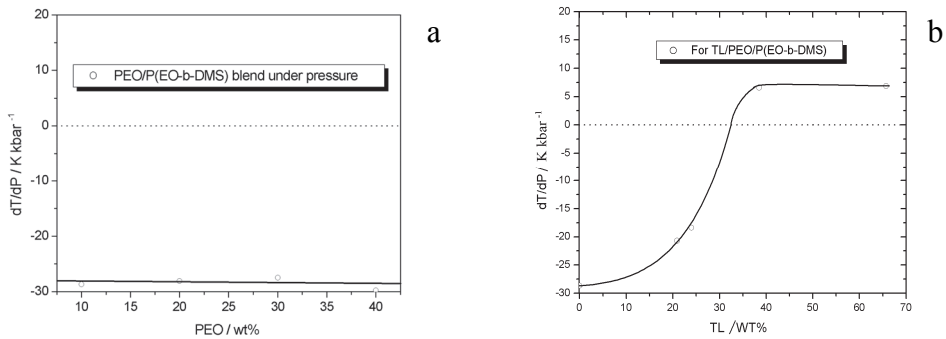


Fig. 27. Pressure dependence of the transition temperatures (dT/dP) vs referred composition(a: for binary mixtures; b: for ternary solutions).

The same as classic thermodynamics, polymer thermodynamics is function of pressure, temperature and composite. But in many cases, pressure effects on polymer thermodynamics was neglected, because polymer thermodynamics were often studied under atmosphere. The classic theory of polymer thermodynamics is Flory-Huggins hard lattice theory. In this theory, the hard lattice is incompressible. A rigorously incompressible system should be unaffected by pressure. However, since experimental results show that the critical temperature for polymer demixing system is strongly affected by pressure, it is clear that polymer containing systems show significant departures from this ideal limit. We wish

above introduction would give readers a rough draft on the thermodynamics of polymer containing systems.

6. Acknowledgements

This work is supported by the National Natural Science Foundation of China (20974077，51073157) and subsidized by the Special Funds for National Basic Research Program of China (2010CB631102). We appreciated helpful discussion provided by Prof. Lijia An (Changhun Institute of Applied Chemistry) and Prof. Bernhard A. Wolf (Mainz University).

7. References

An et al. 1997, Pressure dependence of the miscibility of poly(vinyl methyl ether) and polystyrene: Theoretical representation, *J Chem. Phys*. Vol.107, No.7, PP. 2597-2602.

An &Wolf, 1998, Combined effects of pressure and shear on the phase separation of polymer solutions, *Macromolecules,* Vol.31, No. 14, PP. 4621-4625.

Beiner et al. 1998, Pressure-induced compatibility in a model polymer blend, *Physical Review Letters,* Vol. 81, No. 3, PP. 594-597; 2002, Strong isotopic labeling effects on the pressure dependent thermodynamics of polydimethylsiloxane/polyethylmethysiloxane blends, *J. Chem. Phys.*, Vol. 116, No. 3, PP.1185-1192.

Blaum & Wolf, 1976, Generation of true cosolvency by pressure - solubility limits of high molecular-weight polystyrene in mixtures of acetone and diethyl-ether, *Macromolecules,* Vol. 9, No. 4, PP. 579-582.

Bruhat, 1930, *Chaleur Ind*., 11, 263.

de Loos et al. 1983, Fluid phase-equilibria in the system polyethylene + ethylene .1. systems of linear polyethylene + ethylene at high-pressure, *Macromolecules*, Vol. 16, No. 1, PP. 111-117.

Dudowicz & Freed, 1995, Pressure-dependence of polymer fluids - application of the lattice cluster theory, *Macromolecules,* Vol. 28, No. 19, PP. 6625-6641; 2006, Entropy theory of polymer glass formation revisited, I. Gerneral formulation, *J. Chem. Phys.*, Vol. 124, No. 6, PP. 064901.

Flory, 1953, *Principle of Polymer Chemistry*; Cornell University Press: Ithaca, NY, Chapter 12.

Geerissen et al. 1985, Continuous fractionation and solution properties of PVC. 3. Pressure-dependence of the solubility in single solvents, *Makromol. Chem.,* Vol. 186, No. 4, PP. 769-776

Hajduk et al., 1995, High-pressure effects on the disordered phase of block-copolymer melts, *Macromolecules*, Vol. 28, No. 21, PP. 7148-7258; 1996, High-pressure effects on the order-disorder transition in block copolymer melts, *Macromolecules*, Vol. 29, No. 5, PP. 1473-1481

Hammouda & Bauer, 1995, compressibility of 2 polymer blend mixtures, *Macromolecules,* Vol. 28, No. 13, PP. 4505-4508; 1997, Small-angle neutron scattering from pressurized polyethylbutylene/ polymethylbutylene blends, *Macromolecules,* Vol. 30, No. 18, PP. 5572-5574

Horst, 1995, Calculation of phase-diagrams not requiring the derivatives of the gibbs energy demonstrated for a mixture of 2 homopolymers with the corresponding copolymer, *Macromol. Theory Simul.*, Vol. 4, No. 3, PP. 449-458

Horst & Wolf, 1992, Calculation of the phase-separation behavior of sheared polymer blends, *Macromolecules*, Vol. 25, No. 20, PP. 5291-5296

Hosokawa et al. 1993, Coexistence curve of polystyrene in methylcyclohexane .7. coexistence surface and critical double point of binary-system in t-p-phi space, *J. Chem. Phys.*, Vol. 98, N0. 12, PP. 10078-10084

Imre, 2001, On the effect of pressure on the phase transition of polymer blends and polymer solutions: Oligostyrene-n-alkane systems, *Phys. Chem. Chem. Phys.*, Vol. 3, No. 6, PP. 1063-1066

Janssen et al. 1993, Pressure-dependence of the flory-huggins interaction parameter in polymer blends - a sans study and a comparison to the flory-orwoll-vrij equation of state, *Macromolecules,* Vol. 26, No. 21, PP. 5587-5591; 1995, Coil and melt compressibility of polymer blends studied by sans and pvt experiments, *Macromolecules*, Vol. 28, No. 7, PP. 2555-2560

Jiang et al. 2002, Pressure-induced compatibility in PEO/P(EO-b-DMS) polymer mixtures, *Macromolecules*, Vol. 35, No. 15, PP. 5727-5730

Koningsveld & Staverman, 1968, liquid-liquid phase separation in multicomponent polymer solutions .2. critical state, *J. Polym. Sci., Part A*, Vol. 6, No. 2, PP. 325-347

Konowalow, 1903, The critical area of solutions and the appearance of opalescence, *Ann. Physik.*, Vol. 10, No. 2, PP. 360-392

Kumar, 2000, Pressure effects on the thermodynamics of polymer blends, *Macromolecules,* Vol. 33, No. 14, PP. 5285-5291

Lacombe & Sanchez, 1976, Statistical thermodynamics of fluid mixtures, *J. Phys. Chem.*, Vol. 80, No. 23, PP. 2568-2580

Lefebvre etal 1999, Relationship between internal energy and volume change of mixing in pressurized polymer blends, *Macromolecules,* Vol. 32, No. 16, PP. 5460-5462; 2000, Neutron scattering from pressurized polyolefin blends near the limits of metastability, *Macromolecules,* Vol. 33, No. 21, PP. 7977-7989

Ma, 1982, *Statistical Mechanics;* World Scientific: Singapore

Madbouly & Wolf, 2002, Equilibrium phase behavior of polyethylene oxide and of its mixtures with tetrahydronaphthalene or/and poly(ethylene oxide-block-dimethylsiloxane), *J. Chem. Phys.*, Vol. 117, No. 15, PP. 7357-7363

Maderek et al. 1983, High-temperature demixing of poly(decyl methacrylate) solutions in isooctane and its pressure-dependence, *Makromol. Chem.*, Vol. 184, No. 6, PP. 1303-1309; Lower critical solution temperatures of poly(decyl methacrylate) in hydrocarbons, *Eur. Polym. J.*, Vol. 19, No. 10, PP. 963-965

Patterson & Robard, 1978, Thermodynamics of polymer compatibility, *Macromolecules,* Vol. 11, No. 4, 690-695

Rabeony et al 1998, Effect of pressure on polymer blend miscibility: A temperature-pressure superposition, *Macromolecules,* Vol. 31, No. 19, PP. 6511-6514

Rodgers, 1991, Procedure for predicting lower critical solution temperature behavior in binary blends of polymers, *Macromolecules*, Vol. 24, No. 14, PP. 4101-4109

Rudolf & Cantow, 1995, Description of phase-behavior of polymer blends by different equation-of-state theories .2. excess volumes and influence of pressure on miscibility, *Macromolecules*, Vol. 28, No. 19, PP. 6595-6599

Sanchez & Balasz, 1989, Generalization of the lattice-fluid model for specific interactions, *Macromolecules*, Vol. 22, No. 5, PP. 2325-2331

Sanchez & Lacombe, 1976, Elementary molecular theory of classical fluids - pure fluids, *J. Phys. Chem.*, Vol. 80, No. 21, PP. 2352-2362; 1978, statistical thermodynamics of polymer-solutions, *Macromolecules*, Vol. 11, No. 6, PP. 1145-1156

Schmidt & Maurer, 1998, Pressure-volume-temperature properties and free volume parameters of PEO/PMMA blends, *Polym. Sci. B: Polym. Phys.*, Vol. 36, No. 6, PP. 1061-1080

Schwahn et al. 1995, ginzburg criterion for the mean-field to 3-dimensional ising crossover in polymer blends, *Phys. Rev. E*, Vol. 52, No. 2, PP. 1288-1291; 1996, Temperature and pressure dependence of the order parameter fluctuations, conformational compressibility, and the phase diagram of the PEP-PDMS diblock copolymer, *Physical Review Letters,* Vol. 77, No. 15, PP. 3153-3156; 2001, Abnormal pressure dependence of the phase boundaries in PEE-PDMS and PEP-PDMS binary homopolymer blends and diblock copolymers, *Macromolecules,* Vol. 34, No. 6, PP. 1694-1706

Sun et al. 1999, Phase-separation behavior of the system PES/phenoxy: An application of the Sanchez-Lacombe lattice fluid theory, *Macromol. Sci.-Phys.*, Vol. 38, No. 1-2, PP. 67-74

Tait, 1888, *Physics and Chemistry of the Voyage of H.M.S. Challenger, Vol. 2, Part 4* HMSO, London

van Opstal & Koningsveld, 1992, Mean-field lattice equations of state .4. influence of pressure on the phase-behavior of the system polystyrene cyclohexane, *Polymer*, Vol. 33, No. 16, PP. 3433-3444; Mean-field lattice equations of state .5. influence of pressure on liquid liquid-phase behavior of polymer blends, *Polymer*, Vol. 33, No. 16, PP. 3445-3452

Walsh & Rostami, 1985, Effect of pressure on polymer polymer phase-separation behavior, *Macromolecules*, Vol. 18, No. 2, PP. 216-219

Wolf & Blaum, Pressure influence on true cosolvency - measured and calculated solubility limits of polystyrene in mixtures of acetone and diethylether , *Makromol. Chem.,* 1976, Vol. 177, No. 4, PP.1073-1088; 1977, dependence of oligomer-oligomer incompatibility on chain-length and pressure .1. oligo-isobutene-oligo-propylene glycol and oligo-styrene-oligo-ethylene glycol, *Journal of Polymer Science: Polymer Symposium,* No. 61, PP. 251-270

Wolf & Jend, 1977, Possibilities for determination of enthalpies and volumes of mixing from molecular-weight dependence of critical demixing temperatures and pressures, demonstrated for system trans-decahydronaphthalene-polystyrene, *Makromol. Chem.*, Vol. 178, No. 6, PP. 1811-1822; 1979, Pressure and temperature-dependence of the viscosity of polymer-solutions in the region of phase-separation, *Macromolecules,* Vol. 12, No. 4, PP. 732-737

Zeman et al. 1972, Pressure effects in polymer-solution phase equilibria .1. lower critical solution temperature of polyisobutylene and polydimethylsiloxane or lower alkanes, *J. Phys. Chem.*, Vol. 76, No. 8, PP. 1206-1213

Zeman &Patterson, 1972, Pressure effects in polymer-solution phase equilibria .2. systems showing upper and lower critical solution temperatures, *J. Phys. Chem.*, Vol. 76, No. 8, PP. 1214-1219

9

Potential-pH Diagrams for Oxidation-State Control of Nanoparticles Synthesized via Chemical Reduction

Shunsuke Yagi
Nanoscience and Nanotechnology Research Center,
Osaka Prefecture University
Japan

1. Introduction

In the past, many synthesis methods of nanoparticles have been reported, but the synthesis processes have not been well discussed from the viewpoint of thermodynamics. In this chapter, a general concept using potential-pH diagrams is described for oxidation-state control of nanoparticles synthesized via chemical reduction (also called electroless deposition or liquid-phase reduction). By comparing kinetically determined mixed potential measured in reaction solution and thermodynamically drawn potential diagrams, e.g., potential–pH diagram, it is possible to know "what chemical species is stable in the reaction solution?." It is predicted from potential-pH diagrams that nanoparticles in different oxidation states can be selectively synthesized by controlling the mixed potential. This concept is verified by selectively synthesizing Cu and Cu_2O nanoparticles from CuO aqueous suspension via chemical reduction using the concept as an example. The advantage of this chemical reduction method is that abundant nanoparticles can be obtained for a short time by a simple operation. An extremely small activity of Cu^{2+} aquo ion is achieved by using insoluble CuO powder as a Cu(II) ionic source, which is a key for the synthesis of nanosized particles. The dependency of mixed potential on pH and temperature is discussed in the verification process. This chapter is written based on the result of the authors' paper (Yagi et al., 2009) with further detailed information on practical thermodynamic calculation and drawing procedure of potential-pH diagrams.

2. Brief overview of objective materials and their synthesis

2.1 Copper (Cu) nanoparticles

Copper (Cu) nanoparticles are of great interest in various fields, specifically that of printable electronics. Cu interconnects less than 20 μm wide can be made with a high resolution screen printer or a super inkjet printer using an ink which contains dense Cu nanoparticles. Cu nanoparticles have been synthesized by various reduction processes from cuprous (Cu(I)) or cupric (Cu(II)) compounds, including direct electrochemical reduction (Han et al., 2006; L. Huang et al., 2006; Yu et al., 2007), chemical reduction (Lisiecki & Pileni, 1993; Lisiecki et al., 1996; H. H. Huang et al., 1997; Qi et al., 1997; Ohde et al., 2001), thermal reduction (Dhas et al.,

1998), sonochemical reduction (Dhas et al., 1998; Kumar et al., 2001), laser irradiation (Yeh et al., 1999), and gamma radiolysis (Joshi et al., 1998). Many of these methods are conducted in liquid phase, which has the important advantage that nanoparticles can be formed in the presence of dispersing agents and no additional process for the addition of dispersing agents is required for the fabrication of inks to avoid the agglomeration.

In recent years, Cu nanoparticle also has drawn a lot of attention due to its novel optical properties from surface plasmon resonance (SPR), which occurs when a light electromagnetic field drives the collective oscillation of free electrons in Cu nanoparticles (J. Li et al., 2011). This resonance creates large local electric field enhancement on the nanoparticle surface, which can be used to manipulate light–matter interactions and boost non-linear phenomena. The SPR of Cu nanoparticles is observed at 560-580 nm, which is slightly changed by a medium and a dispersing/protecting agent (Singh et al., 2010; X.D. Zhang et al., 2011), and Pileni and Lisiecki demonstrated the strong broadening of the plasmon peak for Cu nanoparticles with decreasing the particle size (Pileni & Lisiecki, 1993). In addition, Cu nanoparticles display fluorescence. Siwach and Sen reported fluorescence peaks at ca. 296 nm for excitation wavelength at 213 nm and 270 nm using Cu nanoparticles about 13 nm in diameter in water (Siwach & Sen, 2010).

2.2 Cuprous oxide (Cu_2O) nanoparticles

Cuprous oxide (Cu_2O) is a p-type semiconductor (Young & Schwartz, 1969) with unique optical and magnetic properties (Mahajan et al., 2009) and is also a promising material with potential applications in solar energy conversion, microelectronics, magnetic storage, catalysis, photocatalysis (Gu et al., 2010), and gas sensing (Zhang et al., 2006). Cu_2O is also used for the investigation into Bose–Einstein condensation of excitons (Snoke, 2002). Many methods have been reported to synthesize Cu_2O nanoparticles (Zhang et al., 2006; Muramatsu & Sugimoto, 1997; X. Liu et al., 2007), nanocubes (Gou & Murphy, 2003; R. Liu et al., 2003), octahedral nanocages (Lu et al., 2005), nanorods (Cheng et al., 2011), and nanowires (Wang et al, 2002). For example, Muramatsu and Sugimoto reported that Cu_2O particles with an average diameter of 270 nm were synthesized in large quantities from a CuO aqueous suspension using hydrazine as the reducing agent (Muramatsu & Sugimoto, 1997). Liu et al. formed CuO, Cu_2O, and Cu nanoparticles using copper(II) acetylacetonate [$Cu(acac)_2$] as a precursor in oleylamine by controlling temperature (Liu et al., 2007).

It is noteworthy that nanoparticles of transition-metal oxides including Cu_2O and CuO are also expected as a negative electrode material with a high specific capacity and extraordinary large surface area for lithium ion battery (Poizot et al., 2000). The cycle property of Cu_2O nanoparticles is strongly dependent on the particle size, shape, and structure; there is an optimum condition, displaying the highest electrochemical performance (C.Q. Zhang et al., 2007).

In the next section, a general guideline is introduced for the selective synthesis of Cu and Cu_2O nanoparticles possessing superior properties described above from the same CuO aqueous suspension.

3. Theory for oxidation-state control of nanoparticles

3.1 Mixed potential

At any point within the reaction volume during nanoparticle synthesis via chemical reduction, several partial reactions occur in parallel (e.g. oxidation of a reducing agent,

reduction of metallic ion, reduction of dissolved oxygen, and solvent reduction decomposition). The potential at a point is determined at the value where the total of anodic currents $I_{a,total}$ (currents for oxidation reaction) balances the total of cathodic currents $I_{c,total}$ (currents for oxidation reaction) unless current flows outside (Fig. 1). This potential is called the mixed potential. Stable chemical species in the reaction solution can be determined by comparing measured mixed potential and calculated potential diagrams, and chemical species synthesized can be changed by controlling mixed potential. As an example, considering a reaction system for the synthesis of Cu or Cu_2O nanoparticles from a CuO aqueous "suspension" of pH 9 via hydrazine reduction, anodic reactions are mainly the oxidation reactions of hydrazine as follows

$$N_2H_4 + 4OH^- \rightarrow N_2 + 4H_2O + 4e \quad (1)$$

$$N_2H_4 + OH^- \rightarrow 1/2N_2 + NH_3 + H_2O + e \quad (2)$$

Cathodic reactions are mainly the deposition of Cu or Cu_2O and hydrogen generation as follows

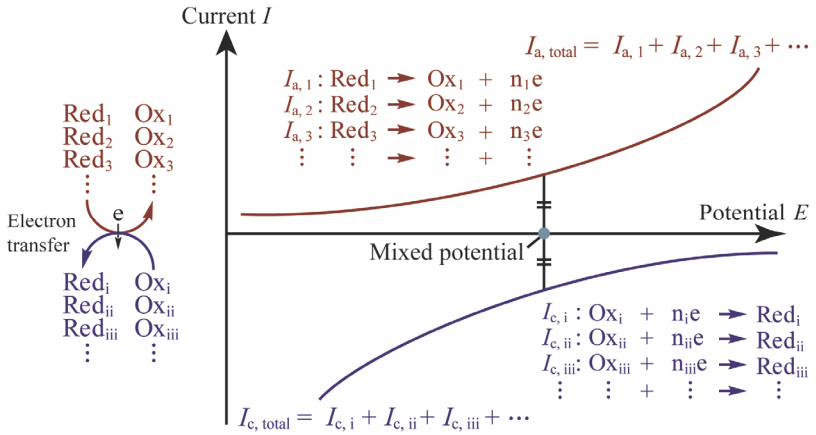


Fig. 1. Schematic current-potential curves: mixed potential is determined at the value where the total of anodic currents $I_{a,total}$, balances the total of cathodic currents $I_{c,total}$, unless current flows outside.

Deposition of Cu

$$Cu(II) + 2e \rightarrow Cu \quad (3)$$

Deposition of Cu_2O

$$2Cu(II) + 2OH^- + 2e \rightarrow Cu_2O + H_2O \quad (4)$$

Hydrogen generation

$$2H_2O + 2e \rightarrow H_2 + 2OH^- \quad (5)$$

where Cu(II) represents all the copper(II) species and is distinguished from a copper(II) aquo ion, Cu^{2+}. The overall deposition reactions of Cu and Cu_2O in this system can be expressed using the following equations

Deposition of copper

$$CuO + 1/2N_2H_4 \rightarrow Cu + 1/2N_2 + H_2O \quad (6)$$

$$CuO + 2N_2H_4 \rightarrow Cu + N_2 + 2NH_3 + H_2O \quad (7)$$

Deposition of cuprous oxide

$$2CuO + 1/2N_2H_4 \rightarrow Cu_2O + 1/2N_2 + H_2O \quad (8)$$

$$2CuO + 2N_2H_4 \rightarrow Cu_2O + N_2 + 2NH_3 + H_2O \quad (9)$$

In practice, the "absolute" mixed potential cannot be measured because of the following experimental reason (Bockris et al., 2000). When one tries to measure the mixed potential using potentiometers, two metallic terminals of the potentiometers must be connected to the two electrodes immersed in the reaction solution (electrolyte). In this case, all one can measure is the potential "difference" across the two electrode/electrolyte interfaces. Each electrode/electrolyte interface has a potential difference due to a double layer between the electrode and electrolyte, and thus, the measured potential difference includes at least two potential differences at different interfaces. In other words, when one tries to measure one potential difference, one must create at least one additional potential difference. This is not favorable, but inevitable. However, by using a nonpolarizable electrode as one electrode, the measured potential difference directly reflects the change in the potential difference of the other electrode/electrolyte interface because the potential difference of the nonpolarizable electrode/electrolyte interface does not change. Therefore, the measured potential difference can be treated as the "relative" mixed potential by using a nonpolarizable electrode as a reference electrode, e.g. standard hydrogen electrode, Ag/AgCl electrode, and saturated calomel electrode. In addition, the other measuring electrode should not have a strong catalytic activity for possible reactions in order not to affect the original balance of anodic currents and cathodic currents so much, which can cause the change in the mixed potential. In the present system, a gold electrode is appropriate as a measuring electrode since gold has a low catalytic activity for hydrazine decomposition or hydrogen generation.

3.2 Potential-pH diagram (Pourbaix diagram)

Potential-pH diagram is a kind of phase diagram showing possible stable phases of an electrochemical system in equilibrium. Potential-pH diagram was invented by Marcel Pourbaix (1904-1998) and is also known as Pourbaix diagram (Pourbaix, 1966). In the potential-pH diagram, the vertical axis is potential normally with respect to the standard hydrogen electrode (SHE) and the horizontal axis is pH(=$-\log a_{H+}$), where a_{H+} is the activity of proton. By comparing the potential-pH diagram and the mixed potential described above, it is possible to know the stable or obtainable chemical species in the reaction solution.

3.3 How to draw potential-pH diagram

In order to draw potential-pH diagram, it is necessary to determine the activity of ionic species. For example, consider an equilibrium reaction of metallic oxide M_xO_y and M^{z+} aquo ion in an aqueous solution as follows;

$$M_xO_y + 2yH^+ = xM^{z+} + yH_2O \quad (10)$$

where M is a metal and M^{z+} is a metallic aquo ion. There is a relationship $2y=xz$, which is derived from the charge balance of the reaction 10. The following equation indicates the change in Gibbs free energy, ΔG_T of the reaction 10 at absolute temperature T in kelvins and 1 atm.

$$\Delta G_T = \Delta G_T^0 + RT\ln\frac{a_{M^{Z+}}^{x} a_{H_2O}^{y}}{a_{M_xO_y} a_{H^+}^{2y}} \tag{11}$$

where ΔG_T^0 is the change in Gibbs free energy at a standard state and a temperature T (K), and R is the gas constant. In equilibrium, the change in Gibbs free energy is zero and thus

$$0 = \Delta G_T^0 + RT\ln\frac{a_{M^{Z+}}^{x} a_{H_2O}^{y}}{a_{M_xO_y} a_{H^+}^{2y}} \tag{12}$$

ΔG_T^0 can be calculated by the following equation using thermodynamic data;

$$\Delta G_T^0 = \Delta H_T^0 - T\Delta S_T^0 = \Delta H_{298}^0 + \int_{298}^{T}\Delta C_P dT - T\left(\Delta S_{298}^0 + \int_{298}^{T}\frac{\Delta C_P}{T}dT\right) \tag{13}$$

where ΔH_T^0 and ΔS_T^0 are the changes in enthalpy and entropy that accompanies the formation of 1 mole of a product from its component elements at a temperature T (K) and 1 atm. ΔC_p is the change in specific heat at constant pressure (1 atm). The absolute value of enthalpy H cannot be measured but chemical reaction system is a closed reaction system, and the change in enthalpy can be obtained by just subtracting the sum of the heat of formation, which is measurable, of the components in the left-hand side from that of the components in the right-hand side. Thermodynamic databooks include the list of the standard heat of formation, standard entropy at 298 K, and specific heat (Kubaschewski & Alcock, 1979; Latimer 1959; Stull & Prophet, 1971). Going back to the equation 12, the activities of water and oxide are both 1, and the equation 12 is simplified into the relationship between pH and the activity of the metallic aquo ion.

$$2y\mathrm{pH} = -\frac{\Delta G_T^0}{\ln(10)RT} - x\log a_{M^{Z+}} \tag{14}$$

Thus, the pH value, at which M_xO_y and M^{z+} are in equilibrium, is determined by the activity of the metallic aquo ion. Conversely, the activity of the metallic aquo ion in equilibrium can be determined by measured or maintained pH of a metallic oxide reaction suspension.

In the present system, CuO powder is dispersed in the reaction solution. When the pH of the reaction solution is 9 and the temperature is 298 K, the activity of Cu^{2+} aquo ion is calculated to be about 7.5×10^{-11} using the thermodynamic data shown in Table 1 (Criss & Cobble, 1964; Kubaschewski & Alcock, 1979; Latimer 1959; Stull & Prophet, 1971). This value itself does not have a precise meaning; the calculated value of the activity of Cu^{2+} aquo ion can be significantly affected by an experimental error of thermodynamic data, e.g. the standard entropy of CuO. Normally, this value just indicates that CuO does not dissolve into the solution in a practical sense. However using this value, one can calculate other important thermodynamic values such as the oxidation-reduction (redox) potential of Cu^{2+}/Cu redox pair in this reaction system using the same thermodynamic data set.

Chemical species	Standard heat of formation (kJ mol^{-1})	Standard entropy (J K^{-1} mol^{-1})	Specific heat at constant pressure (J K^{-1} mol^{-1}) $C_p = a + b\text{x}10^{-3}T + c\text{x}10^{5}T^{2}$		
H^+(aq)	0.0	0.0	129.7	–	–
Cu^{2+}(aq)	64.4	–98.7	267.8	–	–
H_2O(l)	–285.8	70.0	75.4	–	–
H_2(g)	0.0	130.6	27.3	3.3	0.50
Cu(s)	0.0	33.1	22.6	6.3	–
Cu_2O(s)	–167.4	93.1	62.3	23.9	–
CuO(s)	–155.2	42.7	38.8	20.1	–
N_2(g)	0.0	191.5	27.9	4.3	–
NH_3(g)	–45.9	192.7	37.3	18.7	–6.49
N_2H_4(l)	50.6	121.6	72.7	87.2	0.60

Table 1. List of standard heat of formation, entropy at 298 K and 1 atm, and specific heat at constant pressure (1 atm) considered and used for thermodynamic calculation. The units of the original data are converted to match the SI units using the thermochemical calorie, 1 cal_{th}=4.184 J. All the data considered and used for thermodynamic calculation are listed in this table. Only the constant *a* and coefficients *b* and *c* for the specific-heat capacity of hydrazine are estimated by fitting the discrete data of the specific heat capacity of hydrazine in the temperature range from 100 to 600 K. The data of liquid hydrazine N_2H_4(l), are used instead of those of hydrated hydrazine N_2H_4(aq), because of the lack of the data of N_2H_4(aq).

The procedure to draw potential–pH diagram is shown in Fig. 2. In this system, the equilibrium pH between Cu^{2+} aquo ion with the activity of 7.5 x 10^{-11} and CuO is 9, which is obvious and can also be checked by substituting the activity of Cu^{2+} ions into the equation 14. This equilibrium relationship can be represented by the vertical line 1 (Fig. 2a). The equilibrium activity of Cu^{2+} ions exceeds the determined value (7.5 x 10^{-11}) in the left region of the line 1 and underruns it in the right region. On the other hand, the redox potential of Cu^{2+}/Cu redox pair with respect to the standard hydrogen electrode can be determined by considering the electromotive force between the following half-cell reactions:
Copper deposition

$$Cu^{2+} + 2e = Cu \tag{15}$$

Hydrogen generation at a standard state (the activities of proton and hydrogen gas are both 1)

$$2H^+ + 2e = H_2 \tag{16}$$

The total cell reaction is obtained by subtracting the equation 16 from the equation 15 to erase the electron term.

$$Cu^{2+} + H_2 = Cu + 2H^+ \tag{17}$$

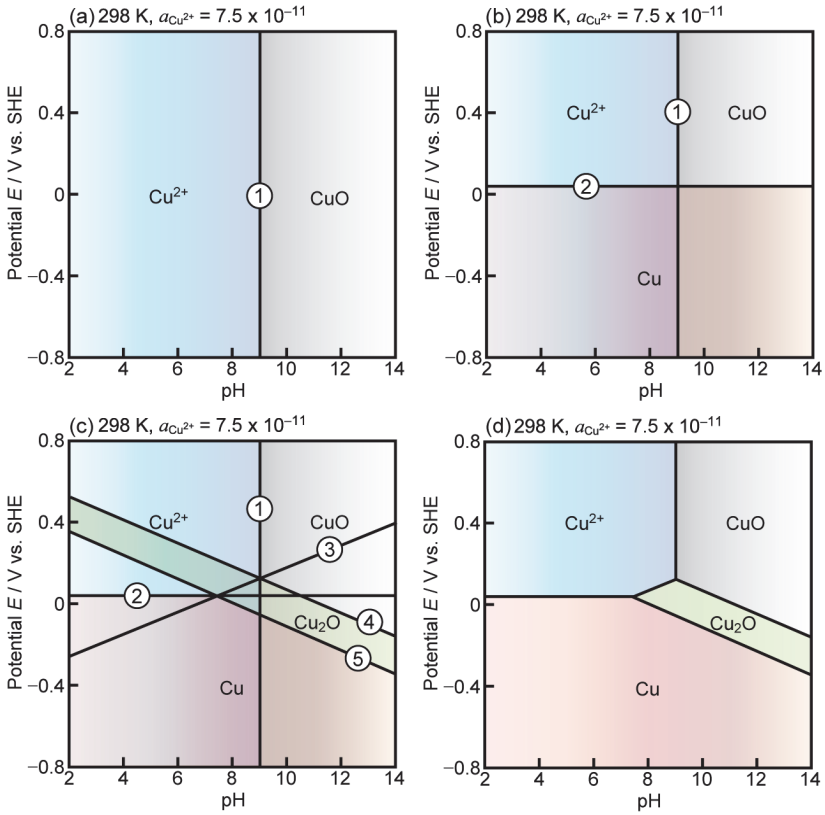


Fig. 2. Procedure to draw potential–pH diagram at pH 9 and 298 K in a CuO aqueous suspension. Vertical line indicates the equilibrium relationship of an acid-base reaction and horizontal line indicates that of a redox reaction. Diagonal line indicates an equilibrium relationship of the reaction including both an acid-base reaction and a redox reaction. In this procedure, the equilibrium activities of Cu^{2+} aquo ion was set to be 7.5 x 10^{-11} considering the following copper species; Cu^{2+}, CuO, Cu_2O, and Cu. Oxyanions of copper, $HCuO_2^-$ and CuO_2^{2-}, or hydroxyanions of copper, $Cu(OH)_3^-$ and $Cu(OH)_4^{2-}$ (Beverskog, 1997), are not considered for simplicity. In order to include such ions, one should consider all the additional possible equilibrium relationships.

The redox potential E_T at a temperature T and 1 atm of Cu^{2+}/Cu redox pair is calculated using the following Nernst equation

$$\begin{aligned} E_T &= -\frac{\Delta G_T}{nF} = -\frac{1}{nF}\left(\Delta G_T^0 + RT\ln Q\right) \\ &= -\frac{1}{nF}\left[\left(\Delta H_T^0 - T\Delta S_T^0\right) + RT\ln Q\right] \\ &= -\frac{1}{nF}\left[\left(\Delta H_{298}^0 + \int_{298}^{T}\Delta C_P dT\right) - T\left(\Delta S_{298}^0 + \int_{298}^{T}\frac{\Delta C_P}{T}dT\right) + RT\ln Q\right] \end{aligned} \quad (18)$$

where, n is the number of electrons transferred in the reaction, F is the Faraday constant. Q is the reaction quotient, which is expressed as the following equation 19 for the calculation of the redox potential of Cu^{2+}/Cu redox pair considering that the activities of metallic copper, proton and hydrogen gas in the standard hydrogen electrode are all 1:

$$Q = \frac{a_{Cu} a_{H^+}^2}{a_{Cu^{2+}} a_{H_2}} = \frac{1}{a_{Cu^{2+}}} \tag{19}$$

It should be noted that ΔG_T is the "total" Gibbs free energy change, including hydrogen oxidation as in the total cell equation 17. When CuO and Cu^{2+} ions are in equilibrium at pH 9 (the activity of Cu^{2+} ions is 7.5×10^{-11}), the redox potential is calculated to be 0.036 V vs. SHE using the equation 18. This equilibrium relationship is represented by the horizontal line 2 (Fig. 2b). The region above the line is the stability region of Cu^{2+} ions (with a higher oxidation number +II) and the region below the line is the stability region of Cu (with a lower oxidation number 0).

Equation 10 is a simple acid-base reaction (proton exchange reaction), and a unique equilibrium pH is determined. Similarly, equation 15 is a simple redox reaction (electron exchange reaction), and a unique equilibrium potential is determined. There is also a combination reaction of an acid-base reaction and a redox reaction (proton and electron exchange reaction). In this case, the equilibrium relationship is expressed also by Nernst equation (equation 18), which provides a relational expression between potential E_T and pH. In the present system, equilibrium relationships of Cu^{2+}/Cu_2O, CuO/Cu_2O, and Cu_2O/Cu are all categorized as this kind of reaction (proton and electron exchange reaction), corresponding to the diagonal lines 3, 4, and 5 in Fig. 2c, respectively.

After all of the possible equilibrium lines are drawn in the diagram, a part of the equilibrium lines are deleted to leave the common stability regions. Figure 2d is a potential-pH diagram drawn by the above procedure. As can be seen from this procedure, potential-pH diagram can be altered by a set of chemical species and conditions considered, e.g. temperature, activity of ions. Therefore, the selection of chemical species in addition to the activity settings is important and should be done according to an intended purpose: for example, dissolution of chemicals, corrosion proof, electrolysis, and electroless deposition.

3.4 Potential-pH diagrams for Cu and Cu_2O nanoparticles formation system

Figure 3 shows potential–pH diagrams drawn at an equilibrium activity of the Cu^{2+} aquo ion in the presence of abundant CuO at pH 9 in accordance with the procedure written in the section 3.3. The equilibrium activity of Cu^{2+} aquo ion is determined by considering an equilibrium reaction of Cu^{2+} and CuO, where the activities of CuO and H^+ are assigned to 1 and 10^{-9}, respectively. The species Cu^{2+}, CuO, Cu_2O, and Cu are only considered, and oxyanions of copper, such as $HCuO_2^-$ and CuO_2^{2-}, or hydroxyanions of copper, $Cu(OH)_3^-$ and $Cu(OH)_4^{2-}$, are not considered. This is only for simplicity because the redox potentials of CuO/Cu_2O and Cu_2O/Cu redox pairs at pH 9 are important for the prediction of synthesized chemical species in the present system. For instance, the redox potentials of Cu_2O and any Cu(II) ionic species in equilibrium are the same at a constant pressure, pH, and temperature in the presence of abundant solid CuO powder, resulting in the same redox potential of the Cu_2O/Cu redox pair. The equilibrium activity of Cu^{2+} aquo ion changes with pH and therefore, the potential–pH diagram is only valid at the constant pH considered. Fortunately, pH of the solution barely changes during the reaction in the

present case, and it is possible to determine the most stable chemical species throughout the reaction by comparing the measured mixed potential to the vertical line of a constant pH (in this case, pH 9) in the potential–pH diagrams. In other words, the kind of chemical species synthesized can be controlled if the mixed potential is changed as expected.

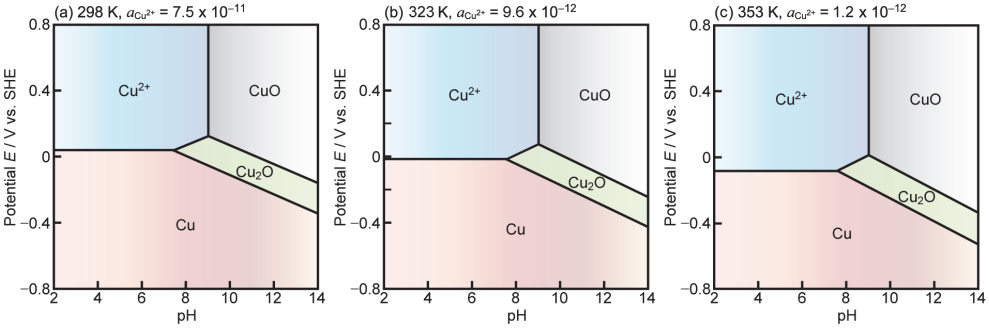


Fig. 3. Potential–pH diagrams drawn at the equilibrium activities of Cu^{2+} aquo ion at pH 9 and (a) 298, (b) 323, and (c) 353 K considering the species Cu^{2+}, CuO, Cu_2O, and Cu. Oxyanions of copper, $HCuO_2^-$ and CuO_2^{2-}, or hydroxyanions of copper, $Cu(OH)_3^-$ and $Cu(OH)_4^{2-}$, are not considered for simplicity. These potential–pH diagrams are only valid at the vertical line of pH 9 because the equilibrium activity of Cu^{2+} aquo ion changes with pH in this system.

	Redox potential at pH 9 (V vs. SHE)			Redox potential at pH 12 (V vs. SHE)			Reaction no.
Partial reaction	298 K	323 K	353 K	298 K	323 K	353 K	
$N_2 + 4H_2O + 4e = N_2H_4 + 4OH^-$	–0.92	–0.98	–1.06	–1.09	–1.17	–1.27	(1)
$1/2N_2 + NH_3 + H_2O + e = N_2H_4 + OH^-$	–2.59	–2.72	–2.88	–2.77	–2.91	–3.09	(2)
Cu(II) + 2e = Cu	0.04	–0.02	–0.08	–0.14	–0.21	–0.29	(3)
$2Cu(II) + 2OH^- + 2e = Cu_2O + H_2O$	0.13	0.07	0.01	–0.05	–0.12	–0.20	(4)
$2H_2O + 2e = H_2 + 2OH^-$	–0.18	–0.19	–0.21	–0.35	–0.38	–0.42	(5)

Table 2. List of redox potentials calculated for partial reactions considered in this system. Activity of each chemical species is substituted by the actual molar concentration in the present reaction system, and specifically, activities of H_2 and NH_3 are both substituted by 10^{-6} to obtain reference values.

3.5 How to control the mixed potential?

As described in the section 3.1, the mixed potential is determined by the balance of each partial current. Therefore, the mixed potential is altered when the redox potential of each partial reaction is changed. The redox potential E_T at a temperature T and 1 atm is calculated using the Nernst equation (equation 18). Table 2 shows calculated results of the redox potentials of all the partial reactions considered. Activity of each chemical in the reaction quotient is

substituted by the same value as the actual molar concentration of the chemical in the present reaction system, and specifically, both activities of H_2 and NH_3, which cannot be determined exactly, are substituted by 10^{-6} to obtain referential values. As shown in Table 2, all the redox potentials shift to the negative direction with increasing both pH and temperature, which can result in a negative shift of the mixed potential. Even the redox potentials of reactions 3 and 4 shift to the negative direction with increasing pH; the former is normally constant with pH and the latter increases with the increase in pH at a constant activity of Cu(II) ions. This is because the equilibrium activity of Cu^{2+} aquo ion decreases with increasing pH under the coexistence of abundant solid CuO in equilibrium, which brings about the negative shift of the redox potentials of reactions 3 and 4. Thus, possibly the mixed potential is shifted to the negative direction with increasing pH and temperature because the redox potentials of all the partial reactions shift to the negative direction. Consequently, factors to change the mixed potential are pH and temperature, which are easily controllable.

4. Experimental confirmation

Is it actually possible to control oxidation state of nanoparticles by controlling the mixed potential? This query was substantiated through the following experimental procedure.

4.1 Procedure

Reaction solutions were prepared using cupric oxide (CuO) (Kanto Chemical, Inc.), sodium hydroxide (NaOH), hydrazine monohydrate ($N_2H_4 \cdot H_2O$) (Nacalai Tesque, Inc.), and gelatin (Jellice Co. Ltd., P459) as received. All chemicals except for gelatin were of reagent grade. Gelatin was added as a dispersing agent to avoid agglomeration and for the suppression of particle growth. Gelatin blocked oxygen and was also effective in preventing the oxidation of the resulted nanoparticles. Reactions were conducted in a Pyrex beaker 250 cm^3 in capacity by the following procedure. First, a CuO colloidal aqueous suspension was prepared by dispersing 0.060 moles of CuO particles in 42.0 cm^3 of distilled water using ultrasound. Next, 18.0 cm^3 of 10 wt % gelatin aqueous solution was added into the solution as a dispersing agent. The initial pH of the mixed solution was adjusted to the reaction pH (9.0–12.0) at 298 K by 1.0 mol dm^{-3} sodium hydroxide aqueous solution with a pH meter (Horiba, Ltd., D-21). The temperature of the solution was kept at the reaction temperature (298–353 K) in a water bath with nitrogen gas bubbling (50 cm^3 min^{-1}), which started 30 min before the reaction and lasted throughout the reaction to eliminate the effect of dissolved oxygen. The solution was agitated at a rate of 500 rpm with a magnetic stirring unit. Then, 42.0 cm^3 of 1.43 mol dm^{-3} hydrazine aqueous solution was added to 18.0 cm^3 of 10 wt % gelatin aqueous solution, and this solution was kept at the reaction temperature (298–353 K) with nitrogen gas bubbling (50 cm^3 min^{-1}) for 30 min. The initial pH of the hydrazine solution was adjusted to the reaction pH (9.0–12.0) by 1.0 mol dm^{-3} sodium hydroxide aqueous solution at 298 K. Next, the hydrazine solution was added to the CuO aqueous suspension as a reducing agent to start the reaction. The total amount of the reaction suspension was 120.0 cm^3, and thus the reaction suspension was 0.50 mol dm^{-3} hydrazine aqueous solution with 0.50 mol dm^{-3} dispersed CuO. Gelatin was not added in the reaction suspension at 298 K because gelatin became a gel at 298 K, inhibiting the reaction. The crystalline structure of precipitates was investigated by X-ray diffraction (XRD: MAC Science Co., Ltd., M03XHF22) using a molybdenum X-ray tube. The morphology of precipitates was observed using a field-emission scanning electron microscope (JEOL Ltd., JSM-6500F). The immersion potential of Au-sputtered round quartz-crystal substrates 5.0 mm

diameter was measured during several experimental runs by a potentiostat/ galvanostat (Hokuto Denko Co., Ltd., HA-151), and it was assumed that the measured immersion potential was almost the same as the mixed potential in reaction suspension in this system. A Ag/AgCl electrode (Horiba, Ltd., 2565A-10T) was used as a reference electrode, and the internal liquid, 3.33 mol dm^{-3} KCl aqueous solution, was replaced for each experimental run. The measured potential was converted to values vs. SHE using the following empirical equation given by Horiba, Ltd. The potential of the reference electrode could be expressed as a function of temperature T (K) as E (mV) vs. SHE = +206 − 0.7(T – 298).

4.2 Results

In the present method, pH of the reaction suspension is initially adjusted at 298 K, and actual pH at reaction temperatures differs from the adjusted value. This is ascribable to the change in the ionic product for water, $K_w = [H^+][OH^-]$; for example, the value of K_w (mol^2 dm^{-6}) is 1.008 x 10^{-14} at 298.15 K and 5.476 x 10^{-14} at 323.15 K. Figure 4 shows the change with temperature in pH, which is initially adjusted to 9.0 at 298 K. pH linearly decreases with increasing temperature, and the values are 9.0 at 298 K, 8.5 at 323 K, and 7.6 at 353 K. Figure 5 shows the change in the mixed potential during the reaction for 2 h. At the same temperature of 323 K, the mixed potential at pH 10.4 (Fig. 5d) is lower than that at pH 8.5 (Fig. 5b). This agrees well with the discussion in the previous theory section.

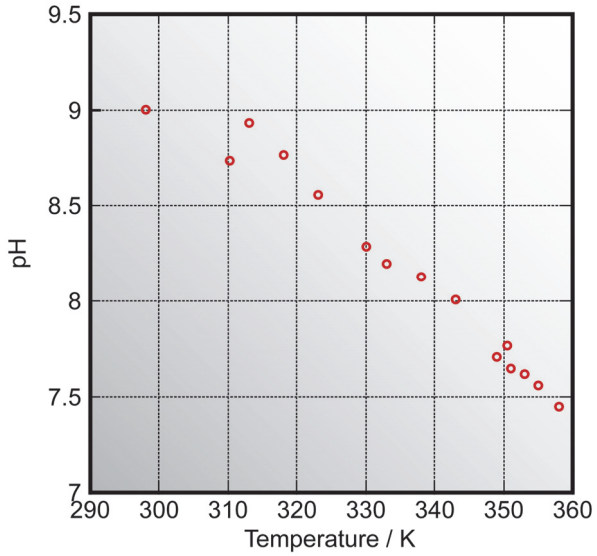


Fig. 4. Change in pH of the reaction suspension with temperature; pH of the reaction suspension was initially adjusted to 9.0 at 298 K.

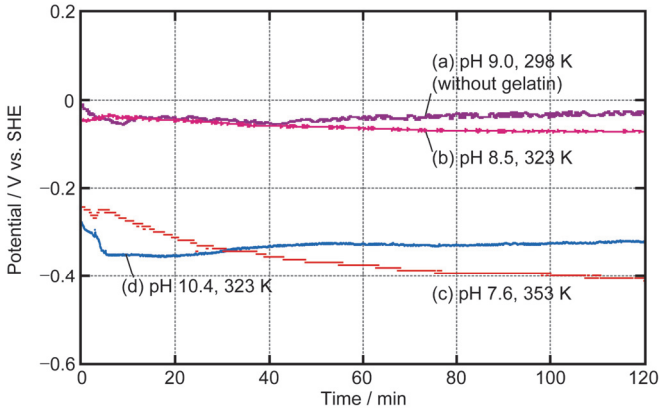


Fig. 5. Mixed potentials measured during the reaction at (a) pH 9.0, 298 K (without gelatin), (b) pH 8.5, 323 K, (c) pH 7.6, 353 K, and (d) pH 10.4, 323 K.

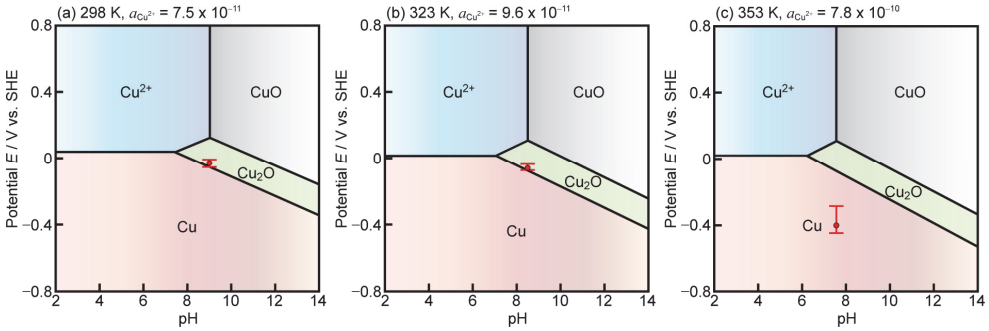


Fig. 6. Potential–pH diagrams drawn at the equilibrium activities of Cu^{2+} aquo ion at (a) pH 9.0, 298 K, (b) pH 8.5, 323 K, and (c) pH 7.6, 353 K considering the species Cu^{2+}, CuO, Cu_2O, and Cu. Bars and circles indicate the range and average value of the measured mixed potential, respectively.

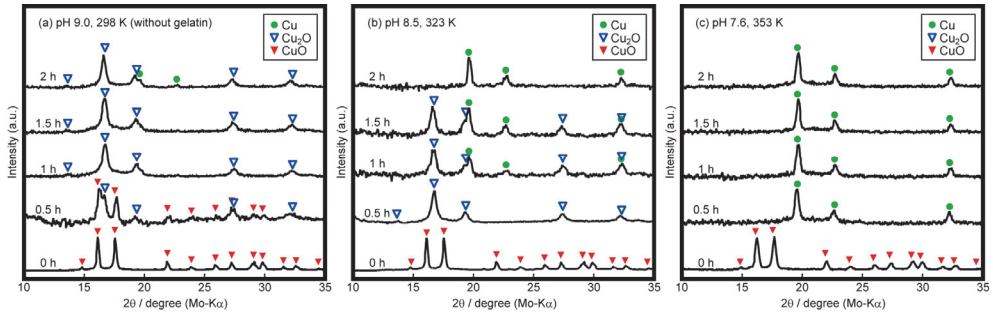


Fig. 7. XRD patterns of precipitates obtained during the reaction at (a) pH 9.0, 298 K (without gelatin), (b) pH 8.5, 323 K, and (c) pH 7.6, 353 K for 2 h.

Moreover, the mixed potential shifts to the negative direction with increasing temperature although pH slightly decreases, indicating that the mixed potential can be lowered just by increasing temperature without fixing pH. By comparing the mixed potential and the potential–pH diagrams shown in Fig. 6, the mixed potential was in the stability region of Cu_2O both at pH 9.0, 298 K (Fig. 6a) and pH 8.5, 323 K (Fig. 6b), whereas the mixed potential was in the stability region of metal Cu at pH 7.6, 353 K (Fig. 6c). Figure 7 shows the change with time in XRD patterns of precipitates during the reaction for 2 h. CuO powder was reduced to Cu_2O after 1 h at pH 9.0, 298 K although slight Cu peaks are recognized in the XRD pattern taken at 2 h. At pH 7.6, 353 K, CuO powder was completely reduced to metal Cu after 0.5 h. These results are almost consistent with the result that the mixed potentials measured at pH 7.6, 353 K and pH 9.0, 298 K were always in the stability region of Cu and Cu_2O, respectively. In contrast, at pH 8.5, 323 K, all CuO powder was reduced to Cu_2O after 0.5 h and then the resulting Cu_2O was gradually reduced to Cu metal. This is not consistent with the measurement result of the mixed potential. By carefully observing the change in the mixed potential at pH 8.5, 323 K, the initial mixed potential is −0.045 V vs. SHE and in the stability region of Cu_2O, and the mixed potential gradually decreases to −0.072 V with reaction time, which is just beside the calculated redox potential of Cu_2O/Cu, −0.077 V vs. SHE; i.e., as CuO is reduced to Cu_2O, the mixed potential approaches the redox potential of Cu_2O/Cu redox pair, where Cu_2O and Cu coexist. Nevertheless, the mixed potential is 5 mV higher than the calculated redox potential of Cu_2O/Cu and it is difficult to adequately elaborate on the result that the deposition of Cu completely proceeded after 2 h. Here, pH of the solution tends to slightly decrease during the reaction, although ideally, pH would not be changed by taking into account the overall reaction in the present reaction system. For example, pH actually decreased from 8.5 to approximately 8 at 323 K, resulting in a slight increase in the redox potential of a Cu_2O/Cu redox pair. The extra pH change is caused by the decrease in the equilibrium concentration of protonated hydrazine $N_2H_5^+$ due to the decrease in the total amount of hydrazine species by decomposition or oxidization during the reaction. Furthermore, there can be a slight difference between the mixed potential measured using Au electrode and the actual mixed potential in the reaction suspension. Therefore, the actual mixed potential is slightly lower than the redox potential of the Cu_2O/Cu redox pair. Nonetheless, a Au electrode is preferable for the measurement of the mixed potential in this reaction system because a Au electrode does not have intense catalytic activity for a specific partial reaction possible in the reaction system, and the effect of the immersion of the Au substrate is extremely low (Iacovangelo, 1991; Iacovangelo & Zarnoch, 1991).

The scanning electron microscopy (SEM) images of the precipitates obtained by the reaction are shown in Fig. 8. Relatively large Cu_2O particles with an average diameter of 475 nm are observed in the precipitates obtained after 2 h at pH 9.0, 298 K. These large particles are attributable to the absence of a dispersing agent, gelatin. At pH 8.5, 323 K, Cu_2O particles with an average diameter of 97 nm are observed after 0.5 h and Cu particles with an average diameter of 82 nm are observed after 2 h. Cu particles with an average diameter of 55 nm are observed after 2 h at 353 K. As above, the mean particle size is decreased with the increase in temperature because nucleation site increases with the increase in temperature and an inordinate number of nucleation sites at high temperatures leads to the absence of reactant ions (Cu(II) or Cu(I) ions) for the growth.

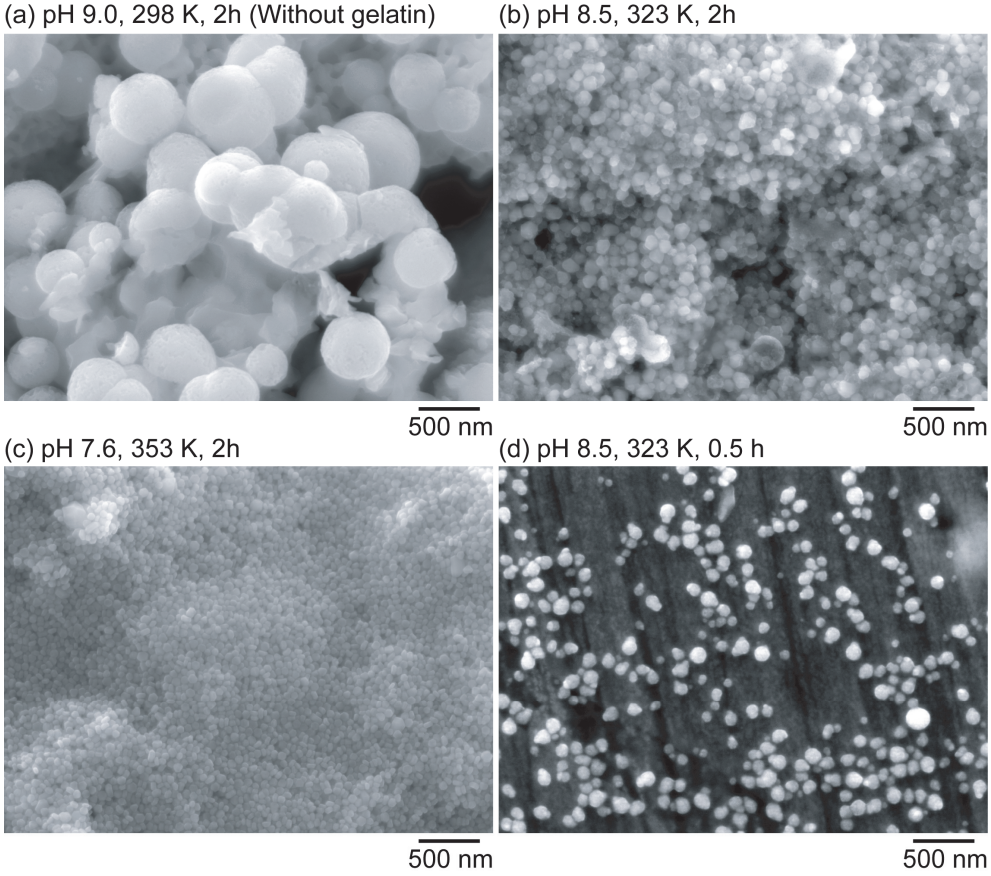


Fig. 8. SEM images of precipitates obtained by the reaction at (a) pH 9.0, 298 K (without gelatin), (b) pH 8.5, 323 K, (c) pH 7.6, 353 K for 2 h, and (d) pH 8.5, 323 K for 0.5 h.

5. Conclusions

Mixed potential is the most effective indicator of solution circumstances, indicating the most stable chemical species in the reaction solution, i.e., thermodynamic considerations in conjunction with monitoring of mixed potential is useful for the in situ prediction of "what chemical species will be synthesized." Solution pH and temperature are the factors which affect the mixed potential, and thus, the most stable chemical species can be changed by pH and temperature. Specifically in this system, the mixed potential decreases with the increase both in pH and temperature. Using this concept, the selective synthesis of Cu and Cu_2O nanoparticles was demonstrated. This concept is also effective in other reaction systems where thermodynamic data are available. Furthermore in another journal (Yagi et al, 2010), the authors also reported that this concept can be widely applied and helps us to thermodynamically consider even those systems where insufficient or no thermodynamic data are available by experimentally measuring the redox potential with voltammetry combined with quartz crystal microbalance.

6. Acknowledgment

The author thanks Mr. Shohei Shiomi and Mr. Hidetaka Nakanishi from Kyoto University for their experimental help. The author is also grateful to Professor Tetsu Ichitsubo, Professor Eiichiro Matsubara from Kyoto University, Professor Emeritus Yasuhiro Awakura at Kyoto University, and Professor Takuya Iida from Osaka Prefecture University for fruitful discussions. This research was supported by a grant-in-aid for Knowledge Cluster Initiative (Kyoto Nanotechnology Cluster), Grant-in-Aid for the Global COE Program (International Center for Integrated Research and Advanced Education in Materials Science), and Grant-in-Aid for Young Scientists (B 20760505), all from the Japan Society for the Promotion of Science, Ministry of Education, Culture, Sports, Science and Technology (MEXT) of Japan. This research was also supported by the Kurata Memorial Hitachi Science and Technology Foundation and Shorai Foundation for Science and Technology. The publication cost of this article was assisted by Special Coordination Funds for Promoting Science and Technology commissioned by MEXT of Japan.

7. References

Beverskog B. & Puigdomenech I. (1997), Revised Pourbaix Diagrams for Copper at 25 to 300°C, *Journal of the Electrochemical Society*, Vol.144, No. 10, pp.3476-3483, ISSN 0013-4651

Bockris, J. O'M; Reddy, A. K. N. & Gamboa-Aldeco, M. (2000). *Modern Electrochemistry 2A: Fundamentals of Electrodics* (2nd ed.), Kluwer Academic/Plenum Publishers, ISBN 0-306-46167-6, New York, USA.

Cheng, Z.; Xu, J, Zhong, H., Chu, X.Z., & Song J. (2011). Repeatable Synthesis of Cu_2O Nanorods by a Simple and Novel Reduction Route, *Materials Letters*, Vol.65, pp.1871-1874, ISSN 0167-577X

Criss, C. M. & Cobble, J. W. (1964), The Thermodynamic Properties of High Temperature Aqueous Solutions. IV. Entropies of the Ions up to 200° and the Correspondence Principle, *Journal of the American Chemical Society*, Vol.86, pp.5385-5393, ISSN 0002-7863

Dhas, N. A.; Raj, C. P. & Gedanken, A. (1998). Synthesis, Characterization, and Properties of Metallic Copper Nanoparticles, *Chemistry of Materials*, Vol.10, pp.1446-1452, ISSN 0897-4756

Gou, L. & Murphy, C. J. (2003). Solution-Phase Synthesis of Cu_2O Nanocubes, *Nano Letters*, Vol.3, No.2, pp.231-234, ISSN 1530-6984

Gu, Y.-E; Zhang, Y., Zhang, F., Wei, J., Wang, C., Du, Y., & Ye, W. (2010). Investigation of Photoelectrocatalytic Activity of Cu_2O Nanoparticles for p-Nitrophenol using Rotating Ring-Disk Electrode and Application for Electrocatalytic Determination, *Electrochimica Acta*, Vol.56, pp.953-958, ISSN 0013-4686

Han, W. K.; Choi, K. W., Hwang, G. H., Hong, S. J., Lee, J. S., & Kang, S. G. (2006). Fabrication of Cu Nano Particles by Direct Electrochemical Reduction from CuO Nano Particles, *Applied Surface Science*, Vol.252, pp.2832-2838, ISSN 0169-4332

Huang, H. H.; Yan, F. Q., Kek, Y. M., Chew, C. H., Xu, G. Q., Ji, W., Oh, P. S., & Tang, S. H. (1997). Synthesis, Characterization, and Nonlinear Optical Properties of Copper Nanoparticles, *Langmuir*, Vol.13, pp.172-175, ISSN 0743-7463

Huang, L.; Jiang, H., Zhang, J., Zhang, Z., & Zhang, P. (2006). Synthesis of Copper Nanoparticles Containing Diamond-Like Carbon Films by Electrochemical Method, *Electrochemistry Communications*, Vol.8, pp.262-266, ISSN 1388-2481

Iacovangelo, C. D. & Zarnoch, K. P. (1991). Substrate-Catalyzed Electroless Gold Plating, *Journal of the Electrochemical Society*, Vol.138, No.4, pp.983-988, ISSN 0013-4651

Iacovangelo, C. D. (1991). Autocatalytic Electroless Gold Deposition Using Hydrazine and Dimethylamine Borane as Reducing Agents, *Journal of the Electrochemical Society*, Vol.138, No. 4, pp.976-982, ISSN 0013-4651

Joshi, S. S.; Patil, S. F., Iyer, V., & Mahumuni, S. (1998). Radiation Induced Synthesis and Characterization of Copper Nanoparticles, *Nanostructured Materials*, Vol.10, No.7, pp.1135-1144, ISSN 0965-9773

Kubaschewski, O. & Alcock, C. B. (1979), *Metallurgical Thermochemistry* (5th ed. Revised and enlarged), Elsevier, ISBN 0-08-022107-6, New York, USA

Kumar, R. V.; Mastai, Y., Diamant, Y., & Gedanken, A. (2001). Sonochemical Synthesis of Amorphous Cu and Nanocrystalline Cu_2O Embedded in a Polyaniline Matrix, *Journal of Materials Chemistry*, Vol.11, pp.1209-1213, ISSN 0959-9428

Latimer, W. M. (1959), *The Oxidation States of the Elements and Their Potentials in Aqueous Solutions* (2nd ed.), Prentice-Hall, Englewood Cliffs, New Jersey, USA.

Li J.; Liu, C.-Y. & Xie, Z. (2011). Synthesis and Surface Plasmon Resonance Properties of Carbon-coated Cu and Co Nanoparticles, *Materials Research Bulletin*, Vol.46, pp.743-747, ISSN 0025-5408

Lisiecki, I. & Pileni, M. P. (1993). Synthesis of Copper Metallic Clusters Using Reverse Micelles as Microreactors, *Journal of the American Chemical Society*, Vol.115, pp.3887-3896, ISSN 0002-7863

Lisiecki, I.; Billoudet, F. & Pileni, M. P. (1996). Control of the Shape and the Size of Copper Metallic Particles, *Journal of Physical Chemistry*, Vol.100, pp.4160-4166, ISSN 0022-3654

Liu, R.; Oba, F., Bohannan, E. W., Ernst, F., & Switzer, J. A. (2003). Shape Control in Epitaxial Electrodeposition: Cu_2O Nanocubes on InP(001), *Chemistry of Materials*, Vol.15, pp.4882-4885, ISSN 0897-4756

Liu, X.; Geng, B., Du, Q., Ma, J., & Liu, X. (2007). Temperature-Controlled Self-Assembled Synthesis of CuO, Cu_2O and Cu Nanoparticles Through a Single-Precursor Route, *Materials Science and Engineering: A*, Vol.448, pp.7-14, ISSN 0921-5093

Lu, C.; Qi, L., Yang, J., Wang, X., Zhang, D., Xie, J., & Ma, J. (2005). One-Pot Synthesis of Octahedral Cu_2O Nanocages via a Catalytic Solution Route, *Advanced Materials*, Vol.17, pp.2562-2567, ISSN 1521-4095

Mahajan, M.B.; Pavan, M.S. & Joy, P.A. (2009). Ferromagnetic Properties of Glucose Coated Cu_2O Nanoparticles, *Solid State Communications*, Vol.149, pp.2199-2201, ISSN 0038-1098

Muramatsu, A. & Sugimoto T. (1997). Synthesis of Uniform Spherical Cu_2O Particles from Condensed CuO Suspensions, *Journal of Colloid and Interface Science*, Vol.189, pp.167-173, ISSN 0021-9797

Ohde H.; Hunt, F. & Wai, C. M. (2001). Synthesis of Silver and Copper Nanoparticles in a Water-in-Supercritical-Carbon Dioxide Microemulsion, *Chemistry of Materials*, Vol.13, pp.4130-4135, ISSN 0897-4756

Pileni, M. P. & Lisiecki, I., (1993). Nanometer Metallic Copper Particle Synthesis in Reverse Micelles, *Colloids and Surfaces A: Physicochemical and Engineering Aspects,* Vol.80, pp.63-68, ISSN 0927-7757

Poizot, P.; Laruelle, S., Grugeon, S., Dupont, L., & Tarascon, J-M. (2000). Nano-sizedtransition-metaloxidesas negative-electrode materials for lithium-ion batteries, *NATURE*, Vol. 407, pp.496-499, ISSN 0028-0836

Pourbaix, M. (1974). *Atlas of Electrochemical Equilibria in Aqueous Solutions* (2nd ed.), NACE International, Cebelcor, Brüssel, Belgium

Qi, L.; Ma, J. & Shen J. (1997). Synthesis of Copper Nanoparticles in Nonionic Water-in-Oil Microemulsions, *Journal of Colloid and Interface Science*, Vol.186, pp.498-500, ISSN 0021-9797

Singh, M.; Sinhaa, I., Premkumarc, M., Singhc, A.K., & Mandal, R.K. (2010). Structural and Surface Plasmon Behavior of Cu Nanoparticles using Different Stabilizers, *Colloids and Surfaces A: Physicochemical and Engineering Aspects*, Vol.359, pp.88-94, ISSN 0927-7757

Siwach, O. P. & Sen, P., (2010). Study of fluorescence properties of Cu nanoparticles, *Solid State Sciences,* Vol.12, pp.1107-1111, ISSN 1293-2558

Snoke D. (2002). Spontaneous Bose Coherence of Excitons and Polaritons, *Science*, Vol.298, pp.1368-1372, ISSN 0036-8075

Stull, D. R. & Prophet, H. (1971), *JANAF Thermochemical Tables* (2nd ed.), NSRDS-NBS, Washington, DC, USA

Wang, W.; Wang, G., Wang, X., Zhan, Y., Liu, Y., & Zheng, C. (2002). Synthesis and Characterization of Cu_2O Nanowires by a Novel Reduction Route, *Advanced Materials*, Vol.14, No. 1, pp.67-69, ISSN 1521-4095

Yagi, S.; Nakanishi, H., Ichitsubo, T. & Matsubara, E. (2009). Oxidation-State Control of Nanoparticles Synthesized via Chemical Reduction Using Potential Diagrams, *Journal of the Electrochemical Society*, Vol.156, No.8, pp.D321-D325, ISSN 0013-4651

Yagi, S.; Kawamori, M. & Matsubara, E. (2010). Electrochemical Study on the Synthesis Process of Co–Ni Alloy Nanoparticles via Electroless Deposition, *Journal of the Electrochemical Society*, Vol.157, No.5, pp.E92-E97, ISSN 0013-4651

Yeh, M.-S.; Yang, Y.-S., Lee, Y.-P., Lee, H.-F., Yeh, Y.-H., & Yeh, C.-S. (1999). Formation and Characteristics of Cu Colloids from CuO Powder by Laser Irradiation in 2-Propanol, *The Journal of Physical Chemistry B*, Vol.103, pp.6851-6857, ISSN 1089-5647

Young, A. P. & Schwartz, C. M. (1969). Electrical Conductivity and Thermoelectric Power of Cu_2O, *Journal of Physics and Chemistry of Solids*, Vol.30, pp.249-252, ISSN 0022-3697

Yu, L.; Sun, H., He, J., Wang, D., Jin, X., Hu, X., & Chen, G. Z. (2007). Electro-reduction of Cuprous Chloride Powder to Copper Nanoparticles in an Ionic Liquid, *Electrochemistry Communications*, Vol.9, pp.1374-1381, ISSN 1388-2481

Zhang, C.Q.; Tu, J.P., Huang, X.H., Yuan, Y.F., Chen, X.T., & Mao F. (2007). Preparation and Electrochemical Performances of Cubic Shape Cu_2O as Anode Material for Lithium Ion Batteries, *Journal of Alloys and Compounds*, Vol.441, pp.52–56, ISSN 0925-8388

Zhang, J.; Liu, J., Peng, Q., Wang, X., & Li, Y. (2006). Nearly Monodisperse Cu_2O and CuO Nanospheres: Preparation and Applications for Sensitive Gas Sensors, *Chemistry of Materials*, Vol.18, pp.867-871, ISSN 0897-4756

Zhang, X.D.; Xi, J.F., Shen, Y.Y., Zhang, L.H., Zhu, F., Wang, Z., Xue, Y.H. & Liu, C.L. (2011). Thermal Evolution and Optical Properties of Cu Nanoparticles in SiO_2 by Ion Implantation, *Optical Materials*, Vol.33, pp.570-575, ISSN 0925-3467

10

On the Extremum Properties of Thermodynamic Steady State in Non-Linear Systems

Gy. Vincze and A. Szasz
Department of Biotechnics, St.Istvan University, Hungary

1. Introduction

The starting point of irreversible thermodynamics is the local form of the entropy balance equation. In a wide range of non-equilibrium processes, for instance heat conduction, thermo-diffusion non reactive mixture and thermoelectric effects are taking place in a volume V in static conditions, the entropy balance has the following form

$$\frac{\partial s}{\partial t} + \nabla \cdot \mathbf{j}_s = \sigma \tag{1}$$

where

$$. \mathbf{j}_s = \sum_{i=1}^{N} \Gamma_i \mathbf{j}_i \, . \tag{2}$$

is the entropy current density and

$$\sigma = \sum_{i=1}^{N} \nabla \Gamma_i \cdot \mathbf{j}_i \tag{3}$$

is the spontaneous internal entropy production, which is fixed non-negative by the second law of thermodynamics $\sigma \geq 0$ [1], [2]. The symbols are: $\mathbf{j}_i$ is the current density vector of the i-th component, belonging to the extensive state variable a_i; Γ_i is the i-th intensive scalar variable and $\nabla \Gamma_i$ is the i-th thermodynamic force (gradient of the intensive scalar Γ_i). The state variables $a_i, (i = 1,2,...,N)$ and current densities $\mathbf{j}_i, (i = 1,2,...,N)$ are assumed to obey local balance equations of the form

$$\frac{\partial a_i}{\partial t} + \nabla \cdot \mathbf{j}_i = \sigma_i, (i = 1,2,...,N) \tag{4}$$

where σ_i is the source of a_i.

Let us construct a more concise notation:

$$\mathbf{j} := (j_1 = j_{11}, j_2 = j_{12}, j_3 = j_{13}, .., j_{3i-2} = j_{i1}, .., j_{3N} = j_{N3}) \tag{5}$$

(and) *3N* dimensional thermodynamic flux and the

$$\mathbf{X} := \left(X_1 = \frac{\partial \Gamma_1}{\partial x_1}, X_2 = \frac{\partial \Gamma_1}{\partial x_2}, X_2 = \frac{\partial \Gamma_1}{\partial x_3}, .., X_{3i-2} = \frac{\partial \Gamma_i}{\partial x_1}, .., X_{3N} = \frac{\partial \Gamma_N}{\partial x_3} \right) \tag{6}$$

conjugate thermodynamic force vector from the coordinates of the current densities $\mathbf{j}_i$ and $\nabla \Gamma_i$, respectively. The spontaneous entropy production (3) can then be written as a bilinear form

$$\sigma = \sum_{i=1}^{N} \nabla \Gamma_i \cdot \mathbf{j}_i = \sum_{j=1}^{3N} X_j j_j \tag{7}$$

To make the notation simple, we drop out the summation symbol adopting the standard summation convention:

$$\sigma = \sigma\left(X_j, j_j \right) = X_j j_j \tag{8}$$

where any index repeated twice must be summed from *1* to 3*N*.
The *X's* depend on the thermodynamic fluxes j_j. They can also depend on the intensive thermodynamic scalar variables Γ_k. We accordingly make the constitutive assumptions

$$X_j = X_j\left(j_l, \Gamma_k \right), \left(j = 1, 2, ..., 3N \right) \tag{9}$$

Thus from (7) and (8) we get

$$\sigma = X_j j_j = \sigma\left(j_l; X_j\left(j_l, \Gamma_k \right) \right) = \sigma^*\left(j_l, \Gamma_k \right) \geq 0 \tag{10}$$

The thermodynamic fluxes and the entropy production vanish at thermodynamic equilibrium

$$\sigma^*\Big|_{eq} = \sigma^*\left(0, \Gamma_k \right) = 0 \tag{11}$$

and hence $j_j = 0$ is a minimizing point of $\sigma^*\left(j_l, \Gamma_k \right)$, thus we have

$$\left. \frac{\partial \sigma^*}{\partial j_j} \right|_{j_l = 0} = 0 \tag{12}$$

The differentiation of (10) and the above relations imply the equilibrium conditions

$$X_j\Big|_{eq} = X_j\left(0, \Gamma_k \right) = 0, \left(j = 1, 2, ..., 3N \right) \tag{13}$$

1.1 The Onsager constitutive theory and its non-linear generalization

The Onsager's irreversible thermodynamics [3] states linear constitutive relations of the form

$$X_j = R_{jl}\left(\Gamma_k \right) j_l \tag{14}$$

where the kinetic coefficients $R_{jl}(\Gamma_k)$ are independent of the *j's* and satisfy the well known symmetry, the reciprocity relations

$$R_{jl} = R_{lj} \tag{15}$$

The coefficients $R_{jl}(\Gamma_k)$ must satisfy the second law of thermodynamics too

$$\sigma^* = X_j(j_l, \Gamma_k) j_j = j_j R_{jl} j_l \geq 0 \tag{16}$$

Due to the conditions (15) and (16) R_{jl} is positive definite matrix. Furthermore, there are two independent requirements, for the constitutive relations (13): the reciprocity relations (15) and the second law of thermodynamics $\sigma^* \geq 0$. Additionally, the differentiation of (13) gives

$$\frac{\partial X_j}{\partial j_l} = R_{jl} \tag{17}$$

and hence the reciprocity relations (15) can be written in the equivalent form

$$\frac{\partial X_j}{\partial j_l} = \frac{\partial X_l}{\partial j_j} \tag{18}$$

This form of the reciprocity relations is also independent of the assumption that the constitutive relations are linear in the *j's*. The equilibrium conditions (13) are always satisfied, independently of the actual form of the constitutive relations.

In the following we are going to introduce the generalized Onsager constitutive theory for a non-linear system of constitutive (9) satisfying the (18) reciprocity relations, the (13) equilibrium conditions and the (10) second law of thermodynamics. In the following we shall present that the Edelen's decomposition theorem [4] is valid in every class of the thermodynamic forces, which are two times continuously differentiable with respect to fluxes.

This theorem states that every thermodynamic force can be decomposed into the sum of two unique forces. The first identically satisfies the reciprocity relations and the second gives an identically vanished entropy production. The first type of forces can be a deduction of a so called flux dissipation potential.

Let us denote the thermodynamic forces of any system in general by $X_j = X_j(j_l, \Gamma_k)$. There are existing unique forces $X_j^*(j_l, \Gamma_k)$ and $V_j(j_l, \Gamma_k)$ for each $X_j(j_l, \Gamma_k)$ like

$$\begin{aligned} &X_j(j_l, \Gamma_k) = X_j^*(j_l, \Gamma_k) + V_j(j_l, \Gamma_k), \\ &\frac{\partial X_j^*}{\partial j_l} = \frac{\partial X_l^*}{\partial j_j}, \end{aligned} \tag{19}$$

and

$$j_j V_j(j_l, \Gamma_k) = 0 \tag{20}$$

which is identically satisfied with every variable (j_l,Γ_k), where

$$\begin{aligned} X_j^*(j_l,\Gamma_k) &= \frac{\partial\Phi(j_l,\Gamma_k)}{\partial j_j}, \\ \Phi(j_l,\Gamma_k) &= \int_0^1 j_j X_j(\tau j_l,\Gamma_k)d\tau \\ V_j(j_l,\Gamma_k) &= \int_0^1 j_m\left(\frac{\partial X_j(\tau j_l,\Gamma_k)}{\partial\tau j_m}-\frac{\partial X_m(\tau j_l,\Gamma_k)}{\partial\tau j_j}\right)d\tau \end{aligned} \tag{21}$$

This is the Edelen's decomposition theorem. Its essence is that every force has a definite separate part $X_j^*(j_l,\Gamma_k)$ which satisfies the reciprocity relations (19) and a certain other separate part $V_j(j_l,\Gamma_k)$, which identically satisfies the relation (20). We refer to this second type of force as non dissipative force. The flux dissipation potential $\Phi(j_l,\Gamma_k)$ is uniquely determined to within an additive function of Γ's. A direct consequence of Edelen`s theorem is when the thermodynamic forces are Onsager forces and this assertion is equivalent to the assumption of the existence of physical systems with zero non dissipative forces.

1.2 The second law of thermodynamics and the dissipation potentials

From the Edelen's decomposition theorem follows that the existence of the scalar potential $\Phi(j_l,\Gamma_k)$ and the associated thermodynamic forces

$$X_j(j_l,\Gamma_k)=\frac{\partial\Phi(j_l,\Gamma_k)}{\partial j_j} \tag{22}$$

which satisfy the reciprocity relation, and these also have to have to satisfy all the supplementary conditions of the generalized Onsager constitutive theory. A/One consequence of the second law of thermodynamics is the existence of a so called flux dissipation potential $\Phi(j_l,\Gamma_k)$ which satisfies the inequality (10), in every case, i.e.

$$\sigma^*(j_l,\Gamma_k)=\sigma\left(j_l;X_j(j_l,\Gamma_k)\right)=j_j\frac{\partial\Phi(j_l,\Gamma_k)}{\partial j_j}\geq 0\,. \tag{23}$$

From the equations (11) and (12) follows

$$\left.\frac{\partial\Phi(j_l,\Gamma_k)}{\partial j_j}\right|_{j_i=0}=X_j(j_l=0,\Gamma_k)=0\,. \tag{24}$$

There are two further important properties of the thermodynamic flux potential $\Phi(j_l,\Gamma_k)$, [4], namely

$$\begin{aligned} &\Phi(j_l,\Gamma_k)>0 \quad for\ j_l\neq 0, \\ &\Phi(0,\Gamma_k)=0 \end{aligned} \tag{25}$$

Also, the potential $\Phi(j_l,\Gamma_k)$ is strictly positive with respect to j's, except in equilibrium point $j_l = 0, (l = 1,2,...,3N)$. The above three relations however, do not guarantee that the dissipation-potential $\Phi(j_l,\Gamma_k)$ is a convex function with respect to the thermodynamic fluxes.
Introducing the entropy production into the second equation of (21) like

$$\Phi(j_l,\Gamma_k) = \int_0^1 \tau\, j_j X_j(\tau\, j_l,\Gamma_k)\frac{d\tau}{\tau} = \int_0^1 \sigma^*(\tau\, j_j,\Gamma_k)\frac{d\tau}{\tau} = \\ = \int_0^\infty \sigma^*(e^{-t} j_j,\Gamma_k)dt \tag{26}$$

where we change the integration variable by the substitution $\tau = e^{-t}$. Also, the flux dissipation potential in the (j_l,Γ_k) state is equal to the time integral of the entropy production in the pure relaxation processes $j_j e^{-t}$ from the initial states j_j to the equilibrium state. This is the physical meaning of the dissipation potential.

1.3 On the validity of Onsager's reciprocal relations

The original Onsager's proof of reciprocal relation is based on three assumptions and on the principle of microscopic reversibility. The first is that the constitutive equations are linear. The second is the so called regression hypothesis. The third Onsager's assumption is that the thermodynamic fluxes are time derivatives of state variables [3]. However, in the case of vector flows, this condition is not satisfied. For example the heat flux vector and the viscous pressure tensor cannot be expressed in such simple time-derivative form in classical non-equilibrium thermodynamics. On the other hand, as Casimir pointed out [5]: the proof of this reciprocity relations is based directly on the hypothesis that the regression of fluctuations follows the same laws as macroscopic irreversible processes do. This is the regression hypothesis. Consequently, the use of the Onsager's reciprocal relation describing of macroscopic phenomena must be based on experimental verification or we have to accept Truesdell's opinion [6]: „if the reciprocal relations are correct we have to derive them by pure phenomenology also".
The experimental investigations prove the validity of the reciprocal relations for several types of irreversible processes; moreover, as we show below the linear and generalized reciprocal relations are clear phenomenological consequence of the principle of minimal entropy production. Consequently, the generalized reciprocal relations are reasonably well-established thermodynamic law, however of course it has no such general validity as the basic laws of thermodynamics, (e.g. the energy conservation or the direction of the spontaneous thermodynamic processes).
Berdichevsky [7] showed in the case of non-linear constitutive equations, when the macro parameters of the thermodynamic system as stochastic variables are Markov-type ones and satisfy the Focker-Planck equation, as well as stationer and stochastically reversible in meaning of Kolmogorov's conditions, then the non-linear constitutive equations of the thermodynamic fluxes and forces fulfill the generalized reciprocal relations. It is specially interesting, that Berdichevsky's results are based on the necessary and sufficient conditions of reversibility of Markov-processes obtained by Kolmogorov in 1937 [8]. Kolmogorov had formulated his reversibility relation on α-type macro-variables which do not change their sign by reverse process. The necessary and sufficient conditions of reversibility for stochastic

processes on β-type variables (such variables which change their sign in reverse process) were formulated by Yaglom [9].
The Markov-character of macro-variables was observed in various physical systems [10], consequently, the generalized Onsager's reciprocal relations are expected to be correct in numerous physical systems. This is supported by another result of Berdichevsky, [11], showing that the thermodynamically slow macro variables are parts of these equations when the dynamics of the system is governed by Hamiltonian equations. When we integrate the Liuville-equation of the above system on the rapidly changing variables, the probability density function of the thermodynamic distribution of the slow changing stochastic macro variables satisfy the Focker-Planck equation. Combining this result with the flux-hypothesis on the probability of current densities, the generalized Onsager's reciprocal relation could be obtained. Consequently, the Hamiltonian character of the macro-dynamics together with some further plausible conditions verify the reciprocal relations. This generalized reciprocal relation may be considered as a reasonably well-established law, but less rigorously valid than the Onsager's original one.

1.4 Minimum principle of the generalized Onsager constitutive theory

Limit our investigation on the generalized Onsager's constitutive theory, containing the linear Onsager's theory as well. In this case the thermodynamic fluxes have no non dissipative parts and the thermodynamic forces could be deduced from the flux dissipation potential by the Edelens' decomposition theorem. Assuming a dissipation potential connecting the thermodynamic forces and fluxes like:

$$X_i = \frac{\partial \Phi\left(J_k, \Gamma_l\right)}{\partial J_i}, (i = 1,2,...,3N) \tag{27}$$

In the following part our aim is to introduce the Legendre transform of flux potential $\Phi\left(J_k, \Gamma_l\right)$.
To perform a Legendre transform, it is necessary that the above constitutive equations are invertible to the fluxes. It is possible, when from the derivate functions $\frac{\partial \Phi\left(J_k, \Gamma_l\right)}{\partial J_i}$ as coordinates constructed

$$\mathbf{X} = \mathbf{X}\left(\mathbf{J}, \Gamma_l\right) := \nabla_{\mathbf{J}} \Phi\left(\mathbf{J}, \Gamma_l\right) \tag{28}$$

gradient mapping is a one-to-one and onto (transformation) mapping between the two open sets of fluxes and forces for any $\frac{\partial \Phi\left(J_k, \Gamma_l\right)}{\partial J_i}$'s, where $\mathbf{J} = \left(J_1, J_2, ..., J_{3N}\right)$ and $\mathbf{X} = \left(X_1, X_2, ..., X_{3N}\right)$ are vectors of $3N$ dimensions obtained from the fluxes and forces. To give a criteria to the invertibility of (27) we introduce the so called Hessian matrix [12]:

$$\mathbf{H}\Phi\left(\mathbf{J}, \Gamma_l\right) = \left[\frac{\partial^2 \Phi\left(\mathbf{J}, \Gamma_l\right)}{\partial J_i \partial J_i}\right] \tag{29}$$

Let G be a convex open set and C^2 be the space of all two times continuously differentiable functions.

The flux potential is strictly convex by definition, when $\Phi \in C^2(G)$ and the Hessian matrix of Φ is positive definite on G [13], [14], i.e.

$$\left[\mathbf{H}\Phi(\mathbf{J},\Gamma_l)\mathbf{J}\right]\cdot\mathbf{J} = \frac{\partial^2 \Phi(\mathbf{J},\Gamma_l)}{\partial J_k \partial J_i} J_i J_k > 0, \forall \mathbf{J} \neq 0 \tag{30}$$

When the flux potential is strictly convex, then the gradient mapping (28) is one-to-one and onto. This condition is sufficient to perform the Legendre transform as follows:
The partial Legendre transform of $\Phi(\mathbf{J},\Gamma_l)$ is defined by solving $\mathbf{X} = \mathbf{X}(\mathbf{J},\Gamma_l)$ for $\mathbf{J} = \mathbf{J}(\mathbf{X},\Gamma_l)$, and then setting

$$\begin{aligned} \Psi(\mathbf{X},\Gamma_l) &= \mathbf{X}\cdot\mathbf{J} - \Phi(\mathbf{J},\Gamma_l) = \\ &= \mathbf{X}\cdot\mathbf{J}(\mathbf{X},\Gamma_l) - \Phi\left(\mathbf{J}(\mathbf{X},\Gamma_l),\Gamma_l\right) \end{aligned} \tag{31}$$

In the theory of Legendre transformation we refer to the fluxes $\mathbf{J} = (J_1, J_2, ..., J_{3N})$ as original coordinates, while the forces $\mathbf{X} = (X_1, X_2, ..., X_{3N})$ will be the conjugate coordinates. The symbol $\mathbf{X}\cdot\mathbf{J}$ is the dot product, so

$$\Psi(\mathbf{X},\Gamma_l) = \mathbf{X}\cdot\mathbf{J} - \Phi(\mathbf{J},\Gamma_l) = \Psi(X_j,\Gamma_l) = X_i \cdot J_i - \Phi\left(J_j,\Gamma_l\right) \tag{32}$$

Proposition. The force potential $\Psi(X_j,\Gamma_l)$ could be derived two times continuously by forces, and

$$\nabla_{\mathbf{X}}\Psi(\mathbf{X},\Gamma_l) = \mathbf{J}(\mathbf{X},\Gamma_l) = \mathbf{J},\ \nabla_{\mathbf{\Gamma}}\Phi(\mathbf{J},\Gamma_l) = -\nabla_{\mathbf{\Gamma}}\Psi(X_j,\Gamma_l),\ \mathbf{\Gamma} = (\Gamma_l,\Gamma_l,...,\Gamma_l) \tag{33}$$

By calculating differential of the force potential

$$\begin{aligned} d\Psi(\mathbf{X},\Gamma_l) &= \nabla_{\mathbf{X}}\Psi(\mathbf{X},\Gamma_l)\cdot d\mathbf{X} + \nabla_{\mathbf{\Gamma}}\Psi(\mathbf{X},\Gamma_l)\cdot d\mathbf{\Gamma} = \\ &= d\mathbf{X}\cdot\mathbf{J} + \mathbf{X}\cdot d\mathbf{J} - \nabla_{\mathbf{J}}\Phi(\mathbf{J},\Gamma_l)\cdot d\mathbf{J} - \nabla_{\mathbf{\Gamma}}\Phi(\mathbf{J},\Gamma_l)\cdot d\mathbf{\Gamma} = \\ &= d\mathbf{X}\cdot\mathbf{J} - \nabla_{\mathbf{\Gamma}}\Phi(\mathbf{J},\Gamma_l)\cdot d\mathbf{\Gamma} \end{aligned} \tag{34}$$

since $\nabla_{\mathbf{J}}\Phi(\mathbf{J},\Gamma_l) = \mathbf{X}$. Hence

$$\nabla_{\mathbf{X}}\Psi(\mathbf{X},\Gamma_l) = \mathbf{J} = \mathbf{J}(\mathbf{X},\Gamma_l), \tag{35}$$

$$\nabla_{\mathbf{\Gamma}}\left[\Phi(\mathbf{J},\Gamma_l) + \Psi(X,\Gamma_l)\right] = 0 \tag{36}$$

The equation (36) is referred to further as Donkin's theorem [14].
It can be shown that the Hessian matrix of $\Psi(X_j,\Gamma_l)$ is positive definite, so if the range of gradient mapping (28) is a convex open set, then the force dissipation potential is also strictly convex.
Since the force dissipation potential $\Psi(X_j,\Gamma_l)$ has continuous second derivative with respect to forces it follows that the generalized reciprocal relations:

$$\frac{\partial \Psi^2(X_j,\Gamma_l)}{\partial X_k \partial X_i} = \frac{\partial \Psi^2(X_j,\Gamma_l)}{\partial X_i \partial X_k}, (i,k = 1,2,...,3N) \tag{37}$$

The Legrende transform by these above assumptions is an involution, which means applying the transformation on the force dissipation potential, the result will be the flux potential.

1.4.1 The Onsager-Machlup-function

The Legendre transformation could be formulated as an extremum-task as well [15]. The Onsager-Machlup function (OM-function) play a central role in the extremum principles of the irreversible thermodynamics based on generalized Onsager constitutive theory. The OM-function could be introduced by the spontaneous entropy production and one of dissipation potentials using the Legendre transformation.

Let the open convex sets G and G^* be the domain of $\Phi(\mathbf{J},\Gamma_l)$ and range of the gradient mapping (28) respectively. If $\Phi(\mathbf{J},\Gamma_l)$ is strictly convex then

$$\Psi(\mathbf{X},\Gamma_l)=\max_{\mathbf{J}\in G}\left[\mathbf{X}\cdot\mathbf{J}-\Phi(\mathbf{J},\Gamma_l)\right],\ \forall\mathbf{X}\in G^* \tag{38}$$

Indeed, the necessary condition of the extremum of $\Psi(\mathbf{X},\Gamma_l)$ is $\mathbf{X}(\mathbf{J},\Gamma_l)=\nabla_{\mathbf{J}}\Phi(\mathbf{J},\Gamma_l)$. Its solution for $\mathbf{J}$ substituted into (38), we obtain the definition (31) of Legendre transform. Due to the $\mathbf{X}\cdot\mathbf{J}-\Phi(\mathbf{J},\Gamma_l)$ can only be equal with its maximum, thus in case of strictly convex $\Phi(\mathbf{J},\Gamma_l)$ and $\Psi(\mathbf{X},\Gamma_l)$ dissipation potentials we get the Young-inequality from (38):

$$\mathbf{X}\cdot\mathbf{J}\le\Psi(\mathbf{X},\Gamma_l)-\Phi(\mathbf{J},\Gamma_l),\ \forall\mathbf{X}\in G^*,\forall\mathbf{J}\in G \tag{39}$$

where the equality can be satisfied only when the constitutive equations (28) and (35) are satisfied. Defining the Onsager-Machlup function as:

$$\begin{aligned} o(\mathbf{X},\mathbf{J},\Gamma_l)&=\Psi(\mathbf{X},\Gamma_l)+\Phi(\mathbf{J},\Gamma_l)-\mathbf{X}\cdot\mathbf{J}=\\ &=\Psi(\mathbf{X},\Gamma_l)+\Phi(\mathbf{J},\Gamma_l)-\sigma(\mathbf{X},\mathbf{J})\ge 0, \end{aligned} \tag{40}$$

Also, this contains the entropy production, and was introduced first by Onsager and Machlup in their work about non-equilibrium fluctuation theory [16]. With this we can formulate the minimum theorem of the generalized Onsager constitutive theory: the OM-function is the non-negative function of the fluxes, forces and intensive parameters. It only becomes zero, which is its minimum, when the material equations of the generalized Onsager constitutive theory are satisfied. The OM-function has crucial importance, because it contains all the important constitutive properties of the linear and generalized constitutive theories; these follow from the necessary conditions of the minimum of OM-function:

$$\begin{aligned} &\nabla_{\Gamma}o(\mathbf{X},\mathbf{J})=0\rightarrow\nabla_{\Gamma}\left[\Phi(\mathbf{J},\Gamma_l)+\nabla_{\mathbf{X}}\Psi(\mathbf{X},\Gamma_l)\right]=0,\\ &\nabla_{\mathbf{J}}o(\mathbf{X},\mathbf{J})=0\rightarrow\nabla_{\mathbf{J}}\Phi(\mathbf{J},\Gamma_l)=\mathbf{X},\\ &\nabla_{\mathbf{X}}o(\mathbf{X},\mathbf{J})=0\rightarrow\nabla_{\mathbf{X}}\Psi(\mathbf{X},\Gamma_l)=\mathbf{J} \end{aligned} \tag{41}$$

With little exaggeration we can state the OM-function has similar role in the constitutive theory of the non-equilibrium thermodynamics to the Hamiltonian in the dynamics of reversible processes. An important consequence of the Donkin's theorem is the gradient of

the spontaneous entropy production by the intensive parameters is zero in the generalized Onsager constitutive theory:

$$\nabla_{\mathbf{\Gamma}}\sigma(\mathbf{X},\mathbf{J})=0,\ if\ \nabla_{\mathbf{J}}\Phi(\mathbf{J},\Gamma_l)=\mathbf{X},\nabla_{\mathbf{X}}\Psi(\mathbf{X},\Gamma_l)=\mathbf{J}, \tag{42}$$

We refer to this result as Donkin's theorem of the entropy production.
As we shall see later some dissipation potentials are homogeneous Euler's functions. This means for example, that the $\Phi(\mathbf{J},\Gamma_l)$ is a homogeneous Euler's function of degree k when every scalar of $\lambda\in R$ the following relation is valid:

$$\Phi(\lambda\mathbf{J},\Gamma_l)=\lambda^k\Phi(\mathbf{J},\Gamma_l),\ \forall\mathbf{J}\in G \tag{43}$$

It can be proved [17], that the dissipation potential $\Psi(\mathbf{X},\Gamma_l)$ dual-pair of $\Phi(\mathbf{J},\Gamma_l)$ is also homogeneous Euler's-function, like

$$\Psi(\lambda\mathbf{X},\Gamma_l)=\lambda^m\Psi(\mathbf{X},\Gamma_l),\ \frac{1}{k}+\frac{1}{m}=1 \tag{44}$$

Substituting the (14) quasi-linear Onsager constitutive equations for the second equation of the Edelen's decomposition theorem (21), then we get for the flux potential

$$\begin{aligned}\Phi(j_l,\Gamma_k)&=\int_0^1 j_jX_j(\tau j_l,\Gamma_k)d\tau=\int_0^1 R_{jl}(\Gamma_k)j_j\tau j_l d\tau=\\&=R_{jl}(\Gamma_k)j_j j_l\left[\frac{\tau^2}{2}\right]_0^1=\frac{1}{2}R_{jl}(\Gamma_k)j_j j\end{aligned} \tag{45}$$

Hence, based on the definition of (43), we see that the dissipation potentials are 2nd-degree homogeneous Euler's functions in Onsager's linear constitutive theory:

$$\Phi(\lambda j_l,\Gamma_k)=\frac{1}{2}R_{jl}(\Gamma_k)\lambda j_j\lambda j=\lambda^2\frac{1}{2}R_{jl}(\Gamma_k)j_j j=\lambda^2\Phi(j_l,\Gamma_k) \tag{46}$$

we can see from the definition (43) and the equation (44), that the dissipation potentials are homogeneous Euler's functions of degree two in the quasi-linear Onsager constitutive theory.
Consequently, a homogeneous Euler's function of degree two dissipation potential belongs to the homogeneous Euler's function of degree one constitutive equations. This could be generalized to the homogeneous dissipative functions of any degrees: the constitutive equations describe one degree less homogeneous functions of fluxes and forces than the dissipation potentials associated to them.

1.5 Example

1.5.1 Thermodynamic constrains. Supplementary thermodynamic forces

Numerous differential theorems are well known in the classical mechanics of mass-points, like the theorem of virtual work, the Gauss-theorem, the D'Alambert theorem, etc. These are supporting the construction of the equations of motion and are especially useful to handle

the constrained problems. The minimum theorem of constitutive theory, formulated by OM-function, can be used similarly as above theorem in the non-linear constitutive theory of non-equilibrium thermodynamics.

As an example we study the transport processes of a membrane (e.g. cellular membranes) in the frame of Onsager constitutive theory. For simplicity we study the transport of only three chemical components (for cell membrane: Na, K, Cl) and assume that the T temperature is constant on the membrane. These transports are governed by the corresponding gradients of μ_1, μ_2, μ_3 chemical potentials. Due to the chemical components are ions, with their transport not only the chemical species but electric charges are transported. Assuming that the net current density in every internal point of the membrane is zero. This is a local constraint. When the molar current density of chemical components are $\mathbf{j}_1, \mathbf{j}_2, \mathbf{j}_3$, the thermodynamic forces conjugated to these current densities are $\mathbf{X}_1 := -\nabla\frac{\mu_1}{T}$, $\mathbf{X}_2 := -\nabla\frac{\mu_2}{T}$, $\mathbf{X}_3 := -\nabla\frac{\mu_3}{T}$, then the constrain can be formulated by the equation

$$z_1\mathbf{j}_1 + z_2\mathbf{j}_2 + z_3\,\mathbf{j}_3 = 0 \tag{47}$$

where z_i are the charges on the ions. The force potential knowing that the $\Phi(\mathbf{J},\Gamma_l)$ flux potential can be deduced from the extremum-task determination based on equation (38), which has constraint of (47), i.e.

$$\begin{aligned} &\Psi(\mathbf{X},\Gamma_l) = \max_{\mathbf{J}\in G}\left[\mathbf{X}_i\cdot\mathbf{J}_i - \Phi\left(\mathbf{J}_i,\Gamma_l\right)\right], \forall\mathbf{X}\in G^*, \\ &z_1\mathbf{j}_1 + z_2\mathbf{j}_2 + z_{C3}\,\mathbf{j}_3 = 0 \end{aligned} \tag{48}$$

We solve this constrained extremum problem by Lagrange's method of multipliers using the

$$\begin{aligned} &\max_{\mathbf{J}\in G} F(\mathbf{X}_i,\mathbf{J}_i,\boldsymbol{\lambda},\Gamma_l) \\ &F(\mathbf{X}_i,\mathbf{J},\boldsymbol{\lambda},\Gamma_l) := \mathbf{X}_i\cdot\mathbf{J}_i - \Phi\left(\mathbf{J}_i,\Gamma_l\right) + \boldsymbol{\lambda}\cdot\left(z_1\mathbf{j}_1 + z_2\mathbf{j}_2 + z_3\,\mathbf{j}_3\right) \end{aligned} \tag{49}$$

free task. In this case the necessary conditions of the minima are:

$$\nabla_{\mathbf{J}_i} F(\mathbf{X}_i,\mathbf{J}_i,\boldsymbol{\lambda},\Gamma_l) = 0 \rightarrow \nabla_{\mathbf{J}_i}\Phi\left(\mathbf{J}_i,\Gamma_l\right) = \mathbf{X}_i - \boldsymbol{\lambda}z_i := \mathbf{X}_i^*, \left(i=1,2,3\right) \tag{50}$$

Consequently, the $\mathbf{X}_i = -\nabla\frac{\mu_i}{T}, (i=1,2,3)$ free dissipative forces causing the entropy production are not enough to describe the constrained transport processes on a membrane. Together with these free forces we have to take into account the $\boldsymbol{\lambda}z_i$ supplementary forces, which do not cause entropy production. These forces are called thermodynamic reaction forces. Consequently, the membrane processes in fact are characterized by the

$$\mathbf{X}_i^* := \mathbf{X}_i - \boldsymbol{\lambda}z_i = -\boldsymbol{\lambda}z_i - \nabla\frac{\mu_i}{T}, \left(i=1,2,3\right) \tag{51}$$

thermodynamic forces.

The thermodynamic reaction forces have full analogy with the reaction forces (which do not produce work) rendered to ideal constraints in the mechanics of the mass-points. The

thermodynamic reaction forces are similar to this having no entropy production by their actions, since

$$\mathbf{j}_1 \cdot z_1\boldsymbol{\lambda} + \mathbf{j}_2 \cdot z_2\boldsymbol{\lambda} + \mathbf{j}_3 \cdot z_3\boldsymbol{\lambda} = \left(z_1\mathbf{j}_1 + z_2\mathbf{j}_2 + z_3\ \mathbf{j}_3\right)\cdot\boldsymbol{\lambda} = 0 \tag{52}$$

where we take the equation (47) into account.
It is possible to realize the constrain by an electric field inside of the membrane, which forces the zero value of the overall current density. In case of the cell-membrane this extra field is produced by the free charges on the inner and outer surface of the membrane in stationer state. When the electric field- strength is $\mathbf{E}$, and $\boldsymbol{\lambda} = \mathbf{E}/T$, the force potential is calculated by inversion of equation (50) on fluxes, and substitute these into the second equation of (49). When the inverted constitutive equations are $\mathbf{J}_i = \mathbf{J}_i(\mathbf{X}_j^*, \Gamma_l)$, we get for the force potential

$$\begin{aligned}\Psi(\mathbf{X}_i^*, \boldsymbol{\lambda}, \Gamma_l) &:= \mathbf{X}_i \cdot \mathbf{J}_i - \Phi\left(\mathbf{J}_i, \Gamma_l\right) + \boldsymbol{\lambda}\cdot\left(z_1\mathbf{j}_1 + z_2\mathbf{j}_2 + z_3\ \mathbf{j}_3\right) = \\ &= \mathbf{X}_i^* \cdot \mathbf{J}_i(\mathbf{X}_j^*, \Gamma_l) - \Phi\left(\mathbf{J}_i(\mathbf{X}_j^*, \Gamma_l), \Gamma_l\right)\end{aligned} \tag{53}$$

The constitutive equations for the current densities can be deduced from this potential

$$\mathbf{J}_i = \nabla_{\mathbf{X}_i^*}\Psi(\mathbf{X}_i^*, \boldsymbol{\lambda}, \Gamma_l) \tag{54}$$

The Lagrange multiplier λ in the force potential is determined from the constrain equation (47):

$$\begin{aligned}&z_1\mathbf{j}_1 + z_2\mathbf{j}_2 + z_3\ \mathbf{j}_3 = \\ &= z_1\nabla_{\mathbf{X}_1^*}\Psi(\mathbf{X}_i^*, \boldsymbol{\lambda}, \Gamma_l) + z_2\nabla_{\mathbf{X}_2^*}\Psi(\mathbf{X}_i^*, \boldsymbol{\lambda}, \Gamma_l) + \\ &+ z_3\nabla_{\mathbf{X}_3^*}\Psi(\mathbf{X}_i^*, \boldsymbol{\lambda}, \Gamma_l) = 0\end{aligned} \tag{55}$$

This means a solution of a non-linear equation, except the material equations are quasi-linear. The task would be solved in the frame of quasi-linear Onsager constitutive theory. In this case the flux potential for isotropic membrane is

$$\Phi\left(\mathbf{J}_i, \Gamma_l\right) = \frac{1}{2}R_{ik}\left(\Gamma_l\right)\mathbf{J}_i \cdot \mathbf{J}_k,\ \left(i,k = 1,2,3\right) \tag{56}$$

Which modifies (50) to the case of quasi-linear constitutive equation:

$$\mathbf{X}_i^* = R_{ik}\mathbf{j}_k, \left(i,k = 1,2,3\right) \tag{57}$$

where, $\mathbf{L} = \left[L_{ik}\right]$ are the resistivity coefficients which might depend on the various intensity parameters like temperature, ionic concentrations, etc. Inverting the equations (57) by fluxes, we obtain

$$\mathbf{j}_i = L_{ik}\mathbf{X}_k^* =, \left(i,k = 1,2,3\right), \mathbf{L} = \left[L_{ik}\right] = \mathbf{R}^{-1} = \left[R_{ik}\right]^{-1} \tag{58}$$

where $\mathbf{L} = \left[L_{ik}\right]$ is the matrix composed by the conductivity coefficients of the given membrane. With these

$$\begin{aligned}\Psi(\mathbf{X}_i^*,\boldsymbol{\lambda},\Gamma_l) &:= \mathbf{X}_i\cdot\mathbf{J}_i - \Phi\left(\mathbf{J}_i,\Gamma_l\right) + \boldsymbol{\lambda}\cdot\left(z_1\mathbf{j}_1 + z_2\mathbf{j}_2 + z_3\,\mathbf{j}_3\right) = \\ &= \mathbf{X}_i^*\cdot\mathbf{J}_i - \Phi\left(\mathbf{J}_i,\Gamma_l\right) = L_{ik}\mathbf{X}_i^*\cdot\mathbf{X}_k^* - \frac{1}{2}R_{ik}L_{il}\mathbf{X}_l^*\cdot L_{km}\mathbf{X}_m^* = \frac{1}{2}L_{ik}\mathbf{X}_i^*\cdot\mathbf{X}_k^*\end{aligned} \tag{59}$$

Determining the thermodynamic constrain-force we use the equation (47):

$$\begin{aligned}&z_1\mathbf{j}_1 + z_2\mathbf{j}_2 + z_3\,\mathbf{j}_3 = \\ &= z_i\nabla_{\mathbf{X}_i^*}\Psi(\mathbf{X}_i^*,\boldsymbol{\lambda},\Gamma_l) = L_{ik}z_i\mathbf{X}_k^* = 0\end{aligned} \tag{60}$$

Using equation (51) and the condition $\boldsymbol{\lambda} = \mathbf{E}/T$, we get for the electric field-strength

$$\begin{aligned}&\mathbf{E} = -\frac{L_1\nabla\mu_1 + L_2\nabla\mu_2 + L_2\nabla\mu_3}{L_1z_1 + L_2z_2 + L_2z_3}, \\ &L_i := z_i\sum_{k=1}^{3}L_{ik}\end{aligned} \tag{61}$$

which is necessary for a net current-free state of membrane. Consequently, the constrained transport processes of membrane by constitutive point of view can be characterized by the material equations

$$\begin{aligned}&\mathbf{j}_i = L_{ik}\mathbf{X}_k^*\left(i,k=1,2,3\right), \\ &\mathbf{X}_i^* := \frac{1}{T}\frac{L_1\nabla\mu_1 + L_2\nabla\mu_2 + L_2\nabla\mu_3}{L_1z_1 + L_2z_2 + L_2z_3}z_i - \nabla\frac{\mu_i}{T}\end{aligned} \tag{62}$$

and the following OM-function

$$\begin{aligned}o(\mathbf{X}_i^*,\mathbf{J}_i,\Gamma_l) &= \Psi(\mathbf{X}_i^*,\Gamma_l) + \Phi\left(\mathbf{J}_i,\Gamma_l\right) - \mathbf{X}_i^*\cdot\mathbf{J}_i = \\ &= \frac{1}{2}L_{ik}\mathbf{X}_i^*\cdot\mathbf{X}_k^* + \frac{1}{2}R_{ik}\mathbf{J}_i\cdot\mathbf{J}_k - \mathbf{X}_i^*\cdot\mathbf{J}_i\end{aligned} \tag{63}$$

The above method could be applied to other biologically important problems like ion-selective membranes, semi-permeable membranes, etc. The above example shows, that constraints (at least when they are linear) do not destroy the symmetry relations.

1.6 Exercises

1.6.1. Show the product of L_{ik} (conductivity matrix) and R_{ik} (resistivity matrix) in linear Onsager constitutive theory is the unit matrix.

1.6.2. Show that the flux and force potentials in linear Onsager theory are strictly convex functions.

1.6.3. Prove the Donkin's theorem based on Exercise 1.

1.6.4. Show the product of Hessian matrices of dissipation potentials is the unit matrix.

1.6.5. Determine the thermodynamic reaction forces for a membrane having multiple transports of chemical components forced by a current density from outside source. Determine the OM-function of the process. What is the entropy production resulted by the reaction forces in this case?

1.6.6. Determine the linear constitutive equations of a semi-permeable membrane.

2. Minimum principle of steady-state system in case of vector processes. Deduction of the non-homogeneous transport equations from variational principle

Let us consider a system having N vector-processes with $\mathbf{j}_i$ and $\nabla\Gamma_i$ thermodynamic fluxes and forces. Study these in the frame of the generalized Onsager constitutive theory. Consequently the flux of the i-th vector process can be derived from the strictly convex $\Psi\left(\nabla\Gamma_j,\Gamma_k\right)$ force dissipation potential

$$\mathbf{j}_i = \frac{\partial\Psi}{\partial\nabla\Gamma_i},\ (i=1,2,...,N) \tag{64}$$

Substituting this constitutive equation of i-th flux into the stationer balance equation (4) of i-th extensive

$$\nabla\cdot\mathbf{j}_i = \sigma_i,\ (i=1,2,...,N) \tag{65}$$

we get the following system of partial differential equations on intensive parameters

$$\nabla\cdot\frac{\partial\Psi\left(\nabla\Gamma_j,\Gamma_k\right)}{\partial\nabla\Gamma_i} = \sigma_i,\ (i=1,2,...,N) \tag{66}$$

which are the inhomogeneous transport equations of the steady-state thermodynamic system. To generalize of the task, here we introduced the source σ_i, supposing their prescribed continuous dependence from the space-coordinates. Such source could be for example in heat-conduction process when the system is placed in a radiation field, and absorbs energy from that in all of its volume; or in case of the electric conduction of a semiconductor, in which the radiation field generates electron-hole pairs. It is evident that the equations (64) and (65) are equivalent with (66). Let the boundary of the system be Ω, and order for every transport process one $\Omega_{\Gamma_j},\Omega_{\mathbf{j}_j}$ surface partition such as: $\Omega=\Omega_{\Gamma_j}\cup\Omega_{\mathbf{j}_j},\Omega_{\Gamma_j}\cap\Omega_{\mathbf{j}_j}=\varnothing$. Let's give time-independent mixed boundary conditions for partial differential equations (66). This means that we use Dirichlet boundary condition for Γ_j's on Ω_{Γ_j}'s and a Neumann boundary condition for $\mathbf{j}_j$'s on $\Omega_{\mathbf{j}_j}$'s, i.e.

$$\begin{aligned}&\Gamma_j=\Gamma_j^*\ on\ \Omega_{\Gamma_j}\\&\mathbf{j}_{jn}=j_j^*\ on\ \Omega_{\mathbf{j}_j},\ (j=1,2,...,N)\end{aligned} \tag{67}$$

where Γ_j^* and j_j^* are given functions defined on those partitions of the boundary and $\mathbf{j}_{jn}$ is the normal coordinate of the j-th flux on the boundary.

Let the Γ_j intensives be two times and the $\mathbf{j}_j$ fluxes be continuously differentiable functions with respect to the space coordinates in all the points of the system, i.e. $\Gamma_j\in C^2(V),\mathbf{j}_j\in C^1(V),(j=1,2,...,N)$, where V is the region occuppied by the system.

Suppose that the transport equations (66) have a unique solution. Furthermore, we use the plausible equivalence of the equations (64) and (65) with (66). As we showed above, the OM-function satisfies the Young-inequality in any $\left(\mathbf{j}_1,...,\mathbf{j}_N,\Gamma_1,...,\Gamma_N\right)$ ordered vector-scalar $2N$ sets, so

$$o\left(\nabla\Gamma_j,\mathbf{j}_j,\Gamma_k\right)=\Phi\left(\mathbf{j}_j,\Gamma_k\right)+\Psi(\nabla\Gamma_j,\Gamma_k)-\mathbf{j}_i\cdot\nabla\Gamma_i\geq 0 \tag{68}$$

where the equality in (68) is satisfied only when the constitutive equations expressed by thermodynamic forces:

$$\nabla\Gamma_i=\frac{\partial\Phi\left(\mathbf{j}_j,\Gamma_k\right)}{\partial\mathbf{j}_i},\left(i=1,2,...,N\right) \tag{69}$$

are valid together with (64).
This local minimum principle could be extended to global by integration of OM-function into the region V of the system:

$$\int_V o\left(\nabla\Gamma_j,\mathbf{j}_j,\Gamma_k\right)dV=\int_V\left[\Phi\left(\mathbf{j}_j,\Gamma_k\right)+\Psi(\nabla\Gamma_j,\Gamma_k)-\mathbf{j}_i\cdot\nabla\Gamma_i\right]dV\geq 0 \tag{70}$$

This functional will be minimal, (i.e. zero) when the (64), (65) generalized Onsager constitutive equations are fulfilled. Using the Gauss's divergence theorem on the total entropy production, we obtain

$$\int_V\mathbf{j}_i\cdot\nabla\Gamma_i dV=\int_V\nabla\cdot\left(\mathbf{j}_i\Gamma_i\right)dV-\int_V\Gamma_i\nabla\cdot\mathbf{j}_i dV=\oint_\Omega\left(\mathbf{j}_i\Gamma_i\right)_n dA-\int_V\Gamma_i\sigma_i dV \tag{71}$$

where the n-index denotes the coordinates in direction of the outer normal vector of Ω and we used the stationer balance equations (65). With these the global OM-function of the steady-state system is

$$\begin{aligned}&O\left(\nabla\Gamma_j,\mathbf{j}_j,\Gamma_k\right):=\int_V o\left(\nabla\Gamma_j,\mathbf{j}_j,\Gamma_k\right)dV=\\&=\int_V\left[\Gamma_i\sigma_i+\Phi\left(\mathbf{j}_j,\Gamma_k\right)+\Psi(\nabla\Gamma_j,\Gamma_k)\right]dV-\oint_\Omega\left(\mathbf{j}_i\Gamma_i\right)_n dA\geq 0\end{aligned} \tag{72}$$

Let's construct a conditional extremum principle with this functional. To do this, we shall introduce the boundary conditions (67) into the above OM-functional, then we get

$$\begin{aligned}&O\left(\nabla\Gamma_j,\mathbf{j}_j,\Gamma_k\right)=\int_V\left[\sigma_i\Gamma_i+\Phi\left(\mathbf{j}_j,\Gamma_k\right)+\Psi(\nabla\Gamma_j,\Gamma_k)\right]dV+\\&-\int_{\Omega_{\Gamma_j}}j_{jn}\Gamma_j^* dA-\int_{\Omega_{j_j}}j_j^*\Gamma_j dA\end{aligned} \tag{73}$$

and let its definition domain be

$$D(O)=\left\{\begin{array}{l}\left(\mathbf{j}_1,..,\mathbf{j}_j,..,\mathbf{j}_N,\Gamma_1,..,\Gamma_j,..,\Gamma_N\right)\in\left|\nabla\cdot\mathbf{j}_j=\sigma_j,\ \Gamma_j=\Gamma_j^*\ on\ \Omega_{\Gamma_j}\right.\\ j_{jn}=j_j^*\ \ on\ \Omega_{\mathbf{j}_j}\end{array}\right\} \tag{74}$$

Also the admissible intensive parameters are all of the $C^2(V)$ functions which have prescribed boundary values. As well as the admissible fluxes all of the $C^1(V)$ vector functions which satisfy the stationer balance equations (65) belong to them and their normal coordinates are prescribed on the boundary.

Now, we can see that exclusive restrictions were used only for thermodynamic fluxes, namely they have to satisfy the steady-state balance equations. It is easy to see that $D(O)$ is not empty, and contains the unique solutions $(\mathbf{j}_{10},..,\mathbf{j}_{i0},..,\mathbf{j}_{N0},\Gamma_{10},..,\Gamma_{k0},..,\Gamma_{N0})$ of the steady-state transport equations. Now we can formulate the minimum principle of the steady-state system.

Theorem 1. The $(\mathbf{j}_{10},..,\mathbf{j}_{i0},..,\mathbf{j}_{N0},\Gamma_{10},..,\Gamma_{k0},..,\Gamma_{N0})$ ordered vector-scalar $2N$ set is the solution of the system of transport equations (64), (65) if and only if it minimizes the OM-functional.

$$O\left(\nabla\Gamma_{j0},\mathbf{j}_{j0},\Gamma_{k0}\right)<O\left(\nabla\Gamma_j,\mathbf{j}_j,\Gamma_k\right),\ \forall\left(\mathbf{j}_1,..,\mathbf{j}_j,..,\mathbf{j}_N,\Gamma_1,..,\Gamma_j,..,\Gamma_N\right)\in D(O) \tag{75}$$

Proof. First, we show that if $(\mathbf{j}_{10},..,\mathbf{j}_{i0},..,\mathbf{j}_{N0},\Gamma_{10},..,\Gamma_{k0},..,\Gamma_{N0})\in D(O)$ is solution of the equations (64) and (65) then it minimizes the OM-functional (73). To do this, add and deduct the entropy production to the integrand of (73), and apply the Gauss's divergence theorem like

$$\begin{array}{l}\int\limits_V\mathbf{j}_j\cdot\nabla\Gamma_j dV=-\int\limits_V\nabla\cdot\mathbf{j}_j\Gamma_j dV+\oint\limits_\Omega\mathbf{j}_j\Gamma_j\cdot d\mathbf{A}=\\ =-\int\limits_V\nabla\cdot\mathbf{j}_j\Gamma_j dV+\int\limits_{\Omega_{\Gamma_j}}j_{jn}\Gamma_j^*dA+\int\limits_{\Omega_{\mathbf{j}_j}}j_j^*\Gamma_j dA\end{array} \tag{76}$$

In this case we obtain

$$\begin{array}{l}O\left(\nabla\Gamma_j,\mathbf{j}_j,\Gamma_k\right)=\int\limits_V\left[\left(\sigma_i-\nabla\cdot\mathbf{j}_i\right)\Gamma_i+\left(\Phi+\Psi-\mathbf{j}_j\cdot\nabla\Gamma_j\right)\right]dV=\\ =\int\limits_V\left[\left(\Phi+\Psi-\mathbf{j}_j\cdot\nabla\Gamma_j\right)\right]dV\geq 0\end{array} \tag{77}$$

where we used the fact that the fluxes satisfy the (65) stationer balance equations. Due to (70) this expression is always positive on $D(O)$, and it has its minimum when the constitutive equations are obtained, i.e. at the solution of the transport equations. Hence at the solution of the transport equations $O\left(\nabla\Gamma_j,\mathbf{j}_j,\Gamma_k\right)$ has its minimum.

Second, we show that the necessary conditions of the minimum of the OM-functional on $D(O)$ are identical with the transport equations. For this we introduce the admissible variations. The

$$\left(\mathbf{j}_{10}+\delta\mathbf{j}_1,..,\mathbf{j}_{i0}+\delta\mathbf{j}_i,..,\mathbf{j}_{N0}+\delta\mathbf{j}_N,\Gamma_{10}+\delta\Gamma_1,..,\Gamma_{k0}+\delta\Gamma_k..,\Gamma_{N0}+\delta\Gamma_N\right) \tag{78}$$

ordered vector-scalar $2N$ sets are called varied elements of $D(O)$, when the variations $\left(\delta\mathbf{j}_1,..,\delta\mathbf{j}_i,..,\delta\mathbf{j}_N\right)$ of fluxes satisfy the homogeneous stationer balance equations:

$$\nabla\cdot\delta\mathbf{j}_i=0,(i=1,2,...,N) \tag{79}$$

and they are zero on $\Omega_{\mathbf{j}_j}$'s as well as the $\left(\delta\Gamma_1,..,\delta\Gamma_j,.,\delta\Gamma_N\right)$ variations of intensives are all of the $C^2(V)$ functions, which are zero on Ω_{Γ_j}'s. Then the first variation of OM- functional (73) is

$$\begin{aligned}\delta O&=\int_V\left[\left(\sigma_i-\nabla\cdot\left.\frac{\partial\Psi}{\partial\nabla\Gamma_i}\right|_{\Gamma_{k0},\nabla\Gamma_{j0}}+\left.\frac{\partial\left(\Psi+\Phi\right)}{\partial\Gamma_i}\right|_{\Gamma_{k0},\nabla\Gamma_{j0},j_{j0}}\right)\delta\Gamma_i+\left.\frac{\partial\Phi}{\partial\mathbf{j}_i}\right|_{\Gamma_{k0},\mathbf{j}_{j0}}\cdot\delta\mathbf{j}_i\right]dV=\\&=\int_V\left[\left(\sigma_i-\nabla\cdot\left.\frac{\partial\Psi}{\partial\nabla\Gamma_i}\right|_{\Gamma_{k0},\nabla\Gamma_{j0}}\right)\delta\Gamma_i+\left.\frac{\partial\Phi}{\partial\mathbf{j}_i}\right|_{\Gamma_{k0},\mathbf{j}_{j0}}\cdot\delta\mathbf{j}_i\right]dV=0\end{aligned} \tag{80}$$

where we take Donkin's theorem into account (36), which states that the variations of the sum of dissipation potentials with respect to the intensive parameters are zero. The necessary condition of the extremum of a functional is that its first variation vanishes. Since the variations of intensive parameters $\delta\Gamma_i=0,(i=1,2,3,...,N)$ are arbitrary and independent of the variation of fluxes, it is possible to choose $\delta\Gamma_i=0,(i=1,2,3,...,N)$, then we get:

$$\int_V\left[\left.\frac{\partial\Phi}{\partial\mathbf{j}_i}\right|_{\Gamma_{k0},\mathbf{j}_{j0}}\cdot\delta\mathbf{j}_i\right]dV=0,\ \ (i=1,2,...,N) \tag{81}$$

with these from (80) follows

$$\delta O=\int_V\left(\sigma_i-\nabla\cdot\left.\frac{\partial\Psi}{\partial\nabla\Gamma_i}\right|_{\Gamma_{k0},\nabla\Gamma_{j0}}\right)\delta\Gamma_i dV=0 \tag{82}$$

From this and the fundamental lemma of variational calculus, we obtain the transport equations (66)

$$\sigma_i-\nabla\cdot\frac{\partial\Psi}{\partial\nabla\Gamma_i}=0,\ \left(i=(1,2,...,N\right) \tag{83}$$

because the variations of intensives are arbitrary.
Consequently, the first group of the necessary conditions of the extremum of the OM-functional is equivalent with the transport equations. To discuss the second group of the necessary conditions, which is represented in a form of functional, consider the following vector analytical

$$\begin{aligned}&\int_V\left[\left.\frac{\partial\Phi}{\partial\mathbf{j}_i}\right|_{\Gamma_{k0},\mathbf{j}_{j0}}\cdot\delta\mathbf{j}_i\right]dV=\int_V\left[\nabla\Gamma_{i0}\cdot\delta\mathbf{j}_i\right]dV=\int_V\nabla\cdot\left(\Gamma_{i0}\delta\mathbf{j}_i\right)dV-\\&-\int_V\Gamma_{i0}\nabla\cdot\delta\mathbf{j}_i dV=\oint_A\Gamma_{i0}\delta\mathbf{j}_i\cdot d\mathbf{A}-\int_V\Gamma_{i0}\nabla\cdot\delta\mathbf{j}_i dV=0\end{aligned}\qquad(i=1,2,...,N) \tag{84}$$

identity transform, where we used Gauss's divergence theorem. Now, we see that the second group of necessity condition obtained the constitutive equations (69) and the conditions for that reason, because the variations (79) for the zero-values of the fluxes are source-free and are zero on the boundary.

Short formulation. To keep the further discussion simple, let us assume that thermodynamic forces and fluxes are square-integrable scalar or vector functions, respectively. In this case these functions form a Hilbert space, where the scalar-product can be introduced by volume integral. Denote this space by *H*, and the scalar-product by the brackets $\langle , \rangle$. The solutions of the homogeneous balance equations (79) also form Hilbert space too, which is a sub-space in *H* [18]. Denote this sub-place with H_1, and its orthogonal complement in *H* with H_2, so $H = H_1 \oplus H_2$. In this case any $\mathbf{w} \in H$ can be given in the form $\mathbf{w} = \mathbf{u} + \mathbf{v}$, where $\mathbf{u} \in H_1, \mathbf{v} \in H_2$ and $\langle \mathbf{u}, \mathbf{v} \rangle = 0$. This means that $\mathbf{u}$ and $\mathbf{v}$ are orthogonal with each other. In the light of this, the second group of the necessity of the extremum can be interpreted that in a steady-state system the quantities

$$\left. \frac{\partial \Phi}{\partial \mathbf{j}_i} \right|_{\Gamma_{k0}, \mathbf{j}_{j0}}, \left(i = 1,2,...,N\right) \tag{85}$$

which are the thermodynamic forces and to this canonically conjugated thermodynamic flux variations $\delta \mathbf{j}_i$ satisfying the homogeneous balance equations (79), are orthogonal, so

$$\left\langle \nabla \Gamma_{i0} = \left. \frac{\partial \Phi}{\partial \mathbf{j}_i} \right|_{\Gamma_{k0}, \mathbf{j}_{j0}}, \delta \mathbf{j}_i \right\rangle = 0, \forall \delta \mathbf{j}_i \in H_1, \left(i = 1,2,...,N\right) \tag{86}$$

Hence

$$\nabla \Gamma_{i0} = \left. \frac{\partial \Phi}{\partial \mathbf{j}_i} \right|_{\Gamma_{k0}, \mathbf{j}_{j0}} \in H_2, \left(i = 1,2,...,N\right) \tag{87}$$

This property was proven directly in (84), and we prove it again with the new formalism regarding the further considerations. To proceed in our study the $\langle \nabla \Gamma_i, \delta \mathbf{j}_i \rangle$ scalar-product where $\delta \mathbf{j}_i$ disappears at the boundary of the system. It is easy to show with the help of the definition of the inner product and by the Gauss's divergence theorem, that

$$\langle \nabla \Gamma_i, \delta \mathbf{j}_i \rangle = \langle \Gamma_i, -\nabla \cdot \delta \mathbf{j}_i \rangle \tag{88}$$

Also, the adjoint of the gradient operator is equally of the minus of the divergence operator. Due to the variation of fluxes satisfy the homogeneous stationer balance equations, we get from (88) the ortogonality is described by (87):

$$\langle \nabla \Gamma_i, \delta \mathbf{j}_i \rangle = \langle \Gamma_i, -\nabla \cdot \delta \mathbf{j}_i \rangle = 0 \tag{89}$$

Finally, we can construct from the above restricted extremum principle a free extremum-task by Lagrange's method of multipliers. This free variational principle is the following

$$
\begin{aligned}
&O\left(\mathbf{j}_j,\nabla\mathbf{j}_j,\nabla\Gamma_j,\Gamma_k,\lambda_i\right)=\\
&=\int_V L\left(\mathbf{j}_j,\nabla\mathbf{j}_j,\nabla\Gamma_j,\Gamma_k,\lambda_i\right)dV-\int_{\Omega_{\Gamma_j}} j_{jn}\Gamma_j^* dA-\int_{\Omega_{\mathbf{j}_j}} j_j^*\Gamma_j dA=\min \qquad (90)\\
&L\left(\mathbf{j}_j,\nabla\mathbf{j}_j,\nabla\Gamma_j,\Gamma_k,\lambda_i\right)=\sigma_i\Gamma_i+\Phi\left(\mathbf{j}_j,\Gamma_k\right)+\Psi(\nabla\Gamma_j,\Gamma_k)+\lambda_i\left(\nabla\cdot\mathbf{j}_i-\sigma_i\right)
\end{aligned}
$$

The Euler-Lagrange equations belong to (90) follows by the free variations of intensive parameters, fluxes and multipliers, i.e.

$$
\begin{aligned}
&\frac{\partial L}{\partial\Gamma_i}-\nabla\cdot\frac{\partial L}{\partial\nabla\Gamma_i}=\sigma_i-\nabla\cdot\frac{\partial\Psi}{\partial\nabla\Gamma_i}=0,\ \frac{\partial L}{\partial\mathbf{j}_i}-\nabla\frac{\partial L}{\partial\nabla\cdot\mathbf{j}_i}=\frac{\partial\Phi}{\partial\mathbf{j}_i}-\nabla\lambda_i=0\\
&\frac{\partial L}{\partial\lambda_i}=\nabla\cdot\mathbf{j}_i-\sigma_i=0,\left(i=1,2,...,N\right)
\end{aligned} \qquad (91)
$$

from where the firsts are stationer transport equations, the seconds are constitutive equations expressing the connections between the forces and currents, while the thirds are the stationer balance equations. Now we see that it is possible to choose the λ_i multipliers as the intensive parameters Γ_i of the steady-state system, because the constitutive equations $\nabla\Gamma_i=\frac{\partial\Phi}{\partial\mathbf{j}_i}$ are fulfilled. The boundary conditions can be obtained from the surface terms by detailed calculations

$$
\begin{aligned}
&\left.\frac{\partial\Psi}{\partial\nabla\Gamma_i}\right|_n=j_{jn}^*\ \ on\ \ \Omega_{\mathbf{j}_j}\\
&\Gamma_i=\Gamma_i^*\ \ on\ \Omega_{\Gamma_j},\left(i=1,2,...,N\right)
\end{aligned} \qquad (92)
$$

It is evident that the above formulated free variational principle is a direct consequence of the restricted minimum principle, which is formulated in the Theorem 1. We refer to both of these principles as minimum principle of steady-state system (MPSTS). This principle states the minimum of OM-functional for the real processes of a steady-state thermodynamic system, also which satisfy the balance equations and the constitutive equations as well.

2.1 Minimum principle of global entropy production. Deduction of the homogeneous steady-state transport equations from variational principle

In following we shall show that in case of homogeneous steady-state balance equations the MPSTS guarantees the minimum of the global entropy production.

In case of source-free balance equations the homogeneous transport equations of a steady-state system can be obtained by the substitution of the generalized Onsager's constitutive equations

$$
\mathbf{j}_i=\frac{\partial\Psi}{\partial\nabla\Gamma_i},\left(i=1,2,...,N\right) \qquad (93)
$$

into the homogeneous steady-state balance equations using the of the vector-fluxes (4)

$$\nabla \cdot \mathbf{j}_i = 0, (i = 1,2,...,N) \tag{94}$$

then we get the system of the homogenous stationer transport equations

$$\nabla \cdot \frac{\partial \Psi\left(\nabla \Gamma_j, \Gamma_k\right)}{\partial \nabla \Gamma_i} = 0, (i = 1,2,...,N) \tag{95}$$

Consider the earlier fixed (67) boundary conditions for these partial differential equations and start again with the OM-function of the generalized Onsager constitutive theory to explain the principle of minimal entropy production

$$o\left(\nabla \Gamma_j, \mathbf{j}_j, \Gamma_k\right) = \Phi\left(\mathbf{j}_j, \Gamma_k\right) + \Psi(\nabla \Gamma_j, \Gamma_k) - \mathbf{j}_i \cdot \nabla \Gamma_i \geq 0 \tag{96}$$

Considering the disappearance of the divergences of fluxes by (94), reformulate (96):

$$\begin{aligned} &o\left(\nabla \Gamma_j, \mathbf{j}_j, \Gamma_k\right) = \Phi\left(\mathbf{j}_j, \Gamma_k\right) + \Psi(\nabla \Gamma_j, \Gamma_k) - \mathbf{j}_i \cdot \nabla \Gamma_i = \\ &= \Phi\left(\mathbf{j}_j, \Gamma_k\right) + \Psi(\nabla \Gamma_j, \Gamma_k) - \nabla \cdot \left(\mathbf{j}_i \Gamma_i\right) \geq 0 \end{aligned} \tag{97}$$

The integral of the OM-function on the V volume of the system gives the global OM-function of the steady-state system

$$\begin{aligned} &O\left(\nabla \Gamma_j, \mathbf{j}_j, \Gamma_k\right) := \int_V o\left(\nabla \Gamma_j, \mathbf{j}_j, \Gamma_k\right) dV = \\ &= \int_V \left[\Phi\left(\mathbf{j}_j, \Gamma_k\right) + \Psi(\nabla \Gamma_j, \Gamma_k)\right] dV - \oint_\Omega \left(\mathbf{j}_i \Gamma_i\right)_n dA \geq 0 \end{aligned} \tag{98}$$

Introducing the boundary values of intensive parameters to this global OM-function by the method used in the previous paragraph, we get the final form of the OM-functional

$$\begin{aligned} &O\left(\nabla \Gamma_j, \mathbf{j}_j, \Gamma_k\right) := \int_V \left[\Phi\left(\mathbf{j}_j, \Gamma_k\right) + \Psi(\nabla \Gamma_j, \Gamma_k)\right] dV + \\ &- \int_{\Omega_{\Gamma_j}} j_{jn} \Gamma_j^* dA - \int_{\Omega_{\mathbf{j}_j}} j_j^* \Gamma_j dA \end{aligned} \tag{99}$$

Let its definition domain be

$$D_0(O) = \left\{ \begin{matrix} \left(\mathbf{j}_1,..,\mathbf{j}_j,..,\mathbf{j}_N,\Gamma_1,..,\Gamma_j,..,\Gamma_N\right) \in \left| \nabla \cdot \mathbf{j}_j = 0, \Gamma_j = \Gamma_j^* \; on \; \Omega_{\Gamma_j} \right. \\ j_{jn} = j_j^* \; on \; \Omega_{\mathbf{j}_j} \end{matrix} \right\} \tag{100}$$

where the admissible intensive parameters are all the $C^2(V)$ functions which have definite boundary values, the admissible fluxes are all the $C^1(V)$ vector-functions which satisfy the source-free stationer balance equations and their normal coordinates are prescribed on the boundary. Consider the $\left(\mathbf{j}_{10},..,\mathbf{j}_{i0},..,\mathbf{j}_{N0},\Gamma_{10},..,\Gamma_{k0},..,\Gamma_{N0}\right)$ ordered vector-scalar $2N$ set as the solution of equations (93) and (94), which are tantamount with the transport equations (95). Now we can formulate the minimum principle of the steady-state system in case of source-free balance equations.

Theorem 2. The $(\mathbf{j}_{10},..,\mathbf{j}_{i0},..,\mathbf{j}_{N0},\Gamma_{10},..,\Gamma_{k0},..,\Gamma_{N0})$ ordered vector-scalar *2N* set is the solution of the system of the transport equations (93), (94), if and only if when it minimizes the OM-functional.

$$O\left(\nabla\Gamma_{j0},\mathbf{j}_{j0},\Gamma_{k0}\right)<O\left(\nabla\Gamma_{j},\mathbf{j}_{j},\Gamma_{k}\right),\ \forall\left(\mathbf{j}_{1},.,\mathbf{j}_{j},.,\mathbf{j}_{N},\Gamma_{1},.,\Gamma_{j}.,\Gamma_{N}\right)\in D_0(O) \tag{101}$$

To prove this, add and deduct the entropy production to the integrand of (99), and apply the Gauss's divergence theorem like

$$\begin{aligned} &O\left(\nabla\Gamma_{j},\mathbf{j}_{j},\Gamma_{k}\right)=\int_V\left[-\nabla\cdot\mathbf{j}_i\Gamma_i+\left(\Phi+\Psi-\mathbf{j}_j\cdot\nabla\Gamma_j\right)\right]dV= \\ &=\int_V\left[\left(\Phi+\Psi-\mathbf{j}_j\cdot\nabla\Gamma_j\right)\right]dV\geq 0 \end{aligned} \tag{102}$$

where we used the fact that the fluxes satisfy the (94) homogeneous balance equations. This expression is always positive on $D_0(O)$, and it has its minimum when the constitutive equations are valid as well, i.e. at the solution of the transport equations. Hence the solution of the transport equations is minimum place of $O\left(\nabla\Gamma_{j},\mathbf{j}_{j},\Gamma_{k}\right)$. We can prove that a part of the necessary conditions of the minimum are identical with the transport equations. For this we define the

$$\left(\mathbf{j}_{10}+\delta\mathbf{j}_1,..,\mathbf{j}_{i0}+\delta\mathbf{j}_i,..,\mathbf{j}_{N0}+\delta\mathbf{j}_N,\Gamma_{10}+\delta\Gamma_1,..,\Gamma_{k0}+\delta\Gamma_k..,\Gamma_{N0}+\delta\Gamma_N\right) \tag{103}$$

vector-scalar ordered *2N sets* as admissible variations, when the flux $(\delta\mathbf{j}_1,..,\delta\mathbf{j}_i,..,\delta\mathbf{j}_N)$ satisfy the homogeneous stationer balance equations

$$\nabla\cdot\delta\mathbf{j}_i=0,(i=1,2,...,N) \tag{104}$$

and they are zero on the boundaries, as well as the $\left(\delta\Gamma_1,.,\delta\Gamma_j.,\delta\Gamma_N\right)$ variations of intensives are two times continuously derivable function with zero values on the boundaries. Then the first variation of (99) is

$$\delta O=\int_V\left[\left(-\nabla\cdot\left.\frac{\partial\Psi}{\partial\nabla\Gamma_i}\right|_{\Gamma_{k0},\nabla\Gamma_{j0}}\right)\delta\Gamma_i+\left.\frac{\partial\Phi}{\partial\mathbf{j}_i}\right|_{\Gamma_{k0},\mathbf{j}_{j0}}\cdot\delta\mathbf{j}_i\right]dV=0 \tag{105}$$

where we take the Donkin's theorem into account, which states the zero value of the variation of the sum of dissipation potentials by intensive parameters and that the necessary condition of the extremum of the functional is the zero value of its first variation.
Due to the variations of intensives are independent from variations of fluxes, we can choose $\delta\Gamma_i=0,(i=1,2,3,...,N)$, then from (105) follows

$$\int_V\left[\left.\frac{\partial\Phi}{\partial\mathbf{j}_i}\right|_{\Gamma_{k0},\mathbf{j}_{j0}}\cdot\delta\mathbf{j}_i\right]dV=0,\left(i=(1,2,...,N\right) \tag{106}$$

Consequently we get from (105):

$$\delta O = \int_V \left(-\nabla \cdot \left. \frac{\partial \Psi}{\partial \nabla \Gamma_i} \right|_{\Gamma_{k0}, \nabla \Gamma_{j0}} \right) \delta \Gamma_i dV = 0 \tag{107}$$

from where we obtain the system of the transport equations (95)

$$\nabla \cdot \frac{\partial \Psi}{\partial \nabla \Gamma_i} = 0, \ \left(i = (1,2,...,N) \right) \tag{108}$$

because the variations of the intensives are arbitrary.
Consequently, one group of the necessary conditions of the extremum is the group of the transport equations. The second group, which is formulated in functional form, were discussed above.
Due to the Lagrangian of the functional (99) is the sum of the dissipation potentials, which is equal to the entropy production in case of every real steady-state physical processes, this extremum theorem involves the minimum principle of global entropy production (MPGEP).
The physical meaning of MPGEP needs a clarification. Consider the variations of the fluxes and of the intensive parameters as fluctuations of the system around their stationer state values. When these fluctuations are small, the fluctuation of the global entropy production of the system is equal to its first approximations and it has a form

$$\delta P := \left\langle \nabla \Gamma_{i0} + \delta \nabla \Gamma_i, \mathbf{j}_{i0} + \delta \mathbf{j}_i \right\rangle - \left\langle \nabla \Gamma_{i0}, \mathbf{j}_{i0} \right\rangle \cong \left\langle \nabla \Gamma_{i0}, \delta \mathbf{j}_i \right\rangle + \left\langle \delta \nabla \Gamma_i, \mathbf{j}_{i0} \right\rangle \tag{109}$$

The first term of the fluctuation of entropy production is zero due to the (86) ortogonality conditions, while the second term, using the

$$\delta P \cong \left\langle \delta \nabla \Gamma_i, \mathbf{j}_{i0} \right\rangle = \left\langle \delta \Gamma_i, -\nabla \cdot \mathbf{j}_{i0} \right\rangle \tag{110}$$

transformation, will also be zero due to the source free balance equation (94). Consequently in first approximation the global entropy production has no fluctuation, so the necessary condition of the minimum of the global entropy production is fulfilled. The sufficient conditions also have to be proved. The sum of the dissipation potentials in case when the forces and fluxes are satisfied the transport- and constitutive equations is equal to the entropy production, due to the minimum of the OM-functional Proceed from (36). Due to this fact we could study the sum of the dissipation potentials instead of the entropy production. Regarding the Donkin's theorem, the second term of the Taylor series at the stationer state will be

$$\begin{aligned} \delta P &= \int_V \left[\Phi \left(\mathbf{j}_{j0} + \delta \mathbf{j}_j, \Gamma_{k0} + \delta \Gamma_k \right) + \Psi \left(\nabla \Gamma_{j0} + \delta \nabla \Gamma_{j0}, \Gamma_{k0} + \delta \Gamma_k \right) \right] dV = \\ &= \int_V \left[\Phi \left(\mathbf{j}_{j0}, \Gamma_{k0} \right) + \Psi \left(\nabla \Gamma_{j0}, \Gamma_{k0} \right) \right] dV + \left\langle \left. \frac{\partial \Phi}{\partial \mathbf{j}_i} \right|_{\Gamma_{k0}, \mathbf{j}_{j0}}, \delta \mathbf{j}_i \right\rangle + \left\langle \left. \frac{\partial \Psi}{\partial \nabla \Gamma_i} \right|_{\Gamma_{k0}, \mathbf{j}_{j0}}, \delta \nabla \Gamma_i \right\rangle + \\ &+ \left\langle \left. \frac{\partial^2 \Phi}{\partial \mathbf{j}_j \partial \mathbf{j}_i} \right|_{\Gamma_{k0}, \mathbf{j}_{j0}} \delta \mathbf{j}_j, \delta \mathbf{j}_i \right\rangle + \left\langle \left. \frac{\partial^2 \Psi}{\partial \nabla \Gamma_j \partial \nabla \Gamma_i} \right|_{\Gamma_{k0}, \mathbf{j}_{j0}} \delta \nabla \Gamma_j, \delta \nabla \Gamma_i \right\rangle \end{aligned} \tag{111}$$

When $\delta^{(1)}P$ is the linear part of the variations of entropy production (first variation) and the quadratic part is the $\delta^{(2)}P$ (second variation), then (111) will have a form

$$\delta P = \delta^{(1)}P + \delta^{(1)}P,\ \delta^{(1)}P = \left\langle \left.\frac{\partial \Phi}{\partial \mathbf{j}_i}\right|_{\Gamma_{k0},\mathbf{j}_{j0}}, \delta\mathbf{j}_i \right\rangle + \left\langle \left.\frac{\partial \Psi}{\partial \nabla\Gamma_i}\right|_{\Gamma_{k0},\mathbf{j}_{j0}}, \delta\nabla\Gamma_i \right\rangle =$$

$$= \left\langle \left.\frac{\partial \Phi}{\partial \mathbf{j}_i}\right|_{\Gamma_{k0},\mathbf{j}_{j0}}, \delta\mathbf{j}_i \right\rangle + \left\langle -\nabla\cdot \left.\frac{\partial \Psi}{\partial \nabla\Gamma_i}\right|_{\Gamma_{k0},\mathbf{j}_{j0}}, \delta\Gamma_i \right\rangle \tag{112}$$

$$\delta^{(2)}P = \left\langle \left.\frac{\partial^2 \Phi}{\partial \mathbf{j}_j \partial \mathbf{j}_i}\right|_{\Gamma_{k0},\mathbf{j}_{j0}} \delta\mathbf{j}_j, \delta\mathbf{j}_i \right\rangle + \left\langle \left.\frac{\partial^2 \Psi}{\partial \nabla\Gamma_j \partial \nabla\Gamma_i}\right|_{\Gamma_{k0},\mathbf{j}_{j0}} \delta\nabla\Gamma_j, \delta\nabla\Gamma_i \right\rangle$$

where we consider the Donkin's theorem (36) in the deduction of (36) for second variation, (for details see Appendix 7.3.). In case when the stationer balance equations are source-free, the terms containing the first variation of the entropy production, satisfying the balance equations and the homogeneous boundary conditions, have zero values. The second terms are positive, because the dissipation potentials in generalized Onsager constitutive theory are strictly convex functions of fluxes and forces, so their functionals, which are generated from the bilinear forms of their Hessian matrices, are positive definite as well. Consequently, the global entropy production in case of homogeneous transport equations is minimal (to form an idea of the role and the significance of the above variational principles see Appendix 7.1.).

2.2 Examples

2.2.1 Deduction of the equation of steady-state heat conduction in anisotropic media. Onsager's problem [3]

In this part we study the stationer heat conduction in anisotropic solid, e.g. in triclinic crystal. We suppose certain energy deposit σ_u in all the volume elements of the system, and let the heat conduction tensor be temperature dependent. The stationer balance equation is

$$\nabla\cdot\mathbf{j}_u = \sigma_u \tag{113}$$

where $\mathbf{j}_u$ is the internal energy (heat) flux. From this, by multiplication with $1/T$ we obtain the entropy balance

$$\nabla\cdot\frac{\mathbf{j}_u}{T} = \mathbf{j}_u\cdot\nabla\left(\frac{1}{T}\right) + \frac{\sigma_u}{T} \tag{114}$$

where the $\frac{\sigma_u}{T}$ is the source generated (e.g. originated from the outer radiation field) entropy production, while the spontaneous entropy production is

$$\sigma = \mathbf{j}_u\cdot\nabla\left(\frac{1}{T}\right) \tag{115}$$

Supposing quasi-linear constitutive equation, so the heat flux and the $\nabla\left(\frac{1}{T}\right)$ thermodynamic force are connected by the constitutive equation

$$\mathbf{j}_u = \mathbf{\Lambda}(T)\nabla\left(\frac{1}{T}\right) \tag{116}$$

The second order symmetric tensor $\mathbf{\Lambda}(T)$ is the specific heat conduction tensor. Its inverse tensor is the specific resistance tensor: $\mathbf{R}(T) = \left[\mathbf{\Lambda}(T)\right]^{-1}$. Substitute this with the entropy production (115), then we get

$$\begin{aligned} \sigma &= \mathbf{j}_u \cdot \nabla\left(-\frac{1}{T}\right) = \left[\mathbf{\Lambda}(T)\nabla\left(\frac{1}{T}\right)\right]\cdot\nabla\left(\frac{1}{T}\right) = \\ &= \left[\mathbf{R}(T)\mathbf{j}_u\right]\cdot\mathbf{j}_u \geq 0 \end{aligned} \tag{117}$$

From this become the second degree homogenous dissipation potentials:

$$\Psi = \frac{1}{2}\left[\mathbf{\Lambda}(T)\nabla\left(\frac{1}{T}\right)\right]\cdot\nabla\left(\frac{1}{T}\right),\ \Phi = \frac{1}{2}\left[\mathbf{R}(T)\mathbf{j}_u\right]\cdot\mathbf{j}_u \tag{118}$$

Consequently the partially restricted variation-task of (73) in this actual case is

$$\begin{aligned} &\int_V L\left(\mathbf{j}_u, \nabla\left(\frac{1}{T}\right), \left(\frac{1}{T}\right)\right)dV = \\ &= \int_V \frac{\sigma_u}{T} + \frac{1}{2}\left[\mathbf{\Lambda}(T)\nabla\left(\frac{1}{T}\right)\right]\cdot\nabla\left(\frac{1}{T}\right) + \frac{1}{2}\left[\mathbf{R}(T)\mathbf{j}_u\right]\cdot\mathbf{j}_u dV = \min \end{aligned} \tag{119}$$

The necessary condition of the minimum is the zero value of the first variation of the above (119) functional

$$\begin{aligned} &\delta\int_V L\left(\mathbf{j}_u, \nabla\left(\frac{1}{T}\right), \left(\frac{1}{T}\right)\right)dV = \\ &= \int_V \left(\sigma_u - \nabla\cdot\left[\mathbf{\Lambda}(T)\nabla\left(\frac{1}{T}\right)\right]\right)\delta\left(\frac{1}{T}\right)dV + \int_V \left[\mathbf{R}(T)\mathbf{j}_u\right]\cdot\delta\mathbf{j}_u dV = 0 \end{aligned} \tag{120}$$

Due to the variations of the temperature are arbitrary and independent from the variations of the heat flux the first term in above equation is zero, consequently we obtain the stationer transport equation

$$\sigma_u - \nabla\cdot\left[\mathbf{\Lambda}(T)\nabla\left(\frac{1}{T}\right)\right] = 0 \tag{121}$$

The second term in (120) is zero as well, because of the ortogonality relation, which means in this actual case, that the $\nabla\left(\frac{1}{T}\right)$ thermodynamic force is orthogonal to the variation of the divergence-free heat flux. The free variation-task of the problem is

$$\begin{aligned}&\int_V L\left(\mathbf{j}_u, \nabla\left(\frac{1}{T}\right), \left(\frac{1}{T}\right), \lambda\right) dV = \\ &= \int_V \left\{ \frac{\sigma_u}{T} + \frac{1}{2}\left[\mathbf{\Lambda}(T)\nabla\left(\frac{1}{T}\right)\right]\cdot\nabla\left(\frac{1}{T}\right) + \frac{1}{2}\left[\mathbf{R}(T)\mathbf{j}_u\right]\cdot\mathbf{j}_u + \lambda\left(\nabla\cdot\mathbf{j}_u - \sigma_u\right)\right\} dV = \min\end{aligned} \tag{122}$$

In this case the following Euler-Lagrange equations belong to the extremum of the functional as necessary conditions

$$\begin{aligned}&\frac{\partial L}{\partial\left(\frac{1}{T}\right)} - \nabla\cdot\frac{\partial L}{\partial\nabla\left(\frac{1}{T}\right)} = 0 = \sigma_u - \nabla\cdot\left[\mathbf{\Lambda}(T)\nabla\left(\frac{1}{T}\right)\right], \\ &\frac{\partial L}{\partial\mathbf{j}_u} - \nabla\frac{\partial L}{\partial\nabla\cdot\mathbf{j}_u} = 0 = \mathbf{R}(T)\mathbf{j}_u - \nabla\lambda, \\ &\frac{\partial L}{\partial\lambda} = 0 = \nabla\cdot\mathbf{j}_u - \sigma_u\end{aligned} \tag{123}$$

showing that the multiplier is the temperature. It is an interesting data of history of non-equilibrium thermodynamics, that Onsager can solve this task only partially, namely, it can deduce only the constitutive equation (116) (for details see Appendix 7.2.)

2.2.2 Deduction of the transport equation of steady-state electric conduction in inhomogeneous media

We mean term non-homogeneity in the chemical composition of the system, consequently the mechanism of the electric conduction is place dependent as well. The non-homogeneity in electric DC circuits could be regarded as compartmentation of homogeneous subparts composing the inhomogeneous system step-by step. Let the temperature of the system be constant and have stationer current conduction as a single transport process. The role of the three space-dependent quantities: electric field-strength (**E**), electric current density (**j**), and EMF-producing electric field $\mathbf{E}^{(i)}$ generated by other causes as by a battery. The last one is the thermodynamic reaction force, which has to be considered within the current continuity constraint. We suppose, that all of the above physical quantities are functions of the space coordinates. These quantities are dependent on each other; the differential Ohm-law connects them in the linear Onsager constitutive theory

$$\mathbf{j} = \gamma\left(\mathbf{E} + \mathbf{E}^{(i)}\right) \tag{124}$$

where γ is the electric conductivity. The field equations of electrodynamics of the system are [19]

$$\nabla \cdot \mathbf{j} = 0, \nabla \times \left(\frac{\mathbf{j}}{\gamma} \right) = -\nabla \times \mathbf{E}^{(i)},$$
$$\nabla \times \mathbf{E} = 0, \nabla \cdot \left(\gamma \mathbf{E} \right) = -\nabla \cdot \left(\gamma \mathbf{E}^{(i)} \right) \tag{125}$$

To simplify the problem, assume $\nabla \times \mathbf{E}^{(i)} = 0, \nabla \times \mathbf{j} = 0$.
The stationer transport equation is the consequence of the differential Ohm-law and the continuity equation of the current

$$\mathbf{j} = -\gamma \nabla \left(\varphi + \varphi^{(i)} \right), \nabla \cdot \mathbf{j} = 0 \rightarrow \tag{126}$$

$$\rightarrow \nabla \cdot \left(\gamma \nabla \varphi \right) = -\nabla \cdot \left(\gamma \nabla \varphi^{(i)} \right) \tag{127}$$

Hence the first two equations are equivalent with the third one, where we used the curl-free property of the field-strength and introduced the electric potentials $\varphi, \varphi^{(i)}$.
Prescribe the coordinate of the current density in direction of outer normal vector on $\Omega_{\mathbf{j}}$ part of the boundary $\Omega = \Omega_{\mathbf{j}} \cup \Omega_{\varphi}$ and the electric potential on Ω_{φ}: $j_n = j^* \ on \ \Omega_{\mathbf{j}}$, $\varphi = \varphi^* \ on \ \Omega_{\varphi}$. Also we suppose that the $(\mathbf{j}_0, \varphi_0)$ vector-scalar ordered pair is the unique solution of the equations (126), which are identical(126), and uniform with the (127) transport equation. The entropy production of the system could be calculated by the Joule-law, [19].

$$\sigma = \frac{1}{\gamma} |\mathbf{j}|^2 = \mathbf{j} \cdot \left(\mathbf{E} + \mathbf{E}^{(i)} \right) = -\mathbf{j} \cdot \nabla \left(\varphi + \varphi^{(i)} \right) \tag{128}$$

From this the thermodynamic forces and fluxes as well as the dissipation functions of the linear Onsager constitutive theory and its OM-function follow

$$\Phi = \frac{1}{2\gamma} |\mathbf{j}|^2, \Psi = \frac{\gamma}{2} \left| \nabla \left(\varphi + \varphi^{(i)} \right) \right|^2,$$
$$o\left(\nabla \varphi, \mathbf{j}, \varphi \right) = \Phi + \Psi - \sigma \geq 0 \tag{129}$$

where in the second equation the equality is valid, when the

$$\mathbf{j} = -\gamma \nabla \left(\varphi + \varphi^{(i)} \right) \tag{130}$$

Ohm-law is valid.
Integrate the OM-function on the volume V of the system, while the entropy production is transformed by the surface integral taking into account the continuity equation of the electric current density. Take into account also the boundary conditions prescribing the values of the electric potential and current densities on the boundaries of the system. Than we get the following form of the OM-functional

$$O\left(\nabla \varphi, \mathbf{j}, \varphi \right) := \int_V \left[\Phi \left(\mathbf{j}, \varphi \right) + \Psi \left(\nabla \varphi, \varphi \right) \right] dV +$$
$$- \int_{\Omega_{\varphi}} j_n \varphi^* dA - \int_{\Omega_{\mathbf{j}}} j^* \varphi dA \tag{131}$$

Let its definition domain be

$$D_0(O)=\left\{\begin{array}{l}(\mathbf{j},\varphi)\in\left|\nabla\cdot\mathbf{j}=0,\ \varphi=\varphi^*\ on\ \Omega_\varphi\right.\\ j_n=j^*\ on\ \Omega_{\mathbf{j}}\end{array}\right\} \tag{132}$$

Let us consider the admissible electric potential as a $C^2(V)$ function, and the admissible current density is $C^1(V)$. We prove in the following that the $(\mathbf{j}_0,\varphi_0)$ pair is the solution to the transport equations (126) if and only if when it is the extremal point of the(131) functional (131), i.e.

$$O(\nabla\varphi_0,\mathbf{j}_0,\varphi_0)<O(\nabla\varphi,\mathbf{j},\varphi),\ \ \forall(\mathbf{j},\varphi)\in D_0(O) \tag{133}$$

To prove this statement we add and deduct again to the integrand the entropy production in (131) and apply the Gauss's divergence theorem, then we get

$$\begin{aligned}&O(\nabla\varphi,\mathbf{j},\varphi)=\int_V\left[-\varphi\nabla\cdot\mathbf{j}+\left(\Phi+\Psi-\mathbf{j}\cdot\nabla\varphi\right)\right]dV=\\&=\int_V\left[\left(\Phi+\Psi-\mathbf{j}\cdot\nabla\varphi\right)\right]dV\geq 0\end{aligned} \tag{134}$$

where we take into account that the current density satisfies the homogeneous balance equation of (126). The functional (134) is always positive on $D_0(O)$, and it has its minimum when the Ohm-law is valid as well, also when the $(\mathbf{j},\varphi)$ pair of functions is the solution to the transport equations (126). Consequently the solution of the transport equations is the minimum place of $O(\nabla\varphi,\mathbf{j},\varphi)$. Because the sum of the dissipation potentials is equal to the entropy production, the minimum principle of global entropy production is valid in this case.

We can also proof, that one of the necessary conditions belongs to the minimum of the functional are equivalent with the transport equations (126). To do this, we define the $(\mathbf{j}+\delta\mathbf{j},\varphi_0+\delta\varphi)$ type vector-scalar ordered pairs as admissible, when the variation of the flux $\delta\mathbf{j}$ satisfies the condition of

$$\nabla\cdot\delta\mathbf{j}=0 \tag{135}$$

homogeneous stationer balance equation, and it is zero on the boundary surface, while the variations of the electric potential $\delta\varphi$ are arbitrary two times continuously differentiable functions. With these assumptions the first variation of (131) is

$$\begin{aligned}&\delta O=\int_V\left[\left(-\nabla\cdot\left.\frac{\partial\Psi}{\partial\nabla\varphi}\right|_{\varphi_0,\nabla\varphi_0}\right)\delta\varphi+\left.\frac{\partial\Phi}{\partial\mathbf{j}}\right|_{\varphi_0,\mathbf{j}_0}\cdot\delta\mathbf{j}\right]dV=\\&=\int_V\left[\left(-\nabla\cdot\frac{1}{\gamma}\nabla\left(\varphi+\varphi^{(i)}\right)\Big|_{\varphi_0,\nabla\varphi_0}\right)\delta\varphi+\left.\frac{\partial\Phi}{\partial\mathbf{j}}\right|_{\varphi_0,\mathbf{j}_0}\cdot\delta\mathbf{j}\right]dV=0\end{aligned} \tag{136}$$

As since the necessary condition of the extremum of functional is the zero value of its first variation. Due to the variations of the potential are arbitrary and independent of the variations of the flux, than we get with $\delta\varphi = 0$

$$\int_V \left[\left. \frac{\partial \Phi}{\partial \mathbf{j}} \right|_{\varphi_0, \mathbf{j}_0} \cdot \delta \mathbf{j} \right] dV = 0 \tag{137}$$

Using this, for arbitrary $\delta\varphi$ it follows from (136):

$$\delta O = \int_V \left(-\nabla \cdot \frac{1}{\gamma} \nabla \left(\varphi + \varphi^{(i)} \right) \Big|_{\varphi_0, \nabla\varphi_0} \right) \delta\varphi dV = 0 \tag{138}$$

and from this by the fundamental lemma of variational calculus follows the transport equation (127):

$$\nabla \cdot \frac{1}{\gamma} \nabla \left(\varphi + \varphi^{(i)} \right) = 0 \tag{139}$$

Consequently one of the necessary conditions of the extremum is the transport equation. The second one has to be considered as the functional form used in the discussions above, regarding on $\left. \frac{\partial \Phi}{\partial \mathbf{j}_i} \right|_{\Gamma_{k0}, \mathbf{j}_{j0}} = -\nabla \left(\varphi + \varphi^{(i)} \right)$, from what the direct proof of orthogonality is obtained

$$\int_V \left(\left. \frac{\partial \Phi}{\partial \mathbf{j}} \right|_{\varphi_0, \mathbf{j}_0} \cdot \delta \mathbf{j} \right) dV = \int_V \left(-\left(\varphi + \varphi^{(i)} \right) \cdot \nabla \cdot \delta \mathbf{j} \right) dV = 0 \tag{140}$$

2.2.3 Murray's law of mammalian circulatory systems

Murray C.D. suggested [20], that the inner radius of blood vessel increases as an outcome of a compromises between the advantage of increasing of the lumen, (which decreases the viscous dissipation of blood flow) and the disadvantage of increasing blood volume, (which increases the metabolic demand to maintain the physiological quality of blood) [21]. For example, because the red blood cells must be produced continuously to replace vanishing cells. The Murray's "cost function" is the total power required to maintain the blood flow and the prescribed stationer quality of blood [19], [20]. According to the first law of thermodynamics which is valid on the stationer cases, the total demand of power is identical with the sum of the viscous dissipation of the stationer flow of blood, and the energy dissipation of the irreversible chemical, physical, biological changes in blood. Due to these, the "cost function" as the energy-dissipation will be discussed in the next part, and use the following entropy production. Let the thermodynamic system be a part of vessel with length *L*, radius *r* and has constant *T* temperature. Consider the flow stationer, neglecting the changes by the living pulsation of the blood. The blood is not compressible, and we do not consider the dilatation of the vessels. In this case the homogeneous balances and the minimum principle of global entropy production are valid. Let us study the constitutive equations of the system, when the blood is Newtonian liquid, and the pressure difference is

Δp, the volume-current of the vessel is I. Their connection is described by the Hagen-Poiseuille-law

$$\Delta p = \frac{8\eta L}{\pi r^4} I \tag{141}$$

from where the viscous energy dissipation per unit time is

$$\Delta pI = \frac{8\eta L}{\pi r^4} I^2 \tag{142}$$

We assume that power of the irreversible processes which change the blood quality are proportional with its volume

$$P_{\text{deg}} = kr^2 \pi L \tag{143}$$

where k is the specific energy dissipation per unit time. Consequently, the global entropy production is

$$P = \frac{\frac{8\eta L}{\pi r^4} I^2 + kr^2 \pi L}{T} \tag{144}$$

It could be shown that the entropy production is a strictly convex function of the radius of the vessel. The necessary condition of the minimum gives the Murray-law

$$\frac{dP}{dr} = \frac{d}{dr} \frac{\frac{8\eta L}{\pi r^4} I^2 + kr^2 \pi L}{T} = 0 \rightarrow I = \frac{\pi}{4} \left(\frac{k}{\eta} \right)^{\frac{1}{2}} r^3 \tag{145}$$

This law expresses that the volume current of blood is proportional with the third power of the radius of the vessel when the entropy production is minimal.
Let us study now the areas of bifurcations. One of the characteristic properties of vasculature is that it divides into smaller and smaller vessels down to the level of capillaries. These divisions are the bifurcations. When the mother-vessel with radius r_0 divides into two daughter vessels with r_1 and r_2 radii we can introduce the ratios

$$\alpha = \frac{r_2}{r_1}, \beta = \frac{r_1^2 + r_2^2}{r_0^2} \tag{146}$$

characterizing the bifurcation. The volume currents fulfill the Kirchhoff's first law

$$I_0 = I_1 + I_2 \tag{147}$$

From this and from the Murray's law for the above introduced ratios follow

$$\frac{r_1}{r_0} = \frac{1}{\left(1+\alpha^3\right)^{\frac{1}{3}}}, \frac{r_2}{r_0} = \frac{\alpha}{\left(1+\alpha^3\right)^{\frac{1}{3}}}, \beta = \frac{1+\alpha^2}{\left(1+\alpha^3\right)^{\frac{2}{3}}} \tag{148}$$

When the bifurcation is symmetric (also $\alpha = 1$), then

$$\frac{r_1}{r_0} = \frac{r_2}{r_0} = \frac{\alpha}{\left(1+\alpha^3\right)^{\frac{1}{3}}} \cong 0.8, \beta = \frac{1+\alpha^2}{\left(1+\alpha^3\right)^{\frac{2}{3}}} \cong 1,26 \tag{149}$$

This clearly shows that the vessel's cross section increases and the average velocity of the blood decreases by 26% in symmetric bifurcation following the Murray's law.
Substituting the Murray's law into the entropy production (144), we get the global entropy production in optimal case for the vessel

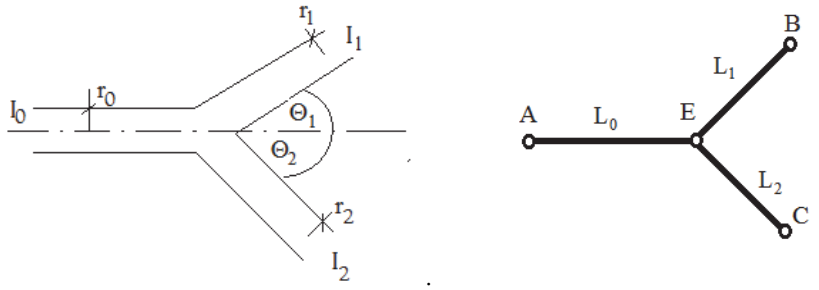
$$P = \frac{\frac{8\eta L}{\pi r^4} I^2 + kr^2 \pi L}{T} = \frac{1}{T}\frac{3}{2} kr^2 \pi L \tag{150}$$

The minimal entropy production has general consequences on the geometry of the vascular network. The most important one that the supply-network must not have loops, otherwise we could neglect this hypothetic loop without changing the supply with blood in any of the junction points. Also, we could decrease the entropy production without the missing supply of the junctions. The network, which has no loops called tree. So the minimal entropy production guarantees the tree-character of the vascular geometry.
The other important characteristic of bifurcation pattern is the bifurcation angle (see Fig. 1.a.), which also satisfies the minimal entropy production principle. To prove this consider the junction geometry determined by the *A, B*, and *C* points in Fig. 1.b. Its entropy production is

$$P = \frac{1}{T}\frac{3}{2}\pi k\left(r_0^2 L_0 + r_1^2 L_1 + r_2^2 \pi L_2\right) \tag{151}$$

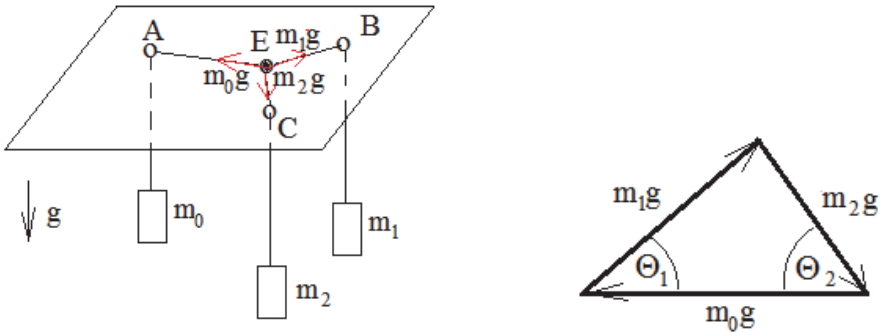
When we move the junction point *E*, the length of the sections will be varied as well, which changes the entropy production too. This problem could be elucidated by a mechanical analogue, which helps to show that the function (151) has a minimum. Draw the junction of Fig. 2.a. on a simple sheet of paper, and punch it at the points *A, B* and *C*. Let we load three strings with equal length with weights $m_A = \frac{1}{T}\frac{3}{2}\frac{\pi k}{g} r_0^2$, $m_B = \frac{1}{T}\frac{3}{2}\frac{\pi k}{g} r_1^2$ and $m_C = \frac{1}{T}\frac{3}{2}\frac{\pi k}{g} r_2^2$. Reeve the strings on corresponding *A, B* and *C* punches, and lift up the free ends of the strings by L_0, L_1, L_2 lengths at the m_A, m_B, and m_C weights, respectively. Knot the ends of the strings in one. The value of the potential energy of this system will be equal with the value of the entropy production of (151). Leaving the knot free, and neglect the friction, due to the Dirichlet principle the system will occupy its equilibrium state which has minimal potential energy. The geometry of the strings on the sheet will produce the optimal junction geometry. Due to this, the equilibrium is the resultant of the forces in the knot is zero (also the forces form a triangle; see Fig. 2.b). Following the analogy, from this the junction angles belonging to the minimum of global entropy production can be calculated by the law of cosines

$$\cos\Theta_1 = \frac{r_0^4 + r_1^4 - r_2^4}{2r_0^2 r_1^2}, \cos\Theta_2 = \frac{r_0^4 + r_2^4 - r_1^4}{2r_0^2 r_1^2} \tag{152}$$



(a) (b)

Fig. 1. Sketch for determination of the bifurcation angle.



(a) (b)

Fig. 2. Mechanical analogue for determination of bifurcation angle.

According to the experiments the Murray's law is not always accurate enough to describe the bifurcation properties of the vascular network. Recently Huo and Kassab [22] offered a correction of the (143) constitutive equation, based on the allometric scaling law with ¾ power [23]. The essence of their proposal is that they regard the cumulative vessel as thermodynamic system. The V volume of this cumulative vessel by the allometric scaling depends on the radius by $V = k_v \pi r^{\frac{2}{3}} L$, where k_v is constant. Hence the constitutive law (143) will be

$$P_{\deg r} = k k_v \pi r^{\frac{2}{3}} L \tag{153}$$

and the entropy production has a form

$$P = \frac{\frac{8\eta L}{\pi r^4} I^2 + k k_v r^{\frac{2}{3}} \pi L}{T} \tag{154}$$

2.3 Exercises

2.3. 1. Show that the Euler-Lagrange equations are identical with the inhomogeneous transport equations, when the variation principle (73) is a free-task varying only the intensive parameters. (Such type of non-exact principles of thermodynamics are called the partial variational principles [24].

2.3. 2. Show that in the case of inhomogeneous steady-state balance equations the fluctuation of GEP is not zero in the first order.

2.3. 3. Show with the help of the short calculus that fluctuation of the sum of the global spontaneous entropy production $\langle \nabla \Gamma_i, \mathbf{j}_i \rangle$ and the entropy production $\langle \sigma_i, \Gamma_i \rangle$ generated by the sources is zero in the first order in a steady-state system.

2.3. 4. Show that from the minimum principle of the global entropy production (which as we seen, is a restricted variation principle) can be constructed by Lagrange's method of multipliers in the following free variational-task:

$$\begin{aligned} &\int_V L\left(\mathbf{j}_j, \nabla \mathbf{j}_j, \nabla \Gamma_j, \Gamma_k, \lambda_i\right) dV - \int_{\Omega_{\Gamma_j}} j_{jn} \Gamma_j^* dA - \int_{\Omega_{j_j}} j_j^* \Gamma_j dA = \min \\ &L\left(\mathbf{j}_j, \nabla \mathbf{j}_j, \nabla \Gamma_j, \Gamma_k, \lambda_i\right) = \Phi\left(\mathbf{j}_j, \Gamma_k\right) + \Psi(\nabla \Gamma_j, \Gamma_k) + \lambda_i \nabla \cdot \mathbf{j}_i \end{aligned} \tag{155}$$

2.3. 5. Show with short calculus, that in the chapter of 2.2.2. the first variation of the global entropy production is zero, the second one is positive definite.

2.3. 6. Make a free variation-task from the partially restricted variation (131) by Lagrange's method of multipliers. Give the Euler-Lagrange equations.

2.3. 7. Determine the law corresponding with Murray's one when the global entropy production of the vessel has a form of (154).

2.3. 8. Determine the bifurcation angles, when the global entropy production of vessel has a form of (154).

3. Deduction of a class of constitutive equations from the principle of global entropy production

Up to this section we have used the minimum principle of the steady-state system and the connected minimum principle of the global entropy production to deduce the transport equations. In the following we shall give an equivalent functional formulation of the minimum principle of the global entropy production. We don't assume the validity of generalized Onsager constitutive theory and that the dissipation potentials are homogeneous Euler's functions. As opposed to this, we shall investigate the restrictions by minimum principle of global entropy production in the possible mathematical form of the constitutive equations.

We introduce from the minimum principle of global entropy production

$$P = \int_V L\left(X_j, j_j\right) dV = \int_V X_j j_j dV = \min . \tag{156}$$

It is well known from the variational calculus that the mathematical character and the necessary condition of a extremum are invariant to the transformations of the Langrangian

$$L\left(X_j, j_j\right) \to L\left(X_j, j_j\right) + \nabla \cdot \mathbf{F}\left(X_j, j_j\right), \tag{157}$$

where **F** is an arbitrary vector function with respect to Γ's, X's and j's.
Assume that the generator function **F** has the following bilinear form

$$\mathbf{F}\left(X_j, j_j\right) := -\mu \mathbf{J}_\mathbf{s} = -\mu \Gamma_i \mathbf{j}_i , \tag{158}$$

where μ is an arbitrary positive scalar. With this, the variation principle will have the following form

$$\begin{aligned} &S = \int_V L\left(X_j, j_j\right) dV = \min, \\ &L\left(X_j, j_j\right) = -(\lambda + 1)\Gamma_j \nabla \cdot \mathbf{j}_j - \lambda \nabla \Gamma_j \mathbf{j}_j \end{aligned} . \tag{159}$$

where we take into account that in the case of irreversible vector processes the $\nabla\Gamma_j$ are the conjugated thermodynamic force to the $\mathbf{j}_j$ flux, and we introduced the $\lambda := \mu - 1$ notation. To determine the possible form of the constitutive equations deduced from the above variation principle, starting with the thermodynamic equilibrium. This could be regarded as stationer state, where the global entropy production has its absolute minimum. As we know from irreversible thermodynamics this is zero and the forces and currents are independent and also have zero values.
Supposing there are missing connections between the forces and currents, the variational principle (159) is the necessary condition for the extremum at the free variation of the intensives and fluxes we get the Euler-Lagrange equations belonging to the variation principle (159) are the following

$$\nabla \cdot \mathbf{j}_j = 0, \quad \nabla \Gamma_j = 0, (j = 1,2,...,N) . \tag{160}$$

which are compatible with thermodynamic equilibrium of the system.
Supposing the constitutive equations between the forces and fluxes as

$$\begin{aligned} &\nabla \Gamma_i = \mathbf{f}_i(\mathbf{j}_j, \Gamma_k), \ (i, j = 1,2,...,N), \\ &\frac{\partial \Gamma_i}{\partial x_u} = f_{iu}(j_{ju}, \Gamma_k), (i, j, k = 1,2,...,N), (u = 1,2,3) \end{aligned} . \tag{161}$$

and their exact mathematical form is unknown. We study how the Euler-Lagrange equations belonging to the variation principle (159) limit the possible mathematical form of the constitutive equations.
Substituting the constitutive equations (161) into the Lagrangian of (159), then we get the following Lagrangian

$$L = -(\lambda + 1)\Gamma_i \frac{\partial j_{iu}}{\partial x_u} - \lambda f_{iv} j_{iv}, (i = 1,2,...,N), (u,v = 1,2,3) . \tag{162}$$

In consequence, the necessary condition of the extremum gives the Euler-Lagrange equations

$$\begin{aligned} &\frac{\partial L}{\partial \Gamma_i} = 0 - (\lambda + 1)\frac{\partial j_{iu}}{\partial x_u} - \lambda \frac{\partial f_{iu} j_{iu}}{\partial \Gamma_i} = 0, \\ &\frac{\partial L}{\partial f_{iu}} = 0 - (\lambda + 1)\frac{\partial \Gamma_i}{\partial x_u} - \lambda f_{iu} - \lambda j_{iv}\frac{\partial f_{iv}}{\partial j_{iu}} = \\ &= f_{iu} - \lambda j_{iv}\frac{\partial f_{iv}}{\partial j_{iu}} = 0, (i = 1,2,...,N), (u,v = 1,2,3) \end{aligned} \tag{163}$$

The Euler-Lagrange equations belonging to the variation of the thermodynamic fluxes are

$$f_{iu} = \lambda j_{iv}\frac{\partial f_{iv}}{\partial j_{iu}} (i = 1,2,...,N), (u,v = 1,2,3) \tag{164}$$

Assume that f_{iu} 's are two times continuously differentiable functions of the fluxes, and then the following derivatives are

$$\begin{aligned} &\frac{\partial f_{iu}}{\partial j_{kv}} = \lambda \frac{\partial f_{kv}}{\partial j_{iu}} + \lambda j_{kv}\frac{\partial^2 f_{kv}}{\partial j_{kv}\partial j_{iu}}, \\ &\frac{\partial f_{kv}}{\partial j_{iu}} = \lambda \frac{\partial f_{iu}}{\partial j_{kv}} + \lambda j_{iu}\frac{\partial^2 f_{iu}}{\partial j_{iu}\partial j_{kv}}, (i = 1,2,...,N), (u,v = 1,2,3) \end{aligned} \tag{165}$$

From the equations (164) and the Young' theorem of the continuous second partial derivatives, we get

$$\frac{\partial f_{iu}}{\partial j_{kv}} - \frac{\partial f_{kv}}{\partial j_{iu}} = \lambda(\frac{\partial f_{kv}}{\partial j_{iu}} - \frac{\partial f_{iu}}{\partial j_{kv}}), (i = 1,2,...,N), (u,v = 1,2,3) \tag{166}$$

Since these equations are obtained for every λ, we conclude the generalized reciprocal relations

$$\frac{\partial f_{iu}}{\partial j_{kv}} = \frac{\partial f_{kv}}{\partial j_{iu}}, (i = 1,2,...,N), (u,v = 1,2,3) \tag{167}$$

This very important result states that the necessary condition of the minimum principle of the global entropy production is that the constitutive equations must belong to the generalized Onsager constitutive theory. This property implies the existence of the dissipation potentials, which are constructed by Legendre transform to each other. Furthermore, the Donkin's theorem of entropy production (42) is valid, so the derivatives of the entropy production by the intensive parameters vanish. Consequently, the first group of the Euler-Lagrange equations (163) reduced to the homogeneous steady-state balance equations:

$$\frac{\partial j_{iu}}{\partial x_u} = 0, \left(i = 1,2,...,N\right), \left(u = 1,2,3\right) \tag{168}$$

which are necessary to the minimum of GEP as we showed above. Consequently, the proposed variation principle is equivalent with a class of stationer transport equations, where the constitutive equations are generalized Onsager-like. There are more constitutive properties resulting from this aspect of minimum entropy production. Studying these, substitute the constitutive equations into the entropy production

$$\sigma^* = f_{iv} j_{iv}, \left(i = 1,2,...,N\right), \left(u = 1,2,3\right) \tag{169}$$

and derive it with respect to the fluxes

$$\frac{\partial \sigma^*}{\partial j_{iv}} = f_{iv} + j_{iu} \frac{\partial f_{iu}}{\partial j_{iv}}, \left(i = 1,2,...,N\right), \left(u = 1,2,3\right) \tag{170}$$

From these equations using the equation (164), we get

$$\frac{\partial \sigma^*}{\partial j_{iv}} = \frac{\lambda + 1}{\lambda} f_{iv}, \left(i = 1,2,...,N\right), \left(u = 1,2,3\right) \tag{171}$$

and from this we obtain

$$j_{iv} \frac{\partial \sigma^*}{\partial j_{iv}} = \frac{\lambda + 1}{\lambda} j_{iv} f_{iv} = \frac{\lambda + 1}{\lambda} \sigma^* \tag{172}$$

So the entropy production is homogeneous Euler's function of degree $\frac{\lambda + 1}{\lambda}$ with respect to the fluxes. From this property of entropy production and from the physical meaning of flux potential described in (26), we get the connection between the flux dissipation potential and entropy production, like

$$\begin{aligned} \Phi\left(j_l, \Gamma_k\right) &= \int_0^\infty \sigma^*\left(e^{-t} j_j, \Gamma_k\right) dt = \sigma^*\left(j_j, \Gamma_k\right) \int_0^\infty e^{-t\frac{\lambda+1}{\lambda}} dt = \\ &= \frac{\lambda}{\lambda + 1} \sigma^*\left(j_j, \Gamma_k\right) \end{aligned} \tag{173}$$

Hence the flux potential is also a homogeneous Euler's function of degree $\frac{\lambda + 1}{\lambda}$, and it is a $\frac{\lambda}{\lambda + 1}$-th part of the entropy production. Since the force dissipation potential is the Legrende transform of the flux potential, so it follows that it (44) is homogeneous Euler's function of degree $\lambda + 1$ of the forces and $\frac{1}{\lambda + 1}$-th part of the entropy production. When $\lambda = 1$ then the dissipation potentials are the same as it is predicted in the linear Onsager's theory.

In summary, the previously given form of the minimal principle of entropy production leads to a class of generalized Onsager constitutive theory, which is also direct generalization of the linear Onsager's theory having their dissipation potentials as homogeneous Euler's functions.

3.1 Examples

3.1.1 Thermo-diffusion in triclinic system

In case of thermo-diffusion the global entropy production [25], [1], [2] is

$$\begin{aligned} P &= \int_V \left[\mathbf{j}_q \cdot \nabla\left(\frac{1}{T}\right) - \mathbf{j}_\alpha \cdot \nabla\left(\frac{\mu_\alpha}{T}\right) \right] dV = \\ &= \int_V \left[j_{qi} \frac{\partial}{\partial x_i}\left(\frac{1}{T}\right) - j_{\alpha i} \frac{\partial}{\partial x_i}\left(\frac{\mu_\alpha}{T}\right) \right] dV, (\alpha = 1,2,3,...,K), (i = 1,2,3) \end{aligned} \tag{174}$$

where $\mathbf{j}_q$ is the heat-current density vector, $\mathbf{j}_\alpha$ is the mass-current density of the α-th component, T is the temperature and μ_α is the chemical potential of the α-th component.
Let us consider the following linear constitutive equations

$$\begin{aligned} \frac{\partial}{\partial x_i}\left(\frac{1}{T}\right) &= R_{ik} j_{qk} - \mathrm{K}_{ik}^{(\alpha)} j_{\alpha k} \\ \frac{\partial}{\partial x_i}\left(\frac{\mu_\alpha}{T}\right) &= -D_{ik}^{(\alpha\beta)} j_{\beta k} - D_{ik}^{(\alpha)} j_{qk} \end{aligned} \tag{175}$$

where $[R_{ik}]$ is the specific heat-resistivity tensor, $\left[\mathrm{K}_{ik}^{(\alpha)}\right]$'s are second order tensors connect the temperature gradient and the fluxes of chemical components, $\left[D_{ik}^{(\alpha\beta)}\right]$ is the diffusion tensor, and $\left[D_{ik}^{(\alpha)}\right]$'s are second order tensors connect the chemical potential gradients and the heat-flux.
With these equations the Lagrangian of the problem is (162) :

$$\begin{aligned} L = &-\frac{\lambda+1}{T}\frac{\partial j_{qi}}{\partial x_i} - (\lambda+1)\frac{\mu_\alpha}{T}\frac{\partial j_{\alpha i}}{\partial x_i} - \lambda(R_{ik} j_{qi} j_{qk} - \mathrm{K}_{ik}^{(\alpha)} j_{qi} j_{\alpha k} \\ &- D_{ik}^{(\alpha)} j_{qi} j_{\alpha k} - D_{ik}^{(\alpha\beta)} j_{\alpha i} j_{\beta k}) \end{aligned} \tag{176}$$

It is possible to determine the relations among the kinetic coefficients from the Euler-Lagrange- equations belonging to the variation of i-th coordinate j_{qi} and $j_{\alpha i}$ of the heat and α-th chemical component current densities

$$\begin{aligned} j_{qi}\left(\lambda R_{ik} - R_{ki}\right) + j_{\alpha i}\left(\lambda\, \mathrm{K}_{ik}^{(\alpha)} - D_{ki}^{(\alpha)}\right) &= 0 \\ j_{qi}\left(\lambda\, \mathrm{K}_{ki}^{(\alpha)} - D_{ik}^{(\alpha)}\right) + j_{\alpha i}\left(\lambda\, D_{ik}^{(\alpha\beta)} - D_{ki}^{(\beta\alpha)}\right) &= 0 \end{aligned} \tag{177}$$

From these (due to the arbitrary current densities) we receive

$$\begin{aligned} j_{qi}\left(\lambda R_{ik} - R_{ki}\right) &= 0 \\ j_{\alpha i}\left(\lambda\, \mathrm{K}_{ki}^{(\alpha)} - D_{ik}^{(\alpha)}\right) &= 0 \\ j_{\alpha i}\left(\lambda\, D_{ik}^{(\alpha\beta)} - D_{ki}^{(\beta\alpha)}\right) &= 0 \end{aligned} \tag{178}$$

Since in the case of linear constitutive equations the λ=1, and hence we get the symmetry relations as follows

$$R_{ik} = R_{ik}, \quad \mathrm{K}_{ki}^{(\alpha)} = D_{ik}^{\alpha}, \quad D_{ik}^{(\alpha\beta)} = D_{ki}^{(\beta\alpha)} \tag{179}$$

With the application of the Curie principle from this general result we have the reciprocal relations for isotropic case, [1], [2], [25]. Due to the Curie-principle the $\left[R_{ik}\right]$, $\left[\mathrm{K}_{ik}^{(\alpha)}\right]$, $\left[D_{ik}^{(\alpha)}\right]$, and $\left[D_{ik}^{(\ \alpha\beta)}\right]$ tensors are simplified in isotropic case in the form of

$$R_{ik} = R\delta_{ik}, \quad \mathrm{K}_{ik}^{(\alpha)} = K^{(\alpha)}\delta_{ik}, D_{ik}^{\alpha} = D^{(\alpha)}\delta_{ik}, D_{ik}^{(\alpha\beta)} = D^{(\alpha\beta)}\delta_{ik} \tag{180}$$

where δ_{ik} is a Kronecker-delta. From this we receive the well-known [26] reciprocal relations

$$K^{(\alpha)} = D^{(\alpha)}, D^{(\alpha\beta)} = D^{(\beta\alpha)}, \left(\alpha, \beta = 1,2,3,...,K\right) \tag{181}$$

3.1.2 Electric conductivity in triclinic system in presence of external magnetic field

For simplicity we suppose the investigated process is isotherm. In this case the entropy production can be written as [1], [2]:

$$\sigma = \frac{1}{T}\mathbf{j}\cdot\mathbf{E} = -\frac{1}{T}\mathbf{j}\cdot\nabla\varphi \tag{182}$$

where $\mathbf{j}$ is the electric current density, and the intensive parameter φ is the electric potential. Let the constitutive equation be quasi-linear

$$\nabla\varphi = -\mathbf{R}\mathbf{j} - \kappa\,\mathbf{j}\times\mathbf{B} := \mathbf{R}(\mathbf{B})\mathbf{j}, \quad \frac{\partial\varphi}{\partial x_i} = -R_{ij}j_j - \kappa\varepsilon_{ijk}j_jB_k \tag{183}$$

where $\mathbf{R}$ is the specific resistivity tensor of the given crystal, κ is a constant, characterizing of the material, and ε_{ijk} is a permutation symbol, equals *1* when the series of *i,j,k* is even or *-1* when odd permutation of numbers *1,2,3*. The corresponding Lagrangian is

$$L = -(\lambda+1)\varphi\frac{\partial j_i}{\partial x_i} + \lambda R_{ij}j_ij_j + \lambda\kappa\varepsilon_{ijk}j_ij_jB_k \tag{184}$$

With variation of j-th coordinate j_j of electric current density we get the following Euler-Lagrange-equations

$$\begin{aligned} &\lambda\left(R_{ij} + R_{ji}\right)j_i + \lambda\kappa\left(\varepsilon_{ijk}j_iB_k + \varepsilon_{jik}j_iB_k\right) = \\ &= (\lambda+1)\left(R_{ij}j_i + \varepsilon_{ijk}j_iB_k\right) \end{aligned} \tag{185}$$

which by algebraic rearrangement leads to the equations

$$\lambda R_{ji} j_i + \lambda \kappa \varepsilon_{jik} j_i B_k = R_{ij} j_i + \varepsilon_{ijk} j_i B_k \tag{186}$$

Regarding the λ=1 in case of linear constitutive equations, we receive the relations

$$R_{ji} j_i + \kappa \varepsilon_{jik} j_i B_k = R_{ij} j_i + \kappa \varepsilon_{ijk} j_i B_k \,, \tag{187}$$

which is valid for every j_j and B_k. However, it could be satisfied in case only when the elements of the resistivity tensor are as follows

$$R_{ij}(\mathbf{B}) = R_{ij} + \kappa \varepsilon_{ijk} B_k \,, (i, j, k = 1,2,3) \,, \tag{188}$$

and $R_{ij}(\mathbf{B})$ satisfy the well-known

$$R_{ij}(\mathbf{B}) = R_{ji}(-\mathbf{B}), (i, j, k = 1,2,3) \,, \tag{189}$$

reciprocity relations.

4. Evolution toward steady-state. Glansdorff-Prigogine general evolution criterion. The Glansdorff - Prigogine theorem of the minimum entropy production

The above parts show the minimum principle for vector processes in the frame of the generalized Onsager constitutive theory by the directions of Onsager's last dissipation of energy principle. We had seen above that in case of source-free balances, this principle is equivalent with the principle of minimal entropy production. The equivalence of the two theorems in the frame of the linear constitutive theory was proven by Gyarmati [2] first. Furthermore, we showed that in case when the principle of minimal entropy production is used for the determination of the possible forms of constitutive equations, the results are similar to the linear theory in the frame of the Onsager's constitutive theory, where the dissipation potentials are homogeneous Euler's functions.

Glansdorff and Prigogine had chosen different approach for the minimum entropy production in the frame of linear constitutive theory [27], [28]. In the following we shall fit this theory into the generalized Onsager constitutive theory.

The original Glansdorff-Prigogine theorem of the minimum global entropy production may be described as follows. In every system satisfying the reciprocity relations of the linear Onsager's irreversible thermodynamics, where the convective velocity is zero and where the local value of the intensive parameters $\Gamma_k, (k = 1,2,...,N)$ at the boundary of the system are fixed, then the stationary state corresponds to a minimum of the global entropy production of the system. During the evolution of the system from its initial state to its stationary state, the rate of the global entropy production (RGEP)

$$\frac{dP}{dt} = \frac{d}{dt}\int_V \sigma \, dV = \int_V \frac{\partial X_j}{\partial t} j_j dV + \int_V X_j \frac{\partial j_j}{\partial t} dV < 0 \,. \tag{190}$$

decreases monotony and at the stationary state

$$\frac{dP}{dt} = 0 \,. \tag{191}$$

It can be shown that these properties of the global entropy production (GEP) guarantees the stability of non-equilibrium stationer state.
If the assumptions of the linear Onsager constitutive theory are not fulfilled, during the evolution of the system a more general theorem can still be derived in non-stationer state, namely,

$$\frac{dP_X}{dt} = \int_V \frac{\partial X_j}{\partial t} j_j dV < 0 \,. \tag{192}$$

and

$$\frac{dP_X}{dt} = 0 \,. \tag{193}$$

at a stationer state. This theorem is the Glansdorff-Prigogine general evolution criterion which is also a universal necessary condition to the existence of steady state [27], [28]. Consequently, the range of this theorem is independent of the assumption that the constitutive relations are linear, also it is valid in case of the generalized Onsager constitutive theory. In the followings we assume that the thermodynamic forces X_l and fluxes j_l belong to generalized Onsager constitutive theory. In this case, as a consequence of the (31) Legendre transformation the global entropy P produced inside the system can be formulated with the dissipation potentials, i.e.

$$P = \int_V \sigma dV = \int_V X_j j_j dV = \int_V \left[\Psi\left(X_l, \Gamma_k\right) + \Phi\left(j_l, \Gamma_k\right) \right] dV \,, \tag{194}$$

while the rate of the GEP can be written as

$$\frac{dP}{dt} = \int_V \frac{\partial X_j}{\partial t} \frac{\partial \Psi}{\partial X_j} dV + \int_V \frac{\partial j_j}{\partial t} \frac{\partial \Phi}{\partial j_j} dV \,, \tag{195}$$

where V is the volume of the system, which is at rest in the frame of reference.
The Glansdorff-Prigogine general evolution criterion, in case of time independent boundary conditions, which are valid during the evolution of the thermodynamic steady state, can be formulated as

$$\frac{dP_X}{dt} = \int_V \frac{\partial X_j}{\partial t} \frac{\partial \Psi}{\partial X_j} dV \le 0 \tag{196}$$

No similar conditions exist for RGEP and the flux part of RGEP in generalized Onsager constitutive theory. Also, the assumptions of this theory, contrary to the linear constitutive theory, do not guarantee the stabile extremum of the GEP at a thermodynamic steady-state. It is well known that in the linear constitutive theory the dissipation potentials are second order homogeneous Euler's functions. This is a central property of the potentials, which guarantees that the two parts of the GEP are proportional to the GEP and are equal with

each other. Consequently, if we assume that the dissipation potentials are arbitrary homogeneous Euler' functions in generalized Onsager constitutive theory, then the GEP have similar properties. Consider the force potential as Euler's function of degree k with respect to the X's, the GEP in this case from (43) is

$$P = \int_V \sigma dV = \int_V X_j j_j dV = \int_V X_j \frac{\partial \Psi}{X_j} dV = k \int_V \Psi dV \tag{197}$$

and the RGEP is

$$\frac{d}{dt} P = k \int_V \frac{\partial \Psi}{\partial X_j} \frac{\partial X_j}{\partial t} dV \tag{198}$$

According to the Glansdorff-Prigogine general evolutions criterion (192), from the above equation we get the necessary conditions of the minimum of the entropy production at the thermodynamics steady state

$$\frac{d}{dt} P = k \int_V \frac{\partial \Psi}{\partial X_j} \frac{\partial X_j}{\partial t} dV < 0 \quad away\ from\ steady\ state \tag{199}$$

and

$$\frac{d}{dt} P = 0 \quad at\ steady\ state\ . \tag{200}$$

Also the Glansdorff-Prigogine principle of the minimum of the global entropy production is valid in the generalized Onsager constitutive theory, when the dissipation functions are homogeneous Euler's functions with respect to thermodynamic forces and fluxes. It is evident, that the above criteria fulfilled for RGEP are sufficient for a stabile stationer state and it is not necessary. So the stationer state can exist not meeting this criteria.
It is important to note that the Glansdorff-Prigogine principle of the minimum of the global entropy production is more general than our above formulated minimum principle for vector processes even in the class of constitutive equations where the dissipation potentials are homogeneous Euler's functions. The reason of the limited sphere of operation that the general evolution criteria in (192) is valid for chemical reactions [29], [1], [2] which are crucial for the biological applications. The present generalization however is important, because any dissipation potential and the entropy production could be used as kinetic potential, which are important in the studies of the stability of chemical reactions [28], [29].

4.1 Stability of steady-state

The Glansdorff-Prigogine general evolution criterion involves the minimum of global entropy production in such a constitutive theory where the potentials are homogeneous Euler's functions. We show below the strictly convex property of dissipation potentials guarantee the minimum, and the function

$$P = \int_V \sigma dV, \int_V \Psi dV, \int_V \Phi dV \tag{201}$$

representing the global dissipations are simultaneous Lyapunov functions as well.
With this the Glansdorff-Prigogine's principle could be summarized in mathematical forms like

$$\begin{aligned}
&\frac{dP}{dt}=\frac{d}{dt}\int_V \sigma\, dV \le 0, \frac{d}{dt}\int_V \Psi dV \le 0, \frac{d}{dt}\int_V \Phi dV \le 0 \,(\textit{evolution criterion}),\\
&\delta^{(1)}P=0, \delta^{(1)}\int_V \Psi dV = 0,\ \ \delta^{(1)}\int_V \Phi dV = 0 \qquad (\textit{stationerity criterion}),\\
&\delta^{(2)}P\ge 0, \delta^{(2)}\int_V \Psi dV \ge 0,\ \ \delta^{(2)}\int_V \Phi dV \ge 0 \qquad (\textit{stability criterion})
\end{aligned} \tag{202}$$

where $\delta^{(1)},\delta^{(2)}$ are the first and second variations introduced in the first part of this article.
In the following we will consider only the second variations, showing that the criteria in the third row of (202) are equivalent. For simplicity we use the short calculus, which we introduced above and show that the strictly convex character of potentials assures the positive definite second variation $\delta^{(2)}P=\langle \delta X_i,\delta j_i\rangle \ge 0$ of GEP

$$\begin{aligned}
&j_i=\frac{\partial\Psi}{\partial X_i}\rightarrow \delta j_i=\delta\frac{\partial\Psi}{\partial X_i}=\frac{\partial^2\Psi}{\partial\Gamma_j\partial X_i}\delta\Gamma_j+\frac{\partial^2\Psi}{\partial X_j\partial X_i}\delta X_j=\frac{\partial^2\Psi}{\partial X_j\partial X_i}\delta X_j\rightarrow\\
&\rightarrow \delta^{(2)}P=\langle \delta j_i,\delta X_i\rangle=\left\langle \frac{\partial^2\Psi}{\partial X_j\partial X_i}\delta X_j,\delta X_i\right\rangle=\delta^{(2)}\int_V \Psi dV\ge 0\\
&X_i=\frac{\partial\Phi}{\partial j_i}\rightarrow \delta X_i=\delta\frac{\partial\Phi}{\partial j_i}=\frac{\partial^2\Phi}{\partial\Gamma_j\partial j_i}\delta\Gamma_j+\frac{\partial^2\Phi}{\partial j_j\partial j_i}\delta j_j=\frac{\partial^2\Phi}{\partial j_j\partial j_i}\delta j_j\rightarrow\\
&\rightarrow \delta^{(2)}P=\langle \delta j_i,\delta X_i\rangle=\left\langle \frac{\partial^2\Phi}{\partial j_j\partial j_i}\delta j_j,\delta j_i\right\rangle=\delta^{(2)}\int_V \Phi dV\ge 0
\end{aligned} \tag{203}$$

where we used the Donkin's theorem (36) and that the Hessian matrices of potentials are positive definite, (see for further proof the Appendix 7.3.).
Possibilities of the various stabilities of a steady-state system is shown on Fig. 3. A spontaneous fluctuation or perturbation can deviate the system from its stable steady-state in both directions. In consequence, the system has a new state by the fluctuation or perturbation, which would be the initial-state of the evolution after the disturbance. If, in this new state the entropy production is larger, then the evolution of the system will decrease the entropy production according to the first relation of (202) and the system returns to its original state. This case is shown in case of large perturbations by the X_{i1} state of the Figure 3., or in stationer case by X_{i2} state, when the perturbations are small enough. In case of the perturbed state, when the entropy production is smaller than it was in its original state, the evolution of the system will further decrease the entropy production, and continuously moves away from the initial stationer state governed by the first relation of (202). This happens in stationer state X_{i3}.
The stability of a stationer state against any perturbations could be determined from dissipation potentials, because for a strictly convex potential, the availability of global

minimum is guaranteed by availability of the local minimum; (for a proof see Appendix 7.4.). This condition guarantees that the state like X_{i2} cannot appear in constitutive theory. This circumstance has two important consequences. First is that the stationer state is stable, and second is that the system starting in any neighboring states of the final stationer one always reaches the same final state by its evolution. Shortly: these systems (in this general formulation also) show the Bertalanffy's equifinal behavior [30], [31].

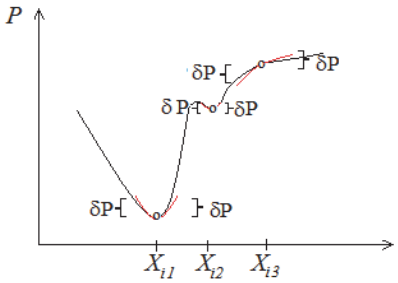


Fig. 3. Stable stationer state at X_{i1}. A state at X_{i2} which is stable to small fluctuations but unstable to fluctuations which exceed some threshold amplitudes. Instable state at X_{i3}.

4.2 Examples

4.2.1 Homeostasis, equifinality and 1/f scaling

It is proven above, when the dissipation potentials are homogeneous Euler's functions in the frame of generalized Onsager constitutive theory, then the GEP is governed by the Glansdorff-Prigogine's general evolution criterion. This guarantees two conditions:

1. From one side the stationer state with minimal global entropy production is stable, and
2. From other side, the Bertalanffy's equifinality principle is valid, due to the dissipation potentials are strictly convex functions. The arbitrary large fluctuations (in the frame of reasonable limits) the system remains stable, returns to its stationer state.

Let us study the homeostasis of the living systems as its stationer state, accepting the concept of Nicolis and Prigogine, that the reference state of homeostasis should be regarded as a non-equilibrium steady- state [32]. The homeostasis as thermodynamic steady-state is controlled by hypothalamus, regulated by the sympathetic and the parasympathetic nervous system, as the action processor [33]. More precisely, the homeostatic state is caused by the intricate balance between antagonistic activity of the sympathetic and the parasympathetic nervous system [34], [33]. The actual value of the physical character is governed by the homeostasis $X(t)$ composed of a time-independent steady state X_S and a time-dependent stochastic fluctuation $V(t)$. This last one is created commonly by the outside perturbations and the controller is independent from the initial conditions according to the Bertalanffy's equifinality principle.

In the following we discuss further two characters of the homeostasis, which could be obtained form the generalized results of [34]. They had shown, that a healthy heart rate obeys 1/f scaling. When this scaling property is valid for all the homeostasis; then

1. the necessary consequences of the homeostasis are the large fluctuations, and on the other hand,

2. the 1/f scaling guarantees the constant (time independent) value of the deviation when the fluctuations could be characterized by their autocorrelation function in statistical point of view.

We shall show these points in the next parts. Supposing the homeostatic signal around the stationer state is

$$V(t) = X(t) - X_S \tag{204}$$

And its fluctuation is a self-similar stochastic signal, can be characterized by its autocorrelation function form statistical point of view. The self-similarity in this case means that the deviation of the fluctuations is

$$\left\langle V^2(t) \right\rangle = t^{2H} \tag{205}$$

where H is the Hurst' exponent of the self-similarity [35], [36]. For example when $H = 1/2$, than

$$\left\langle V^2(t) \right\rangle = t \tag{206}$$

In this case the fluctuation is the noise of the random walk (Brownian) motion. The scaling law from (205) is

$$\left\langle V^2(rt) \right\rangle = r^{2H} \left\langle V^2(t) \right\rangle \tag{207}$$

The noise could be characterized not only by the exponent of self-similarity, but by its spectral power density as well [37]:

$$S(f) = F\left[G(\Delta t)\right] \tag{208}$$

where

$$G(\Delta t) = \left\langle V(t)V(t+\Delta t) \right\rangle - \left\langle V(t)^2 \right\rangle \tag{209}$$

is the autocorrelation function of the homeostasis noise and $F\left[G(\Delta t)\right]$ is the Fourier transform of this autocorrelation function. The spectral power density of the self-similar noises is [36], [37]:

$$S(f) \propto \frac{1}{f^{\beta}} \tag{210}$$

where f is the frequency. Consequently, the power density of the self-similar signal is also self-similar, by exponent of $(-\beta)$. In case of $1/f$ scaling law, from (210):

$$\lim_{f \to 0} S(f) \propto \lim_{f \to 0} \frac{1}{f} \to \infty \tag{211}$$

Hence the amplitude of the low frequency components of the noise could be extremely large, pushing far away the system from the X_S state, which anyway is required to control the homeostasis. This is why the system must have the property of Bertalanffy's equifinality. The noise "colority" exponent β, and the self-similarity exponent H are not independent [34], [35]:

$$\beta = 2H + 1 \tag{212}$$

Consequently, in the case of $1/f$ scaling law

$$H = 0 \tag{213}$$

Hence, in consequence of (205), the deviation of this noise is time-independent. From this point of view, especially interesting the results of Struzik et al [34]: „modifying the relative importance of either of the two branches of the autonomic nervous system leads to a substantial decrease in 1/f scaling. In particular, the relative parasympathetic suppression, both by congestive heart failure and by the parasympathetic blocker atropine, results in a substantial increase in the Hurst exponent H and a shift of the power density spectrum $S(f)$ from 1/f towards random walk scaling:$1/f^{\beta}, \beta > 1$." If $\beta > 1$, than, based on (212) the Hurst's exponent will be positive, and so the deviation of the fluctuation by (205) will be monotonic increasing function of time. This is/can be predicted by Chebishev's theorem [38], [39] increasing probability leaving the domain where the Bertalanffy's equifinality property is valid.

4.2.2 Universality of entropy production of living systems

It is a widely accepted opinion, that by evolution the living systems are developing towards such optimal state, where the energy used for sustenance of the life and operation of the energy-distribution system is the smallest. This last fact/point was also mentioned in connection with Murray's law before. However, the evolution of the living systems has a universal thermodynamic behavior (Andresen et al [40]). Regarding the living system as a set of cells, where the cells born and die, and which gets its energy from an energy-distributor network to fulfill its functions. The state of the system could be characterized by the extensive parameters of number of the cells N and internal energy U, for which the balance equations are

$$\frac{dN}{dt} = I_N + \Sigma_N, \quad \frac{dU}{dt} = I + \Sigma_U \tag{214}$$

Disregarding the cell-motility, the number of cells has no flux, so it has only an augment (cell born) and attenuate (cell death) factors

$$\frac{dN}{dt} = K - D \tag{215}$$

where K is the born number rate of cells, and D is the death number rate of cells. This last one obtains the constitutive equation [41],

$$N = \lambda N \tag{216}$$

where λ is the specific rate of cell-death, and its reciprocal value $\lambda^{-1} = T$ is the average lifetime of the cells. With these the balance equation of the cell-number is

$$\frac{dN}{dt} = K - \lambda N \tag{217}$$

The I energy delivered in a unit-time by the blood-flow to the organ and distributed by the capillaries is the metabolic flux. This energy serves mainly a double task: a part for the sustaining the system and a part for the generation of new cells

$$I = NB + EP = NB + E\left(\frac{dN}{dt} + \lambda N\right) \tag{218}$$

where B is the metabolic rate of a single cell, and E is the metabolic energy for the production of a single cell. From this the energy-balance follows

$$\begin{aligned} &E\frac{dN}{dt} := \frac{dU}{dt} = I - N\left(B + \lambda E\right) = I - U\left(\frac{B}{E} + \lambda\right) = \\ &= I + \Sigma_U,\ U := EN,\ \Sigma_U := U\left(\frac{B}{E} + \lambda\right) \end{aligned} \tag{219}$$

A constitutive equation is valid for the metabolic flux [41]:

$$I = B_0 N^{\alpha} = \frac{B_0}{E^{\alpha}} U^{\alpha} \tag{220}$$

where B_{0*} is the metabolic flux of a single cell, $|1/2 < \alpha < 1$ is a scaling exponent, which characterizes the fractal structure of the vascularization of the actual system. This last $\alpha = 3/4$, supposed to create optimal distribution network according to West [42], [43], [44]. With this the energy-balance will be a non-linear differential equation for internal energy

$$\frac{dU}{dt} = aU^{\alpha} - bU\ ,\ a := \frac{B_0}{E},\quad b := \left(\frac{B}{E} + \lambda\right) \tag{221}$$

The first term on the right side of the equation is the energy flux, while the second term is the dissipation of the energy (Σ_U) in the unit time. (The solution of the energy-balance is a sigmoid function [37] can be given in the form [41]:

$$\begin{aligned} &\left(\frac{U}{U_\infty}\right)^{1-\alpha} = 1 - \left(1 - \left(\frac{U_0}{U_\infty}\right)^{1-\alpha}\right) e^{-at/4U_\infty^{1-\alpha}} = 1 - e^{-at/4U_\infty^{1-\alpha} + \ln\left(1 - \left(\frac{U_0}{U_\infty}\right)^{1-\alpha}\right)} = 1 - e^{-\tau}, \\ &\tau(t) := at/4U_\infty^{1-\alpha} - \ln\left(1 - \left(\frac{U_0}{U_\infty}\right)^{1-\alpha}\right) \end{aligned} \tag{222}$$

where U_∞ is the energy in stationer final state, U_0 is the energy in the time of birth. According to (222) the $\left(\frac{U}{U_\infty}\right)^{1-\alpha}$ quantity is a universal function of $\tau(t)$ time (so called biological self-time), in each biological systems. Let us study the dissipative part of the energy-balance in stationer state. The entropy production is stationer state from (221) is

$$P = \frac{bU_\infty}{T} = \frac{aU_\infty^\alpha}{T} \tag{223}$$

where T the temperature of the system, is supposed to be constant. Consequently, the scaling law of the entropy-production is

$$P = \frac{aU_\infty^\alpha}{T} = \frac{aE^\alpha}{Tm^\alpha}\left(m\ N_\infty\right)^\alpha = k_p M^\alpha, \quad M := m\ N_\infty \tag{224}$$

where m is the mass of a single cell, N_∞ is the number of cells in stationer state, and k_p is a constant.

This means that the entropy change of the system due to the entropy production during dt time is

$$dS = Pdt = k_p M^\alpha dt \tag{225}$$

Having the same entropy change in biological self-time scale the entropy production is: $P_b = P\frac{dt}{d\tau}$. We get that in the scale of the biological self-time the entropy production per unit mass of the system is constant [40]:

$$\begin{aligned} &\frac{1}{M}P_b = \frac{P}{M}\frac{dt}{d\tau} = \frac{P}{M}\frac{4U_\infty^{1-\alpha}}{a} = \frac{4P}{aM}\left(\frac{E}{m}\right)^{1-\alpha} M^{1-\alpha} = \frac{4k_p M^\alpha}{aM}\left(\frac{E}{m}\right)^{1-\alpha} M^{1-\alpha} = \\ &= \frac{4}{a}\left(\frac{E}{m}\right)^{1-\alpha} k_p = k \end{aligned} \tag{226}$$

Supposing that, due to the evolution process, this specific entropy production is optimal, also minimum, then we can conclude that the state, in which the entropy production is minimal, is realized on the biological self-time scale. This optimal state is independent of the optimum of the vascularization, because the result does not depend on α. Due to the equality of the metabolic flux and the rate of energy dissipation in stationer state, the metabolic entropy flux (belonging to the metabolic flux) per unit mass is constant on the biological self-time scale. Also, we can suppose the nourishment intake (food consumption per unit time) is proportional to the metabolic flux, so the nourishment rate per unit mass is also a universal constant in the biological self-time scale.

Furthermore, supposing the universal life-span in the biological self-time scale for every living individuals, the life-span (τ_{life}) is also a universal constant on this scale. Based on (222) the life-span on the physical time-scale is

$$t_{life} := \tau_{life} \frac{4}{a} U_{\infty}^{1-\alpha} = \tau_{life} \frac{4}{a} \left(\frac{m}{E} \right)^{1-\alpha} M^{1-\alpha} = k_{life} M^{1-\alpha} \tag{227}$$

Supposing that the nourishment rate (NR) is proportional with the metabolic flux, so $TR \propto M^{\alpha}$ then the energy demand by the nourishment on the physical life-span of the living organisms is

$$\begin{aligned} W_{tot} = NRt_{life} := \tau_{life} \frac{4}{a} U_{\infty}^{1-\alpha} = \tau_{life} \frac{4}{a} \left(\frac{m}{E} \right)^{1-\alpha} M^{1-\alpha} = \\ = NRk_{life} M^{1-\alpha} = k_{tot} M^{\alpha} M^{1-\alpha} = k_{tot} M \end{aligned} \tag{228}$$

This means that each unit mass of all the living organisms consumes the same amount of energy in their life-span. The nourishment rate is proportional to the product of the mass flux of nutrients and its chemical potential, so proportional to the flux of negentropy [45]. Consequently, the small living organisms have large, the large organisms have small negentropy flux, because according to (228) their mass density has the same negentropy consumption in their actual life-span.

4.2.3 Autocatalytic chemical reactions

The chemical reaction kinetics is a class of irreversible processes, which remains outside of the applicability area of the Onsager constitutive theory. Hence, the reaction kinetics is an ideal discipline to test the applicability of the generalized Onsager constitutive theory. We shall/are going to study the autocatalytic reactions in a different way other than [46], [47] did it before, emphasizing the following:

- We show that these types of reactions could be described by a dissipation potential which are homogeneous Euler's function of degree 3 of the concentration of chemical components.
- We show that this potential is a Lyapunov function of the differential equation of reaction kinetics.
- We give criteria of the stability of the stationer state by means of the dissipation potential (as Lyapunov function).
- We show that in the case of small perturbations this stability criteria is identical with the stability criteria deduced from the positive definite property of the second variation of the global entropy production.

Consider the following homogeneous autocatalytic reaction in a volume V

$$A + X \underset{k_{-1}}{\overset{k_1}{\rightleftarrows}} 2X \tag{229}$$

where the chemical component A has C_A constant concentration which is necessary to proceed in/to a stationer state. The rate of the reaction $\frac{d\xi}{dt}$ as a thermodynamic flux is proportional with the rate of the C concentration of the chemical component X [48], [46].

$$\frac{dC}{dt} = \frac{1}{V} \frac{d\xi}{dt} \tag{230}$$

The parameter conjugated to the rate of reaction in the entropy production is the affinity A has the form

$$\begin{aligned} A &= \mu_A + \mu_X - 2\mu_X = const - RT\ln\gamma C, \\ \mu_X &= \mu_{X0} + RT\ln\gamma C, \end{aligned} \tag{231}$$

where γ the affinity coefficient [48]. The reaction rate can be described by the reaction kinetic equation

$$\frac{1}{V}\frac{d\xi}{dt} = \frac{dC}{dt} = k_1 C_A C - k_{-1} C^2 \tag{232}$$

and it is homogeneous Euler's function of degree 2. From this, the concentration in stationer state is

$$C_{stac} = \frac{k_1 C_A}{k_{-1}} \tag{233}$$

This steady-state is an equilibrium state in thermodynamic point of view, where the entropy production has its absolute minimum which is zero. Let the difference of the actual concentration from its stationer state of component X be c. In this case, following a simple calculation, the reaction kinetic equation (232) can be written in the following form

$$\frac{dc}{dt} = -k_1 C_A c - k_{-1} c^2 \tag{234}$$

Denote the deviation of the actual affinity from its stationer value by α , then

$$\alpha = -RT\ln\gamma\left(1 + \frac{c}{C_{stac}}\right) \cong -\gamma\frac{RT}{C_{stac}}c \tag{235}$$

where we assumed that the deviation is small. In this approach the entropy production is

$$P = \frac{1}{T}\frac{d\xi}{dt}A \cong \frac{V}{T}\frac{dc}{dt}\frac{\gamma RT}{C_{stac}}(-c) = \frac{\gamma RV}{C_{stac}}\frac{dc}{dt}(-c) \tag{236}$$

which shows that the concentration rate and the concentration itself could be regarded as thermodynamic flux and thermodynamic force, respectively. The reaction kinetic equation (232), which makes a connection between these variables, can be considered as constitutive equation of the system. With this, on the analogy of the second equation of (21) the force dissipation potential is

$$\begin{aligned} &\Psi(C_A, c) = \int_0^1 (-c)\frac{dc}{dt}(C_A, \tau c)d\tau = \int_0^1 c k_1 C_A(\tau c) + k_{-1} c(\tau c)^2 d\tau = \\ &= \frac{1}{2}k_1 C_A c^2 + \frac{1}{3}k_{-1}c^3 \end{aligned} \tag{237}$$

We show, that this dissipation potential is a Lyapunov function for definite range of concentrations. (Properties of Lyapunov function is discussed in [49], [47], and in Appendix

7.8.) The dissipation function (237) is positive in the $-\frac{3}{2}C_{stac} < c < C_{\lim}$ interval of the concentration, where $C_{\lim}$ is a positive limit-concentration till the model is applicable, and it has its zero value only at $c = 0$ steady-state point. The second variation of (237) and its positivity interval are

$$\delta^{(2)}\Psi = \frac{\partial^2 \Psi\left(C_A, c\right)}{\partial c^2}\left(\delta c\right)^2 = \left(k_1 C_A + 2k_{-1}c\right)\left(\delta c\right)^2 \geq 0, \quad -\frac{1}{2}C_{stac} < c < C_{\lim} \tag{238}$$

Taking into account the reaction kinetic equation (232) in the time-derivative of the potential, then we get

$$\frac{d\Psi}{dt} = \frac{\partial \Psi\left(C_A, c\right)}{\partial c}\frac{dc}{dt} = -\left(\frac{dc}{dt}\right)^2 \leq 0 \tag{239}$$

It is also seem from the reaction kinetic equation (232), that except the $c = 0$ steady-state point the rate of concentration has no more zero values in the interval of $-\frac{1}{2}C_{stac} < c < C_{\lim}$. Consequently, the dissipation potential is a Lyapunov function of the reaction kinetic equation (232) in the $-\frac{1}{2}C_{stac} < c < C_{\lim}$ interval, and it guarantees that its solution is asymptotically stable, i.e.

$$\lim_{t \to \infty} c(t) = 0 \tag{240}$$

In case of small perturbations the dissipation potential can be replaced by its second variation at the steady-state point

$$\Psi(C_A, \delta c) = \delta^{(2)}\Psi = \left.\frac{\partial^2 \Psi\left(C_A, c\right)}{\partial c^2}\right|_{c=0}\left(\delta c\right)^2 = k_1 C_A \left(\delta c\right)^2 \geq 0, \tag{241}$$

and so the reaction kinetic equation will be

$$\frac{d\delta c}{dt} = \frac{\partial \Psi(C_A, \delta c)}{\partial\left(-\delta c\right)} = -2k_1 C_A \left(\delta c\right) \rightarrow \delta c = \delta c_0 e^{-\frac{t}{\tau}}, \quad \tau := \frac{1}{2k_1 C_A} = \frac{1}{2k_{-1}C_{stac}} \tag{242}$$

Consequently, the system is stable for small perturbations too. From this equations and the (230) relation follows that

$$\delta\frac{d\xi}{dt} = V\frac{d\delta c}{dt} = -V2k_1 C_A \delta c \tag{243}$$

On the other hand, the variation of affinity from (235) is

$$\delta\alpha = -\delta\left[RT\ln\gamma(1+\frac{c}{C_{stac}}) \right] = -\gamma\frac{RT}{C_{stac}}\delta c \tag{244}$$

Hence the second variation of the global entropy production is positive definite for all small perturbations, i.e.

$$\delta^{(2)}P = \frac{1}{T}\delta\alpha\delta\frac{d\xi}{dt} = V2\gamma R\frac{k_1 C_A}{C_{stac}}(\delta c)^2 \geq 0 \tag{245}$$

Consequently, the stability criteria (202) is satisfied in case of small perturbations. This can be used to decide about the stability of the system, but only in case of small perturbations.

4.3 Exercises

4.3.1. Let ΔX be the deviation of the homeostasis signal from its stationer value X_S, till the equifinality is valid. Calculate the P probability by Chebishev theorem, when the homeostasis signal exceeds the value $P < \frac{\langle V^2(t)\rangle}{\Delta X^2}$, where the ΔX is the deviation from the homeostatic X_S stationer state. Explain the result.

4.3.2. Accomplish the above estimation when the homeostasis signal has Gauss' distribution. Compare the results.

4.3.3. Calculate the energy-balance equation when the living organism performs extra mechanical work.

4.3.4. According to H. Lin [50] the maximal mechanical work of an animal (e.g. when they must get away from enemies) has the same scaling law as the metabolic rate [51] has. Show in this case, that the validity of universality law of specific entropy -production is valid at the biologic self-time scale.

4.4 Boundary layer of homogeneous systems with two state-variables, as steady-state system

We shall study the system in Fig. 4. by means of the minimum principle of the global entropy production. Let the system be a so called current-tube, which is a tube where the boundary surface is constructed by those vector-lines of the thermodynamic fluxes, which fit on a closed line (for example circle, polygon).

Moreover, let us assume a common current-tube for all the fluxes in the system. (This is always the case in the isotropic system.) As shown above, the minimum principle of the global entropy production is valid when the fluxes fulfill the source-free balance equations

$$\nabla\cdot\mathbf{j}_i = 0, (i = 1,2,3,...,N) \tag{246}$$

Integrating (246) on the V volume of the current tube and applying the Gauss's divergence theorem

$$\int_V \nabla\cdot\mathbf{j}_i dV = \oint_\Omega j_{in} dA = 0, (i = 1,2,3,...,N) \tag{247}$$

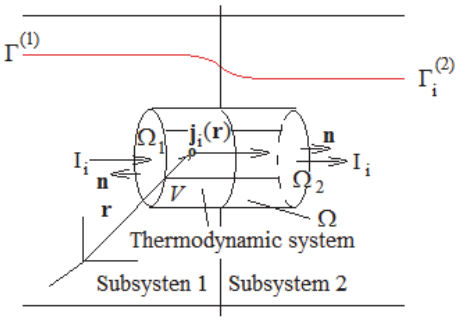


Fig. 4. Current-tube as thermodynamic system.

where Ω is the boundary of the system. Counting the vanished normal coordinates of fluxes on the boundary, the currents of the fluxes in the tube are:

$$\begin{aligned} &-\oint_{\Omega_1} j_{in}dA + \oint_{\Omega_2} j_{in}dA = 0, \rightarrow -\oint_{\Omega_1} j_{in}dA = -\oint_{\Omega_2} j_{in}dA \\ &\rightarrow I_i := -\oint_{\Omega_1} j_{in}dA = -\oint_{\Omega_2} j_{in}dA, \left(i = 1,2,3,...,N\right) \end{aligned} \tag{248}$$

The global entropy production of the current-tube is

$$\begin{aligned} &\int_V \nabla\Gamma_i \cdot \mathbf{j}_i dV = \int_V \nabla\cdot\left(\Gamma_i \mathbf{j}_i\right)dV = \oint_\Omega \Gamma_i j_{in}dA = \\ &= -\oint_{\Omega_1} \Gamma_i^{(1)} j_{in}dA + \oint_{\Omega_2} \Gamma_i^{(2)} j_{in}dA = \left(\Gamma_i^{(1)} - \Gamma_i^{(2)}\right) I_i = \\ &= X_i I_i,\ X_i := \left(\Gamma_i^{(1)} - \Gamma_i^{(2)}\right),\ (i = 1,2,3,...,N) \end{aligned} \tag{249}$$

where we assumed that the fluxes have no sources, and the intensive parameters are space-wise constants at the ends of the tube as we introduced the X_i thermodynamic forces, which are the differences of the intensive parameters of the subsystems. The global entropy production is a bilinear form again, where the currents and forces are connected by the constitutive equations. Assume these constitutive equations belong to that type of generalized Onsager constitutive theory; where the dissipation potentials are homogeneous Euler's functions. In this case

$$\begin{aligned} &I_i = \frac{\partial\Phi\left(X_j\right)}{\partial X_i}, \quad X_i = \frac{\partial\Psi\left(I_j\right)}{\partial I_i}, (i,j = 1,2,3,...,N) \\ &O(X_j, I_j) := \Phi + \Psi - X_i I_i \geq 0 \end{aligned} \tag{250}$$

where the equality in the global OM-function is valid when the material equations of the first row are fulfilled. Moreover, for the relations among the global entropy production and the dissipation potentials, as we had seen before, we get

$$P = X_i(I_j)I_i = \frac{\partial\Phi\left(I_j\right)}{\partial I_i}I_i = k_1\Phi\left(I_j\right) =$$
$$= X_i I_i(X_j) = X_i\frac{\partial\Psi\left(X_j\right)}{\partial X_i} = k_2\Psi\left(X_j\right), (i,j = 1,2,3,...,N) \quad (251)$$
$$k_1^{-1} + k_2^{-1} = 1$$

where we used the second relation of (44).
Let us study two important cases:

- one is when we fix the first k forces from all N thermodynamic forces of the system. According to DeGroot referring on this stationer state of the system as the force-constrained stationer state of degree k [52],
- the second that we fix the first k currents from all N thermodynamic currents of the system. According to DeGroot referring on this stationer state of the system as the current-constrained stationer state of degree k [52]

4.5 Force-constrained stationer state

Let the forces $X_1,...,X_k$ be fixed and allow the forces $X_{k+1},...,X_N$ to vary. According to (251), in this case the force dissipation potential and the GEP depend on these variables and

$$P(X_j) = k_2\Psi\left(X_j\right), (j = k+1,2,3,...,N) \quad (252)$$

Differentiate the GEP with respect to $X_{k+1},...,X_N$, then we obtain

$$\frac{\partial P(X_j)}{\partial X_j} = k_2\frac{\partial\Psi(X_j)}{\partial X_j} = k_2 I_j = 0, (j = k+1,2,3,...,N) \quad (253)$$

which are the necessary conditions of the minimum of the global entropy production in a steady-state. Also fluxes conjugated to free thermodynamic forces vanish in steady-state, however the conjugated currents to the fix forces are not zero. These could be determined form the equations

$$\left.\frac{\partial\Phi(I_j)}{\partial I_j}\right|_{I_{k+1}=...I_{N=0}} = X_j = 0, (j = 1,2,3,...,k) \quad (254)$$

4.6 Current constrained stationer state

Let the fluxes $I_{k+1},...,I_N$ be prescribed and allow the forces $I_{k+1},...,I_N$ to vary. According to (251), in this case the current dissipation potential and the GEP depend on these variables and

$$P(I_j) = k_1\Phi\left(I_j\right), (j = k+1,2,3,...,N \quad (255)$$

and

$$\frac{\partial P(I_j)}{\partial I_j} = k_1 \frac{\partial \Phi(I_j)}{\partial I_j} = k_1 X_j = 0, (j = k+1, 2, 3, ..., N) \tag{256}$$

Differentiate the GEP with respect to free fluxes $I_{k+1}, ..., I_N$, then we get the necessary conditions of the minimum of the GEP

$$\left.\frac{\partial \Psi(I_j)}{\partial I_j}\right|_{I_{k+1}=...=I_N=0} = X_j = 0, (j = 1, 2, 3, ..., k) \tag{257}$$

Also, in this steady state the forces conjugate to free currents vanish. However, the conjugated currents to the fix forces are not zero. These could be determined form the equations

$$\left.\frac{\partial \Psi(I_j)}{\partial I_j}\right|_{I_{k+1}=...=I_N=0} = X_j = 0, (j = 1, 2, 3, ..., k) \tag{258}$$

There are two extreme cases:

- when $k = 0$ and so all the currents are zero in stationer case. This is the thermodynamic equilibrium, and the global entropy production has its absolute minimum.
- when $k = N$. This is a stationer state without freedom (no free parameters exist).

4.7 Example: serial and parallel connections of current tubes

4.7.1 Serial circuit

Consider the serial array shown in Fig. 5., where both the thermodynamic subsystems have identical thermodynamic currents, and its thermodynamic forces are additive values, because they are originated from the gradients of intensive parameters by line-integration). The dissipation potentials are given for each current-tube.

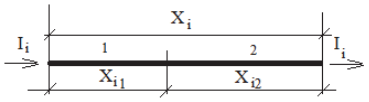


Fig. 5. Serial connection of the current-tubes.

The thermodynamic currents are identical in each current tube of the subsystems, so:

$$\frac{\partial \Psi_2(X_{j2})}{\partial X_{i2}} = -\frac{\partial \Psi_2(X_j - X_{j1})}{\partial X_{j1}}, (j = 1, 2, 3, ..., N) \tag{259}$$

Due to the fixed sums of the forces, the dissipation potentials depend on the thermodynamic forces of one of subsystems only; let us say the 1st one. Using this

$$\frac{\partial \left[\Psi_1(X_{j1}) + \Psi_2(X_j - X_{j1})\right]}{\partial X_{i1}} = 0, \ (i, j = 1, 2, 3, ..., N) \tag{260}$$

Substituting these to (259) the result is

$$X_{i2} = X_i - X_{i1}, \quad (i = 1,2,3,...,N) \tag{261}$$

The forces belonging to the stationer state of the 1st subsystem are the solutions of the above generally non-linear system of N equations and with them the forces of the 2nd subsystem are

$$\begin{aligned} &\Psi_1\left(X_{j1}\right) + \Psi_2\left(X_{j2}\right) = \min, \\ &X_{j1} + X_{j2} = X_j \end{aligned} \tag{262}$$

This result shows how the subsystems share the thermodynamic forces. What is the situation with the extremum property of GEP? The equations to solve this problem could be derived from the following constrained extremum-task

$$P = I_i X_i = I_i X_{i1} + I_i X_{i2} = P_1 + P_2 \tag{263}$$

and due to the additivity of the forces, the entropy production of the system is

$$P_l = I_{il} X_i = X_i \frac{\partial \Psi_l}{\partial X_i} = k_{\Psi l} \Psi_l\left(X_j\right), \left(l = 1,2\right), \left(i, j = 1,2,3,...,N\right) \tag{264}$$

Using the equation expressing the connection between the entropy production and dissipation potentials

$$P\left(X_{j1}\right) = k_{\Psi 1} \Psi_1\left(X_{j1}\right) + k_{\Psi 2} \Psi_2\left(X_j - X_{j1}\right) \tag{265}$$

We get from (265)

$$\begin{aligned} &\frac{\partial P\left(X_{j1}\right)}{\partial X_{i1}} = k_{\Psi 1} \Psi_1\left(X_{j1}\right) + k_{\Psi 2} \Psi_2\left(X_j - X_{j1}\right) = \\ &= k_{\Psi 1} \frac{\partial\left[\Psi_1\left(X_{j1}\right) + \Psi_2\left(X_j - X_{j1}\right)\right]}{\partial X_{i1}} + \quad \left(k_{\Psi 2} - k_{\Psi 1}\right) \frac{\partial \Psi_2\left(X_j - X_{j1}\right)}{\partial X_{i1}} = \\ &= \quad \left(k_{\Psi 2} - k_{\Psi 1}\right) \frac{\partial \Psi_2\left(X_j - X_{j1}\right)}{\partial X_{i1}}, \quad \left(i, j = 1,2,3,...,N\right) \end{aligned} \tag{266}$$

where $k_{\Psi l}$ are the degrees of the homogeneous Euler's functions. From this we get

$$\begin{aligned} &\frac{\partial P\left(X_{j1}\right)}{\partial X_{i1}} = k_{\Psi 1} \Psi_1\left(X_{j1}\right) + k_{\Psi 2} \Psi_2\left(X_j - X_{j1}\right) = \\ &= k_{\Psi 1} \frac{\partial\left[\Psi_1\left(X_{j1}\right) + \Psi_2\left(X_j - X_{j1}\right)\right]}{\partial X_{i1}} + \quad \left(k_{\Psi 2} - k_{\Psi 1}\right) \frac{\partial \Psi_2\left(X_j - X_{j1}\right)}{\partial X_{i1}} = \\ &= \quad \left(k_{\Psi 2} - k_{\Psi 1}\right) \frac{\partial \Psi_2\left(X_j - X_{j1}\right)}{\partial X_{i1}}, \quad \left(i, j = 1,2,3,...,N\right) \end{aligned} \tag{267}$$

where we used the (261) equations.
These calculations have two important consequences:

1. In the serial connection of the current tubes, the sum of the force potentials will be minimal in the stationer state of the system.
2. The GEP of the system in stationer case could be minimal then and only then when the dissipation potentials of the subsystems are homogeneous Euler's functions with identical degree.

4.7.2 Parallel connection

Consider a parallel array shown in Fig. 6., where both the thermodynamic subsystems have identical thermodynamic forces, and its thermodynamic currents are additive values. The dissipation potentials are given for each current-tube.

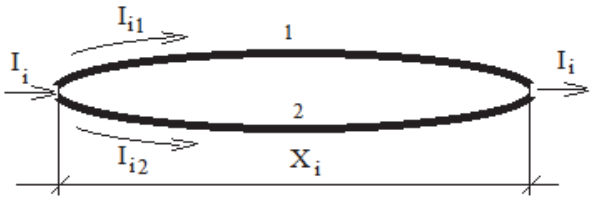


Fig. 6. Parallel connection of the current-tubes.

Because the thermodynamic forces are identical in the current tubes of the subsystems, so

$$\frac{\partial \Phi_2\left(I_{j2}\right)}{\partial I_{i2}} = -\frac{\partial \Phi_2\left(I_j - I_{j1}\right)}{\partial I_{i1}}, \left(i, j = 1,2,3,...,N\right) \tag{268}$$

Due to the fixed sum of the currents, the dissipation potentials depend on the thermodynamic current of one of subsystems only; let us say this is the 1st one. Using this, we get

$$\frac{\partial\left[\Phi_1\left(I_{j1}\right) + \Phi_2\left(I_j - I_{j1}\right)\right]}{\partial I_{j1}} = 0, \left(j = 1,2,3,...,N\right) \tag{269}$$

Substituting these to (268) the result is

$$I_{i2} = I_i - I_{i1}, \quad (i = 1,2,3,...,N) \tag{270}$$

The currents belonging to the stationer state of the 1st subsystem are the solutions of the above (270) generally non-linear system of N equations and with them the currents of the 2nd subsystem are

$$\begin{aligned} &\Phi_1\left(I_{j1}\right) + \Phi_2\left(I_{j2}\right) = \min, \\ &I_{j1} + I_{j2} = I_j \end{aligned} \tag{271}$$

This result shows how the subsystems share the thermodynamic currents. What is the situation with the extremum property of GEP? The equations to solve this problem could be derived from the following constrained extremum-task

$$P = I_i X_i = I_{i1} X_i + I_{i2} X_i = P_1 + P_2 \tag{272}$$

The global entropy production of the system using the additivity of currents is

$$\begin{aligned} P_l &= I_{il} X_i = X_i \frac{\partial \Psi_l}{\partial X_i} = k_{\Psi l} \Psi_l \left(X_j \right) = \\ &= \frac{k_{\Psi l}}{k_{\Psi l} - 1} \Phi_l \left(I_{jl} \right) = k_{\Phi l} \Phi_l \left(I_{jl} \right), \left(l = 1,2 \right), \left(i, j = 1,2,3,...,N \right) \end{aligned} \tag{273}$$

Using the equations expressing the connection between the entropy production and dissipation potentials

$$P = k_{\Phi 1} \Phi_1 \left(I_{j1} \right) + k_{\Phi 2} \Phi_2 \left(I_{j2} \right) \tag{274}$$

and so the GEP of the system is

$$P = k_{\Phi 1} \Phi_1 \left(I_{j1} \right) + k_{\Phi 2} \Phi_2 \left(I_{j2} \right) \tag{275}$$

Fix all the thermodynamic currents for the system. Due to the sums of the currents are prescribed, the entropy production depends on thermodynamic currents of one of subsystems only. Choose the 1st one. Hence

$$P(I_{j1}) = k_{\Phi 1} \Phi_1 \left(I_{j1} \right) + k_{\Phi 2} \Phi_2 \left(I_j - I_{j1} \right) \tag{276}$$

From this we get the following result

$$\begin{aligned} \frac{\partial P}{\partial I_{j1}} &= k_{\Phi 1} \frac{\partial \Phi_1 \left(I_{j1} \right)}{\partial I_{j1}} + k_{\Phi 2} \frac{\Phi_2 \left(I_j - I_{j1} \right)}{\partial I_{j1}} = \\ &= k_{\Phi 1} \frac{\partial \left[\Phi_1 \left(I_{j1} \right) + \Phi_2 \left(I_j - I_{j1} \right) \right]}{\partial I_{j1}} + \left(k_{\Phi 2} - k_{\Phi 1} \right) \frac{\Phi_2 \left(I_j - I_{j1} \right)}{\partial I_{j1}} = \\ &= \left(k_{\Phi 2} - k_{\Phi 1} \right) \frac{\Phi_2 \left(I_j - I_{j1} \right)}{\partial I_{j1}}, \left(j = 1,2,3,...,N \right) \end{aligned} \tag{277}$$

Where we used the equations (270).

These calculations have two important consequences:

1. In the parallel connection of the current tubes the sum of the current potentials will be minimal in the stationer state of the system.
2. The GEP of the system in stationer case could be minimal if and only if when the dissipation potentials of the subsystems are homogeneous Euler's functions with identical degree.

The behavior of the sum of the dissipation potentials and the minimum property of GEP in generalized circuits are given in Appendix 7.5.

4.8 Example

Simple DC networks. The DC networks are typical steady-state systems where the minimum principle of global entropy production is valid. It is originated from two sources:

- The source-free stationer balances are valid for the fluxes (electric currents) expressed by the first Kirchhoff's law of.
- The thermodynamic forces could be originated from the gradients of the electric potential by line-integral, expressed by the Kirchhoff's second law.

However, further conditions are necessary also for the application of the constitutive theory.

- When the Ohm law is valid for the dissipative elements, then the dissipation potentials are Onsager's like.
- In case of non-linearity the dissipation potentials have to fulfill the requests of the generalized Onsager constitutive theory.
- Only two simple arrays will be discussed below: the serial and parallel DC-circuits in the frame of linear Onsager's theory.

4.8.1 Potential distribution in serial circuit

Our task is to determine the potential distribution in the serial circuit shown in Fig. 7., where the electric potentials are fixed in the N_1, N_2 points of the boundary so the sum of the thermodynamic forces is also fixed.

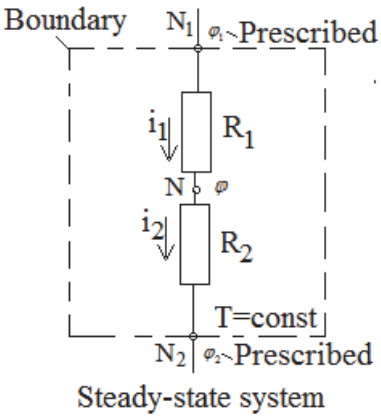


Fig. 7. Simple serial DC-network, for what the minimum principle of global entropy production is valid.

Supposing that the system is isotherm, the dissipation potentials of the subsystems are

$$\Psi_1(\varphi) = \frac{1}{2T}\frac{(\varphi_1 - \varphi)^2}{R_1}, \Psi_2(\varphi) = \frac{1}{2}\frac{(\varphi - \varphi_2)^2}{R_2} \tag{278}$$

Knowing from the previous examples; the sum of the force potentials of the system will be the minimum in stationer state

$$\Psi(\varphi)=\Psi_1(\varphi)+\Psi_2(\varphi)=\frac{1}{2T}\frac{(\varphi_1-\varphi)^2}{R_1}+\frac{1}{2T}\frac{(\varphi-\varphi_2)^2}{R_2} \tag{279}$$

Hence the necessary condition for the minimum is

$$\frac{dP(\varphi)}{d\varphi}=\frac{2}{T}\left(\frac{(\varphi-\varphi_2)}{R_2}-\frac{(\varphi_1-\varphi)}{R_1}\right)=0 \tag{280}$$

We can eliminate the potential φ with the constraint on the electric potential differences like: $(\varphi_1-\varphi)+(\varphi-\varphi_2)=(\varphi_1-\varphi_2)$. Consequently, the potential distribution on the resistors will be

$$u_1:=\varphi_1-\varphi=(\varphi_1-\varphi_2)\frac{R_1}{R_1+R_2},\ u_2:=\varphi-\varphi_2=(\varphi_1-\varphi_2)\frac{R_2}{R_1+R_2} \tag{281}$$

In this state the GEP of the system is minimal, because both the dissipation potentials are homogeneous Euler's functions of degree 2.

4.8.2 Current distribution in parallel circuit

Supposing again, that the system shown in Fig. 8. is isotherm, and on its boundary the current is fixed.

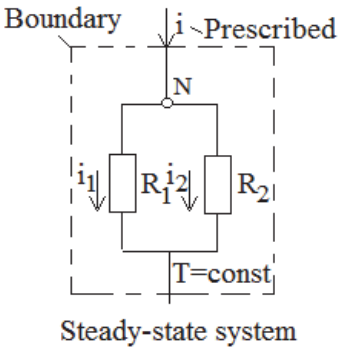


Fig. 8. Simple parallel DC-network, for what the minimum principle of global entropy production is valid.

The current potentials of the system are

$$\Phi_1(i_1)=\frac{1}{2T}i_1^2R_1,\ \Phi_2(i_2)=\frac{1}{2T}i_2^2R_2 \tag{282}$$

And so from the previous example the sum of the current potentials of the system will be the minimum in stationer case

$$\Phi(i_1) = \Phi_1(i_1) + \Phi_2(i - i_1) = \frac{i_1^2 R_1 + (i - i_1)^2 R_2}{2T} \tag{283}$$

From this the current distribution will be determined by the necessary condition of the minimum

$$\frac{d\Phi(i_1)}{di_1} = 0, \rightarrow i_1 R_1 - (i - i_1) R_2 = 0, \rightarrow i_1 = i\frac{R_2}{R_1 + R_2}, i_2 = i\frac{R_1}{R_1 + R_2} \tag{284}$$

Due to both dissipation potentials are homogeneous Euler's functions of degree 2, the GEP of the system is minimal in this state.

4.9 Exercises

4.9.1. Show in case of the current- tube, that $X_i := \Gamma_i^{(1)} - \Gamma_i^{(2)} = r_{ik} I_k$, $r_{ik} = \int_1^2 \frac{R_{ik}}{A(s)} ds$, when $\mathbf{X}_i = R_{ik}\mathbf{j}_k$, $\mathbf{X}_i = \nabla\Gamma_i$, and $A(s)$ is the surface of the cross-section of the tube at the *s* arc-length and *ds* is the arc-element of the axis of the tube.

4.9.2. At which resistance value of R_2 will have the power its maximal value on it, when it is a consumptive element of the circle in Figure7.? Calculate the efficacy as well.

4.9.3. At which resistance value of R_2 will be the power maximal, when it is a consumptive element of the circle in Figure 8.? Calculate the efficacy as well.

4.9.4. Calculate the potential-drop on a given resistivity in an arbitrary DC circuit. Prove that this task could be solved by the minimization of the sum of dissipation potentials or by the principle of minimal entropy production, like we did in the serial connection of the Figure 5., when the circuit regarding the contacts of the resistor is replaced with the Thevenin's potential source equivalent circuit [53].

4.9.5. Calculate the current-flow on a given resistivity on an arbitrary DC circuit. Prove that this task could be solved by the minimization of the sum of dissipation fluxes or by the principle of minimal entropy production, like we did in the case of parallel connection of the Figure 6., when the circuit regarding the contacts of the resistor is replaced with the Norton's current source equivalent circuit [53].

4.9.6. Calculate the results in the above two exercises, when the given resistor is non-linear. Which part of the above results remains correct?

5. Dissipative energy converters; characteristic functions

This topic is discussed only from the point of view of extremum principles. The couplings among the thermodynamic fluxes in standard and generalized Onsager constitutive theories make it possible that the energy could have numerous carriers simultaneously in its degradation process in a dissipative structure; and by this the various processes become energetically coupled, and the structure could work like an energy converter. Such converters in the biology for example are the muscle processes, the photochemical reactions in photosynthesis, the chemical reactions for ATP synthesis. Based on the constitutive theory, in the following we describe the general characterization of the energy converters. Based on the constitutive theory we show that the efficacy of the energy conversion depends on the ratio of the fluxes or the forces in a system which has two carriers only.

A dissipative system could be energy converter, when not all the terms of the entropy production are positive, their sum is positive definite only. The problem will be studied in a dissipative system having two thermodynamic forces and currents and in case when the dissipation potentials are homogeneous Euler's function of degree *k*. In this case the currents and forces are connected with the following constitutive equations (see Appendix 7.6.)

$$\begin{aligned} I_1 &= L_{11}X_1 + L_{12}X_2 \\ I_2 &= L_{21}X_1 + L_{22}X_2, \quad L_{12} = L_{21} \end{aligned} \tag{285}$$

where the conduction coefficients are homogeneous Euler's functions of degree *k-2* of the forces. The entropy production is always positive, except in equilibrium, when the currents and forces have equally zero values. In this case the GEP can be presented as

$$\begin{aligned} P &= X_1 I_1 + X_2 I_2 = L_{11}X_1^2 + 2L_{12}X_1X_2 + L_{22}X_2^2 = \\ &= \left(\sqrt{L_{11}}X_1 + \frac{L_{12}}{\sqrt{L_{11}}}X_2\right)^2 + (L_{22} - \frac{L_{12}^2}{L_{11}})X_2^2 \geq 0 \end{aligned} \tag{286}$$

and thus

$$L_{11} > 0, \quad L_{22} > 0, \quad L_{12}^2 < L_{11}L_{22}, \textit{ for all } X_1, X_2 \tag{287}$$

Hence, the straight coefficients are positive, but the cross-coefficient may be both, positive or negative. This is in agreement with the experimental observation that coefficients like thermal or electrical conductivity are always positive while, for example thermo-electric conductivity coefficient is not definite. Supposing the L_{12} cross-coefficient negative and the driver process have the index 2. When the process is characterized by the I_2 current and by X_2 force, and the temperature T of the system is constant (isotherm process); then this process takes $P_{in} := T X_2 I_2 > 0$ power from an energy source, while transforms from this a power of $T L_{12} X_1 X_2 < 0$ value to the driven process denoted by 1. This process will dissipate $T L_{11} X_1^2$ power, and the remaining $P_{out} := T X_2 I_2 < 0$ power could be used as source for other processes, it could be a supply of outside systems. The efficacy of the energy converter is defined in the standard manner

$$\eta := \frac{P_{out}}{P_{in}} = \frac{-T X_1 I_1}{T X_2 I_2} \tag{288}$$

Introducing the entropy production and the dissipation potential into (286), then we get

$$\eta = \frac{-X_1 I_1}{X_2 I_2} = 1 - \frac{P}{X_2 I_2} = 1 - \frac{k\Psi(X_1, X_2)}{X_2 I_2} \tag{289}$$

The efficacy is maximal at constant primer power when the GEP is minimal in stationer state. However, this cannot be fulfilled in general, as we will show it later. Let us define the force ratio and coupling parameter to approach the classical discussion of Kedem and Caplan [54] in linear constitutive cases

$$x = \sqrt{\frac{L_{11}}{L_{22}}}\frac{X_1}{X_2},\ q := \frac{L_{12}}{\sqrt{L_{11}L_{22}}} < 1,\ x \in [-1,0] \tag{290}$$

Due to the conduction coefficients are homogeneous Euler's function of degree *k-2* of the forces, we may request the universal dependence on the conduction coefficients as well, which are

$$L_{ik} = L_{ik}^0 g(X_1, X_2) \tag{291}$$

The force ratio in this case linear function of the force ratio, and the coupling parameter is constant like Kedem-Caplan's theory. With the new denotations the (285) currents are

$$\begin{aligned} I_1 &= \sqrt{L_{11}L_{22}}X_2(x+q) \\ I_2 &= L_{22}X_2(qx+1) \end{aligned} \tag{292}$$

And so the efficacy of the converter is

$$\eta(x,q) = \frac{-X_1 I_1}{X_2 I_2} = \frac{x(x+q)}{qx+1} \tag{293}$$

Consequently, the efficacy depends on the force ratio and the coupling parameter only. The entropy production or the dissipation function (because our system is isotherm) is

$$\Theta(x,q) := TP = T(X_1 I_1 + X_2 I_2) = TL_{22}X_2^2(x^2 + 2qx + 1) \tag{294}$$

This is another function of force ratio and the coupling parameter, so we do not expect the extremes in the two functions at the same value of the force ratio. This is one of the reasons some experts discuss, that not all biological systems maximize their efficacy, instead they try to optimize other "cost-function" depending its complexity and its connection on the surroundings [55]. In the following we shall try to summarize these functions and their roles. Additionally to the above mentioned two characteristic functions, the isotherm *Power output function (P_{out})* and the *Ecological function* (E_c) are used. Their definitions are [55].

$$\begin{aligned} P_{out}(x,q) &:= -T X_1 I_1 = -TL_{22}X_2^2(x+q)x, \\ E_c(x,q) &:= P_{out} - \Theta = -TL_{22}X_2^2(2x^2 + 3qx + 1) \end{aligned} \tag{295}$$

These functions are universal functions of the force ratio and coupling parameter refer to reference power $TL_{22}X_2^2$. We are going to show in the following, that the reference power does not depend on the force ratio in the linear Onsager constitutive theory, while it depends in the generalized Onsager constitutive theory. Using the (291) and the representation principle of homogeneous Euler's functions (see Appendix 7.7), we get

$$L_{22} = L_{22}^0 g(X_1, X_2) = L_{22}^0 \left(\sqrt{\frac{L_{22}^0}{L_{11}^0}}\right)^{k-2} \left(\sqrt{\frac{L_{11}^0}{L_{22}^0}}\right)^{k-2} g(X_1, X_2) =$$

$$
\begin{aligned}
&= L_{22}^0 \left(\sqrt{\frac{L_{22}^0}{L_{11}^0}} \right)^{k-2} g(\sqrt{\frac{L_{11}^0}{L_{22}^0}} X_1, \sqrt{\frac{L_{11}^0}{L_{22}^0}} X_2) = \\
&= L_{22}^0 \left(\sqrt{\frac{L_{22}^0}{L_{11}^0}} \right)^{k-2} X_2^{k-2} g(\sqrt{\frac{L_{11}^0}{L_{22}^0}} \frac{X_1}{X_2}, \sqrt{\frac{L_{11}^0}{L_{22}^0}}) = \\
&= L_{22}^0 \left(\sqrt{\frac{L_{22}^0}{L_{11}^0}} \right)^{k-2} X_2^{k-2} g(x, \sqrt{\frac{L_{11}^0}{L_{22}^0}})
\end{aligned}
\tag{296}
$$

where we used that $g(X_1, X_2)$ is a homogeneous Euler's function of degree *k-2*. So the reference power is

$$
T L_{22} X_2^2 = T L_{22}^0 \left(\sqrt{\frac{L_{22}^0}{L_{11}^0}} \right)^{k-2} X_2^k g(x, \sqrt{\frac{L_{11}^0}{L_{22}^0}}) \tag{297}
$$

In the Onsager's linear constitutive theory $k = 2$, so $g(X_1, X_2)$ does not depend on *x*, also it is constant, which we choose in 1. In this case we receive the classical equation of the energy converters in which the efficacy does not depend on the force ratio, so the characteristic functions could be scaled by them without distortion. In general the reference power depends on the force ratio, so it is not suitable for the distortion-free scaling till separating the term which depends on the force ratio. In the proposed dissuasion the characteristic functions are

$$
\begin{aligned}
&P_{out}(x,q,m) = -T L_{22}^0 \left(\sqrt{\frac{L_{22}^0}{L_{11}^0}} \right)^{k-2} X_2^k g(x, \sqrt{\frac{L_{11}^0}{L_{22}^0}})(x+q)x \\
&\Theta(x,q,m) = T L_{22}^0 \left(\sqrt{\frac{L_{22}^0}{L_{11}^0}} \right)^{k-2} X_2^k g(x, \sqrt{\frac{L_{11}^0}{L_{22}^0}})\left(x^2 + 2qx + 1\right) \\
&E_c(x,q,m) = P_{out} - \Theta = -T L_{22}^0 \left(\sqrt{\frac{L_{22}^0}{L_{11}^0}} \right)^{k-2} X_2^k g(x, \sqrt{\frac{L_{11}^0}{L_{22}^0}})\left(2x^2 + 3qx + 1\right), \\
&m := k - 2
\end{aligned}
\tag{298}
$$

where *m* is the degree of the *g*-function. From physical point of view it is trivial, that *g* is positive homogeneous function of force ratio, which anyway has a negative value. Due to the degree *m* is not necessarily integer, it could be a fraction, so we may assume that the *g* depends on the power-function of the absolute value of the force ratio *x* by

$$
g(x, \sqrt{\frac{L_{11}^0}{L_{22}^0}}) \propto |x|^m \tag{299}
$$

With this the characteristic functions by the new scaling concept are

$$
\begin{aligned}
P^*_{out}(x,q,m) &= -x(x+q)|x|^m \\
\Theta^*(x,q,m) &= \left(x^2+2qx+1\right)|x|^m \\
E_c{}^*(x,q,m) &= -\left(2x^2+3qx+1\right)|x|^m \\
m &:= k-2
\end{aligned}
\tag{300}
$$

In the case of the classical Onsager constitutive theory the maxima of the characteristic functions do not appear at the same value of the force ratio as it was shown by Arias–Hernandez et al. in details [55]. Also, in this case the very simple explanation of the evolutionary selection would be that a biologic system evolves to the direction of the maximal efficacy together with the minimal dissipation is not fulfilled. However, in a work about the mitochondrial ATP oxidative phosphorysation Stucky [56] shows the way, which is in harmony with the above simple principle of evolutionary selection. Stucky assumed that mitochondrial ATP oxidative phosphorysation is in steady-state corresponding to minimum of GEP which simultaneously corresponds to a maximum efficiency state. To satisfy these conditions, Stucky introduces a third term into the GEP, corresponding to the attached cellular load, to where he fits a maximal output power. We have to study how we can satisfy the principle of evolutionary selection in generalized Onsager constitutive theory without extending of the GEP. Simple computing experiments can show the advantage of the new scaling concept proposed in (298), which "gives possibility" for evolution movements optimizing all the characteristic functions with respect to the force ratio at maximal efficacy by the proper choice of the power *m*. This is shown by the Fig. 9. in the case of two different coupling parameters.

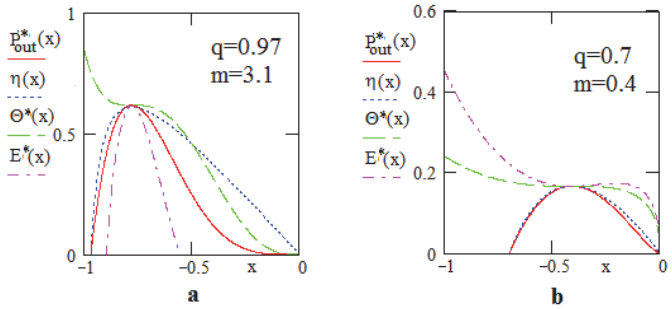


Fig. 9. The characteristic functions of the non-linear energy converters. P is the output power, η is the efficacy, Θ is the dissipation and E is the ecologic function (the difference between the output power and the dissipation), q is the coupling parameter, and m is the power of the universal Euler's function in (297). These functions are not given on a common scale, they are on certain self-scales to show their behavior near the optimum state.

When *m=0*, the theory is linear, the classical theory is valid. This is the classical Onsager constitutive theory, which allows the quasi-linearity as well, when the coefficients could depend on the intensives, but not on their gradients. In this theory the minimum principle of global entropy production is not in harmony with the maximal efficacy, because the functions are independent. However, in non-linear constitutive theory the characteristic functions, as we see in Figure 9., are not in conflict with respect of simultaneous optimum.

The Figure 9. shows, that in case of small coupling parameter a less non-linear system could be in a state of simultaneous optimum, and if the coupling parameter is large the simultaneous optimum comes only when the system is far away form the linearity. Finally, we note that this result is in agreement with the experimental studies of Nath [57].

6. On the principle of maximum entropy production

There have been successful applications of the so called principle of maximum entropy production [58] for meteorological models, biological, environmental and ecological problems in the recent literature. H.Ziegler was the first who formulated this principle phenomenologically [59], substituting with this the Onsager's last dissipation of energy, which was proven in that time only for linear constitutive equations. The intention of Ziegler was to extend the non-equilibrium thermodynamical descriptions to the systems with non-linear constitutive equations. The validity of this maximum principle was proven by Zupanovi'c et.al for DC networks with well generalized conditions [60]. They applied this maximum entropy production principle in their model for bacterial photosynthesis for building the non-linear constitutive equation into their theory, [61].Next, we shall discuss the maximum entropy production principle from the aspect of phenomenological thermodynamics only, and refer to its other applications, see for example [62]. The maximum entropy production principle does not contradict the principle of Onsager's last dissipation of energy in linear cases; on the contrary, they have equivalent value, [63], [64]. Some researchers state this equivalence as general (e.g. [65]), regarding two different, but identical formulations of the same natural principle. The equivalence is debated, and has not been proved till now. We would like to make steps forward the solution, without the final solution of the question; namely we show below the equivalence of the Ziegler's principle with the generalized Onsager's last dissipation of energy principle when the dissipation potentials are homogeneous Euler's functions. Like we did in the Appendix 7.2, let us start from the introduction of the force potential with the minimum task (38) of Legendre transform. Namely, using the definition (38), we get

$$\Psi(X_i) = \max_{J_i}\left[X_i \cdot J_i - \Phi(J_i) \right] \tag{301}$$

Based on this, the original Onsager's last dissipation of energy principle ensures the maximum of the function of the force potential, when the flux potential is homogeneous Euler's function of degree two. Due to the fact that (301) is valid in case of an arbitrary strictly convex flux dissipation potential, the generalized Onsager's last dissipation of energy principle could be formulated as

$$\max_{J_i}\left[X_i \cdot J_i - \Phi(J_i) \right] \tag{302}$$

This is a differential extremum principle. Let us prove the Ziegler's maximum entropy production principle, finding the possible mathematical form of the

$$X_i = X_i(J_1, J_2, ..., J_j, ..., J_n), (i, j = 1, 2, 3, ..., n) \tag{303}$$

constitutive equations. If we substitute these equations into the entropy production, we get

$$\sigma(J_i) = J_i X_i = J_i X_i(J_j) \tag{304}$$

which is a function defined on the linear vector space of fluxes $\mathbf{J} = (J_1, J_2, ..., J_j, ..., J_n)$.
Ziegler proposed the following principle to find the possible form of the constitutive equations (303).
Ziegler's principle. If the forces X_i are prescribed, the actual fluxes J_i, (which are satisfied the entropy production $\sigma(J_i) = J_i X_i$), maximize the function of entropy production (304), when the possible form of constitutive equations (303) are fulfilled as well. This principle can be formulated as a constrained maximum task, i.e.

$$\begin{aligned} &\sigma(J_i) = J_i X_i(J_j) = extr, \\ &\sigma(J_i) = J_i X_i \end{aligned} \tag{305}$$

By the Lagrange's theorem of multipliers can be constructed a free one

$$\sigma(J_i) - \mu\left(\sigma(J_i) - J_i X_i\right) = extr \tag{306}$$

Necessary conditions belonging to this free extremum task are

$$\begin{aligned} &\frac{\partial\left[\sigma(J_j) - j\left(\sigma(J_i) - J_i X_i\right)\right]}{\partial J_j} = 0, (i, j = 1,2,3...,n) \\ &\frac{\partial\left[\sigma(J_j) - \mu\left(\sigma(J_j) - J_i X_i\right)\right]}{\partial \mu} = 0 \end{aligned} \tag{307}$$

From these equations we obtain the possible form of the constitutive equations and the multiplier

$$\begin{aligned} &X_i = \lambda \frac{\partial \sigma(J_j)}{\partial J_i} (i, j = 1,2,3...,n) \\ &\lambda = \sigma(J_j)\left(J_i \frac{\partial \sigma(J_j)}{\partial J_i}\right)^{-1}, \lambda := \frac{\mu - 1}{\mu} \end{aligned} \tag{308}$$

Assume that the entropy production function two times continuously differentiable function with respect to fluxes. Then from the previous first equations and the Young-theorem of the analysis we get the generalized reciprocal relations

$$\frac{\partial X_i}{\partial J_k} = \frac{\partial X_k}{\partial J_i} = \lambda \frac{\partial^2 \sigma(J_j)}{\partial J_k \partial J_i} = \lambda \frac{\partial^2 \sigma(J_j)}{\partial J_i \partial J_k}, (i, j, k = 1,2,3,...,n) \tag{309}$$

Also the entropy production as function of the fluxes shows a property like flux dissipation potential. Because of this we could choose the flux potential as

$$\Phi(J_j) := \lambda \sigma(J_j) \tag{310}$$

From the first equations of (308) the entropy production function is satisfied the following functional equation

$$\sigma(J_j) = J_i X_i(J_j) = \lambda J_i \frac{\partial \sigma(J_j)}{\partial J_i} \tag{311}$$

From this take into account Appendix 7.7., follows that the entropy production, as function of fluxes, is homogeneous Eulers's function of degree λ. Because of this property for the flux dissipation potential we obtain

$$J_i \frac{\partial \Phi(J_j)}{\partial J_i} = J_i X_i(J_j) = \lambda^{-1} \Phi(J_j) \tag{312}$$

Thus the flux potential is homogeneous Euler's function of degree λ^{-1} as well. The flux potential in (301) is strict convex, thus $0 < \lambda < 1, \mu > 0$ in (308). Substitute the flux potential to the Ziegler's principle (306) then after some algebra we obtain the generalized Onsager's last dissipation of energy principle (302), i.e.

$$\begin{aligned} &\sigma(J_i) - \mu\big(\sigma(J_i) - J_i X_i\big) = \max = \\ &= \frac{\Phi(J_i)}{\lambda} - \mu\left(\frac{\Phi(J_i)}{\lambda} - J_i X_i\right) = \mu\big[J_i X_i - \Phi(J_i)\big] \rightarrow J_i X_i - \Phi(J_i) = \max \end{aligned} \tag{313}$$

Also the Ziegler's principle is resulted in the same class of Onsager constitutive theory as was resulted by the minimum principle of GEP and in case of this class the generalized Onsager's last dissipation of energy principle and the (306) form of the Ziegler's principle is equivalent.

7. Appendixes

7.1 The non-existence theorem of variational principle of a steady-state thermodynamic system

The non-existence theorem of Gage-Schiffer-Kline-Reynolds [66] had been neglected in the literature of irreversible thermodynamics like many other "non-convenient" complicated tasks, and on what Richardson focused again [67]. Formulate the problem with Richardson's original wordings:

They demonstrated [66] the impossibility of finding a universal extremal functional of the form (in one dimension),

$$\Theta = \int_A^B \Phi\left(L_{ik}, \Gamma_i, \frac{d\Gamma_i}{dx}\right) dx \tag{314}$$

where Φ is a specific function of its arguments, such as

$$\delta\Theta = 0 \tag{315}$$

is equivalent to the steady-state conservation equation

$$\frac{dj_i}{dx} = \frac{d}{dx} L_{ik} \frac{d\Gamma_k}{dx} = 0 \tag{316}$$

Standard notation is used; x denotes physical extent".
We note, the "non-existence theorem" is formulated for quasi-linear Onsager's constitutive theory, in which the L_{ik} conduction coefficients could depend on the intensive parameters. Jaynes has the same conclusion in his discussion on the minimum principle of global entropy production [46]. However, the theory presented above allows such extremal functional which is free variation principle by the thermodynamic forces and, however restricted by the fluxes; and in which the Euler-Lagrange equations belonging to the variation of intensive parameters are equivalent with the stationer transport -equations. We may go further, construct from this partially restricted variational-task a free one by Lagrange's method of multipliers. The solution of the problem is available by the Donkin's theorem, for what firstly Gyarmati had focused on [68], and he was the first who recognized [2], that for the solution the variations must be done in the unified vector space of the intensives, forces and fluxes.

7.2 Principle of last dissipation of energy

The original Onsager's principle of last dissipation of energy is derivable by the presentation of the force potential with the minimum task (38) of Legendre's transform. Namely, using the definition (38) which was actualized on vector -processes, we get:

$$\Psi(\nabla\Gamma_i, \Gamma_l) = \max_{\mathbf{J}\in G} \left[\nabla\Gamma_i \cdot \mathbf{j}_i - \Phi\left(\mathbf{j}_i, \Gamma_l\right) \right] \tag{317}$$

The functional produced with the above differential principle of (317) is

$$\begin{aligned} &\int_V L\left(\nabla\Gamma_i, \mathbf{j}_i, \Gamma\right) dV, \\ &L\left(\nabla\Gamma_i, \mathbf{j}_i, \Gamma\right) := \nabla\Gamma_i \cdot \mathbf{j}_i - \Phi\left(\mathbf{j}_i, \Gamma_l\right) \end{aligned} \tag{318}$$

Due to the Onsager's principle of last dissipation of energy this functional is maximal with respect to variations of the fluxes $\mathbf{j}_i$. The necessary condition of the maximum are the constitutive equations, which are presented as Euler-Lagrange equations:

$$\frac{\partial L}{\partial \mathbf{j}_i} = \nabla\Gamma_i - \frac{\partial \Phi\left(\mathbf{j}_i, \Gamma_l\right)}{\partial \mathbf{j}_i} = 0 \tag{319}$$

The original Onsager's idea declared this as a variation principle of the constitutive theory.

7.3 Second variation

Calculating the second variation of the sum of the volume-integrated dissipation potentials, we get the following terms:

$$\left\langle \left. \frac{\partial^2 \Phi}{\partial\Gamma_j \partial \mathbf{j}_i} \right|_{\Gamma_{k0}, \mathbf{j}_{j0}} \delta\Gamma_j, \delta\mathbf{j}_i \right\rangle, \quad \left\langle \left. \frac{\partial^2 \Psi}{\partial\Gamma_j \partial \nabla\Gamma_i} \right|_{\Gamma_{k0}, \mathbf{j}_{j0}} \delta\Gamma_j, \delta\nabla\Gamma_i \right\rangle \tag{320}$$

Because Φ and Ψ are two times continuously derivable functions we obtain:

$$\frac{\partial^2 \Phi}{\partial \Gamma_j \partial \mathbf{j}_i} = \frac{\partial^2 \Phi}{\partial \mathbf{j}_i \partial \Gamma_j} = \frac{\partial}{\partial \mathbf{j}_i}\frac{\partial \Phi}{\partial \Gamma_j} = \frac{\partial}{\partial \mathbf{j}_i}\left(-\frac{\partial \Psi}{\partial \Gamma_j}\right) = 0 \tag{321}$$

where we had taken into account the Donkin's theorem (36). Also, we can see that the terms (320) do not give contributions to the second variation of GEP.

7.4 Proof of the equivalence the local and global minimum of a strictly convex potential

Proceed from the potential $\Psi(\mathbf{X})$, where $\mathbf{X}$ is a super-vector constructed by the thermodynamic forces as coordinates. When the $\mathbf{X}_0$ is a local optimum of $\Psi(\mathbf{X})$, we have $\Psi(\mathbf{Z}) \geq \Psi(\mathbf{X}_0)$ for any $\mathbf{Z}$ in some neighborhood $U(\mathbf{X})$ of $\mathbf{X}$. For any $\mathbf{Y}$ $\mathbf{Z} = \lambda\mathbf{X} + (1-\lambda)\mathbf{Y}$ belongs to $U(\mathbf{X})$ for $\lambda < 1$ sufficiently close to 1, and it follows from strictly convexity $\Psi(\mathbf{X})$ of that. From this after some algebra we get

$$\lambda\Psi(\mathbf{X}) + (1-\lambda)\Psi(\mathbf{Y}) \geq \Psi\left(\lambda\mathbf{X} + (1-\lambda)\mathbf{Y}\right) = \Psi(\mathbf{Z}) \geq \Psi(\mathbf{X}) \tag{322}$$

7.5 Current tube n-port as steady-state thermodynamic system

Until now we studied such current-tube networks, where every current or every force was common. This limit is abandoned, introducing the current-tube n-port, as thermodynamic system. This system is symbolically shown in Fig. 10.a., where the positive measuring directions of the forces and currents are also indicated. The request of source-free currents is satisfied by assuming loop-currents for each input two-poles. This condition means for example in the i-th two-pole, that the current flowing in the upper connection, flows out through the lower one. In the following we shall study a thermodynamic system, which is created by the connection of an *n*-port and an *n-k*-port. Fix the first *k* forces of the *n*-ports by the force sources as usual in the network-theory [69] (see Fig. 10.b.).

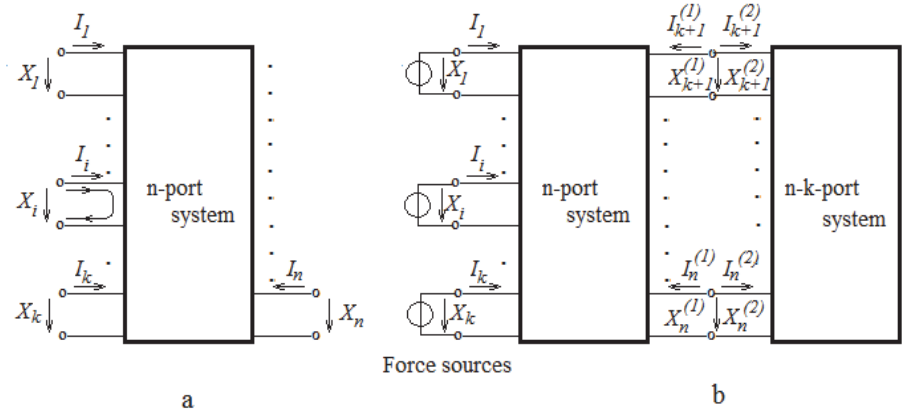


Fig. 10. Current tube n-port as steady-state thermodynamic system. Connection of an n-port and a n-k-port.

Give the force-potential for both the sub-systems. In this case the connection of the free-forces and currents can be formulated by the following non-linear equation

$$\begin{aligned} I_j^{(1)} &= \frac{\partial \Psi^{(1)}\left(X_1,.,X_i,.,X_k,X_{k+1},...,X_l,...,X_n\right)}{\partial X_j} = \\ &= -I_j^{(2)} = \frac{\partial \Psi^{(2)}\left(X_{k+1},...,X_l,...,X_n\right)}{\partial X_j}, \left(j = k+1,...,n\right) \end{aligned} \tag{323}$$

Rearranging its order, we get:

$$\begin{gathered} \frac{\partial\left[\Psi^{(1)}\left(X_1,.,X_i,.,X_k,X_{k+1},...,X_l,...,X_n\right)+\Psi^{(2)}\left(X_{k+1},...,X_l,...,X_n\right)\right]}{\partial X_j} = 0, \\ \left(j = k+1,...,n\right) \end{gathered} \tag{324}$$

Hence the free forces could be determined by the following constrained extremum- task:

$$\begin{aligned} &\Psi^{(1)}\left(X_1,.,X_i,.,X_k,X_{k+1},...,X_l,...,X_n\right)+\Psi^{(2)}\left(X_{k+1},...,X_l,...,X_n\right) = \min \\ &X_i = prescribed\,, \left(i = 1,...,k\right) \end{aligned} \tag{325}$$

Using the following relation of the homogeneous Euler's function potentials and of the GEP of the sub-systems:

$$P^{(l)} = k_{\Psi_l}\Psi^{(l)}, \left(l = 1,2\right) \tag{326}$$

then we get for the total entropy production of the system

$$P = k_{\Psi_1}\Psi^{(1)} + k_{\Psi_2}\Psi^{(2)} \tag{327}$$

From this follows:

$$\begin{aligned} \frac{\partial P}{\partial X_j} &= k_{\Psi_1}\frac{\partial\left[\Psi^{(1)}+\Psi^{(2)}\right]}{\partial X_j} + \left(k_{\Psi_2}-k_{\Psi_1}\right)\frac{\partial \Psi^{(2)}}{\partial X_j} = \\ &= \left(k_{\Psi_2}-k_{\Psi_1}\right)\frac{\partial \Psi^{(2)}}{\partial X_j} \end{aligned} \tag{328}$$

Consequently, again we obtained the result that the sum of the dissipation potentials is minimal in stationer state and to the extremum of GEP is requested the identical degrees of Euler's function dissipation potentials. This means also, that using dissipation potentials has advantage in the thermodynamics of steady-state systems.

7.6 Notes about the Hessian matrix

A direct consequence of the possibility to represent the homogeneous Euler's functions by their partial derivative functions is that the thermodynamic currents and dissipation

potentials can be described by the linear and quadratic form of the Hessian matrices. For example, if the force potential is a homogeneous Euler's function of degree *k*, then:

$$\begin{aligned}
&\Psi\left(X_1,...,X_i,...,X_n\right)=\frac{1}{k}X_i\frac{\partial\Psi\left(X_1,...,X_i,...,X_n\right)}{\partial X_i},\\
&\frac{\partial\Psi\left(X_1,...,X_i,...,X_n\right)}{\partial X_i}=\frac{1}{k-1}X_j\frac{\partial^2\Psi\left(X_1,...,X_i,...,X_n\right)}{\partial X_j\partial X_i}=I_i,\\
&\Psi\left(X_1,...,X_i,...,X_n\right)=\frac{1}{k(k-1)}\frac{\partial^2\Psi\left(X_1,...,X_i,...,X_n\right)}{\partial X_j\partial X_i}X_iX_j
\end{aligned}\tag{329}$$

where we used that the partial derivative functions are one degree less homogeneous Euler's functions. When the dissipation potential has two variables only, then the currents:

$$\begin{aligned}
&I_1=\frac{\partial\Psi\left(X_1,X_2\right)}{\partial X_1}=\frac{1}{k-1}X_j\frac{\partial^2\Psi\left(X_1,X_2\right)}{\partial X_j\partial X_1}=\frac{1}{k-1}\left(\frac{\partial^2\Psi}{\partial X_1^2}X_1+\frac{\partial^2\Psi}{\partial X_2\partial X_1}X_2\right)=\\
&=L_{11}X_1+L_{22}X_2,\\
&I_2=\frac{\partial\Psi\left(X_1,X_2\right)}{\partial X_2}=\frac{1}{k-1}X_j\frac{\partial^2\Psi\left(X_1,X_2\right)}{\partial X_j\partial X_2}=\frac{1}{k-1}\left(\frac{\partial^2\Psi}{\partial X_1\partial X_2}X_1+\frac{\partial^2\Psi}{\partial X_2^2}X_2\right)=\\
&=L_{21}X_1+L_{22}X_2,\\
&L_{11}:=\frac{1}{k-1}\frac{\partial^2\Psi}{\partial X_1^2},\ L_{12}:=\frac{1}{k-1}\frac{\partial^2\Psi}{\partial X_2\partial X_1}=\frac{1}{k-1}\frac{\partial^2\Psi}{\partial X_1\partial X_2}=:L_{21},\ L_{22}:=\frac{1}{k-1}\frac{\partial^2\Psi}{\partial X_2^2}
\end{aligned}\tag{330}$$

where we introduced the conductivity coefficients, which are homogeneous Euler's functions of degree k-2.

7.7 Representation theorem of homogeneous Euler's function

The representation theorem is a simple consequence of the definition of the homogeneous Euler's functions:

$$\Psi\left(\lambda X_1,...,\lambda X_i,...,\lambda X_n\right)=\lambda^k\Psi\left(X_1,...,X_i,...,X_n\right)\tag{331}$$

when $\lambda=X_1^{-1}$, $X_1\neq 0$, the result is the representation theorem (for a rigorous proof see [70]) theorem:

$$\Psi\left(X_1,...,X_i,...,X_n\right)=X_1{}^k\Psi\left(1,...,\frac{X_i}{X_1},...,\frac{X_n}{X_1}\right)\tag{332}$$

The proof that the partial derivative functions are one degree lower homogeneous Euler's functions could be shown as below. When the function is:

$$I_i\left(X_1,...,X_i,...,X_n\right)=\frac{\partial\Psi\left(X_1,...,X_i,...,X_n\right)}{\partial X_i},\left(i=1,2,3,...,N\right)\tag{333}$$

then from this, by a few simple steps, we get the result

$$I_i\left(\lambda X_1,...,\lambda X_i,...,\lambda X_n\right)=\frac{\partial\Psi\left(\lambda X_1,...,\lambda X_i,...,\lambda X_n\right)}{\partial\lambda X_i}=\lambda^k\frac{\partial\Psi\left(X_1,...,X_i,...,X_n\right)}{\partial\lambda X_i}=$$
$$=\lambda^{k-1}\frac{\partial\Psi\left(X_1,...,X_i,...,X_n\right)}{\partial X_i}=\lambda^{k-1}I_i\left(\lambda X_1,...,\lambda X_i,...,\lambda X_n\right) \tag{334}$$

7.8 Lyapunov's theorem in the case of thermodynamic steady-state

In this Appendix we study the Lyapunov's theorem in case of thermodynamic steady-state according to considerations of Schlogl [47] and Lassale&Lefschetz [49].
When the changes of the steady-state thermodynamic system after the perturbations is the following system of differential equations:

$$\frac{dX_i}{dt}=f\left(X_1,..,X_j,..,X_N\right),\quad\left(i,j=1,2,3,...,N\right) \tag{335}$$

and $\Psi\left(X_j\right)$ is a real function which exists in the whole state space $\mathbf{X}=\left(X_1,..,X_j,..,X_N\right)$ and is positive definite, i.e.

$$\Psi\left(\mathbf{X}\right)\geq 0 \tag{336}$$

and the equality sign is valid at the steady state only. Moreover, let the second variation of $\Psi\left(\mathbf{X}\right)$ be positive

$$\delta^{(2)}\Psi\left(\mathbf{X}\right)=\frac{\partial^2\Psi}{\partial X_j\partial X_i}\delta X_i\delta X_j \tag{337}$$

in any state point , then we can state:
1. If exists a region in the state space which contains the steady-state point in its interior and if

$$\frac{d\Psi\left(\mathbf{X}\right)}{dt}\leq 0 \tag{338}$$

on the solution of differential equation (335), then the solution is stable whenever it starts in the interior of this region. The term *stable* means that the solution does not lead out of the limited neighborhood of the steady state point.
2. The solution is *asymptotically stable* if the equality sign in (338) is valid in the steady state point only. In this case the solution of the equation (335) leads finally into the steady state point.
In that case if (336) and (338) are valid in this region, then $\Psi\left(\mathbf{X}\right)$ is called a *Lyapunov function*.

8. Summary

We showed above that rigorous conditions are necessary for the extremum principles of non-equilibrium steady-state. The limiting conditions could be categorized into two classes.

The first one is the strong physical principles, expressed mathematically with balance equations. The other class is formed by the material or constitutive equations, describing the material behavior of the system. These equations however are always approximate. The possible general is only the reciprocal relation-principle, which was discovered on the linear constitutive equations by Onsager, and this is accepted in case of the non-linear processes as well. This was the way to the generalized Onsager constitutive theory. When this theory is applicable for the non-equilibrium steady-state system, then the Onsager's last dissipation of energy extremum principle could be generalized when the dissipation potentials derived from the reciprocal relations are strictly convex functions. We showed above that the minimum of entropy production could be derived from this generalized theory only when the balance equations are source-free. In this case however, no further conditions exist for dissipation potentials.

It is shown another way also to obtain the principle of minimal entropy production. Namely, we used the Glansdorff-Prigogine general evolution criterion, as Prigogine originally did. According to this principle the term of entropy production connected to the velocity of the thermodynamical forces monotonically decreases in time in non-equilibrium systems with time-independent boundary conditions, until the system reaches its equilibrium state , when that state exists. This is a strict principle because its validity does not depend on which constitutive theory governs the mentioned equations of the system. When the generalized Onsager constitutive theory in the point of view of the material equation of the given system belongs to such a sub-category where the dissipation potentials are homogeneous Euler's functions, then as it was shown above, Glansdorff-Prigogine general evolution criterion will be valid for all dissipation potentials and so for the entropy production too. Consequently, when steady-state exists, this belongs to the minimal entropy production. This is a generalization of the minimal entropy production principle, which was recognized for linear constitutive equations by Prigogine. It is clear that this principle is weaker than the original, but well useable for decision of the stability of the steady state. We showed above the existence of such steady-state systems where despite one of the global dissipative potentials is minimal, the global entropy production has no extremum properties.

At the end we proved the equivalence of the Ziegler's maximum entropy production principle with the generalized Onsager's last dissipation energy principle in cases when the dissipation potentials are homogeneous Euler's functions.

9. Chart of symbols

Symbolum	**Referencia**
$\nabla\Gamma_i$	i-th thermodynamic force (gradient of the intensive scalar Γ_i)
A_f	affinity
a_i	extensive state variable
B	magnetic flux density
B	metabolic rate of a single cell
C_{stac}	concentration in stationer state
D	death number rate of cells
$D(O)$	definition domain

$D_{ik}^{(\alpha)}$	tensors elements connect the chemical potential gradients and the heat-flux
$D_{ik}^{(\alpha\beta)}$	elements of diffusion tensor
$\mathbf{E}$	electric field-strength
E	metabolic energy for the production
E_c	ecological function
$G(\Delta t)$	autocorrelation function
$\mathbf{H}\Phi(\mathbf{J},\Gamma_l)$	Hessian matrix
$\mathbf{j}$	electric current density
$\mathbf{j}_i$	current density vector of the i-th extensive
$\mathbf{j}_q$	heat-current density vector
$\mathbf{j}_s$	entropy current denstiy
$\mathbf{j}_\alpha$	mass-current density of the α-th component
k	specific energy dissipation per unit time
K	born number rate of cells
$\mathrm{K}_{ik}^{(\alpha)}$	elements of heat conductivity tensor of α-th of chemical components
L	Lagrangian
$\mathbf{L}=[L_{ik}]$	conductivity tensor
L_{ik}	conduction coefficients
m	mass
m_i	mass of i-th component
N	number of the cells
NR	nourishment rate
$O(\nabla\Gamma_j,\mathbf{j}_j,\Gamma_k)$	global OM-function
$o(\mathbf{X},\mathbf{J},\Gamma_l)$	Onsager-Machlup function
P	global entropy production
P_{deg}	power of the irreversible processes
P_{out}	Power output function of disipative energy converter
$\mathbf{R}$	specific resistance tensor
r_0	mother-vessel with radius
r_1	daughter wessels
r_i	radius
$R_{ij}(\mathbf{B})$	resistivity tensor
R_{ik}	specific heat-resistivity tensor
$R_{\mathrm{jl}}(\Gamma_k)$	kinetic coefficients
s	entropy density
$S(f)$	spectral power density of the homesotatic noise
T	absolute temperature
U	internal energy
U_0	energy in the time of birth
u_1	potential distribution
U_∞	energy in stationer final state
V	volume
$V(t)$	homeostatic signal

W_{iot}	energy demand
$\mathbf{X}$	$3N$ dimensional thermodynamic force
$X_j=X_j(j_l,\Gamma_k)$	j-th cordinates of thermodynamic forces
a	ratio of radius
β	"colority" exponent of homeostatic noise
γ	affinity coefficient
Γ_i	i-th intensive scalar variable
δ_i	production term of the i-th extensive
δ_{ik}	Kronecker-delta
ε_{ijk}	permutation symbol
η	efficacy
Θ	dissipation function of dissipative energy converter
λ	special rate of cell-death
$\mathbf{\Lambda}(T)$	specific heat conduction tensor
$\lambda^{-1}=T$	average lifetime
σ	spontaneous internal entropy production
τ_{life}	life-span
φ	electric potential
$\Phi(j_l,\Gamma_k)$	flux dissipation potential
$\Phi_1(i_1)$	current dissipation potentials
$\Psi(\mathbf{X},\Gamma_l)$	force dissipation potential
Ω	boundary of the system

10. References

[1] Meixner J, Reik H (1959) Thermodynamik der irreversiblen Prozesse. Handbuch der Physik, III/2. Springer-Verlag Berlin-Göttingen-Heidelberg

[2] Gyarmati I (1970) Non-equilibrium Thermodynamics. Springer-Verlag Berlin-Göttingen-Heidelberg

[3] Onsager L (1931) Reciprocal Relation in irreversible processes I. Phys. Rev. 37:405-426, Reciprocal Relation in irreversible processes II. Phys. Rev. 37: 2265–2279

[4] Edelen DGB (1937) On the existence of symmetry relations and dissipations potentials. Arch. Ration. Mech. Analysis 51218

[5] Casimir HBG (1945) On Onsager's principle of microscopic reversibility. Rev. Mod. Phys. 17:343-350

[6] Truesdell CA (1960) Zu den Grundlagen der Mechanik und Thermodynamik. Physikalische Bletter 16:512-517

[7] Berdichevsky VL (2003) Structure of equations of macrophysics. Physical Rewiew E 68, 066126

[8] Kolmogorov AN (1936) Zur Umkehrbarkeit der statistiechen Naturgesetze, Math. Ann. 113, 766-772

[9] Yaglom AM (1962) An introduction to the theory of stationary random functions, Prentice-Hall, Englewood Cliffs, N. J.,

[10] Spohn H (1991) Large Scale Dynamics of Interacting Particles. Springer-Verlag, Berlin

[11] Berdichevsky V. (1997) Thermodynamics of Chaos and Order, Addison-Wesley-Longman,
[12] Binmore & Davies, (2007), Calculus Concepts and Methods, Cambridge University Press, p.190
[13] Rockefellar RT (1996) Convex Analysis. Princeton University Press
[14] Gantmacher F (1970) Lectures in Analytical Mechanics. Mir Publichers Moscow
[15] Arnold VI (1978) Mathematical Methods of Classical Mechanics. Springer-Verlag, New York
[16] Onsager L, Machlup S (1953) Fluctuations and irreversible Processes. Physical Rewiew 91(6):1505-1512
[17] Abramowitz, M. Stegun, IA. (1964), Handbook of Mathematical Functions, New York: Dover Publications
[18] Michlin SG (1962) Variationsmethoden der Mathematischen Physik. Akademie-Verlag-Berlin
[19] Becker R (1951) Theorie Der Elektrizitat, Band 1. Teubner Verlagsgesellschaft-Leipzig, (1982) Electromagnetic field and Intaractions. Dover Publications, New-York
[20] Murray CD (1926) The Physiological Principle of Minimum Work. I. The Vascular System and Cost of Blood Volume. PNAS 12 pp 207-214
[21] Murray CD (1926) The Physiological Principle of Minimum Work Applied to the Angle of Branching of Arteries. J. Gen. Physiol. 9:835-841
[22] Yunlong Huo, Kassab GS (2009) A Scaling Law of Vascular Volume. Biophysical Journal 96:347–353
[23] West GB, Brown JH, Enquist BJ (1997) A general model for the origin of allometric scaling laws in biology. Science. 276:122–126
[24] Finlyason BA, Scriven LC (1967) On the search for variational principles. Int. J. Heat Mass Transfer 10:799-821
[25] Lebon G, Jou D, Casas-Vazquez J (2008) Understanding Non-equilibrium Thermodynamics. Springer-Verlag-Berlin-Heidelberg
[26] Gyarmati I (1971) Application of the Gyarmati principle to multicomponent nonisothermal and chemically reacting diffusional systems. Acta chim. Hung. 67: 303–320
[27] Jeans JH (1923) The Mathematical Theory of Electricity and Magnetism. Cambridge University Press, Cambridge, England
[28] Glansdorff P, Prigogine I (1954) Sur les Sur les propriétés différentielles de la production d'entropie. Physica 20:773-773
[29] Glansdorff P, Prigogine I (1971) Thermodynamic Theory of Structure, Stability and Fluctuations. John Wiley & Sons, London, New York
[30] Bertalanffy L, (1968) General System theory: Foundations, Development, Applications, George Braziller, New York
[31] Bertalanffy L, (1962), Modern Theories of Development, New York: Harper
[32] Nicolis G, Prigogine I (1989) Exploring Complexity. Freeman, New-York
[33] Recordati G, Bellini TG (2004) Definition of internal constancy and homeostasis in the contex of non-equilibrium thermodynamics. Exp. Physiol. 89:27-38
[34] Struzik ZR, et al (2005) Dual Antagonistic Autonomic Control Necessary for 1/f Scaling in Heart Rate. In Fractals in Biology and Medicine Volume IV. Editors:G.A. Losa et al. Birkhauser Basel-Boston-Berlin

[35] Mandelbrot BB (1983) The Fractal Geometry of Nature. Freeman, San Francisco
[36] Badii B, Politi A (1997) Complexity, Hierarchical structures and scaling in physics. Cambrige University Press, UK
[37] Musha, T., Sawada, Y. (eds.) (1994). Physics of the living state. IOS Press, Amsterdam
[38] Rényi A, (1966) Wahrscheinlichkeitsrechnung. VEB Deutscer Verlag der Wissenschaften Berlin
[39] Rozanov YA, (1985) Probability Theory, A Concise Course. Dover Publications, Inc. New York
[40] Andresen B, Shiner JS, Uehlinger DE (2002) Allometric scaling and maximum efficiency in physiological eigen time. 99(9):5822–5824
[41] West GB, Brown J.H, Enquist B.J (2001) A general model for ontogenic groth. Nature 413, 628-631
[42] West GB, Woodruff WH, Brown JH (2002) Allometric scaling of metabolic rate from molecules and mitochondria to cells and mammals. 99(1):2473-2478
[43] Brown JH, West JB (2000) Scaling in Biology. Oxford University Press
[44] West GB, Brown JH, Enquist BJ (1999) The Fourt Dimension of Life: Fractal Geometry and Allometric Scaling of Organisms. Science 284(4):1677-1679
[45] Schroedinger E. (1944) What is Life? Cambridge University Press, Cambridge , New York
[46] Ebeling W (1976) Strukturbildung bei Irreversiblen Prozessen. Eine Enführung in die Theory dissipativer Strukturen. Teubner-Verlag-Leipzig
[47] Schlögl F (1971) On Stability of Steady State. Z. Physik 243:303-310
[48] Prigogine I, Defay R (1954) Chemical Thermodynamics. Longman, Green & Co., London
[49] Lassalle J, Lefschetz S (1961) Stability by Lyaponov's direct method. Academic Press Inc. New-York
[50] Lin H (1982) Foundamentals of zoological scaling. Am. J. Phys. 50(1)
[51] McMahon T (1973) Size and Shape in Biology. Science 179:1209-1209
[52] von SR de Groot (1951) Thermodynamics of irreversible processes. North-Holland Publ. Co. Amsterdam
[53] Don H Johnson (2003) Equivalent circuit concept: the voltage-source equivalent. Proceedings of the IEEE 91(4):636-640
[54] Caplan SR, Essig A (1983) Bioenergetics and Linear Non-equilibrium Thermodynamics. 1st edn. Harvard University Press, Cambridge
[55] Arias LA, Hernandez et al (2007) A first–order irreversible thermodynamic approach to a simple energy converter. arXIV: 0709.0321v1 [cond – math.stat - mech]
[56] Stucki JW (1989) The optimal efficiency and the economic degrees of coupling of oxidative phosphorisation. Eur. J. Biochem. 109:269-283
[57] Nath S (1998) A thermodynamic principle for the coupled bioenergetic processes of ATP synthesis. Pure & Appl. Chem. 70(3):639-644
[58] Kleidon A, Lorenz RD (2004) Entropy Production by Earth System Processes. Non-equilibrium thermodynamics and the production of entropy: life, Earth, and beyond. Springer Verlag, Heidelberg
[59] Ziegler H (1963) An attemp to generalize Onsager's principle, and its significance for rheological problems. ZAMP, 9(5-6):748-763

[60] Županovi´CP, Jureti´cD, Botrati´cS (2004) Kirchhoff's loop law and the maximum entropy production principle. Physical Review E 70:056108
[61] Jureti´c D, Županovi´c P (2003), Photosynthetic models with maximum entropy production in irreversible charge transfer steps. Computational Biology and Chemistry 27:541–553
[62] Dewar RC (2005) Maximum entropy production and the fluctuation theorem. J. Phys. A: Math. Gen. 38
[63] Martyushev LM, Seleznev VD (2006), Maximum entropy production principle in physics, chemistry and biology. Physics Reports 426:1-45
[64] Zupanovic´P et al (2010), The Maximum Entropy Production Principle and Linear. Irreversible Processes. Entropy 12:996-1005
[65] Kleidon A, Malhi Y, Cox PM (2010) Introduction Maximum entropy production in environmental and ecological systems. Phil. Trans. R. Soc. B 365:1297-1302
[66] Gage DH, Schiffer M, Kline SJ, Reynolds WC (1966) In Non-Equilibrium Thermodynamics: Variational Techniques and Stability. R. J. Donnelly (ed) University of Chicago
[67] Richardson IW (1969) On the Principle of Minimum Entropy Production. Biophysical Journal Vol.9
[68] Gyarmati I (1869) On the "Governing Principle of Dissipative Processes" and its extension to Non-linear Problemss. Ann. D. Phys, 353-378
[69] Reza FM, Seely S (1959) Modern Network Analysis. McGraw-Hill Book Company. Inc. New York-Toronto-London
[70] Acel J (1961) Vorlesungen über Funktionalgleichungen und Ihre Anwendungen. VEB Deutscher Verlag der Wissenschaften Berlin

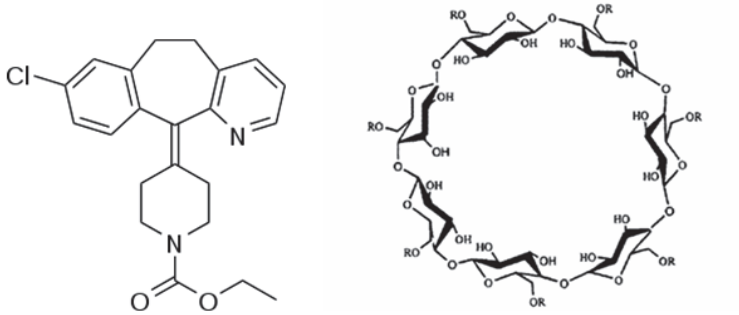
11

Thermodynamic Study of Grinding-Induced Loratadine Inclusion Complex Formation Using Thermal Analysis and Curve-Fitted FTIR Determination

Shan-Yang Lin, Hong-Liang Lin, Chih-Cheng Lin,
Cheng-Hung Hsu, Tieh-kang Wu and Yu-Ting Huang
Department of Biotechnology, Yuanpei University, Hsin Chu, Taiwan, Republic of China

1. Introduction

Cyclodextrins (CDs) have been considerably raised interest and attention in the pharmaceutical sciences due to their specific ability to form inclusion complexes with many drugs, leading to significant enhancements of the solubility, stability and bioavailability, as well as decrease in the bitterness and tissue irritation of the drugs (Carrier et al., 2007; Challa et al., 2005; Miller et al., 2007). Thus, CDs have been extensively applied to the pharmaceutical dosage form development for various drugs. Up-to-date, over 30 CD-containing drug products have been commercialized on the global pharmaceutical market for clinical use (Davis & Brewster, 2004; Loftsson & Brewster, 2010).



Chemical structures of loratadine, β-cyclodextrin (β-CD) and 2-hydroxypropyl-β-cyclodextrin (HP-β-CD)

Loratadine (LOR) is the most commonly used second-generation antihistamine drug for the symptomatic relief of allergic disorders without significant side effects on the central and autonomic nervous systems (Kaiser et al., 2008; Kay & Harris, 1999). One major limited use

of LOR is its water insoluble property. It has been reported that LOR belongs to a BSC (Biopharmaceutical Classification System) class II low solubility and high permeation drug (Khan et al., 2004). Due to the poor water-soluble property and low dissolution rate of LOR, the oral bioavailability of LOR exhibited high intra and inter subject variabilities (Khan et al., 2004; Ramirez et al., 2010). Therefore, how to improve the solubility of LOR in an aqueous solution plays an important role in the pharmaceutical formulation design.

A lot of methods such as grinding, kneading, freeze drying, spray drying and slow evaporation have been widely reported to prepare the inclusion complexes for numerous drugs (Carrier et al., 2007; Challa et al., 2005; Miller et al., 2007; Singh et al., 2010). Grinding is one of the most effective processes not only to alter the molecular behavior of the ground drugs with or without additives but also to improve the molecular interaction between drugs and additives in the ground mixture (Morris et al., 2001; Zhang et al., 2004). The inclusion complex of LOR with β-cyclodextrin (β-CD), hydroxypropyl-β-CD (HP-β-CD) or dimethyl-β-CD had been studied by kneading, spray-drying, freeze-drying and microwave irradiation (Nacsa et al., 2008; Omar et al., 2007; Nacsa et al., 2009), but the conventional solid-state grinding usually used in the pharmaceutical industry for preparing inclusion complex of LOR is scanty investigated (Lin et al., 2010).

Although the grinding has been used to prepare the inclusion complex of the drug, the grinding-induced inclusion mechanism of the drug into the cavity of β-CD or HP-β-CD to form an inclusion complex is less studied. In the present study, the progressive steps of inclusion complex formation caused by grinding the LOR with β-CD were quantitatively investigated by differential scanning calorimetry (DSC) and Fourier transform infrared (FTIR) microspectroscopy with curve-fitting analysis, as compared with the ground mixture of LOR/HP-β-CD (Lin et al., 2010). The thermodynamic kinetics and inclusion efficiency of LOR entering the cavity of β-CD or HP-β-CD to form an inclusion complex were explored.

2. Materials and methods

2.1 Materials

A pharmaceutical grade of loratadine (LOR, Jai Radhe Sales, Gujarat, India) was used without further purification. Beta-cyclodextrin (β-CD) and hydroxypropyl-β-cyclodextrin (HP-β-CD) were obtained from Chinoin Pharm. & Chem. Works Ltd, Budapest, Hungary and Roquette Freres, Lestrem, France, respectively. All the other materials were of analytical reagent grade.

2.2 Preparation of ground mixtures of LOR with β-CD or HP-β-CD

Each ground mixture of LOR and β-CD (LOR/β-CD) or LOR and HP-β-CD (LOR/HP-β-CD) with a 1:1 molar ratio was respectively prepared in an oscillatory ball mill (Mixer Mill MM301, Retsch GmbH & Co., Germany) with 15 Hz oscillation frequency. About 0.2 g of powder sample was placed in a 25 ml volume stainless steel milling jar containing two 15 mm diameter stainless steel balls, and then oscillated. In the grinding process, the sample was withdrawn at the prescribed intervals for further examination (Lin et al., 2010).

2.3 Thermal analysis

Each sample was determined by using a differential scanning calorimetry (DSC, Q20, TA Instruments, Inc., New Castle, DE, USA) at a heating rate of 3°C/min with an open pan

system in a stream of N_2 gas. The enthalpy of endothermic peak in the DSC curve was calculated. There was no oxidation or decomposition phenomenon observed in these determining conditions before the melting of LOR (Gao & Lin, 2010; Lin et al., 2006, 2010).

2.4 Curve-fitted FTIR determination

The infrared spectrum of each sample was analyzed with a Fourier transform infrared (FTIR) microspectroscopy (IRT-5000-16/FTIR-6200, Jasco Co., Tokyo, Japan) equipped with a mercury cadmium telluride (MCT) detector via a transmission technique (Gao & Lin, 2010; Lin et al., 2006, 2010). All the FTIR spectra were obtained at a 4 cm^{-1} resolution and at 100 scans. The components and relative compositions of each sample were estimated quantitatively within the 1740-1600 cm^{-1} region of FTIR spectra by a curve-fitting algorithm with a Gaussian-Lorenzian function (Cheng et al., 2008; Hu et al., 2002). The best curve-fitting procedure was performed by iterative fits toward a minimum standard error. The relative composition of the component was computed to be the fractional area of the corresponding peak, divided by the sum of the area for all the peaks.

3. Results and discussion

Mechanical grinding is well known to be one of the important manufacturing procedures in the pharmaceutical industry (Morris et al., 2001; Zhang et al., 2004). The energetic input of solid-state grinding has been reported not only to alter the particle size, crystallinity, and solid-state polymorphic conversion of the drug but also to produce the drug-excipient interaction, resulting in the modification of physical and chemical properties of drug to finally influence the drug bioavailability (Carrier et al., 2007; Challa et al., 2005; Davis & Brewster, 2004; Loftsson & Brewster, 2010; Miller et al., 2007; Morris et al., 2001; Zhang et al., 2004).

3.1 Grinding-induced changes in LOR crystallinity in the presence of β-CD or HP-β-CD

The grinding-induced enthalpy changes in the DSC curves of the physical and ground mixtures of LOR/β-CD or LOR/HP-β-CD with a 1:1 molar ratio after different grinding times are shown in Figs. 1 and 2, respectively. It clearly indicates that a sharp endothermic peak at 135.1 or 134.6°C was observed in the DSC curve of the physical mixture of LOR/β-CD or LOR/HP-β-CD, which was close to the endothermic peak at 136.1°C for intact LOR (Lin et al., 2010; Ramos & Cavalheiro, 2007). This endothermic peak at 135.1 or 134.6°C was due to the melting point of LOR in the presence of β-CD or HP-β-CD, because the temperature range within 30-160oC lacked an endothermic peak for intact β-CD or HP-β-CD. A sharp endothermic peak at 135.1°C for LOR in the physical mixture of LOR/β-CD became smaller with the increase of the grinding time, and this endothermic peak disappeared after grinding for 15 min. The disappearance of this endothermic peak might be due to the amorphization and/or inclusion complex formation after grinding (Chowdary & Srinivas, 2006; Giordano et al., 2001; Willart & Descamps, 2008). By grinding the physical mixture of LOR/HP-β-CD, on the other hand, the DSC endothermic peak disappeared after a 7 min-grinding course. The disappearance of this endothermic peak was also similar to that of LOR/β-CD, due to the grinding-induced amorphization and/or inclusion complex formation. It is evident that HP-β-CD was more efficient than that of the β-CD to reduce the peak intensity of LOR in the DSC curve.

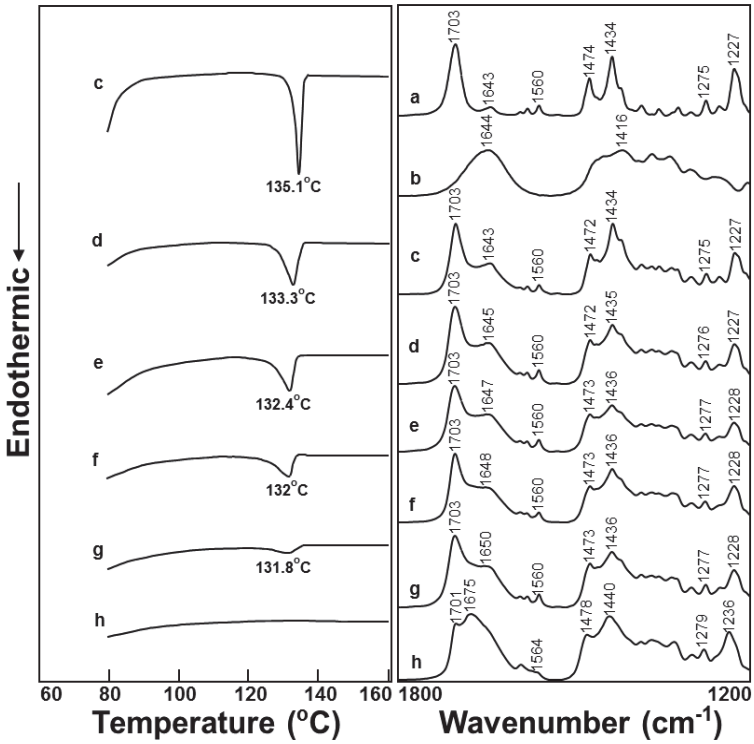


Fig. 1. The grinding-induced enthalpy changes and FTIR spectral shifts of the physical and ground mixtures of LOR/β-CD with a 1:1 molar ratio. Key, a, LOR; b, β-CD; Grinding time: c, o min; d, 3 min; e, 5 min, f, 7 min; g, 10 min, h, 15 min.

Once the physical mixture of drug and CDs was ground, two possible mechanisms can occur. One is the reduction in drug crystallinity due to the cogrinding effect with additive; another is the intact formation of inclusion complex. It is well-known that when the drug was amorphized or formed an inclusion complex, its endothermic peak in the DSC curve commonly disappeared (Chowdary & Srinivas, 2006; Giordano et al., 2001; Willart & Descamps, 2008). From the explanation of DSC results alone, the decrease in the value of DSC enthalpy was difficult to attribute whether the actual CD inclusion complex formation or the amorphization of drug after grinding process. However, the total changes in DSC enthalpy induced by grinding process could be easily evaluated. The relative degree of LOR crystallinity (RDC_{LOR}, %) in the physical or ground mixture might be estimated from the DSC curve by the ratio between the melting enthalpy of the LOR calculated in each sample and that of the melting enthalpy of starting physical mixture or pure LOR (Mura et al., 2003; Sauceau et al., 2008), according to the Eq. (1)

$$\mathrm{RDC}_{\mathrm{LOR}}(\%) = \frac{\Delta H_{MIX}}{\Delta H_{LOR}} \mathrm{x}\ 100 \tag{1}$$

where ΔH_{MIX} is the melting enthalpy of LOR calculated in the physical or ground mixture, and ΔH_{LOR} is the melting enthalpy of starting physical mixture or pure LOR.

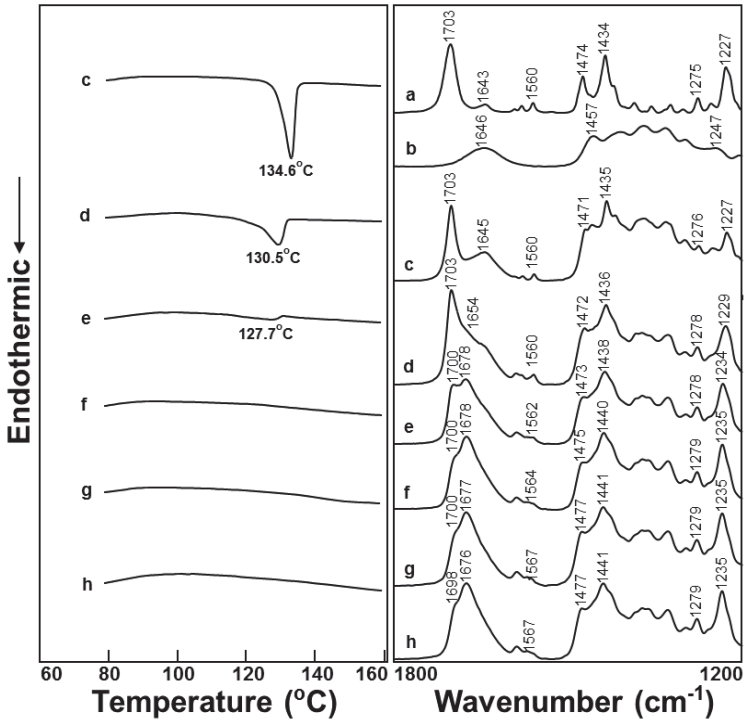


Fig. 2. The grinding-induced enthalpy changes and FTIR spectral shifts of the physical and ground mixtures of LOR/HP-β-CD with a 1:1 molar ratio. Key, a, LOR; b, HP-β-CD; Grinding time: c, o min; d, 3 min; e, 5 min, f, 7 min; g, 10 min, h, 15 min.

The effect of grinding time on the RDC_{LOR} in the absence or presence of β-CD or HP-β-CD is displayed in Fig. 3. Obviously, the crystallinity of LOR was reduced with the increase of grinding time. In addition, the HP-β-CD seems to be faster than that of the β-CD causing the reduction in crystallinity of LOR. The excess amount of HP-β-CD might improve the grinding effect than that of the β-CD used although there was the same molar ratio. Fig. 3 also demonstrates the apparent zero-order kinetics for the reduction in crystallinity of LOR in the presence of β-CD or HP-β-CD with grinding time.

3.2 Grinding-induced FTIR spectral shifts of LOR in the presence of β-CD or HP-β-CD

The FTIR spectral shifts for the physical and ground mixtures of LOR/β-CD or LOR/HP-β-CD at 1:1 molar ratio after grinding are also shown in Figs. 1 and 2. Several characteristic IR absorption bands for LOR, β-CD or HP-β-CD may be assigned as follows: 1703 cm^{-1} (C=O of ester), 1560, 1474 and 1434 cm^{-1} (stretching vibrations of benzene ring), 1227 cm-1 (C-O stretching) for LOR; 1644 or 1646 cm^{-1} (OH groups of the glucose moieties) for β-CD or HP-

β-CD, respectively (Nacsa et al., 2008, 2009; Zhao et al., 2003; Pedersen et al., 2005). In the LOR/β-CD mixture, four new peaks at 1675, 1478, 1440 and 1236 cm^{-1} were clearly observed after grinding for 15 min. The appearance of these new IR peaks might be due to the inclusion complex formation between LOR and β-CD. Once β-CD was replaced by HP-β-CD, however, the changes in FTIR spectra for LOR/HP-β-CD ground mixture were more pronounced.

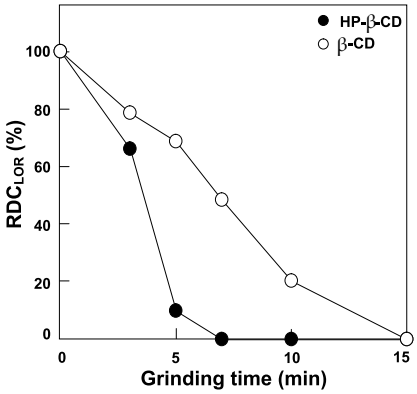


Fig. 3. The effect of grinding time on the RDC_{LOR} of LOR in the absence or presence of β-CD or HP-β-CD.

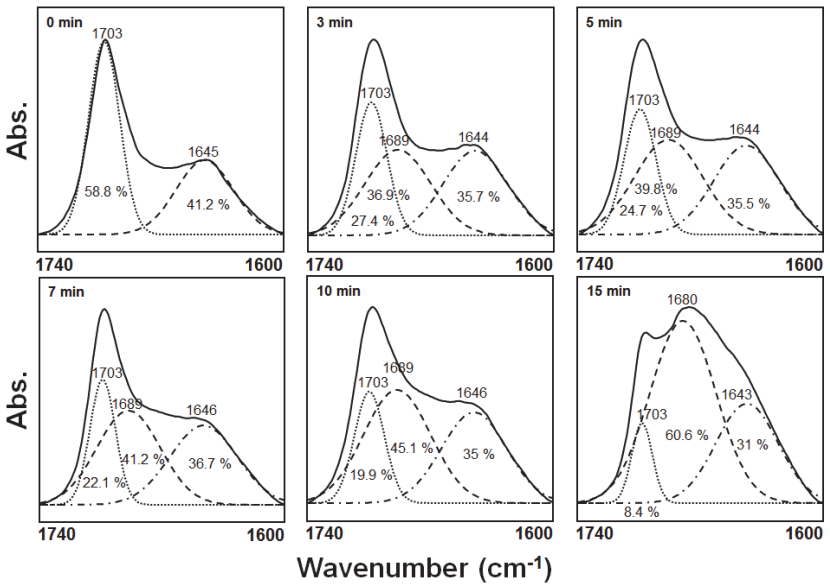


Fig. 4. The representative best-fitted FTIR spectra and the deconvoluted components for the physical mixture of LOR with β-CD before and after various grinding times.

After grinding for 5 min, a shoulder at 1678 cm^{-1} was observed. By increasing the grinding time, the shoulder at 1678 cm^{-1} was shifted to a predominant peak at 1676 cm^{-1}. In addition, the peak at 1229 cm^{-1} assigned to C-O stretching vibration was also shifted to 1235 cm^{-1}. The IR spectrum of 15 min-ground mixture was the same as the IR spectrum of the inclusion complex prepared by co-evaporation method reported in our previous study (Lin et al., 2010). Due to the inclusion complex formation, the characteristic bands at 1703 and 1227 cm^{-1} of LOR were down-shifted to 1676 cm^{-1} and up-shifted to 1235 cm^{-1}, respectively. Moreover, the structural benzene ring-related specific peaks at 1434, 1474, 1560 cm^{-1} were also shifted to 1441, 1477 and 1567 cm^{-1}, respectively. The progressive inclusion mechanism of LOR entering into the cavity of β-CD or HP-β-CD via grinding process was clearly evidenced.

The time-dependent inclusion complex formation of LOR/β-CD or LOR/HP-β-CD with a 1:1 molar ratio in the grinding process was estimated from the IR spectra by using the curve-fitting technique, as shown in Figs. 4 and 5. Obviously, the changes in IR spectra and the relative compositions of three components were obtained for both ground mixtures with the increase of grinding time. Three components were certainly estimated from the IR spectrum of each ground mixture, as compared with two components for LOR and one component for β-CD or HP-β-CD (Lin et al., 2010). Excluding the inherent peaks of LOR (1700-1703 cm^{-1}) and β-CD or HP-β-CD (1643-1646 cm^{-1}), the new peak formed after grinding might be attributed to the inclusion complex formation. In the LOR/β-CD ground mixtures, the best-fitting results indicate that the relative composition of the peak at 1703 cm^{-1} due to LOR was clearly reduced with the grinding time.

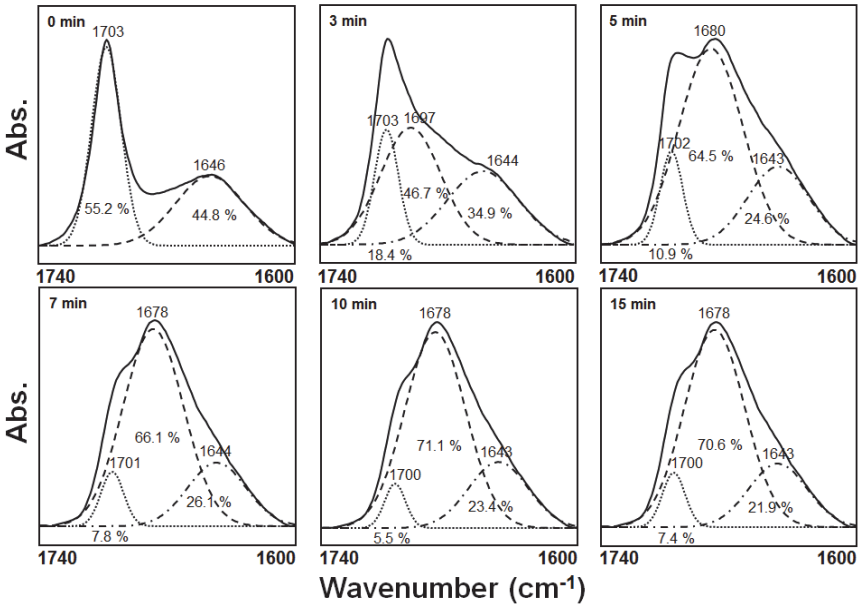
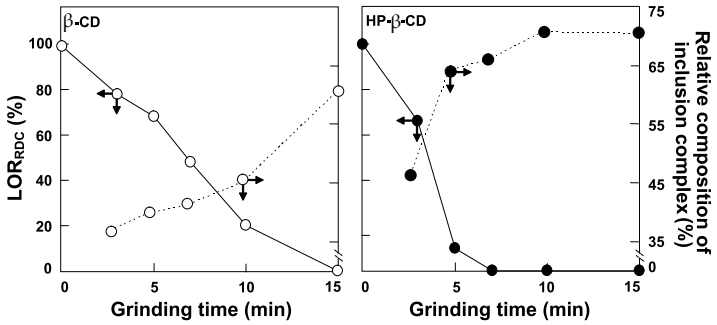


Fig. 5. The representative best-fitted FTIR spectra and the deconvoluted components for the physical mixture of LOR with HP-β-CD before and after various grinding times.

The unique IR peak at 1689 cm^{-1} was evidenced from the 3 min-ground mixture and also increased its relative composition slightly with grinding time. After grinding for 15 min, the peak position of inclusion complex formed was shifted from 1689 to 1680cm^{-1} and its relative compositions were increased from 36.9% to 60.6%. While in the LOR/HP-β-CD ground mixtures, the IR spectral peak at 1697 cm^{-1} was shifted to 1678 cm^{-1} in the grinding course from 3 to 15 min. Moreover, the relative composition of LOR/HP-β-CD inclusion complex was increased from 46.75 to 70.6% with the increase of grinding time.
Figure 6 indicates the relationship between the relative degree of crystallinity of LOR and the relative composition of inclusion complex in the grinding course. Obviously, the formation of inclusion complex for LOR/β-CD or LOR/HP-β-CD was significantly correlated with the reduction in LOR crystallinity under different grinding times. With the increase of grinding time, the crystallinity of LOR was gradually reduced but the inclusion complex formation was conversely enhanced. The possible inclusion mechanism of LOR/β-CD or LOR/HP-β-CD with the increase of grinding time might be proposed via the progressive reduction in LOR crystallinity at the initial and then gradual increase of amorphization of LOR, and finally the amorphized LOR molecule was included in the cavity of β-CD or HP-β-CD to form the inclusion complex.



Solid line: determined by DSC; dotted line: determined by FTIR

Fig. 6. The relationship between the relative degree of crystallinity of LOR and the relative composition of inclusion complex in the grinding course.

4. Conclusions

The results of present study indicate that the formation of inclusion complex for LOR/β-CD or LOR/HP-β-CD was significantly correlated with the reduction in LOR crystallinity. The apparent zero-order kinetics for the reduction in crystallinity of LOR in the presence of β-CD or HP-β-CD was obtained by increasing the grinding time. The inclusion complex of LOR/β-CD or LOR/HP-β-CD was easily formed with the increase of grinding time via the progressive changes in crystallinity and amorphization of LOR, and then entering the cavity of the β-CD or HP-β-CD molecule.

5. Acknowledgement

This work was supported by National Science Council, Taipei, Taiwan, Republic of China (97-2628-B-264-001-MY3).

6. References

Carrier, R.L.; Miller, L.A. & Ahmed, I. (2007). The Utility of Cyclodextrins for Enhancing Oral Bioavailability. Journal of Controlled Release, Vol. 123, No. 2, (November 2007), pp. 78-99, ISSN 0168-3659

Challa, R.; Ahuja, A.; Ali, J. & Khar, R.K. (2005). Cyclodextrins in Drug Delivery: an Updated Review. AAPS PharmSciTech, Vol. 6, No. 2, (October 2005), pp. E329-E357, ISSN 1530-9932

Cheng, W.T.; Lin, S.Y. & Wang, S.L. (2008). Differential Scanning Calorimetry with Curve-Fitting Program Used to Quantitatively Analyze the Polymorphic Transformation of Famotidine in the Compressed Compact. Drug Development and Industrial Pharmacy, Vol. 34, No. 12, (December 2008), pp. 1368-1375, ISSN 0363-9045

Chowdary, K.P. & Srinivas, S.V. (2006). Influence of Hydrophilic Polymers on Celecoxib Complexation with Hydroxypropyl Beta-Cyclodextrin. AAPS PharmSciTech, Vol. 7, No. 3, (September 2006), pp. E184–E189, ISSN 1530-9932

Davis, M.E. & Brewster, M.E. (2004). Cyclodextrin-Based Pharmaceutics: Past, Present and Future. Nature Reviews Drug Discovery, Vol. 3, No. 12, (December 2004), pp. 1023-1035, ISSN 1474-1776

Gao, G.Y. & Lin, S.Y. (2010). Thermodynamic Investigations of Nitroxoline Sublimation by Simultaneous DSC-FTIR Method and Isothermal TG Analysis. Journal of Pharmaceutical Sciences, Vol. 99, No. 1, (January 2010), pp. 255-261, ISSN 0022-3549

Giordano, F.; Novak, C. & Moyano, J.R. (2001). Thermal Analysis of Cyclodextrins and Their Inclusion Compounds. Thermochimica Acta, Vol. 380, No. 2, (December 2001), pp. 123-151, ISSN 0040-6031

Hu, T.C.; Wang, S.L.; Chen, T.F. & Lin, S.Y. (2002). Hydration-Induced Proton Transfer in the Solid State of Norfloxacin. Journal of Pharmaceutical Sciences, Vol. 91, No. 5, (May 2002), pp. 1351-1357, ISSN 0022-3549

Kaiser, H.B.; Gopalan, G. & Chung, W. (2008). Loratadine Provides Early Symptom Control in Seasonal Allergic Rhinitis. Allergy and Asthma Proceedings, Vol. 29, No. 6, (November-December 2008), pp. 654-658, ISSN 1088-5412

Kay, G.G. & Harris, A.G. (1999). Loratadine: A Non-Sedating Antihistamine. Review of its Effects on Cognition, Psychomotor Performance, Mood and Sedation. Clinical & Experimental Allergy, Vol. 29, No. Suppl 3, (July 1999), pp. 147-150, ISSN 0954-7894

Khan, M.Z.; Rausl, D.; Zanoski, R.; Zidar, S.; Mikulcić, J.H.; Krizmanić, L.; Eskinja, M.; Mildner, B. & Knezević, Z. (2004). Classification of Loratadine Based on the Biopharmaceutics Drug Classification Concept and Possible in Vitro-in Vivo Correlation. Biological and Pharmaceutical Bulletin, Vol. 27, No. 10, (October 2004), pp. 1630-1635, ISSN 0918-6158

Lin, S.Y.; Cheng, W.T. & Wang, S.L. (2006). Thermodynamic and Kinetic Characterization of Polymorphic Transformation of Famotidine During Grinding. International Journal of Pharmaceutics, Vol. 318, No. 1-2, (August 2006), pp. 86-91, ISSN 0378-5173

Lin, S.Y.; Hsu, C.H. & Sheu, M.T. (2010). Curve-Fitting FTIR Studies of Loratadine/ Hydroxypropyl-Beta-Cyclodextrin Inclusion Complex Induced by Co-Grinding Process. Journal of Pharmaceutical and Biomedical Analysis, Vol. 53, No. 3, (November 2010), pp. 799-803, ISSN 0731-7085

Loftsson, T. & Brewster, M.E. (2010). Pharmaceutical Applications of Cyclodextrins: Basic Science and Product Development. Journal of Pharmacy and Pharmacology, Vol. 62, No. 11, (November 2010), pp. 1607-1621, ISSN 0022-3573

Miller, L.A.; Carrier, R.L. & Ahmed, I. (2007). Practical Considerations in Development of Solid Dosage Forms that Contain Cyclodextrin. Journal of Pharmaceutical Sciences, Vol. 96, No. 7, (July 2007), pp. 1691-1707, ISSN 0022-3549

Morris, K.R.; Griesser, U.J.; Eckhardt, C.J. & Stowell, J.G. (2001). Theoretical Approaches to Physical Transformations of Active Pharmaceutical Ingredients During Manufacturing Processes. Advanced Drug Delivery Reviews, Vol. 48, No. 1, (May 2001), pp. 91–114, ISSN 0169-409X

Mura, P.; Maestrelli, F.; Cirri, M.; Furlanetto, S. & Pinzauti, S. (2003). Differential Scanning Calorimetry as An Analytical Tool in the Study of Drug-Cyclodextrin Interactions. Journal of Thermal Analysis and Calorimetry, Vol. 73, No. 2, (August 2003), pp. 635-645, ISSN 1388-6150

Nacsa, A.; Ambrus, R.; Berkesi, O.; Szabó-Révész, P. & Aigner, Z. (2008). Water-Soluble Loratadine Inclusion Complex: Analytical Control of the Preparation by Microwave Irradiation. Journal of Pharmaceutical and Biomedical Analysis, Vol. 48, No. 3, (November 2008), pp. 1020-1023, ISSN 0731-7085

Nacsa, Á.; Berkesi, O.; Szabó-Révész, P. & Aigner, Z. (2009). Achievement of pH-Independence of Poorly-Soluble, Ionizable Loratadine by Inclusion Complex Formation with Dimethyl-β-Cyclodextrin. Journal of Inclusion Phenomena and Macrocyclic Chemistry, Vol. 64, No. 3-4, (February 2009), pp. 249-254, ISSN 0923-0750

Omar, L.; El-Barghouthi, M.I.; Masoud, N.A.; Abdoh, A.A.; Al Omari, M.M.; Zughul, M.B. & Badwan, A.A. (2007). Inclusion Complexation of Loratadine with Natural and Modified Cyclodextrins: Phase Solubility and Thermodynamic Studies. Journal of Solution Chemistry, Vol. 36, No. 5, (April 2007), pp. 605-616, ISSN 0095-9782

Pedersen, N.R.; Kristensen, J.B.; Bauw, G.; Ravoo, B.J., Darcy, R., Larsen, K.L. & Pedersen, L.H. (2005). Thermolysin Catalyses the Synthesis of Cyclodextrin Esters in DMSO. Tetrahedron Asymmetry, Vol. 16, No. 3, (February 2005), pp. 615-622, ISSN 0957-4166

Ramirez, E.; Laosa, O.; Guerra, P.; Duque, B.; Mosquera, B.; Borobia, A.M.; Lei, S.H.; Carcas, A.J. & Frias, J. (2010). Acceptability and Characteristics of 124 Human Bioequivalence Studies with Active Substances Classified According to the Biopharmaceutic Classification System. British Journal of Clinical Pharmacology, Vol. 70, No. 5, (November 2010), pp. 694-702, ISSN 0306-5251

Ramos, L.A. & Cavalheiro, E.T.G. (2007). Thermal Behavior of Loratadine. Journal of Thermal Analysis and Calorimetry, Vol. 87, No. 3, (June 2007), pp. 831-834, ISSN 1388-6150

Sauceau, M. ; Rodier, E. & Fages, J. (2008). Preparation of Inclusion Complex of Piroxicam with Cyclodextrin by Using Supercritical Carbon Dioxide. The Journal of Supercritical Fluids, Vol. 47, No. 2, (December 2008), pp. 326-332, ISSN 0896-8446

Singh, R.; Bharti, N.; Madan, J. & Hiremath, S.N. (2010). Characterization of Cyclodextrin Inclusion Complexes – A Review. Journal of Pharmaceutical Science and Technology, Vol. 2, No. 3, (November 2010), pp. 171-183, ISSN 0975-5772

Willart, J. F. & Descamps, M. (2008). Solid State Amorphization of Pharmaceuticals. Molecular Pharmaceutics, Vol. 5, No. 6, (October 2008), pp. 905-920, ISSN 1521-0111

Zhang, G.G.; Law, D.; Schmitt, E.A. & Qiu, Y. (2004). Phase Transformation Considerations During Process Development and Manufacture of Solid Oral Dosage Forms. Advanced Drug Delivery Reviews, Vol. 56, No. 3, (February 2004), pp. 371–390, ISSN 0169-409X

Zhao, J.; Feng, J.; Chen, G.; Liu, Q.; Hu, G. ; Pavol, K. & Alexander, P. (2003). Structure Analysis of Loratadine. Chinese Journal of Analytical Chemistry, Vol. 31, No. 3, (March 2003), pp. 311-314, ISSN 1872-2040

12

Three-Dimensional Constitutive Viscoelastic Model for Isotropic Materials

Donald Picard[1,2] and Mario Fafard[2]
[1]*Department of Civil Engineering,*
Laval University, Quebec,
[2]*NSERC/Alcoa Industrial Research Chair MACE*[3]
and Aluminium Research Centre – REGAL
Canada

1. Introduction

1.1 Content

The characterization of materials behaviour is of importance in many research fields. From an analytical point of view, the mechanical behaviour of materials could be described either by using empirical laws based on experimental observations or by using a framework to develop constitutive laws. In the latter case, the thermodynamic of irreversible processes could be used as the framework. These constitutive laws are generally dedicated to complex problems and are thus developed in a three-dimensional context. Materials behave in different ways under loading but, under specific conditions, they all will generally exhibit instantaneous and time-dependant deformations. Instantaneous deformation could be elastic, plastic and so forth, while time-dependant deformation generally refers to the viscosity of the material.

Moreover, as for the Navier-Stokes equation in fluid mechanics which is only valid for Newtonian fluids, constitutive laws are developed considering the nature of the materials. Hence, in this chapter, targeted materials are isotropic granular materials such as concrete, carbon block materials and ramming paste (Sørlie and Øye, 2010). Contrary to metals, this implies that the tensile behaviour is different from the compressive one. Models are also meant to represent various phenomena such as creep and relaxation, which are mainly those of interest in this chapter. The upcoming sections will give a brief overview of the mechanical deformations such as creep/relaxation that are involved in viscoelasticity. Some basic principles governing the development of constitutive law modelling within a thermodynamic framework will be discussed, as well.

1.2 Viscoelasticity

Viscoelasticity of materials could be expressed in many ways. In this chapter, it will be described as the ability of the material to deform elastically, viscously and/or a combination of those. The interaction between the elastic and viscous behaviours could be explained with the help of the rheology of materials, as discussed in section 3.4.

For solids, the elastic behaviour is related to the instantaneous deformation of a material and is expressed using Young's modulus. This is generally measured by two different methods.

The first one consists of calculating the slope of the stress-strain curve resulting from a uniaxial loading test as described by the one-dimensional (1D) Hooke's law (1):

$$E = \frac{\sigma}{\varepsilon} = \frac{F/A_0}{\Delta L/L_0} = \frac{FL_0}{A_0 \Delta L} \quad where \quad \Delta L = \left| L_f - L_0 \right| \tag{1}$$

where $E, \sigma, \varepsilon, F, A_0, L_f$ and L_0 are respectively the Young's modulus, the engineering stress, the engineering strain, the maximum load applied, the initial surface area on which the force is applied, the final length and the initial length of the sample. In the case of hyper elasticity, the true stress and true strain should be used but won't be discussed here. Both static and dynamic loading tests could be performed and thus leading to the static and dynamic Young's modulus. The second method calculates the Young's modulus using acoustic wave technique (Kinsler, 2000). A relation is established between the elasticity of the material, i.e., the Young's modulus, and the speed of sound travelling through it, as described by Eq. (2):

$$E = c^2 \rho \frac{(1+\nu)(1-2\nu)}{(1-\nu)} \tag{2}$$

where c, ρ and ν are respectively the speed of sound in the material, the density and the Poisson's ratio. For a cylindrical sample, the Poisson's ratio is defined as the ration of the radial strain (perpendicular to the applied load) over the axial strain (in the direction of the applied load):

$$\nu = -\frac{\varepsilon_{radial}}{\varepsilon_{axial}} \tag{3}$$

Young's modulus should therefore be calculated adequately with either Eq. (1) or (2) according to the context. Materials discussed in this chapter generally undergo static loading and thus elasticity of materials will be defined as the slope of the stress-strain curve resulting from a uniaxial static loading test, as defined in Eq. (1).

The viscous behaviour of material could be rather complex. As aforementioned, it may be described with time-dependant functions. Time-dependant deformations can be both reversible and/or irreversible but in either case dissipation of energy is involved. This point will be clarified in section 3. According to material rheology, a reversible mechanism means that the deformations would be recovered once the material is unloaded. For example, the elastic deformation is fully reversible and instantaneous. On the other hand, an irreversible mechanism means that the deformations will not be recovered once the material is unloaded. However, emphasis will be put on a reversible mechanism such as the creep/relaxation phenomenon.

Creep in materials is generally defined as a three stage phenomenon, as shown in Figure 1. Under constant loading, the creep strain rate will decrease (stage I) up to a steady state (stage II) before increasing up to the failure (stage III). The three stages are also called the primary creep, secondary creep and tertiary creep. Tertiary creep may or may not happen depending on the stress level applied. If the applied stress is low (e.g. below $0.4\,\sigma_u$ for the concrete, where σ_u is the compressive strength), tertiary creep won't happen and a steady-state could be reached instead. Creep mechanisms are also function of the material. For metals, the mechanisms could be related to the movements of dislocations. For granular

materials, mechanisms causing the creep are very different. In carbon cathode block for example, creep may originate from the cleavage and slip of basal planes within the carbon based filler particles (Picard et al., 2008).

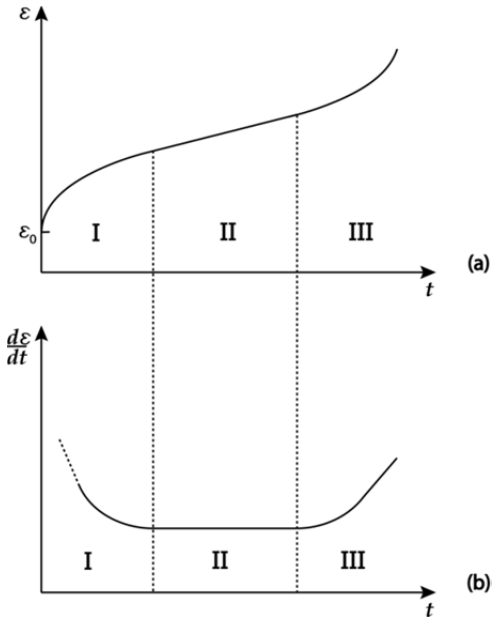


Fig. 1. Scheme of creep strain (a) and creep strain rate (b) during a constant loading test.

In most engineering applications, however, models are developed considering the macroscopic behaviour of materials and thus the creep mechanisms (movement of dislocations, cleavage and slip of basal planes, etc.) are not directly taken into consideration. Nevertheless, the creep mechanisms could be used to give a physical meaning to parameters of the constitutive laws, as discussed in section 3.4.
Also, with the increase of numerical power calculation, complex problems are more and more solved within a three-dimensional context. Simplification of 3D complex problems to 1D ones may then be irrelevant in some cases to get approximated solutions. This may however increase the complexity of the experimental characterisation in order to identify model parameters, as discussed in the following section. This requires that the way models are established be carefully performed in order to make parameters easily identifiable in laboratory.

1.3 Three-dimensional context

As discussed in the previous section, strains and stresses are closely linked together through material characteristics and deformation mechanisms. In solid mechanic, it is useful to mathematically express this relation through constitutive laws. Also, as aforementioned, numerical modelling now requires three-dimensional constitutive models. One of the simplest expressions of a constitutive law is the Hooke's law (4) in Voigt notation (Mase and Mase, 1999), for isotropic materials. In this case, assuming the small strain theory, only two

parameters which are the Young's modulus E and the Poisson's ratio ν have to be identified.

$$\begin{bmatrix} \varepsilon_{11} \\ \varepsilon_{22} \\ \varepsilon_{33} \\ 2\varepsilon_{23} \\ 2\varepsilon_{31} \\ 2\varepsilon_{12} \end{bmatrix} = \begin{bmatrix} \varepsilon_{11} \\ \varepsilon_{22} \\ \varepsilon_{33} \\ \gamma_{23} \\ \gamma_{31} \\ \gamma_{12} \end{bmatrix} = \frac{1}{E} \begin{bmatrix} 1 & -\nu & -\nu & 0 & 0 & 0 \\ -\nu & 1 & -\nu & 0 & 0 & 0 \\ -\nu & -\nu & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2(1+\nu) & 0 & 0 \\ 0 & 0 & 0 & 0 & 2(1+\nu) & 0 \\ 0 & 0 & 0 & 0 & 0 & 2(1+\nu) \end{bmatrix} \begin{bmatrix} \sigma_{11} \\ \sigma_{22} \\ \sigma_{33} \\ \sigma_{23} \\ \sigma_{31} \\ \sigma_{12} \end{bmatrix} \tag{4}$$

Constitutive laws for viscoelastic materials are however more complex than the Hooke's law and consequently the number of parameters to identify increases. Frameworks could also be used to ensure that all parameters of the proposed laws are physically admissible. To do so, it is possible to use the thermodynamic of irreversible processes as the framework. Based on the concepts of continuum mechanics and irreversible thermodynamics, the Clausius-Duhem inequality is obtained for given problems where dissipation mechanisms are of importance, e.g., viscous deformation. Fundamental equations leading to a generic form of the Clausius-Duhem inequality have been well covered by many authors (Bazarov, 1964; Coussy, 2010; Lemaitre and Chaboche, 1990; Mase and Mase, 1999) and thus will be only summarized later in section 3. Based on the generic form of the Clausius-Duhem inequality, models or constitutive laws are further developed considering various assumptions closely related to materials of interest.

2. Materials of interest

The needs of the constitutive laws presented in this chapter originated from projects with the aluminium industry. Those laws have been used in numerical simulations (D'Amours et al., 2003; Picard et al., 2008; Richard et al., 2005; Richard et al., 2006) of the Hall-Héroult electrolysis cell (Figure 2) used for aluminum production (Sørlie and Øye, 2010). The materials of interest in these projects were mainly carbon cathode blocks, carbon anode blocks and the ramming paste. All those materials are enclosed in the potshell, thereby restraining their thermal expansion and/or chemical expansion (swelling) in some cases. In fact, the Hall-Héroult electrolysis cell is built at room temperature while its operational temperature is near 1000 °C. A preheating phase prior to the aluminum production phase is thus performed to avoid or minimize thermal shocks in materials. The preheating of Hall-Héroult electrolysis cell is achieved using various methods (Sørlie and Øye, 2010). The one of interest in this work is commonly named as the electrical preheating that uses the Joule effect to heat the cell lining materials. This method is performed by passing a high density current (up to 1 A/cm^2 at the cathode) from the anodes to the cathode through a coke bed as shown in Figure 3.

This preheating phase is critical for the industry since it can have a major impact on the cell lifespan (Tessier et al., 2011; Tessier et al., 2010) and was thus closely investigated over the past years through numerical simulations (D'Amours et al., 2003; Marceau et al., 2011; Richard et al., 2006; Sun et al., 2004). It is therefore required to develop constitutive laws for the three aforementioned materials to feed numerical models. The targeted materials are briefly presented below.

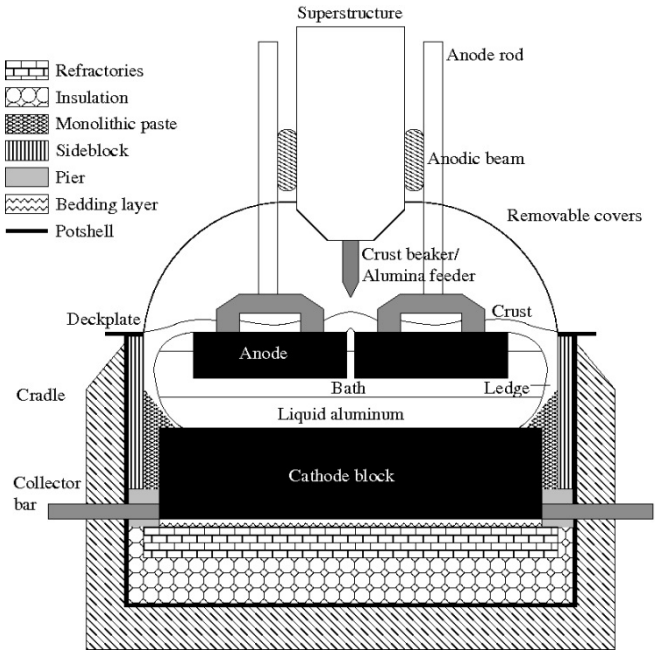


Fig. 2. Diagram of the Hall-Héroult electrolysis cell - from (Richard, 2004).

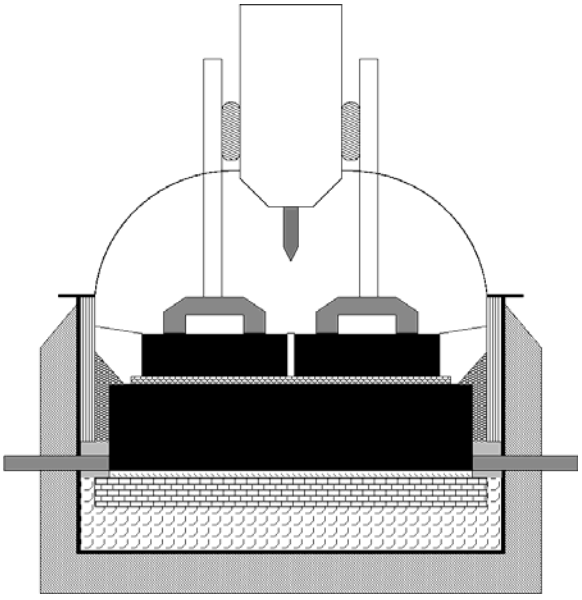


Fig. 3. Diagram of the electrical preheating stage of the Hall-Héroult electrolysis cell - from (Goulet, 2004).

The carbon cathode consists of several carbon blocks joined together by ramming paste. These blocks are located in the lining of the electrolysis cell (Sørlie and Øye, 2010). The cathode block, as well as the ramming paste, consists of filler particles (derived from coke and graphite) mixed up with a binder (e.g. pitch). However, before the cells are started, the cathode blocks are baked while the interblock ramming paste remains unbaked (green) and will only be partially baked during the pot start-up heating (Sørlie and Øye, 2010). Depending on the block type used, heat treatment can be applied to the filler particles and/or blocks at different stages in the manufacturing process.

As aforementioned, during the preheating phase, temperature in the electrolysis cell starts from ambient and slowly increases to the desired operational value, usually around 960 °C. Thermal expansion of all materials of the lining (steel shell, carbon cathode, ramming paste, refractory concrete, etc.) will thus induce mechanical stresses that could lead to material failures. Elasticity, plasticity, damaging, hardening, softening, creep/relaxation, etc. are different mechanisms inducing stresses/strains that may eventually lead to mechanical failure of the carbon materials (ramming paste and cathode). This implies that constitutive laws must take into account temperature evolution as well as many other phenomena (e.g. chemical contamination). Also, the carbon cathode block and ramming paste may exhibit damaging at very low stresses down to $0.2\sigma_u$. Linear creep is thus mainly observed at stress levels under $0.2\sigma_u$ (Picard et al., 2008) where no other phenomenon takes place.

Even though carbon anodes are not confined within a steel shell like cathode, thermal expansion of the anodic beams as well as thermal shocks may lead to mechanical failure of anodes. Creep/relaxation could also play an important role, mainly in the stub-hole region where high localized stresses could be induced by irregular geometries. Anode properties are very similar to carbon cathode and ramming paste, therefore, constitutive laws developed for one of the materials can be easily adapted to the others.

3. Constitutive laws

3.1 Mathematical notation

In this work, the following mathematical notation is used for scalars, vectors and tensors:

$a, A:$	Scalars
$\{a\}, \langle a \rangle, \vec{a}:$	Vectors
$[A]:$	Matrix
$[\mathbf{A}]:$	Second-order tensor
$[\underline{\underline{\mathbf{A}}}]:$	Fourth-order tensor

The double contraction product will also be used in this work and for second-order tensors is defined as:

$$[\mathbf{A}] : [\mathbf{B}] = \mathrm{tr}(\mathbf{A}^{\mathbf{T}}\mathbf{B}) = \mathrm{tr}(\mathbf{B}\mathbf{A}^{\mathbf{T}}) = \mathrm{tr}(\mathbf{A}\mathbf{B}^{\mathbf{T}}) = \mathrm{tr}(\mathbf{B}^{\mathbf{T}}\mathbf{A}) \tag{5}$$

3.2 Thermodynamic framework

The methodology presented in this section is well covered in the literature (Bazarov, 1964; Coussy, 1995; Lemaitre and Chaboche, 1990; Mase and Mase, 1999). So, it will only be summarized here.

The thermodynamic framework implies that all systems must be based on the first and second thermodynamic principles. In continuum mechanic, the first principle which is the energy conservation can be written as:

$$\frac{dE}{dt} + \frac{dC}{dt} = P_{ext} + \sum other \tag{6}$$

$$C = \int_v \rho \vec{v} \cdot \vec{v} dv \tag{7}$$

$$E = \int_v \rho e dv \tag{8}$$

where e is the specific internal energy, E is the internal energy, C is the kinematic energy, P_{ext} is the power of external forces and $\sum other$ is related to all other kinds of energy (thermal, electrical, magnetic, etc.). Also, in equations (7) and (8), $\vec{v}$ is the velocity and v refers to the volume element. All materials also exhibit energy dissipation under external forces. If the dissipation level is of importance for a given problem, the second principle must then be used to obtain the Clausius-Duhem inequality. The inequality is obtained by first establishing a relation between the specific internal energy e, as described in Eq. (8), and the specific entropy density s:

$$e = \psi + Ts \tag{9}$$

while considering

$$S = \int_v \rho s dv \tag{10}$$

where T is the temperature in Kelvin and ψ is the Helmholtz free energy. In solid mechanic while considering a thermomechanical problem, the first principle, i.e. Eq. (6), can be rewritten as:

$$\frac{dE}{dt} + \frac{dC}{dt} = \frac{\partial W_{ext}}{\partial t} + \frac{\partial Q}{\partial t}, \quad \frac{\partial W_{ext}}{\partial t} = P_{ext} \tag{11}$$

where $\partial Q / \partial t$ is the heat rate received by the domain considered. Starting from equation (11) the Clausius-Duhem inequality has to be defined either in initial configuration or deformed configuration. In the latter case, i.e., deformed configuration, by using equations (7) and (8), equation (11) becomes:

$$\frac{d}{dt}\left(\frac{1}{2} \int_v \rho \vec{v} \cdot \vec{v} + \rho e dv \right) = \int_v (\vec{f} \cdot \vec{v} + r) dv + \int_s (\vec{t} \cdot \vec{v} - \vec{q} \cdot \vec{n}) ds \tag{12}$$

with

$$P_{ext} = \int_v \vec{f} \cdot \vec{v} dv + \int_s \vec{t} \cdot \vec{v} ds \tag{13}$$

and

$$\frac{\partial Q}{\partial t} = \int_v r dv - \int_s \vec{q} \cdot \vec{n} ds \tag{14}$$

where $\vec{f}$ is a body load, $\vec{t}$ is a prescribed stress , r is a heat source, $\vec{q}$ is a prescribed heat flux and $\vec{n}$ is a normal vector to a surface, considering that:

$$\vec{t} = [\boldsymbol{\sigma}]\vec{n} \tag{15}$$

and

$$\underbrace{\frac{d\rho}{dt} + \rho div(\vec{v}) = 0}_{\text{mass conservation}} \tag{16}$$

where $[\boldsymbol{\sigma}]$ is a second order stress tensor, Eq. (12) can be rewritten, after few other considerations, as:

$$r = \rho \frac{de}{dt} - [\boldsymbol{\sigma}] : [\mathbf{d}] + div(\vec{q}) \tag{17}$$

here

$$[\mathbf{d}] = sym(grad(\vec{v})) = [\dot{\boldsymbol{\varepsilon}}] \tag{18}$$

In order to get the Clausius-Duhem equation, the second principle must then be introduced and take the following form:

$$\frac{dS}{dt} - \dot{S}_{ext} \geq 0 \tag{19}$$

with

$$\dot{S}_{ext} = \int_r \frac{r}{T} dv - \int_s \frac{\vec{q} \cdot \vec{n}}{T} ds \tag{20}$$

where S was already defined in Eq. (10). After some mathematical manipulations, equation (20) becomes:

$$\rho \frac{ds}{dt} - \frac{r}{T} + \frac{1}{T} div(\vec{q}) - \frac{1}{T^2} \vec{q} \cdot grad(T) \geq 0 \tag{21}$$

Finally, combining the two equations obtained with the first and second principles, i.e., equations (17) and (21), and the Helmholtz free energy (9), a general form of the Clausius-Duhem inequality is obtained:

$$[\boldsymbol{\sigma}] : [\dot{\boldsymbol{\varepsilon}}] - \rho \left(\frac{d\psi}{dt} + s \frac{dT}{dt} \right) - \frac{\vec{q}}{T} \cdot grad(T) \geq 0 \tag{22}$$

One of the most important parts in modelling is then to establish a link using the Clausius-Duhem inequality to materials parameters, e.g., by defining the form of Helmholtz free energy ψ. The parameters will be identified later in laboratory. Different approaches can be used to establish the form of ψ and the choice is of course function of the problem. In some cases, it is also possible to work with the Gibbs free energy instead of the Helmholtz, e.g. (Haslach, 2009; Lion et al., 2010; Schapery, 1997). For hyperelastic materials, the Ogden model, e.g. (Holzapfel, 1996; Li and Lua, 2009; Ogden, 1972), can also be used. Two different approaches will be presented in this work. The first one will link materials properties to the Clausius-Duhem inequality through generic internal state variables (Fafard et al., 2001) and the second one will use phenomenological approach (Picard et al., 2008) based on rheology of materials. The two approaches have already been published and will thus be only summarized here in sections 3.3 and 3.4.

3.3 Generic internal state variables

The Clausis-Duhem inequality (22) has established a relation between the stress and the strain through the Helmholtz free energy ψ. Considering a viscoelastic material and constant and uniform temperature, it can be assumed that the Helmholtz free energy will only be function of the total strain and an undetermined number of second order tensors internal state variables $[\mathbf{q}_\alpha]$ ($\alpha = 1,...,n$). So, Eq. (22) becomes:

$$[\boldsymbol{\sigma}] : [\dot{\boldsymbol{\varepsilon}}] - \dot{\psi}([\boldsymbol{\varepsilon}], [\mathbf{q}_\alpha]) \geq 0 \tag{23}$$

The time derivation of Helmholtz free energy and the substitution into (23) gives

$$\left([\boldsymbol{\sigma}] - \frac{\partial \psi}{\partial \varepsilon}\right) : [\dot{\boldsymbol{\varepsilon}}] - \frac{\partial \psi}{\partial [\mathbf{q}_\alpha]} : \frac{\partial [\mathbf{q}_\alpha]}{\partial t} \geq 0 \tag{24}$$

The inequality (24) must be satisfied for any thermodynamic process. According to the local state law (Lemaitre and Chaboche, 1990) , the Clausius-Duhem inequality leads to the following state equation:

$$[\boldsymbol{\sigma}] = \frac{\partial \psi}{\partial [\boldsymbol{\varepsilon}]} \tag{25}$$

The Clausius-Duhem inequality thus becomes:

$$-\frac{\partial \psi}{\partial [\mathbf{q}_\alpha]} : \frac{\partial [\mathbf{q}_\alpha]}{\partial t} \geq 0 \tag{26}$$

Assuming the linearity of the dissipative mechanism, the so-called dissipative potential ϕ can be chosen as a quadratic form of its arguments (Coussy, 1995; Lemaitre and Chaboche, 1990; Valanis, 1972). Hence, the dissipative potential is:

$$\phi = \frac{1}{2} \sum [\dot{\mathbf{q}}_\alpha] : \underline{\underline{[\mathbf{b}_\alpha]}} : [\dot{\mathbf{q}}_\alpha] \tag{27}$$

Where the coefficients of $\underline{\underline{[\mathbf{b}_\alpha]}}$ must be identified. In the present case, the coefficients of $\underline{\underline{[\mathbf{b}_\alpha]}}$ will be considered constant, i.e., independent of time t, temperature T, etc. The non-constant coefficients case is well covered by (Fafard et al., 2001). To satisfy the inequality (26), the complementary evolution law can then be written as

$$\frac{\partial \phi}{\partial [\dot{\mathbf{q}}_\alpha]} + \frac{\partial \psi}{\partial [\mathbf{q}_\alpha]} = 0 \tag{28}$$

Furthermore, the following quadratic form of the Helmholtz free energy is assumed for isotropic, orthotropic or anisotropic materials (Valanis, 1972) for reversible dissipative mechanical problem:

$$\psi = \frac{1}{2}[\boldsymbol{\varepsilon}] : \underline{\underline{[\mathbf{E}]}} : [\boldsymbol{\varepsilon}] + \sum_\alpha [\mathbf{q}_\alpha] : \underline{\underline{[\mathbf{B}_\alpha]}} : [\boldsymbol{\varepsilon}] + \frac{1}{2}\sum_\alpha [\mathbf{q}_\alpha] : \underline{\underline{[\mathbf{A}_\alpha]}} : [\mathbf{q}_\alpha] \tag{29}$$

where $\underline{\underline{[\mathbf{E}]}}$, $\underline{\underline{[\mathbf{B}_\alpha]}}$ and $\underline{\underline{[\mathbf{A}_\alpha]}}$ are symmetric tensors. The first part of (29) is the energy due to the instantaneous deformation. The two other terms represent the change in free energy due to the viscoelastic behaviour of the material. The stress tensor can be obtained from (25) and (29) that takes the following form:

$$[\boldsymbol{\sigma}] = \frac{\partial \psi}{\partial [\boldsymbol{\varepsilon}]} = \underline{\underline{[\mathbf{E}]}} : [\boldsymbol{\varepsilon}] + \sum_\alpha \underline{\underline{[\mathbf{B}_\alpha]}}^T : [\mathbf{q}_\alpha] \tag{30}$$

The equations (27) to (30) lead to a set of equations for each value of α:

$$\underline{\underline{[\mathbf{b}_\alpha]}} : [\dot{\mathbf{q}}_\alpha] + \underline{\underline{[\mathbf{A}_\alpha]}} : [\mathbf{q}_\alpha] = -\underline{\underline{[\mathbf{B}_\alpha]}} : [\boldsymbol{\varepsilon}] \tag{31}$$

Assuming isotropic materials, the following assumption is proposed: the topology of all tensors is similar to that of the elastic one. This means that for each tensor there are two unknown coefficients analogous to the Young's modulus and Poisson's coefficient (Fafard et al., 2001; Picard et al., 2008). Tensors of Eq. (30) and (31) can then be expressed as:

$$\begin{aligned} \underline{\underline{[\mathbf{A}_\alpha]}} &= \left[\mathbf{A}_\alpha(E_{A_\alpha}, \mu_\alpha)\right] \quad \underline{\underline{[\mathbf{B}_\alpha]}} = \left[\mathbf{B}_\alpha(E_{B_\alpha}, \zeta_\alpha)\right] \\ \underline{\underline{[\mathbf{b}_\alpha]}} &= \left[\mathbf{b}_\alpha(E_{b_\alpha}, \xi_\alpha)\right] \quad \underline{\underline{[\mathbf{E}_\alpha]}} = \left[\mathbf{E}_\alpha(E_E, \nu)\right] \end{aligned} \tag{32}$$

The analytical solution of Eq. (31) can be obtained using the diagonalization technique or modal projection (Kreyszig, 2006) , This is fully detailed in (Fafard et al., 2001). The specific solution for the uniaxial creep case ($\sigma_{11} = \sigma = \text{constant}$) of a cylindrical sample is

$$\varepsilon_{axial} = \varepsilon_{11} = \frac{1}{\gamma\beta}\left[\frac{1}{E} - \frac{1}{3}\frac{1}{G}\left(\frac{K^d}{K} + \gamma\frac{G^d}{G}e^{-\lambda_b t}\right) - \frac{1}{9}\frac{1}{K}\left(\frac{G^d}{G} + \beta\frac{K^d}{K}e^{-\lambda_a t}\right)\right]\sigma \tag{33}$$

$$\varepsilon_{radial} = \varepsilon_{22} = \varepsilon_{33} = \frac{1}{\gamma\beta}\left[\frac{-\nu}{E} - \frac{1}{6}\frac{1}{G}\left(\frac{K^d}{K} + \gamma\frac{G^d}{G}e^{-\lambda_b t}\right) - \frac{1}{9}\frac{1}{K}\left(\frac{G^d}{G} + \beta\frac{K^d}{K}e^{-\lambda_a t}\right)\right]\sigma \tag{34}$$

where

$$\lambda_a = \lambda_1 \gamma; \qquad \lambda_b = \lambda_2 \beta$$
$$\gamma = 1 - \frac{K^d}{K}; \quad \beta = 1 - \frac{G^d}{G} \tag{35}$$

$$G = \frac{E_E}{2(1+\nu)}; \quad K = \frac{E_E}{3(1-2\nu)} \tag{36}$$

and

$$\lambda_{1_\alpha} = \frac{E_{A_\alpha}}{E_{b_\alpha}} \frac{1-2\xi_\alpha}{1-2\mu_\alpha}; \quad \lambda_{2_\alpha} = \frac{E_{A_\alpha}}{E_{b_\alpha}} \frac{1+\xi_\alpha}{1+\mu_\alpha} \tag{37}$$

Coefficients G and K represent the instantaneous shear and compressibility modulus, respectively. In addition to elastic coefficients, one has to identify G^d, K^d, λ_{1_α} parameters and λ_{2_α}. Considering the hypothesis that $\xi_\alpha = \mu_\alpha$ for 3D isotropic materials leads to $\lambda_{1_\alpha} = \lambda_{2_\alpha} = \lambda_\alpha$ and thus reduces the number of parameters to be identified in laboratory. Finally, a link with a viscoelastic rheological model could thereafter be done by defining the rheological parameters as a function of the internal states variables. The alternative solution would be to define rheological parameters as the internal state variables at the very beginning . This corresponds to the second approach (section 3.4) presented in this chapter.

3.4 Strain Based Internal State Variables

Instead of defining the viscoelastic behaviour with generic internal state variable $\left[\underline{\underline{\mathbf{q}_\alpha}}\right]$, the second approach is based on the rheological behaviour of materials. Combined with the thermodynamic of irreversible processes (TIP), rheology gives the possibility to express the mechanical behaviour with simple parameters having a strong physical meaning (Kuiken, 1994). Models can be rather simple or very complex but they are all built using a combination of spring and dashpot elements. The more common elements, also called bodies, are the Kelvin-Voigt and the Maxwell elements. Creep behaviour of the materials studied in this chapter can be expressed using a simple rheological model called the Burger body. It is possible to extend this to another model using a similar approach.

As for the generic internal state variables approach discussed in section 3.3, the rheological approach also needs internal state variables. Instead of choosing generic ISV (Fafard et al., 2001), the choice of these variables is based on the use of a phenomenological approach by assuming a rheological model representing the material behaviour. To represent the creep behaviour of the studied materials, a Kelvin-Voigt rheological model has been chosen (Picard et al., 2008) as shown in Figure 4. As in section 3.3, the small strain theory is assumed.

The parameters of the Kelvin-Voigt model and the internal are fourth order tensors, while the strain and the stress are second order tensors. An undetermined number of Kelvin-Voigt elements give flexibility to the model without increasing the complexity of the constitutive law as it will be discussed later in this section. Assuming a virgin material, having no permanent strain due to earlier where $\left[\varepsilon^e\right]$ is the instantaneous elastic strain, where

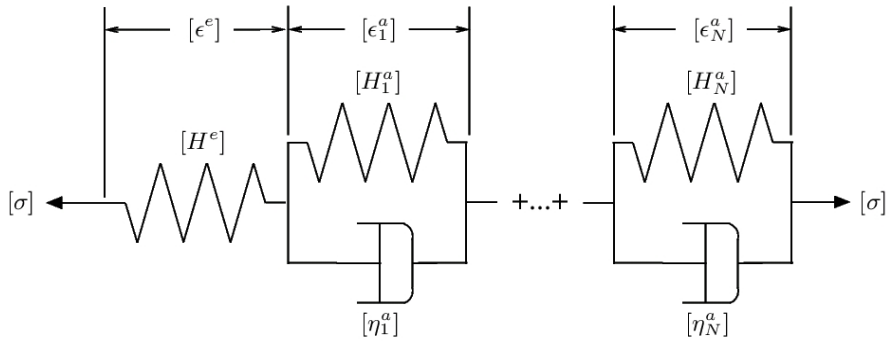


Fig. 4. Diagram of the three-dimensional visco-elastic rheological model.

damage, constant temperature and without other environmental phenomena that can cause additive strains, the total strain $[\varepsilon]$ is reduced to:

$$[\boldsymbol{\varepsilon}] = \left[\boldsymbol{\varepsilon}^{\mathbf{e}}\right] + \sum_{\alpha=1}^{N} \left[\boldsymbol{\varepsilon}_{\boldsymbol{\alpha}}^{\mathbf{a}}\right] \tag{38}$$

$\left[\varepsilon^e\right]$ is the instantaneous elastic strain, $\left[\varepsilon_\alpha^a\right]$ the anelastic strain associated with the Kelvin-Voigt element α, and N is the number of Kelvin-Voigt elements.
Now that the ISV is known, it is thus possible to define the Helmholtz free energy form as:

$$\rho\psi = \frac{1}{2}\left[\boldsymbol{\varepsilon}^{\mathbf{e}}\right] : \underline{\underline{\left[\mathbf{H}^{\mathbf{e}}\right]}} : \left[\boldsymbol{\varepsilon}^{\mathbf{e}}\right] + \frac{1}{2}\left[\boldsymbol{\varepsilon}_{\boldsymbol{\alpha}}^{\mathbf{a}}\right] : \underline{\underline{\left[\mathbf{H}_{\boldsymbol{\alpha}}^{\mathbf{a}}\right]}} : \left[\boldsymbol{\varepsilon}_{\boldsymbol{\alpha}}^{\mathbf{a}}\right] \tag{39}$$

where $\underline{\underline{\left[\mathbf{H}^{\mathbf{e}}\right]}}$ is the Voigt form of the fourth order elasticity tensor and $\underline{\underline{\left[\mathbf{H}_{\boldsymbol{\alpha}}^{\mathbf{a}}\right]}}$ is the fourth order tensor related to the spring of the α Kelvin-Voigt element. The parameters defining those tensors must be identified using appropriate tests. In order to take into account the dissipative mechanism of the model, a dissipative potential, ϕ , per unit of volume, can be chosen as a quadratic form of its arguments:

$$\rho\phi = \frac{1}{2}\sum_{\alpha=1}^{N} \left[\dot{\boldsymbol{\varepsilon}}_{\boldsymbol{\alpha}}^{\mathbf{a}}\right] : \underline{\underline{\left[\boldsymbol{\eta}_{\boldsymbol{\alpha}}^{\mathbf{a}}\right]}} : \left[\dot{\boldsymbol{\varepsilon}}_{\boldsymbol{\alpha}}^{\mathbf{a}}\right] \tag{40}$$

where the fourth-order tensor $\underline{\underline{\left[\boldsymbol{\eta}_{\boldsymbol{\alpha}}^{\mathbf{a}}\right]}}$ is related to the dashpot of the α Kelvin-Voigt element.

The parameters of this tensor must be identified, as well. Based on the thermodynamics of irreversible processes (TIP) and on the choice of the internal state variables (ISV), the Clausius-Duhem inequality can be written as:

$$\varphi_T + \varphi_M \geq 0 \tag{41}$$

where φ_T represents the thermal terms and φ_M the mechanical terms. For the visco-elastic problem, and assuming that the thermal part, φ_T, is satisfied, Eq. (41) can be rewritten using Eq. (39), with the mechanical terms only as:

$$\varphi_M = [\boldsymbol{\sigma}] : [\dot{\boldsymbol{\varepsilon}}] - \dot{\psi} \geq 0 \tag{42}$$

$$\left([\boldsymbol{\sigma}] - \rho \frac{\partial \psi}{\partial [\boldsymbol{\varepsilon}^{\mathbf{e}}]} \right) : [\dot{\boldsymbol{\varepsilon}}^{\mathbf{e}}] + \sum_{\alpha=1}^{N} \left([\boldsymbol{\sigma}] - \rho \frac{\partial \psi}{\partial [\boldsymbol{\varepsilon}_{\boldsymbol{\alpha}}^{\mathbf{a}}]} \right) : [\dot{\boldsymbol{\varepsilon}}_{\boldsymbol{\alpha}}^{\mathbf{a}}] \geq 0 \tag{43}$$

where $[\sigma]$ is the stress tensor. The first term in the Eq. (43) is associated with the elastic strain (Lemaitre and Chaboche, 1990). Since this latter strain is related to a reversible mechanism, it can be assumed that:

$$[\boldsymbol{\sigma}] = \rho \frac{\partial \psi}{\partial [\boldsymbol{\varepsilon}^{\mathbf{e}}]} \tag{44}$$

The Clausius-Duhem inequality (43) can then be rewritten:

$$\sum_{\alpha=1}^{N} \left([\boldsymbol{\sigma}] - \rho \frac{\partial \psi}{\partial [\boldsymbol{\varepsilon}_{\boldsymbol{\alpha}}^{\mathbf{a}}]} \right) : [\dot{\boldsymbol{\varepsilon}}_{\boldsymbol{\alpha}}^{\mathbf{a}}] \geq 0 \tag{45}$$

To satisfy the former equation, the complementary evolution law proposed in the following equation is postulated:

$$[\boldsymbol{\sigma}] - \rho \frac{\partial \psi}{\partial [\boldsymbol{\varepsilon}_{\boldsymbol{\alpha}}^{\mathbf{a}}]} = \frac{\partial \phi}{\partial [\dot{\boldsymbol{\varepsilon}}_{\boldsymbol{\alpha}}^{\mathbf{a}}]} \tag{46}$$

Assuming that $\underline{\underline{[\boldsymbol{\eta}_{\boldsymbol{\alpha}}^{\mathbf{a}}]}}$ are positive, the use of Eq. (40) and Eq. (46) ensures that Eq. (45) is satisfied and thus the Clausius-Duhem equation is also satisfied. Using the definitions of ψ and ϕ in Eq. (46) leads to Voigt notation (Mase and Mase, 1999),

$$[\eta_\alpha^a]\{\dot{\varepsilon}_\alpha^a\} + [H_\alpha^a]\{\varepsilon_\alpha^a\} = \{\sigma\} \tag{47}$$

for a given α. Assuming isotropic materials, tensors of Eq. (39) and (40) could now be expressed as:

$$\begin{aligned} \underline{\underline{[\mathbf{H}^{\mathbf{e}}]}} &= [\mathbf{H}^{\mathbf{e}}(E,\nu)] \\ \underline{\underline{[\mathbf{H}_{\boldsymbol{\alpha}}^{\mathbf{a}}]}} &= [\mathbf{H}_{\boldsymbol{\alpha}}^{\mathbf{a}}(E_{H_\alpha^a}, \mu_\alpha)] \\ \underline{\underline{[\boldsymbol{\eta}_{\boldsymbol{\alpha}}^{\mathbf{a}}]}} &= [\boldsymbol{\eta}_{\boldsymbol{\alpha}}^{\mathbf{a}}(E_{\eta_\alpha^a}, \varsigma_\alpha)] \end{aligned} \tag{48}$$

In order to identify the parameters of tensors of Eq.(48), Eq. (47) must be solved. The analytical solution can be obtained using the diagonalization technique (Kreyszig, 2006) and

the same assumptions aforementioned in section 3.3 will still be valid. The creep case will also be considered here. The solution of Eq. (47) is thus in Voigt notation:

$$\{\varepsilon\} = \left[H^e\right]^{-1}\{\sigma\} + \sum_{\alpha=1}^{N}\left[X_\alpha\right]\begin{Bmatrix} F_{\alpha_1}\int_0^\tau e^{-\lambda_{\alpha_1}(t-\tau)}d\tau \\ F_{\alpha_2}\int_0^\tau e^{-\lambda_{\alpha_2}(t-\tau)}d\tau \\ \vdots \\ F_{\alpha_6}\int_0^\tau e^{-\lambda_{\alpha_6}(t-\tau)}d\tau \end{Bmatrix} \tag{49}$$

with

$$\{F_\alpha\} = \left[X_\alpha\right]^{-1}\left[\eta_\alpha^a\right]^{-1}\{\sigma\} \tag{50}$$

where $\left[X_\alpha\right]$ is a matrix containing the eigenvectors $\{x_\alpha\}$ associated with the eigenvalues λ_α of the following system:

$$\left(\left[\eta_\alpha^a\right]\left[H_\alpha^a\right] - \lambda_\alpha\left[I\right]\right)\{x_\alpha\} = 0 \tag{51}$$

The complete solution is available in (Picard et al., 2008) for further review. Since all tensors have the same topology, it can be shown that $\left[X_\alpha\right] = \left[X\right]$ and that $\lambda_2 = \lambda_3 = \lambda_4 = \lambda_5 = \lambda_6$. Also, the eigenvalues of Eq. (49) are

$$\lambda_{\alpha_1} = \frac{E_{H_\alpha^a}(1-2\varsigma_\alpha)}{E_{\eta_\alpha^a}(1-2\mu_\alpha)}; \quad \lambda_{\alpha_{2,3,4,5,6}} = \frac{E_{H_\alpha^a}(1+\varsigma_\alpha)}{E_{\eta_\alpha^a}(1+\mu_\alpha)} \tag{52}$$

Finally, for the uniaxial creep case (constant loading) of a cylindrical sample, Eq. (49) can be rewritten as

$$\varepsilon_{axial} = \varepsilon_{11} = \underbrace{-\frac{\sigma_1}{E_{H^e}}}_{\text{Elastic strain}} \underbrace{-\sigma_1\left(\frac{1}{9}\sum_{\alpha=1}^{N}K_\alpha^{a^{-1}}\frac{1-e^{-\lambda_{\alpha_1}t}}{\lambda_{\alpha_1}} + \frac{1}{3}\sum_{\alpha=1}^{N}G_\alpha^{a^{-1}}\frac{1-e^{-\lambda_{\alpha_2}t}}{\lambda_{\alpha_2}}\right)}_{\text{Creep strain}} \tag{53}$$

$$\varepsilon_{radial} = \varepsilon_{22} = \varepsilon_{33} = \underbrace{\sigma_1\frac{\nu}{E_{H^e}}}_{\text{Elastic strain}} \underbrace{-\sigma_1\left(\frac{1}{9}\sum_{\alpha=1}^{N}K_\alpha^{a^{-1}}\frac{1-e^{-\lambda_{\alpha_1}t}}{\lambda_{\alpha_1}} - \frac{1}{6}\sum_{\alpha=1}^{N}G_\alpha^{a^{-1}}\frac{1-e^{-\lambda_{\alpha_2}t}}{\lambda_{\alpha_2}}\right)}_{\text{Creep strain}} \tag{54}$$

with

$$K_\alpha^{a^{-1}} = \frac{3(1-2\varsigma_\alpha)}{E_{\eta_\alpha^a}}; \quad G_\alpha^{a^{-1}} = \frac{2(1+\varsigma_\alpha)}{E_{\eta_\alpha^a}} \tag{55}$$

where σ_1 is the constant axial stress, ε_1 is the axial strain and ε_r the hoop strain. All other strain components are null. The coefficients $K_\alpha^{a^{-1}}$ and $G_\alpha^{a^{-1}}$ are related to the hydrostatic and deviatoric creep mechanisms, respectively. Moreover, all λ_{α_i} coefficients are positive and $1/\lambda_{\alpha_i}$ are related to a relaxation time. The elastic parameters E and ν (elastic properties) must be identified through standard compression tests, while the anelastic parameters $E_{H_\alpha^a}$, μ_α, $E_{\eta_\alpha^a}$ and ζ_α must be identified through experimental creep tests.

4. Academic case study

4.1 Finite elements model

Independently of the approach used to get a constitutive law, parameters of the latter must be identified in laboratory. To simplify the identification process, creep tests are preferred (Fafard et al., 2001; Picard et al., 2008) over any other tests to identify the constitutive laws parameters discussed in this work. In such case where the number of parameters to identify is small, methods such as Newton-Raphson algorithm could be used to minimize the error of an objective function through use of the least squares method. If the number of parameters is large, an alternative method like genetic algorithm may be more suitable.

In few cases, parameters identification at a reference state (generally at room temperature) will be sufficient. This implies that in most situations constitutive law parameters must evolve to be representative of the real situation (e.g., the carbon materials have to operate at a temperature near 1000 °C while the concrete may be influenced by the relative humidity of the atmosphere).

The analytical solutions of the proposed models in section 3 were obtained by assuming a reference state, i.e., a virgin material at 25 °C. As aforementioned, parameters evolution was not taken into account until now. Since these models are dedicated to being used in finite element code (Picard et al., 2008; Richard et al., 2005), each parameter could have its own scalar evolution function. To get this function, the methodology used consists of building datasets of creep measures under different states, e.g., at different temperatures. Then, parameters of the three-dimensional model are identified by using the appropriated parameters identification process, at each different state or temperature, for example. The function obtained will then be used in finite element code to take into account changing environment (temperature, relative humidity, etc.). To do this, the analytical solution of the two proposed approaches must be discretized. Regarding this, in the case of the strain based one (section 3.4), the Eq. (47) can be rewritten as:

$$\left[\eta_\alpha^a\right]\left\{\dot{\varepsilon}_\alpha^a\right\}+\left(\left[H_\alpha^a\right]+\left[H^e\right]\right)\left\{\varepsilon_\alpha^a\right\}=\left[H^e\right]\left(\left\{\varepsilon\right\}-\sum_{j=1,j\neq\alpha}^{N}\varepsilon_j^a\right) \tag{56}$$

Assuming a forward difference Euler scheme (Reddy, 1993)

$$\left[\chi_\alpha^a\right]_{t+\Delta t}\left\{\varepsilon_\alpha^a\right\}_{t+\Delta t}=\Delta t\left[H^e\right]_{t+\Delta t}\left\{\varepsilon\right\}_{t+\Delta t}+\left[\eta_\alpha^a\right]_{t+\Delta t}\left\{\varepsilon_\alpha^a\right\}_t-\sum_{j=1,j\neq\alpha}^{N}\Delta t\left[H^e\right]_{t+\Delta t}\left\{\varepsilon_j^a\right\}_{t+\Delta t} \tag{57}$$

where

$$\left[\chi_\alpha^a\right]_{t+\Delta t}=\left[\eta_\alpha^a\right]_{t+\Delta t}+\Delta t\left(\left[H_\alpha^a\right]_{t+\Delta t}+\left[H^e\right]_{t+\Delta t}\right) \tag{58}$$

Moreover, it can be assumed that

$$\left\{\varepsilon_\alpha^a\right\}_{t+\Delta t}=\left\{\varepsilon_\alpha^a\right\}_t+\Delta\left\{\varepsilon_\alpha^a\right\} \tag{59}$$

Equation (57) and (59) lead to a system of equation of dimension $N\times N$ described as follows:

$$\begin{bmatrix}\left[\chi_1^a\right] & \Delta t\left[H^e\right] & \cdots & \Delta t\left[H^e\right]\\ \Delta t\left[H^e\right] & \left[\chi_2^a\right] & \ddots & \vdots\\ \vdots & \ddots & \ddots & \Delta t\left[H^e\right]\\ \Delta t\left[H^e\right] & \cdots & \Delta t\left[H^e\right] & \left[\chi_N^a\right]\end{bmatrix}_{t+\Delta t}\begin{Bmatrix}\left\{\varepsilon_1^a\right\}_{t+\Delta t}\\ \left\{\varepsilon_2^a\right\}_{t+\Delta t}\\ \vdots\\ \left\{\varepsilon_N^a\right\}_{t+\Delta t}\end{Bmatrix}=\begin{Bmatrix}\Delta t\left[H^e\right]_{t+\Delta t}\left\{\varepsilon\right\}_{t+\Delta t}+\sum_{\alpha=1,\alpha\neq1}^{N}\left[\eta_\alpha^a\right]_{t+\Delta t}\left\{\varepsilon_\alpha^a\right\}_t\\ \Delta t\left[H^e\right]_{t+\Delta t}\left\{\varepsilon\right\}_{t+\Delta t}+\sum_{\alpha=1,\alpha\neq2}^{N}\left[\eta_\alpha^a\right]_{t+\Delta t}\left\{\varepsilon_\alpha^a\right\}_t\\ \\ \Delta t\left[H^e\right]_{t+\Delta t}\left\{\varepsilon\right\}_{t+\Delta t}+\sum_{\alpha=1,\alpha\neq N}^{N}\left[\eta_\alpha^a\right]_{t+\Delta t}\left\{\varepsilon_\alpha^a\right\}_t\end{Bmatrix} \tag{60}$$

From Eq. (59) and(60), the strain increment that will be used in finite element software is:

$$\begin{Bmatrix}\Delta\left\{\varepsilon_1^a\right\}\\ \Delta\left\{\varepsilon_1^a\right\}\\ \vdots\\ \Delta\left\{\varepsilon_N^a\right\}\end{Bmatrix}=\begin{bmatrix}\left[\chi_1^a\right] & \Delta t\left[H^e\right] & \cdots & \Delta t\left[H^e\right]\\ \Delta t\left[H^e\right] & \left[\chi_2^a\right] & \ddots & \vdots\\ \vdots & \ddots & \ddots & \Delta t\left[H^e\right]\\ \Delta t\left[H^e\right] & \cdots & \Delta t\left[H^e\right] & \left[\chi_N^a\right]\end{bmatrix}_{t+\Delta t}^{-1}\begin{Bmatrix}\Delta t\left[H^e\right]_{t+\Delta t}\left(\left\{\varepsilon\right\}_{t+\Delta t}-\sum_{j=1}^{N}\left\{\varepsilon_j^a\right\}_t\right)-\Delta t\left[H_1^a\right]_{t+\Delta t}\left\{\varepsilon_1^a\right\}_t\\ \Delta t\left[H^e\right]_{t+\Delta t}\left(\left\{\varepsilon\right\}_{t+\Delta t}-\sum_{j=1}^{N}\left\{\varepsilon_j^a\right\}_t\right)-\Delta t\left[H_2^a\right]_{t+\Delta t}\left\{\varepsilon_2^a\right\}_t\\ \vdots\\ \Delta t\left[H^e\right]_{t+\Delta t}\left(\left\{\varepsilon\right\}_{t+\Delta t}-\sum_{j=1}^{N}\left\{\varepsilon_j^a\right\}_t\right)-\Delta t\left[H_N^a\right]_{t+\Delta t}\left\{\varepsilon_N^a\right\}_t\end{Bmatrix} \tag{61}$$

Other quantities are also needed by finite element software, such as the tangent matrix for the use of Newton-Raphson methods, and are also function of the number N of Kelvin-Voigt elements. In the case of limiting the rheological model to $N=2$, the tangent matrix is:

$$[C]=\left[H^e\right]_{t+\Delta t}-\Delta t\left[H^e\right]_{t+\Delta t}\left(\left[\chi_{12}^a\right]_{t+\Delta t}^{-1}\left[\chi_{22}^a\right]_{t+\Delta t}+\left[\chi_{21}^a\right]_{t+\Delta t}^{-1}\left[\chi_{11}^a\right]_{t+\Delta t}\right)\left[H^e\right]_{t+\Delta t} \tag{62}$$

where

$$\begin{aligned}\left[\chi_{12}^a\right]_{t+\Delta t}&=\left[\chi_1^a\right]_{t+\Delta t}-\Delta t^2\left[H^e\right]_{t+\Delta t}\left[\chi_2^a\right]_{t+\Delta t}^{-1}\left[H^e\right]_{t+\Delta t}\\ \left[\chi_{21}^a\right]_{t+\Delta t}&=\left[\chi_2^a\right]_{t+\Delta t}-\Delta t^2\left[H^e\right]_{t+\Delta t}\left[\chi_1^a\right]_{t+\Delta t}^{-1}\left[H^e\right]_{t+\Delta t}\\ \left[\chi_{11}^a\right]_{t+\Delta t}&=[I]-\Delta t^2\left[H^e\right]_{t+\Delta t}\left[\chi_1^a\right]_{t+\Delta t}^{-1}\\ \left[\chi_{22}^a\right]_{t+\Delta t}&=[I]-\Delta t^2\left[H^e\right]_{t+\Delta t}\left[\chi_2^a\right]_{t+\Delta t}^{-1}\end{aligned} \tag{63}$$

The same methodology can be applied to the generic internal state variables approach to get the corresponding strain increment and tangent matrix.

4.2 Case study

The creep/relaxation behaviour of the ramming paste material hasn't been rigorously investigated yet. However, preliminary numerical results from (D'Amours, 2004) and (Richard, 2004) have clearly shown the importance of this phenomenon. Thus, it is relevant to establish the creep behaviour of this material, even in the form of a qualitative model, in order to take it into account in the pre-heating simulation of the thermo-electro-mechanical behaviour of a Hall-Héroult electrolysis cell. Also, as mentioned in section 2, the ramming paste is initially green and it begins to bake during the pre-heating phase. Thus, the mechanical properties of the paste evolve with the so-called "baking index" ξ proposed by (D'Amours, 2004) which is temperature dependant and irreversible. The viscoelastic rheological model presented in section 3.4 has thus been used to numerically simulate the thermo-mechanical behaviour of ramming paste materials (Picard, 2007) as an academic case study. Based on experimental and extrapolated data (Picard, 2007), the following parameters evolution of the viscoelastic model with two Kelvin-Voigt elements ($N = 2$) is obtained:

$$\begin{aligned}
E_{\eta_1^g}(\mathrm{MPa}\cdot\mathrm{h}) &\quad 271263e^{3.1104\xi} \\
\zeta_1 &\quad 0.1444\xi^2 - 0.3367\xi + 0.3922 \\
E_{H_1^g}(\mathrm{MPa}) &\quad 467276e^{0.2629\xi} \\
\mu_1 &\quad -0.0444\xi^2 + .02267\xi + 0.2078 \\
E_{\eta_2^g}(\mathrm{MPa}\cdot\mathrm{h}) &\quad 440e^{6.9123\xi} \\
\zeta_2 &\quad -0.5056\xi^2 + 0.4783\xi + 0.0972 \\
E_{H_2^g}(\mathrm{MPa}) &\quad 45e^{8.7068\xi} \\
\mu_2 &\quad -0.2889\xi^2 + 0.1733\xi + 0.2456
\end{aligned} \tag{64}$$

Moreover, it is relevant to note that the creep behaviour of the ramming paste in this case study is function of the baking index only, even at various temperatures. In fact, up to now, no creep/relaxation results are available for the ramming paste at various elevated temperatures. Also, it should be mentioned that the baking and the temperature effects are considered independent. The baking effect is related to the microstructure of the paste which is a function of the highest temperature reached by the paste (irreversible process), while the temperature effect is related to the actual temperature of the paste (reversible process).

The proposed model used by (D'Amours, 2004) was adopted to test evolution of the parameters of the ramming paste. This simple model, illustrated in Figure 5, consists of a quarter carbon block (C) and half a ramming paste seam (P) simulating the electrical preheating phase of an Hall-Héroult electrolysis cell (Sørlie and Øye, 2010). A similar model was also used by (Richard et al., 2005). All the details of that method are available in (Picard, 2007). The four elements on each extremity of the cathode (C) do not contribute to the heat transfer related to the anode since they are not covered by the bed of coke. The elements covered by the anode are represented by the "Anode shadow" region on the model. A forced convection ($h = 500\ \mathrm{W/m^2/K}$) was imposed in this "Anode shadow" region, where

the temperature T_∞ was ramped linearly from 20 °C to 1000 °C in a 40 hour interval. All surfaces of the anode except the upper one are insulated. For more information refer to (Picard, 2007).

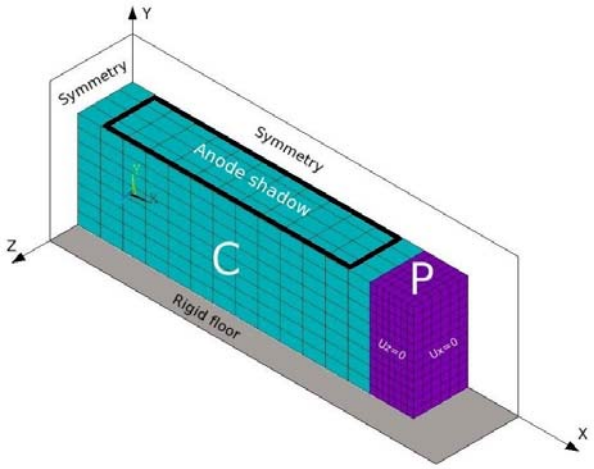


Fig. 5. Simple finite element model for the carbon cell lining materials – Taken from (D'Amours, 2004).

A first simulation was performed without any creep of the ramming paste to get a reference for future comparison. The strain results after a preheating of 40 hours are plotted in Figure 6.

The important point to note here is that a relatively high stress level in the longitudinal direction (X) is obtained in the materials that consequently initiates a high plastic strain amplitude in the ramming paste. This is in contradiction with the experimental observations (D'Amours, 2004; Richard, 2004) where an important creep/relaxation phenomenon should take place. In fact, the role of the ramming paste in the Hall-Héroult electrolysis cell is to let the cathode bottom block expand during the pre-heating phase without leading to any mechanical failure (e.g., cracking where liquid could leak in). It is then relevant to add a viscoelastic behaviour such as the creep/relaxation to the model. The simulation was then run again with the addition of creep/relaxation in the ramming paste and the results are presented in Figure 7.

These last numerical results show that the plastic strain is obviously greatly influenced by the presence of the creep/relaxation phenomenon. In fact, the level of the plastic strain was considerably reduced from 0.004 to 0.0018, i.e a reduction ratio of 2.2 based on the reference case (Figure 6), which does not take into account creep behaviour of the ramming paste. Also, the anelastic strain level at the end of the simulation ($t = 40$ hours) is almost negligible compared to the other strains (e.g., plastic, thermal, etc.). This result directly ensues from the assumption that the baked ($\xi = 1$) ramming paste creep/relaxation behaviour is similar to that of the carbon cathode block (Richard et al., 2006). This case study thus shows the importance of taking all the relevant phenomena including creep behaviour into account in similar problems. A similar analysis could be done for all other deformations (chemical, thermal, plastic, etc.).

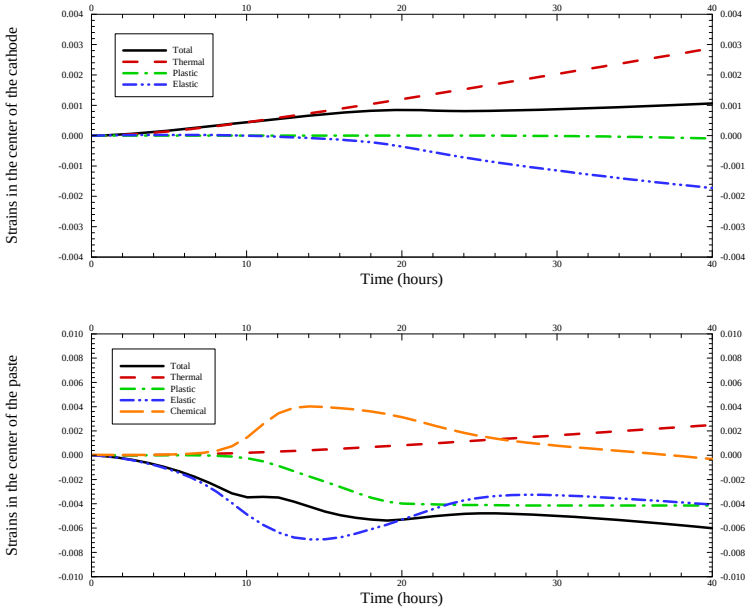


Fig. 6. Strain evolution at the center of each component of the model (Reference case).

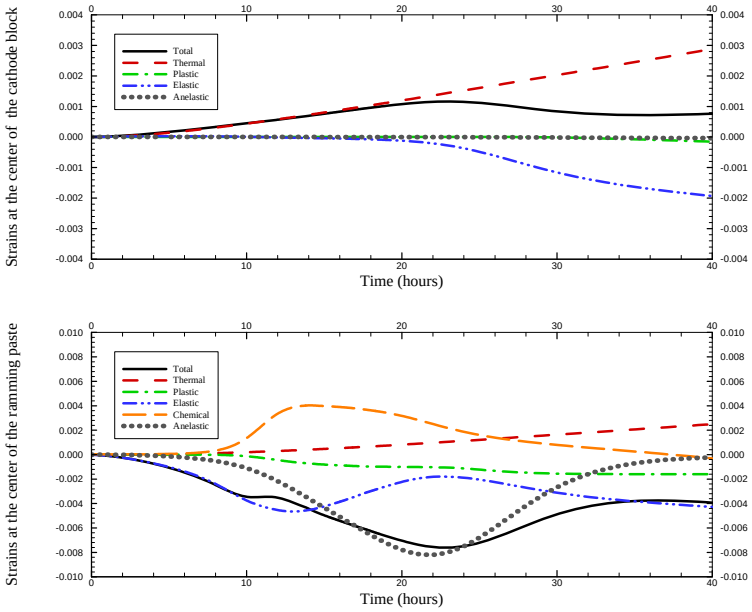


Fig. 7. Normal stresses at the center of each component of the model (creep case).

5. Conclusion

The main contribution of this work is to demonstrate how to use thermodynamics framework to develop three-dimensional viscoelastic constitutive laws. Two different approaches using internal state variables have been presented. The first one used generic internal state variables related to dissipative mechanisms. A link with rheological model could be created by defining the rheological parameters as a function of the internal states variables. The second approach was similar to the first one except that the internal state variables were based on phenomenology with the use of rheological models, i.e., a viscoelastic one in the present case. To achieve this, a classical 1D rheological model was extended to the 3D case.

The two constitutive laws presented in this work were obtained considering an isotropic material at a reference state, i.e., a state without any influence of external parameters such as temperature. The constitutive parameters evolutions are instead taking into account within the numerical simulation and thus the constitutive laws have to be discretized using an Euler scheme. A case study using the strain based internal state variables approach has been presented. This study shows how creep/relaxation could influence the results of an industrial problem such as the "baking" of a carbonaceous ramming paste.

Finally, as aforementioned, both approaches have considered isotropic materials. However, the methodology to obtain constitutive laws of anisotropic or orthotropic materials such as wood would be similar.

6. List of symbols

A : area

$\underline{\underline{[\mathbf{A}_\alpha]}}$: fourth-order tensor

$\underline{\underline{[\mathbf{b}_\alpha]}}$: fourth-order tensor

$\underline{\underline{[\mathbf{B}_\alpha]}}$: fourth-order tensor

c : speed of sound

C : kinematic energy

$[\mathbf{d}]$: second-order strain rate tensor

e : specific internal energy

E : Young's modulus

E_{A_α} : constitutive parameter of $\underline{\underline{[\mathbf{A}_\alpha]}}$

E_{b_α} : constitutive parameter of $\underline{\underline{[\mathbf{b}_\alpha]}}$

E_{B_α} : constitutive parameter of $\underline{\underline{[\mathbf{B}_\alpha]}}$

E_E : constitutive parameter of $\underline{\underline{[\mathbf{E}_\alpha]}}$

$E_{H_\alpha^a}$: constitutive parameter of $\underline{\underline{[\mathbf{H}_\alpha^a]}}$

$E_{\eta_\alpha^a}$: constitutive parameter of $\underline{\underline{[\boldsymbol{\eta}_\alpha^a]}}$

$\underline{\underline{[\mathbf{E}]}}$: fourth-order tensor

$\underline{\underline{[\mathbf{E}_\alpha]}}$: fourth-order tensor

$\vec{f}$: body load

F: force

G: instantaneous shear modulus

$G_\alpha^{a^{-1}}$: deviatoric coefficent

$\underline{\underline{[\mathbf{H^e}]}}$: fourth-order elastic tensor

$\underline{\underline{[\mathbf{H_\alpha^a}]}}$: fourth-order anelastic tensor

K: instantaneous compressive modulus

$K_\alpha^{a^{-1}}$: hydrostatic coefficient

L: length

$\vec{n}$: normal vector

$\vec{q}$: heat flux

$[\mathbf{q_\alpha}]$: second-order internal state variable tensor

Q: heat

r: heat source

s: specific entropy density

S: entropy

t: time

$\vec{t}$: stress vector

T: temperature

v: volume

$\vec{v}$: speed

W_{ext}: external work

$\{x\}$: eigenvector

β: constant

γ: constant

ε: strain

$[\dot{\boldsymbol{\varepsilon}}]$: second-order strain rate tensor

$[\boldsymbol{\varepsilon}^\mathbf{e}]$: second-order elastic strain tensor

$[\boldsymbol{\varepsilon}_\alpha^\mathbf{a}]$: second-order anelastic strain tensor

ζ_α: constitutive parameter of $\underline{\underline{[\mathbf{B}_\alpha]}}$

ς_α: constitutive parameter of $\underline{\underline{[\boldsymbol{\eta}_\alpha^\mathbf{a}]}}$

$\underline{\underline{[\boldsymbol{\eta}_\alpha^\mathbf{a}]}}$: fourth-order anelastic tensor

λ : eigenvalue
μ_α : constitutive parameter of $\underline{\underline{[\mathbf{A}_\alpha]}}$
μ_α : constitutive parameter of $\underline{\underline{\left[\mathbf{H}_\alpha^a\right]}}$
ν : Poisson's ratio
ξ_α : constitutive parameter of $\underline{\underline{[\mathbf{b}_\alpha]}}$
ρ : mass density
σ : stress
$[\boldsymbol{\sigma}]$: second-order stress tensor
φ : dissipative potential
$[\chi]$: matrix of eigenvectors
ψ : Helmholtz free energy

7. References

Bazarov, I. P. (1964). *Thermodynamik*, Deutscher Verlag der Wissenschaften, Berlin.

Coussy, O. (1995). *Mechanics of porous continua*, Wiley, 0-471-95267-2, Chichester.

Coussy, O. (2010). *Mechanics and physics of porous solids*, Wiley, 9780470710388 (online), 9780470721353 (cloth), 0470721359 (cloth), Chichester, West Sussex, U.K.

D'Amours, G. (2004). Développement de lois constitutives thermomécaniques pour les matériaux à base de carbone lors du préchauffage d'une cuve d'électrolyse. Mechanical engineering. Quebec, Canada, Laval University. Ph.D. Thesis, 223 pp.

D'Amours, G., Fafard, M., et al. (2003). "Mechanical behavior of carbon cathode: Understanding, modeling and identification." Light Metals 2003: 633-639.

Fafard, M., Boudjelal, M. T., et al. (2001). "Three-dimensional viscoelastic model with nonconstant coefficients." (ASCE) Journal of Engineering Mechanics 127(8): 808-815.

Goulet, P. (2004). Modélisation du comportement thermo-électro-mécanique des interfaces de contact d'une cuve de Hall-Héroult. Quebec, Canada, Laval University. Ph.D. Thesis, 265 pp.

Haslach, H. W. (2009). "Thermodynamically consistent, maximum dissipation, time-dependent models for non-equilibrium behavior." International Journal of Solids and Structures 46(22-23): 3964-3976.

Holzapfel, G. A. (1996). "On large strain viscoelasticity: Continuum formulation and finite element applications to elastomeric structures." International Journal for Numerical Methods in Engineering 39(22): 3903-3926.

Kinsler, L. E. (2000). *Fundamentals of acoustics* (4th), Wiley, 0-471-84789-5, New York.

Kreyszig, E. (2006). *Advanced engineering mathematics* (9th), Wiley, 0-471-72897-7. 978-0-471-72897-9, Hoboken, N.J.

Kuiken, G. D. C. (1994). *Thermodynamics of irreversible processes - Applications to diffusion and rheology*, J. Wiley & sons, 0-471-94844-6, Chichester New York Brisbane.

Lemaitre, J. and Chaboche, J.-L. (1990). *Mechanics of solid materials*, Cambridge University Press, 0521328535, Cambridge.

Li, C. Y. and Lua, J. (2009). "A hyper-viscoelastic constitutive model for polyurea." Materials Letters 63(11): 877-880.

Lion, A., Liebl, C., et al. (2010). "Representation of the glass-transition in mechanical and thermal properties of glass-forming materials: A three-dimensional theory based on thermodynamics with internal state variables." Journal of the Mechanics and Physics of Solids 58(9): 1338-1360.

Marceau, D., Pilote, S., et al. (2011). "Advanced numerical simulation of the thermo-electro-chemo-mechanical behaviour of Hall-Héroult cells under electrical preheating." Light Metals 2011: 1041-1046.

Mase, G. T. and Mase, G. E. (1999). *Continuum mechanics for engineers* (2nd), CRC Press, 0849318556 (acid-free paper), Boca Raton, Fla.

Ogden, R. W. (1972). Large Deformation Isotropic Elasticity - On the Correlation of Theory and Experiment for Incompressible Rubberlike Solids, Proceedings of the Royal Society of London, Series A, Mathematical and Physical Sciences, Vol. 326, No 1567, London, February 1972.

Picard, D. (2007). Modélisation et caractérisation du fluage/relaxation de matériaux à base de carbone présents dans les revêtements cathodiques des cuves d'électrolyse de l'aluminium. Quebec, Canada, Laval University. Ph.D. Thesis, 212 pp.

Picard, D., D'Amours, G., et al. (2008). "Constitutive laws of carboneous materials of aluminium electrolysis cell: Current knowledge and future development." Light Metals 2008: 987-992.

Picard, D., Fafard, M., et al. (2008). "Room temperature long-term creep/relaxation behaviour of carbon cathode material." Materials Science and Engineering a-Structural Materials Properties Microstructure and Processing 496(1-2): 366-375.

Picard, D., Fafard, M., et al. (2008). "Three-dimensional constitutive creep/relaxation model of carbon cathode materials." Journal of Applied Mechanics-Transactions of the ASME 75, DOI: 10.1115/1.2840044, Artn 031017.

Reddy, J. N. (1993). *An introduction to the finite element method* (Second edition), McGraw-Hill, 0-07-051355-4, Boston, MA.

Richard, D. (2004). Aspects thermomécaniques de la modélisation par éléments finis du préchauffage électrique d'une cuve de Hall-Héroult : lois constitutives, conception orientée-objet et validation. Quebec, Canada, Laval University. Ph.D. Thesis, 183 pp.

Richard, D., D'Amours, G., et al. (2005). "Development and validation of a thermo-chemo-mechanical model of the baking of ramming paste." Light Metals 2005: 733-738.

Richard, D., Goulet, P., et al. (2006). "Thermo-chemo-mechanical modeling of a Hall-Heroult cell thermal bake-out." Light Metals 2006, Vol 3: Carbon Technology: 669-674.

Schapery, R. A. (1997). "Nonlinear viscoelastic and viscoplastic constitutive equations based on thermodynamics." Mechanics of Time-Dependent Materials 1: 209-240.

Sørlie, M. and Øye, H. A. (2010). *Cathode in Aluminium Electrolysis* (3rd edition), Aluminium-Verlag Marketing & Kommunikation GmbH, 978-3-87017-294-7, Düsseldorf, Germany.

Sun, Y., Forslund, K. G., et al. (2004). "3-D modelling of thermal and sodium expansion in soderberg aluminium reduction cells." Light Metals 2004: 587-592.

Tessier, J., Duchesne, C., et al. (2011). "Multiblock monitoring of aluminum reduction cells performance." Light Metals 2011: 407-412.

Tessier, J., Duchesne, C., et al. (2010). "Investigation of the Impact of Preheating, Start-up and Early Operation on Potlife." Light Metals 2010: 1051-1055.

Valanis, K. C. (1972). *Irreversible thermodynamics of continuous media - internal variable theory*, Springer, 0-387-81127-3, New York.

13

Hydrogen Bond Interactions Between Water Molecules in Bulk Liquid, Near Electrode Surfaces and Around Ions

Abhishek Rastogi, Amit K. Ghosh and SJ Suresh
Unilever R & D Bangalore, Whitefield, Bangalore, India

1. Introduction

Recently, a panel of experts, nominated by IUPAC, proposed the following tentative definition for the hydrogen bond: "The hydrogen bond is an attractive interaction between a group X-H and an atom or group of atoms Y in the same or different molecule(s), where there is evidence of bond formation" (Arunan, 2007). The energy of H-bond (~5 Kcal/mole of H-bonds) is intermediate between those of Van der Waals interaction (~0.3 Kcal/mole) and covalent (~100 Kcal/mole) chemical bonds (Stillinger, 1980). Since the energy of H-bond is of the order of a few KTs, thermal energy constantly acts to disrupt H-bonds. One can thus consider the energetics to drive formation of H-bonds, and entropic factors arising from thermal energy to break H-bonds. The result is a time-varying distribution of H-bonds among the different donor-acceptor pairs in the system. Combination of femtosecond 2D IR spectroscopy and molecular dynamics simulations demonstrated that the vast majority of average numbers of H-bonds are part of a H-bonded well of attraction and virtually all molecules return to a H-bonding partner within 200 fs (Eaves et al., 2005). Despite this continuous dynamics, fluctuation in the total number of H-bonds in a system containing a large number of molecules is quite small. Most simulation models suggest that a given H atom in water is H-bonded for 85-90% of the time (Bakker & Skinner, 2010).

H-bonds have been a subject of intense research over several decades owing to the enormous role they play on several physico-chemical properties of interest. As quoted by Buckingham et al., "The concept of the H-bond is a century old but youthful because of its vital role in so many branches of science and because of continued advances in experiment, theory and simulation" (Buckingham et al., 2008). The significance of H-bonds can be best understood by comparing the physical state of water and methane, both of similar size; at room temperature, while methane is supercritical, water exists in liquid state, making it possible for life to sustain on earth. The anomalous expansion of water at 4°C makes it possible for marine life to exist. The high dielectric constant of water opens up the entire field of electrochemistry. The internal structure of water is largely responsible for self-assembly of surfactants, leading to a wide array of liquid crystalline phases. H-bonds are largely responsible for preserving the structure/conformation of several life-supporting biological molecules such as DNA, RNA and proteins in aqueous solutions.

Despite the vast advances made in the recent past, H-bonding in liquid water continues to be one of the most challenging topics to understand. Each water molecule possesses two proton donors and two proton acceptors (lone pairs of electrons). X-ray and neutron diffraction studies reveal a three-dimensional network of H-bonds with a local preference for tetrahedral geometry (Narten & Levy, 1971). *Ab initio* calculations on the water dimer (Morokuma & Pedersen, 1968; Morokuma & Winick, 1970; Yamabe & Morokuma, 1975) reveal the important contribution of electrostatic interactions on H-bonding. Xenides et al (Xenides et al., 2006) have suggested the following parameters for H-bonded water molecules: distance between two oxygen atoms lies between 0.25-0.35 nm, H-bond distance lies between 0.15-0.25 nm, and the angle of the O-H···· O configuration $\theta \geq 100°$. It is, however, important to note that there is no consensus in the literature on the precise values of these parameters; for example, Kumar et al. suggest that $\theta \geq 130°$ (Kumar et al., 2007).

In the absence of a universally agreed mathematical definition for the H-bond, it is not surprising to note that H-bonding between water molecules under the influence of an external field is an even less understood topic. This is nevertheless an extremely important topic given that we commonly observe water being subject to external fields such as that imposed by an uncharged solid wall, a liquid-liquid interface, an electrode surface, a nano-pore or an ion. In-fact, a detailed understanding of this topic probably holds the key in resolving some of the most difficult problems in Chemical Physics, such as: What is the origin of Hofmeister Series? What is the mechanism of gating of ion channels in biological membranes? What is the role played by interfacial water on electrochemical reaction pathways?

The present Chapter aims to elucidate our current understanding of H-bonding in liquid water, and is organized as follows. We start with a description of H-bonding in liquid water in the absence of any external field (Section 2), and then proceed to understand the influence of external fields generated near an electrode surface (Section 3) and around an ion (Section 4).

2. Structure of bulk water

We start with a qualitative discussion on the internal structure of water, followed by a compilation of literature data on a specific quantitative feature of the H-bond, namely the average number of H-bonds per water molecule.

Bernal and Fowler (Bernal & Fowler, 1933) suggested that molecules in liquid water are arranged in a tetrahedral manner, with each water molecule forming approximately four H-bonds with its nearest neighbors. Wall and Horing (Wall & Hornig, 1965) observed the Raman-scattering motions of the HDO molecules in H_2O or D_2O and concluded that liquid water does not have any significant fraction of unbonded molecules. Narten and Levy (Narten & Levy, 1969) supported the viewpoint that water molecules in liquid state are arranged in a predominantly tetrahedral geometry based on the observation that the mean separation between nearest neighbor molecules in ice and water are only marginally different. Stillinger and Rahman (Stillinger & Rahman, 1974) also suggested a similar structure for liquid water and additionally concluded that it contained a large proportion of broken H-bonds. Narten et al. (Narten et al., 1982) showed that most water molecules are connected to their nearest neighbors through nearly straight H-bonds. Gorbaty and Demianets (Gorbaty & Demianets, 1983) and Hoffmann and Conradi (Hoffmann & Conradi, 1997) showed, through independent studies, that the disorder in H-bond structure in liquid

water increases with increasing temperature. Schwegler et al. (Schwegler et al., 2000) showed that the application of external pressure on pure water weakens H-bonds. Lee and Tuckerman (Lee & Tuckerman, 2006) studied the structural property of liquid water using a Car-Parrinello ab initio molecular dynamics (CPAIMD) simulations combined with Kohn-Sham density functional theory and BLYP exchange correlation functional for electronic structure, and found that more than 50% of water molecules are bonded in a tetrahedral geometry. Leetmaa et al. (Leetmaa et al., 2008) confirmed the tetrahedral arrangement with 74% double H-bond donors (DD) and 21% single donors (SD). Lehmann et al. (Lehmann et al., 2009) used quantum cluster theory and found dominance (75%) of two-bonded water molecules at room temperature.

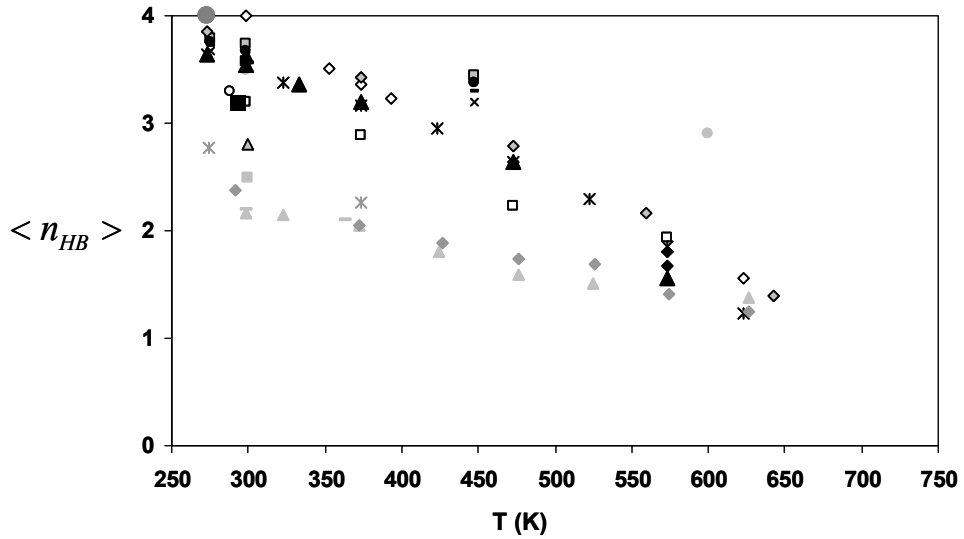


Fig. 1. Average number of H bonds per water molecule at various temperatures, as extracted using different techniques in the past. ▲X ray diffraction study at 100 MPa (Gorbaty & Demianets, 1983); ◇NMR (Hoffmann & Conradi, 1997); ●Water electronic study (Schwegler et al., 2000); ⁄ Quantum cluster equilibrium method (Lehmann et al., 2009); ΔTIP4P (Prendergast & Galli, 2006); + CPAIMD (Lee & Tuckerman, 2006); ♦ Neutron diffraction (Soper et al., 1997); ✳ IR (Luck, 1967); ○TEY-NEXAFS (Smith et al., 2004); ◬SXEER (Myneni et al. 2002); ◆IR at 250 atm (Bondarenko & Gorbaty, 1991); ▲Latent Heat (Haggis et al., 1952); ◇Statistical mechanics (Suresh et al., 2000); ▪Electronic structure study (He et al. 2007); - TIP4P (Nieto-Draghi, 2003); ×TIP5P (Nieto-Draghi et al., 2003); ▫SPC/E (Nieto-Draghi et al., 2003); ● DEC (Nieto-Draghi et al., 2003); □ Monte Carlo (Kalinichev & Bass, 1997); ■ *Ab initio* (Kuo & Mundy, 2004); ■ FPMD (Hetenyi et al., 2004); –X-ray absorption (Wernet et al., 2004).

Figure 1 shows the temperature variation of the average number of H-bonds per water molecule, $< n_{HB} >$. Two points are worth noting. One, irrespective of the source of literature, $< n_{HB} >$ is seen to decrease with increasing temperature. Two, there is a large scatter in data points; for example, $< n_{HB} >$ at room temperature varies from 2 to 4. This

scatter can be attributed to a number of factors, such as differences in the definition of the H-bond, limitations in experimental/simulation techniques, and ambiguities in the interpretation of experimental data in relation to H-bond stoichiometry. Such an in-depth analysis of the factors involved is not within the scope of the present Section. However, in order to provide an appreciation of the ongoing debate in this area of research, we will bring out the key features of two landmark papers published in *Science*, one suggesting that $< n_{HB} >$ ~3.3 and the other suggesting a different value of 2.

We first start with a discussion on the paper published by Smith et al. (Smith et al., 2004) which concluded that $< n_{HB} >$ ~3.3. The authors used oxygen K-edge X-ray absorption technique, wherein a core electron was excited to an unoccupied electronic state. The electronic character of the unoccupied states is known to be sensitive to the local geometric structure. The authors recorded area-normalized spectra of water at 254 K and 288 K. The intensity of pre-edge region (~535 eV), which is a signature of distorted H-bonds configurations, increased with increasing temperature. On the other hand, the intensity of post-edge region (~541 eV), which is a signature of stronger and fully coordinated ice-like bonds, decreased. The authors assumed the relative populations of post-edge (I_{post}) and pre-edge (I_{pre}) intensities to be a function of absolute temperature only, and proposed that a plot of $\ln(I_{post}/I_{pre})$ versus 1/T would yield a straight line with a slope that is proportional to the average difference in energy between the two classes, and determined the rearrangement energy between two classes of H-bonding distributions to be 1.5 ± 0.5 kcal/mol. This small energy difference indicates that molecules contributing to pre-edge intensity are only slightly distorted compared to those contributing to post-edge intensity. The average energy of a fully formed ice-like H bond is known to be ~5.5 kcal/mol (Kuo & Mundy, 2004; Stillinger, 1980). Therefore, the authors attributed the difference in energy (1.5 kcal/mol) between the two H-bonding distributions to a loss of 27 ± 9% of average H-bond energy. Later, Nilsson et al. (Nilsson et al., 2005) questioned the quality of the temperature dependent X-ray spectra presented by Smith et al. (Smith et al., 2004). They pointed that three different sets of measurements displayed an energy difference of approximately 1.2, 1.5 and 1.8 kcal/mol between the two different species, and attributed this lack of reproducibility to energy dependent nonlinear effects in X-ray absorption spectrum measurements. In response to Nilsson comments, Smith et al. (Smith et al., 2005) showed that under constant conditions (e.g., jet size, collection geometry), the X-ray absorption spectrum is highly reproducible.

We now focus our attention on another paper which suggested a different value for $< n_{HB} >$. Wernet et al. (Wernet et al., 2004) employed X-ray absorption (XAS) and X-ray Raman (XRS) spectroscopic techniques to understand the H-bond configurations in bulk ice, ice surface, NH_3 terminated ice surface and bulk liquid water. The spectra were divided into three regions: pre-edge (~535 eV), main edge (537-538 eV), and post-edge (540-541 eV). The bulk ice spectrum was dominated by the intensity in post-edge region and showed a weak main edge structure. Both the surface ice and liquid water spectra had a peak in the pre-edge region, a dominant main edge, and lesser intensity compared with bulk ice in the post-edge region. Termination of ice surface with NH_3 entails a coordination of free O-H groups and caused the pre-edge peak to vanish. It was observed that liquid water spectra closely resembles that of ice surface, which consists of one strong and one non-bonded or weakly-

bonded O-H group, but is very different from that of bulk ice. On the basis of these findings, the authors concluded that liquid water consists of two structural species, one with two H-bonded (at one acceptor and one donor site) and another tetrahedrally coordinated. A theoretical analysis of XAS and XRS spectra, based on density functional theory (DFT) with a small model cluster of 11 molecules, revealed that each molecule has on average 2.2 ± 0.5 H-bonds at 25°C and 2.1 ± 0.5 at 90°C temperature.

3. Structure of water near electrode surfaces

Structure and orientation of water molecules adjacent to charged surfaces play an important role in surface science, electrochemistry, geochemistry and biology (Thiel & Madey, 1987; Henderson, 2002; Guidelli & Schmickler, 2000). Several force fields are operational in such situations. Molecular dynamic simulations (Segura et al., 1997) reveal entropy driven piling-up effect near even an uncharged wall. Additional presence of charge on surface polarizes water molecules. The presence of ions (H^+, OH^-) in the liquid phase further induces formation of an electrical double layer within which the electric field decays with distance from the surface. Advanced experimental and molecular simulation techniques are just beginning to shed light on the influence of a charged electrode surface on the various aspects of the internal structure of water, such as H-bonding, density and dipolar alignment. It is the purpose of this Section to discuss the current understanding of this topic. More specifically, we focus on studies revealing two opposing viewpoints; one suggesting that the H-bond structure of water near a charged surface is significantly disrupted, and the other concluding the opposite, namely that the H-bond network near the surface is largely intact.

Toney et al. (Toney et al., 1994) studied the distribution of water molecules perpendicular to a charged silver (111) electrode interface in NaF solution using X-ray scattering technique. An interface at a single crystal creates additional scattering to the Bragg peaks of crystal, which permits the determination of surface structure and water distribution. Figure 2 shows the distribution of oxygen atoms with distance from the electrode surface. Using Gaussian functions to fit the oxygen distribution function, the first layer density was calculated to be 1.1 and 1.8 water molecules per Ag atom, corresponding to an applied voltage of -0.23 V and +0.52 V, respectively. In contrast, bulk water had a density of ~0.8 water molecules per Ag atom. The conclusion that water density is significantly altered near a charged electrode was confirmed by Danielewicz-Ferchmin (Danielewicz-Ferchmin & Ferchmin, 1996) as well, albeit using a different approach. A simulation study (Zhu & Robinson, 1991) with SPC-FP water model (simple point charge model with flexible bonds and polarization) also showed that water density near a charged surface is higher than that in bulk and increases with field intensity. Moreover, they found that the length of O-H bond (of water) near the electrode surface is smaller than in bulk, indicating weaker H-bonds. Suresh (Suresh, 2007) arrived at a similar conclusion using a statistical thermodynamic model, and showed that the average number of H-bond per molecule near the charged surface decreases from 2.8 at zero electric field to 2 at $E = 2 \times 10^9$ V/m.

We now focus our attention on another set of papers (Schweighofer et al., 1996; Torrie et al., 1988; Yeh & Berkowitz, 2000) that reached a different conclusion, namely that the H-bond structure of water near a charged electrode surface is largely intact. Schweighofer et al. (Schweighofer et al., 1996) performed molecular simulations with SPC/E water molecules

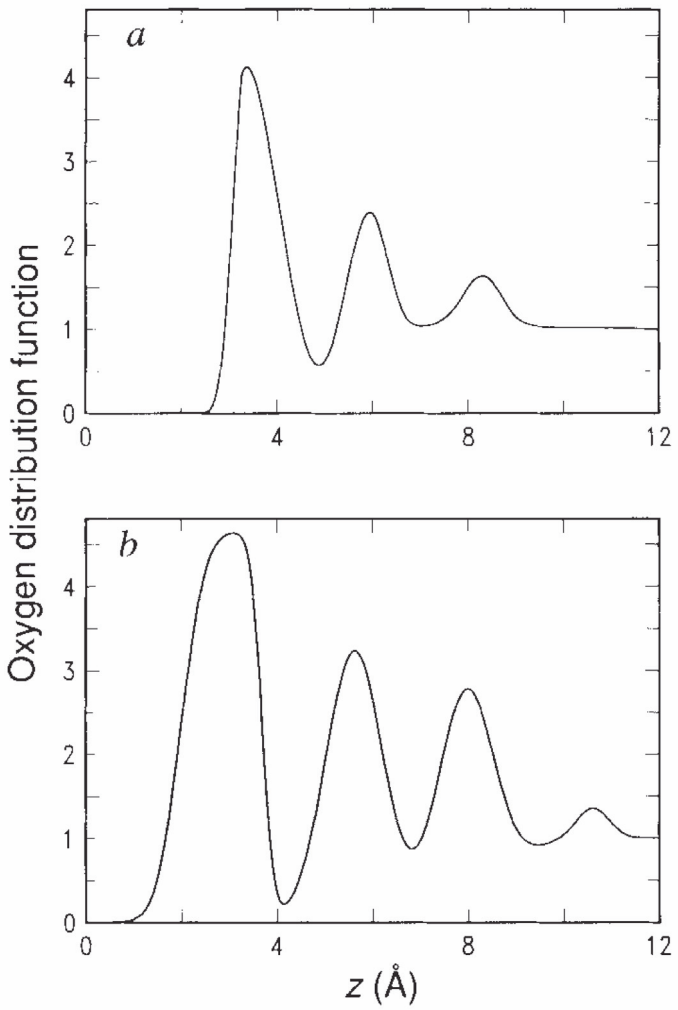


Fig. 2. Oxygen distribution functions of water molecules near silver Ag (111) electrode with potentials of (a) -0.23 V of the p.z.c. and (b) +0.52 V of the p.z.c. Reproduced with permission from Toney, M. F. et al., *Nature*, 1994, 368, 444. © 1994 Nature Publishing Group.

contained between two parallel Ag(111) surfaces with charge densities fixed at 0.0, 8.85, and 26.55 μC/cm^2. They did not observe increase in water density near the electrode even with the highest charge density (+26.5 μC/cm^2); rather, it was found to decrease. Authors explained this observation to strong polarization of water molecules due to the applied electric field resulting in some of the oxygen atoms associated with water molecules to desorb from the surface and move into subsequent layers in order to keep intact the H-bond network. Torrie et al. (Torrie et al., 1988) showed that the H-bond structure near the electrode is resistant to surface charges as high as 17.5 μC/cm^2, apparently because such a

strong field can not compete with the stronger intermolecular forces of water-like models. Yeh (Yeh & Berkowitz, 2000) studied water density near a charged silver electrode surface using polarizable point charge model and found that water density near the surface was not affected by surface charge. Xia and Berkowitz (Xia & Berkowitz, 1995) performed molecular simulations with SPC/E modeled water lamina embedded between two Pt (100) walls, charged with 0, 8.85, 26.55 and 35.40 μC/cm² on the left wall and values opposite in sign on the right wall. Figure 3 shows the density profile of O and H atoms of water as varying from positively charged electrode to negatively charged electrode.

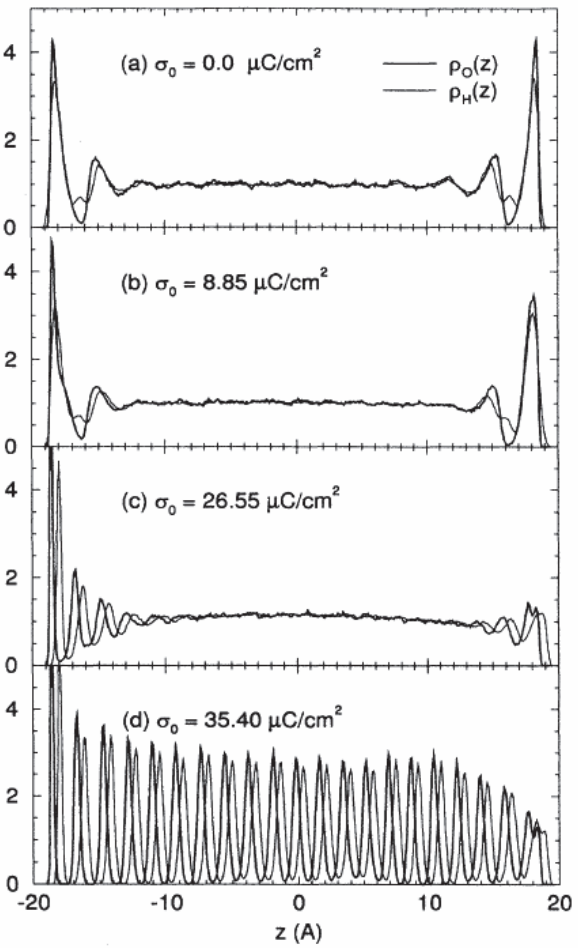


Fig. 3. Density profile of oxygen ρ_O and hydrogen ρ_H atoms for systems with (a) no surface charge density on Pt surface, (b) surface charge density of 8.85 μC/cm², (c) 26.55 μC/cm², and (d) 35.40 μC/cm². The positively charged surface is on the left, the negatively charged surface is on the right. Reproduced with permission from Xia, X.; et al. *Physical Review Letters*, 1995, 74, 3193. © 1995 The American Physical Society.

There was no significant change in density profile up to 8.85 μC/cm² surface charge. When the surface charge density was increased to 26.66 μC/cm², which is close to the surface charge density in experiments performed by Toney et al. (Toney et al., 1994) four distinct water layers were observed near the positively charged surface. When the charge density on electrode was increased to 35.4 μC/cm², water restructured itself in order to adapt to the new environment and eventually crystallized into domains of cubic ice. In other words, the H-bond network not only survives in electric field, but is also responsible for the observed structural changes. Otani et al. (Otani et al., 2008) performed *ab initio* molecular dynamics simulations on water/Pt interface with a negative bias up to -0.81 V. They observed only a small increase in interfacial water density and found that for bias and unbiased conditions, the average number of H-bonds per water molecule in bulk region as well as in contact layer was almost equal. This implied that water molecules rotate their molecular axes in response to the external electric field, but retain their tetrahedral configuration by rearranging the H-bond network. Compared to bulk, however, near the charged interface, the average number of H-bonds per water molecule was lesser.

4. Structure of water near ions

Hofmeister first established that different ions have different efficiency at salting-out egg-white protein (Hofmeister, 1888). The sequence of ions, based on their effectiveness towards enhancing or diminishing the solubility of proteins, is termed as Hofmeister series. This series is generally written as (Marcus, 2009).

Anions:

Kosmotrope (Structure maker) Chaotrope (Structure breaker)

$$CO_3^{2-} > SO_4^{2-} > S_2O_3^{2-} > H_2PO_4^{2-} > OH^- > F^- > HCO_2^- > Cl^- > Br^- > NO_3^- > I^- > SCN^-$$

Cations:

Kosmotrope (Structure maker) Chaotrope (Structure breaker)

$$(CH_3)_4N^+ > (CH_3)_2NH_2^+ > K^+ \sim Na^+ > Cs^+ > Li^+ > NH_4^+ > Mg^{2+} > Ca^{2+} > C(NH_2)^{3+}$$

The molecular origin of salting-out or salting-in effects is not fully understood. But broadly, the current view is that addition of salts alters the internal structure and hence the solubility power of water. In this context, ions are classified in two groups, namely "structure makers" and "structure breakers" (Cox & Wolfenden, 1934). Implicit in this conjecture is that ions influence the long-range structure of water. Whether this conjecture is indeed borne out in experiments is one of the topics of debate in the literature. In this Section, we discuss key aspects of papers on both sides of this debate.

Leberman and Soper (Leberman & Soper, 1995) showed that ions can induce a change in water structure equivalent to that caused by the application of high pressure, and that the extent of this effect is ion-specific. Based on neutron diffraction patterns, the authors determined distribution distances $\Delta g_{HH}(r)$ between water protons in aqueous solutions of salts $(NH_4)_2SO_4$, NH_4Cl, Na_2SO_4, and NaCl), and in pure water at high pressure. Figure 4 shows that the $\Delta g_{HH}(r)$ in all these cases are of qualitatively similar shape. The general trend is a negative region near r = 2 Å, a positive region near r = 3.2 Å and a set of oscillations with a period of ~3 Å. Under ambient conditions, the intermolecular HH pair correlation function for pure water consists of two peaks, one at ~2 Å and another at ~3.8 Å,

with a minimum between them at r = 3.0 Å. In ionic solutions under normal pressure and pure water at high pressure, peaks at ~2.4 Å and 3.8 Å are apparently diminished and a minimum at ~3.0 Å fills in. These results strongly suggest that ions disrupt the water structure comparable in extent to that caused by the application of high pressure to pure water. The possibility of pressure and salt leading to similar effects on the structure of water has also been investigated using molecular simulation techniques. Holzmann et al. performed molecular dynamics simulations of aqueous NaCl solutions, and arrived at three broad conclusions (Holzmann et al., 2007). One, the H-bond network is modified well beyond the first hydration shell. Two, an analysis of free water distribution showed that the effect of salt and pressure might be considered as "two sides of the same coin". Three, the authors cautioned against using the pressure/salt equivalence as a sole factor influencing solubility of biomolecules; adsorption/desorption of ions might also have to be accounted for.

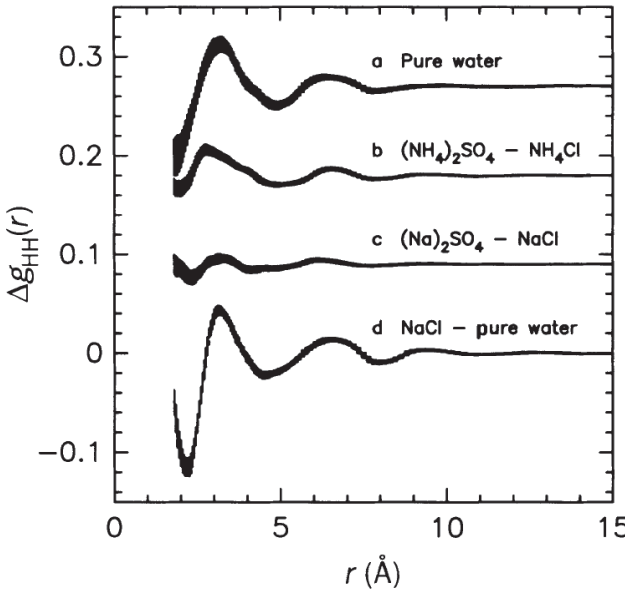


Fig. 4. Changes in the HH pair correlation function: curve a, for pure water for an applied density increase of 0.0017 molecules per Å^3 above ambient density (0.0334 molecules per Å^3); curve b, between a 2 m aqueous solution of Na_2SO_4 and a 4 M solution of NH_4Cl; curve c between 2 M aqueous solution of Na_2SO_4 and 4 m solution of NaCl; and curve d, between a 4 m solution of NaCl and pure water. Reproduced with permission from Leberman, R. et al.; *Nature*. 1995, 378, 364. © 1995 Nature Publishing Group.

Chandra (Chandra, 2000) investigated the specific role of ions on H-bonds between water molecules using molecular dynamics. The systems chosen were NaCl and KCl in water at various concentrations (from 0M to 3.35M). Water molecules were modeled by the extended simple point charge (SPC/E) potential and ions were modeled as charged Lennard-Jones particles. For analyzing the hydrogen bond breaking dynamics, the author calculated the time correlation functions $S_{HB}(t)$ and $S^{(d)}{}_{HB}(t)$. $S_{HB}(t)$ describes the probability that an initially

hydrogen-bonded pair remains bonded at all times up to time t, and $S^{(d)}_{HB}(t)$ describes the dynamics of hydrogen-bond breaking due to translational diffusion. Figure 5A represents the decay of $S_{HB}(t)$ and $S^{(d)}_{HB}(t)$ for the various solutions studied. Decay of both $S_{HB}(t)$ and $S^{(d)}_{HB}(t)$ is found to accelerate with increase in ion concentration. This indicates that the H-bonded water pair in ionic solutions is less strongly held compared to that in pure water. The author also calculated the oxygen-oxygen and oxygen-hydrogen potentials of mean force (pmf) between two neighboring water molecules (figure 5B). The well depth of the pmf decreases with increasing ion concentration, suggesting weakening of the effective strength of water-water hydrogen bonds. This effect is more pronounced in case of NaCl solutions because of smaller size and stronger hydration of sodium ions compared to that of potassium ions.

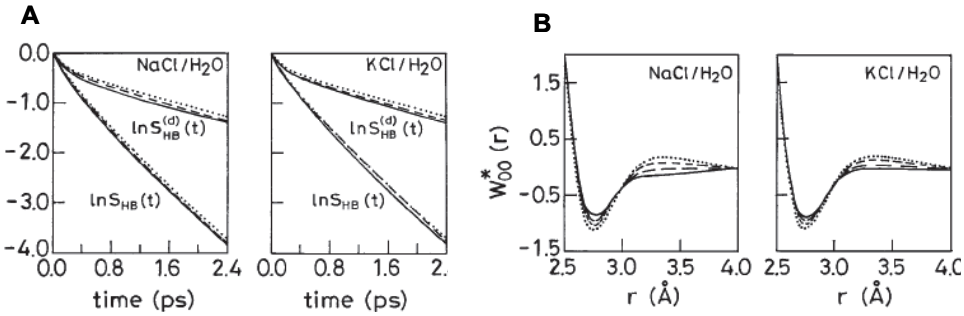


Fig. 5. (A) The time dependence of the logarithm of $S_{HB}(t)$ and $S^{(d)}_{HB}(t)$ for NaCl and KCl solutions. The dotted, dashed, and solid curves are for 0.0*M*, 2.2*M*, and 3.35*M* solutions. (B) The distance dependence of the oxygen-oxygen potential of mean force between two neighboring water molecules in NaCl and KCl solutions [$W^*_{OO}(r) = W_{OO}(r)/k_BT$]. The dotted, short-dashed, long-dashed, and the solid curves are for 0.0*M*, 0.88*M*, 2.2*M*, and 3.35*M* solutions. Reproduced with permission from Chandra, A., *Physical Review Letter*. 2000, 85, 768. © 2000 The American Physical Society.

Hribar et al. (Hribar et al., 2002) employed a two-dimensional MB model, in which each water molecule was represented as a two-dimensional disk that interacted with other water molecules through a Lennard-Jones (LJ) interaction and an orientation-dependent H-bond interaction. Figure 6A shows that the average number of H-bonds per water molecule around the first shell of smaller cations such as Li^+ and Na^+ (kosmotropes) is lesser than that in bulk water, while the corresponding number for molecules around larger ions such as K^+, Rb^+ and Cs^+ are higher. The molecular picture giving rise to this conclusion is shown in Figure 6B. The number of H-bonds per water molecule in the first shell around an ion is governed by two competing ordering effects, one induced by electrostatic interaction with the ion and the other by water-water H-bond interaction. For small ions, the former effect dominates, while for large ions the later effect dominates.

We now turn out attention to papers that concluded the opposite, namely that the presence of ions do not lead to an enhancement or a breakdown of H-bond network in liquid water. This viewpoint, which was originally initiated by Omta et al. (Omta et al., 2003), has profound implications, one of which is that the conjecture of ions being classified as being "structure-makers" and "structure breakers" does not have a molecular basis. The authors

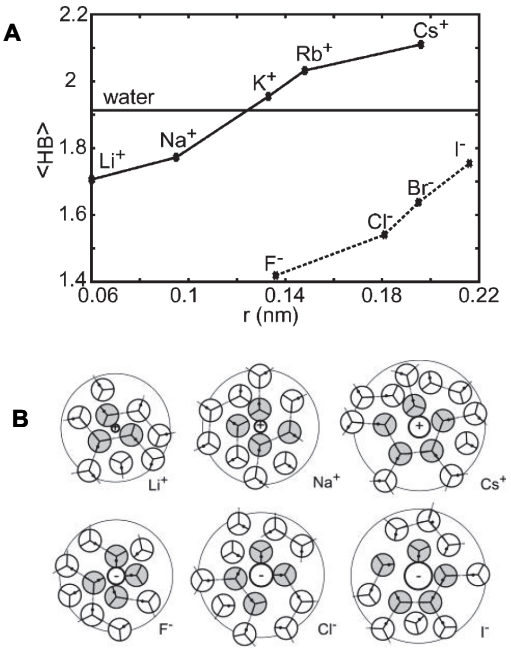


Fig. 6. (A) The average number of the hydrogen bonds per water molecule in the first shell around various ions. (B) Snapshots of waters in the first (shaded) and second shell (white) around an ion (black). Reproduced with permission from Hribar, B. et al., *J. Am. Chem. Soc.* 2002, 124, 12302. © 2002 American Chemical Society.

measured the orientational correlation time of water molecules in $Mg(ClO_4)_2$ solution by using femtosecond pump-probe spectroscopy. The technique was based on measuring the anisotropy parameter (R), which was defined as $[(\Delta\alpha_{\parallel}(\tau) - \Delta\alpha_{\perp}(\tau))/ (\Delta\alpha_{\parallel}(\tau) + 2\Delta\alpha_{\perp}(\tau))]$, wherein $(\Delta\alpha_{\parallel})$ and $(\Delta\alpha_{\perp})$ are absorption changes measured parallel and perpendicular to the pump polarization, and τ is the time delay between pump and probe pulses . To measure R, OH groups in the liquid were anisotropically excited by a linearly polarized pump pulse. The decay time of R represents the orientational correlation time of water molecules, a measure for the stiffness of the H-bond network. The decay of *R* turns out to be independent of the salt concentration (0 M, 1 M, and 3 M $Mg(ClO_4)_2$ in $HDO:H_2O$ solutions), which means the orientational correlation time (τ) for the water-bonded OD groups is not affected by the presence of Mg^{2+} and ClO_4- ions. Moreover, the authors found that the orientational correlation time for salt solutions ($Mg(ClO_4)_2$, $NaClO_4$ and Na_2SO_4) and pure water are not very different, based on which the authors concluded that H-bond interactions between water molecules outside the nearest solvation shell are largely intact.

We now mention two other recent studies whose conclusions are broadly in line with those of Omta et al. (Omta et al., 2003). Guardia et al. (Guardia et al., 2006) carried out extensive molecular dynamics simulations of aqueous alkali metal and halides at ambient and supercritical conditions to explore the effects of ions on the intermolecular connectivity of water in the close vicinity of solutes. From a dynamical perspective, the most relevant feature was that the lifetimes of hydrogen bond do not seem to be affected in an appreciable manner by the presence of ions. More recently, in 2007, Smith et al. (Smith et al., 2007)

compared experimental Raman spectral measurements with classical Monte Carlo simulations and concluded that the change in vibrational spectrum of water by the addition of potassium halides is a direct result of the electric fields that anions exert on adjacent H atoms, and that the halide ions induce only minor HB distortions beyond the adjacent shell of coordinating OH groups.

5. Summary

Given the ongoing debates in our current understanding of liquid structure of water even in the absence of external fields, it is not surprising that there is lack of consensus in the literature on even some of the fundamental aspects of H-bond interactions in the presence of an external field. For example, it is not yet clear whether the H-bond structure near a charged electrode surface is disrupted or not. Similarly, E-fields are generally considered to align water dipoles in the direction of the field; however, a recent study (Suresh et al., 2006) has reported the additional presence of H-bond stabilized water molecules with their dipoles lying perpendicular to that of field. The fraction of such molecules is relatively small; for every thousand molecules with their dipoles aligned in the direction of field, roughly one was found aligned perpendicular to that of the field. Nevertheless, the role of these "defects" on the transport or solvation properties of water remains to be investigated. Whether ions induce long-range changes in the structure of water is still an open question (Bakker, 2008). The study of confined water molecules is equally, if not more interesting. The thermodynamic properties of confined water are generally considered to be different from those of bulk water; however, what gives rise to these differences is still an open question. A recent study (Han et al., 2009) showed reduced extent of H-bonding in a quasi-two-dimensional hydrophobic nanopore slit as compared to that in bulk water, partly due to geometric constraints imposed by confined geometries on the tetrahedral bonding patterns in water; however, the characteristics of H-bond dynamics, such as the Arrhenius dependence of the average H-bond lifetime, have been reported to be largely preserved.

In the above context, it is clear that the structure of water in external fields is a topic that requires much more detailed investigations before unambiguous conclusions can be reached. While this endeavor would require the development of advanced techniques, it is equally if not more important that the scientific community agrees on a single mathematical definition of what constitutes a H-bond. It is hoped that this clarity will help in reducing the confusion prevailing today in relation to design, implementation and interpretation of experimental/simulation data related to the internal structure of water.

6. Reference

Arunan, E. *Current Science,* Vol. 92, 2007, pp. 17-18

Bakker, H. J. *Chemical Reviews*, Vol. 108, 2008, pp. 1456-1473

Bakker, H. J. & Skinner, J. L. *Chemical Reviews*, Vol. 110, 2010, pp. 1498-1517

Bernal, J. D. & Fowler, R. H. *Journal of Chemical physics*, Vol. l, 1933, pp. 515-548

Bondarenko, G. V. & Gorbaty, Yu. E. *Molecular Physics,* Vol. 74, 1991, pp. 639-647

Buckingham, A. D.; Bane, J. E. D. & McDowell, S. A. C. *Chemical Physics letter*, Vol. 463, 2008, pp. 1-10

Chandra, A. *Physical Review Letter*, Vol. 85, 2000, pp. 768-771

Cox, W. M. & Wolfenden, J. H. *Proceedings of the Royal Society A*, Vol. 145, 1934, pp. 475-488

Danielewicz-Ferchmin, I. & Ferchmin, A. R. *Journal of Physical Chemistry*, Vol. 100, 1996, pp. 17281-17286

Eaves, J. D.; Loparo, J. J.; Fecko, C. J.; Roberts, S. T.; Tokmakoff, A. & Geissler, P. L. *Proceedings of the National Academy of Science*, Vol. 102, 2005, pp. 13019-13022

Gorbaty, Yu. E. & Demianets, Yu. N. *Chemical Physics Letters,* Vol. 100, 1983, pp. 450-454

Guardia, E.; Laria, D. & Marti, J. *Journal of Physical Chemistry B*, Vol. 110, 2006, pp. 6332-6338

Guidelli, R. & Schmickler, W. *Electrochimica Acta*, Vol. 45, 2000, pp. 2317-2338

Haggis, G. H.; Hasted, J. B. & Buchanan, T. J. *Journal of Chemical Physics*, Vol. 20, 1952, pp. 1452-1465

Han, S.; Kumar, P. & Stanley, H. E. *Physical Review E*, Vol. 79, 2009, pp. 041202/1-5

He, C.; Lian, J. S. & Jiang, Q. *Chemical Physics Letters*, Vol. 437, 2007, pp. 45-49

Henderson, M. A. *Surface Science Reports,* Vol. 46, 2002, pp. 1-308

Hetenyi, B.; Angelis, F. D.; Giannozzi, P. & Car, R. *Journal of Chemical Physics*, Vol. 120, 2004, pp. 8632-8637

Hoffmann, M. M. & Conradi, M. S. *Journal of American Chemical Society*, Vol. 119, 1997, pp. 3811-3817

Hofmeister, F. *Arch. Exp. Path. Pharmakol*, Vol. 24, 1888, pp. 247-260

Holzmann, J.; Ludwig, R.; Geiger, A. & Paschek, D. *Angewandte Chemie International Edition*, Vol. 46, 2007, pp. 8907-8911

Hribar, B.; Southall, N. T.; Vlachy, V. & Dill, K. A. *Journal of American Chemical Society*, Vol. 124, 2002, pp. 12302-12311

Kalinichev,A. G. & Bass, J. D. *Journal of Physical Chemistry A*, Vol. 101, 1997, pp. 9720-9727

Kumar, R.; Schmidt, J. R. & Skinner, J. L. *Journal of Chemical Physics*, Vol. 126, 2007, pp. 204107/1-12

Kuo, I-F. W. & Mundy, C. J. *Science*, Vol. 303, 2004, pp. 658-660

Leberman, R. & Soper, A. K. *Nature*, Vol. 378, 1995, pp. 364-366

Lee, H.-S. & Tuckerman, M. E. *Journal of Chemical Physics*, Vol. 125, 2006, pp. 154507/1-14

Leetmaa, M.; Wikfeldt, K. T.; Ljungberg, M. P.; Odelius, M.; Swenson, J.; Nilsson, A. & Pettersson, L. G. M. *Journal of Chemical Physics,* Vol. 129, 2008, pp. 084502/1-13

Lehmann, S. B. C.; Spickermann, C. & Kirchner, B. *Journal of Chemical Theory and Computation*, Vol. 5, 2009, pp. 1650-1656

Luck, W. A. P. *Discussions of the Faraday Society*, Vol. 43, 1967, pp. 115-127

Marcus, Y. *Chemical Reviews,* Vol. 109, 2009, pp. 1346-1370

Morokuma, K. & Pedersen, L. *Journal of Chemical physics*, Vol. 48, 1968, pp. 3275-3282

Morokuma, K. & Winick, J. R. *Journal of Chemical physics*, Vol. 52, 1970, pp. 1301-1306

Myneni, S.; Luo, Y.; Naslund, L. A.; Cavalleri, M.; Ojamae, L.; Ogasawara, H.; Pelmanschikov, A.; Wernet, Ph.; Vaterlein, P.; Heske, C.; Hussain, Z. & Pettersson, L. G. M. & Nilsson, A. *Journal of Physics: Condensed Matter*, Vol. 14, 2002, pp. L213-L219

Narten, A. H. & Levy, H. A. *Science*, Vol. 165, 1969, pp. 447-454

Narten, A. H. & Levy, H. A. *Journal of Chemical physics*, Vol. 55, 1971, pp. 2263-2269

Narten, A. H.; Thiessen, W. E. & Blum, L. *Science*,Vol. 217, 1982, pp. 1033-1034

Nieto-Draghi, C.; Avalos, J. B. & Rousseau, B. *Journal of Chemical Physics*, Vol. 118, 2003, pp. 7954-7964

Nilsson, A.; Wernet, Ph.; Bergmann, U.; Nordlund, D.; Cavalleri, M.; Odelius, M.; Ogasawara, H.; Nasulund, L.-A.; Hirsch, T. K. & Ojamae, L. *Science*, Vol. 308, 2005, pp. 793a

Omta, A. W.; Kropman, M. F.; Woutersen, S. & Bakker, H. J. *Science,* Vol. 301, 2003, pp. 347-349

Otani, M.; Hamada, I.; Sugino, O.; Morikawa, Y.; Okamoto, Y. & Ikeshoji, T. *Physical Chemistry Chemical Physics*, Vol. 10, 2008, pp. 3609-3612

Prendergast, D. & Galli, G. *Physical Review Letters*, Vol. 96, 2006, pp. 215502/1-4

Schwegler, E.; Galli, G. & Gygi, F. *Physical Review Letters*, Vol. 84, 2000, pp. 2429-2432

Schweighofer, K. J.; Xia, X. & Berkowitz, M. L. *Langmuir*, Vol. 12, 1996, pp. 3747-3752

Segura, C. J.; Chapman, W. G. & Shukla, K. S. *Molecular Physics,* Vol. 90, 1997, pp. 759-772

Smith, J. D.; Cappa, C. D.; Wilson, K. R.; Messer, B. M.; Cohen, R. C. & Saykally, R. J. *Science*, Vol. 306, 2004, pp. 851-853

Smith, J. D.; Cappa, C. D.; Messer, B. M.; Cohen, R. C. & Saykally, R. J. *Science*, Vol. 308, 2005, pp. 793b

Smith, J. D.; Saykally, R. J. & Geissler, P. L. *Journal of American Chemical Society,* Vol. 129, 2007, pp. 13847-13856

Soper, A. K.; Bruni, F. & Ricci, M. A. *Journal of Chemical Physics*, Vol. 106, 1997, pp. 247-254

Stillinger, F. H. & Rahman, A. *Journal of Chemical Physics*, Vol. 60, 1974, pp. 1545-1557

Stillinger, F. H. *Science*, Vol. 209, 1980, pp. 451-457

Suresh, S. J. & Naik, V. M. *Journal of Chemical Physics*, Vol. 113, 2000, pp. 9727-9732

Suresh, S. J., Satish, A. V. & Choudhary, A. *Journal of Chemical Physics,* Vol. 124, 2006, pp. 074506/1-9

Suresh, S. J. *Journal of Chemical Physics*, Vol. 126, 2007, pp. 204705/1-8

Thiel, P. A. & Madey, T. E. *Surface Science Reports*, Vol. 7, 1987, pp. 211-385

Toney, M. F.; Howard, J. N.; Richer, J.; Borges, G.L.; Gordon, J. G.; Melroy, O. R.; Wiesler, D. G.; Yee, D. & Sorensen, L. B. *Nature*, Vol. 368, 1994, pp. 444-446

Torrie, G. M.; Kusalik, P. G. & Patey, G. N. *Journal of Chemical Physics*, Vol. 88, 1988, pp. 7826-7840

Wall, T. T. & Hornig, D. F. *Journal of Chemical Physics*, Vol. 43, 1965, pp. 2079-2087

Wernet, P.; Nordlund, D.; Bergmann, U.; Cavalleri, M.; Odelius, M.; Ogasawara, H.; Naslund, L. A.; Hirsch, T. K.; Ojamae, L.; Glatzel, P.; Pettersson, L. G. M. & Nilsson, A. *Science*, Vol. 304, 2004, pp. 995-999

Xenides, D.; Randolf, B. R. & Rode, B. M. *Journal of Molecular Liquids,* Vol. 123, 2006, pp. 61-67

Xia, X. & Berkowitz, M. L. *Physical Review Letters*, Vol. 74, 1995, pp. 3193-3196

Yamabe, S. & Morokuma, K. *Journal of American Chemical Society*, Vol. 97, 1975, pp. 4458-4465

Yeh, I.-C. & Berkowitz, M. L. *Journal of Chemical Physics*, Vol. 112, 2000, pp. 10491-10495

Zhu, S.-B. & Robinson, G. W. *Journal of Chemical Physics*, Vol. 94, 1991, pp. 1403-1410

14

The Stability of a Three-State Unfolding Protein

Yang BinSheng
Institute of Molecular Science, Key Laboratory of Chemical Biology and Molecular Engineering of Ministry of Education, Shanxi University, Taiyuan

1. Introduction

The mechanism by which proteins fold to their unique native conformations from an initially disorganized form is one of the fundamental problems in molecular biology. In the study of proteins fold or refold standard Gibbs free energy, $\Delta G^0(\mathrm{H_2O})$ is the single most important parameter for quantitating-protein stability and comparing stabilities of closely related proteins. Nearly all theoretical and experimental aspects of protein folding relate in some way to unfolding free energy changes, and much of the current work involving comparisons of mutant proteins is highly dependent on accurate evaluation of unfolding free energy measurements [1].

The history of evaluation of the quantity known as standard Gibbs free energy spans more than four decades, and at least three procedures involving strong solvent denaturation have been used in evaluation of this quantity. The three procedures are known as the transfer model of Tanford, the denaturant binding model, and the linear extrapolation method. Of the three methods employed in evaluation of $\Delta G^0(\mathrm{H_2O})$ the so-called linear extrapolation method appears to enjoy the greatest acceptance. The virtue of the linear extrapolation method over the other two methods resides in the perception that it is more reliable since it gives $\Delta G^0(\mathrm{H_2O})$ values that appear to be independent of whether urea or guanidinium hydrochloride is used as denaturant [2].

The linear extrapolation method is the most frequently used method of determining protein unfolding free energy changes induced by urea or guanidine hydrochloride (GdnHCl). It is based upon the premise that unfolding free energy changes are linearly dependent on denaturant concentration and that extrapolation to zero denaturant concentration gives $\Delta G^0(\mathrm{H_2O})$, the unfolding free energy change in the absence of denaturant. When only the native and unfolded forms of protein are present in significant concentrations at equilibrium, $F \xleftrightarrow{K} U$, the unfolded process corresponds to one transition, i.e. two-state model. The $\Delta G^0(\mathrm{H_2O})$ can be obtained from $\Delta G = \Delta G^0(\mathrm{H_2O}) + m[D]$, where m reflects the sensitivity of the transition to denaturant concentration [D] by using the linear extrapolation method [3].

When the native, partially unfolded (stable intermediate) and unfolded forms of protein are present in significant concentrations at equilibrium, $F \xleftarrow{K_{\mathrm{FI}}} I \xleftarrow{K_{\mathrm{IU}}} U$, the unfolded process corresponds to two transitions, i.e. three-state model. Moreover, the $\Delta G^0_{\mathrm{total}}(\mathrm{H_2O})$ is

considered to be the sum of $\Delta G^0_{FI}(H_2O)$ and $\Delta G^0_{IU}(H_2O)$ [4-8]. By comparing the two-state and a three-state unfolding processes of proteins, the order of the unfolding free energy was always inconsistent with the order of stability observed. This raises the question about the accuracy of the evaluated unfolding free energy values of three-state models. In this chapter, the linear extrapolation method was considered as a standard method in the two-state unfolding model and particular attention is paid to the $\Delta G^0_{total}(H_2O)$ terms of the three-state unfolding model.

2. Assumptions

1. The protein unfolding process is determined by the structural element, E_i. E_i can change from a native state, $E_i(F)$ to a denatured, unfolded state $E_i(U)$, in two-state model. S^i_F , S^i_U , are the measureable signal of the folded and unfolded states for E_i, respectively, which were assumed to not vary with denaturant concentration to minimize the number of parameters. The equilibrium between $E_i(F)$ and $E_i(U)$ is described in eq (1).

$$E_i(F) \leftrightarrow E_i(U) \tag{1}$$

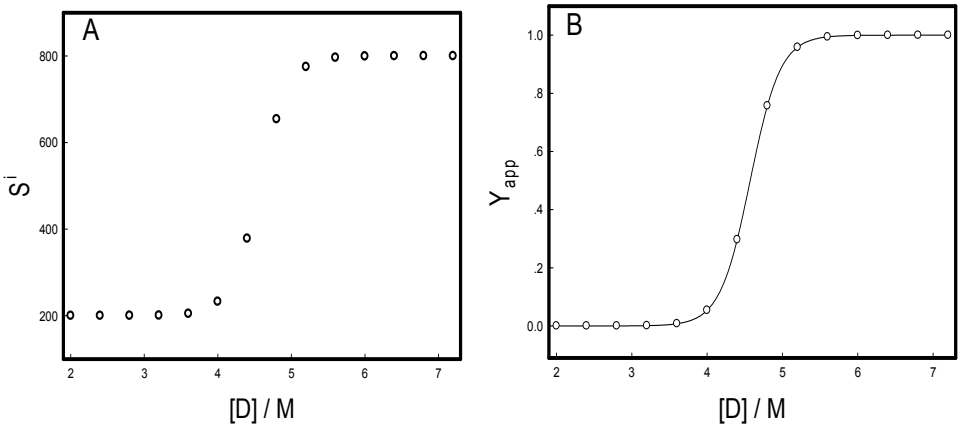


Fig. 1. The variation of S^i (A) and Y^i_{app} (B) of structural element, E_i, with the denaturant concentration. The observed data were fitted based on eq (3)

Figure 1A shows that the observed signals of the various experimental techniques (optical rotation, viscosity, UV-Vis absorption, fluorescence, circular dichroism, dynamic light scattering, nuclear magnetic resonance, etc) for structural element, E_i vary with the denaturant concentrations. At low denaturant concentrations (from [D] = 2.0 to [D] = 3.0) the measured signals are constant and can be considered as S^i_F , the signal of folded state. At high denaturant concentrations (from [D] = 6.5 to [D] = 7.5) the measured signals are constant, too, and can be considered as S^i_U , the signal of unfolded state. The unfolding curve of E_i, which is the plot of apparent fraction of unfolded structural element, Y^i_{app} against the denaturant concentrations, can be prepared as Figure 1B by using eq (2).

$$Y_{app}^{i} = \frac{S_O^i - S_F^i}{S_U^i - S_F^i} \tag{2}$$

where S_O^i is the observed signal at the given denaturant concentration

According to the linear extrapolation method the unfolding curve of E_i can be fitted based on eq (3).

$$Y_{app}^{i} = \frac{\exp(-\Delta G_i / RT)}{1 + \exp(-\Delta G_i / RT)} \tag{3}$$

where ΔG_i is the difference in free energy between the folded and the unfolded states of E_i at the given concentration of denaturant, R is the gas constant, and T is the absolute temperature. The equilibrium constant of unfolding transition is described in eq (4).

$$K_i = \frac{S_O^i - S_F^i}{S_U^i - S_O^i} \tag{4}$$

The free energy difference, ΔG_i can be evaluated following eq (5)

$$\Delta G_i = -RT \ln K_i = \Delta G_i^0(H_2O) + m_i[D] \tag{5}$$

$\Delta G_i^0(H_2O)$ is the single most important parameter for quantitating structural element, E_i stability and comparing stabilities of closely related structural element, E_i. m_i is a measure of the dependence of ΔG_i on denaturant concentration.

From Figure 1B the plot of ΔG_i against the concentration of denaturant is shown in Figure 2. $\Delta G_i^0(H_2O)$ and m_i can be obtained from Figure 2. They are 56.70 kJ·mole^{-1} and -12.41 kJ·mole^{-1}·M^{-1}, respectively.

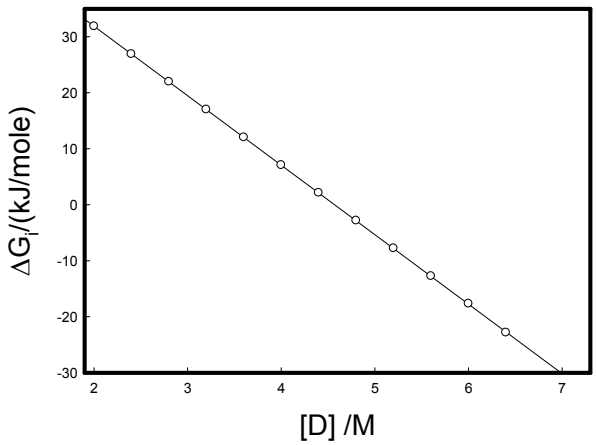


Fig. 2. The plot of ΔG_i against the concentration of denaturant at 25°C

When Y_{app}^{i} is 0.5 or the native and unfolded forms of structural element, E_i are 50%, respectively, the equilibrium constant of unfolding transition, K_i is 1. From eq (5) eq (6) can be obtained.

$$\Delta G_i^0(H_2O) = -m_i[D]_{1/2} \tag{6}$$

where $[D]_{1/2}$ is the midpoint concentration of denaturant required for unfolding of structural element. It means that we can use ether m_i and $[D]_{1/2}$ as two parameters or $\Delta G_i^0(H_2O)$ as single parameter to quantitate the structural element stability and compare the stabilities of closely related structural element. Only one $[D]_{1/2}$ or m_i is not enough to account for the stability of structural element, E_i.

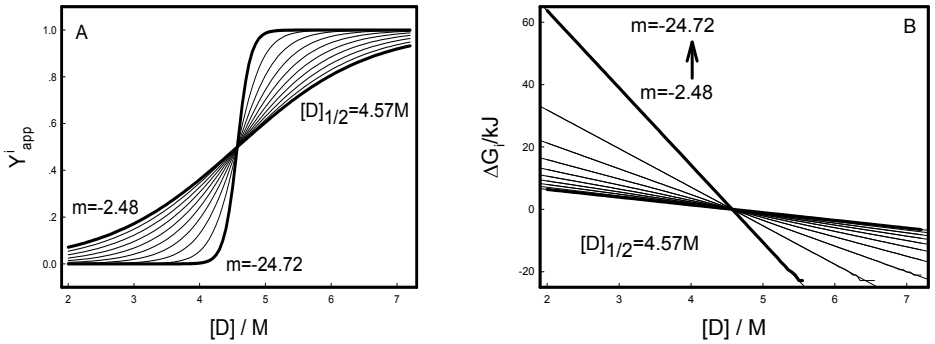


Fig. 3. The plots of Y_{app}^{i} (A) and ΔG_i (B) against the concentration of denaturant at 25°C

At $[D]_{1/2}$ = 4.57M the variations of Y_{app}^{i} and ΔG_i with the concentration of denaturant are shown in Figure 3. At constant m_i the unfolding curves corresponding to different $[D]_{1/2}$ are only shifts of a unfolding curve.

ΔG_i is a function of m_i and $[D]_{1/2}$, and can be written as ΔG_i (m_i, $[D]_{1/2}$). Therefore, Y_{app}^{i} can also be written as Y_{app}^{i} (m_i, $[D]_{1/2}$).

2. Protein is consisted of structural elements connected by chemical bonds that there is no contribution to measured signal. When protein contains two kinds of structural elements, E_i and E_j, the three-state unfolding behavior may be observed. The observed three-state unfolding curve Y_{app} is composed of curves (Y_{app}^{i}) and (Y_{app}^{j}) mixed at a certain molar fraction.

If the number of structural elements is n, the protein can be express as scheme 1. Typical 1 of protein is consisted of same structural elements and typical 2 and 3 of proteins are consisted of different structural elements.

It is known that the unfolding of structural element, E_i obeys two-state model. The stability can be expressed as ΔG_i (m_i, $[D]_{1/2}$) and the unfolding process can be described by Y_{app}^{i} (m_i, $[D]_{1/2}$). If m_1 is -24.72 kJ·M^{-1}·mole^{-1} and $[D]_{1/2}^{1}$ is 2.57 M we can obtain Y_{app}^{1} (-24.72 kJ·M^{-1}·mole^{-1}, 2.57M) from eq(3) and eq(5). Similarly, we can obtain Y_{app}^{2} (-24.72 kJ·M^{-1}·mole^{-1}, 6.57M). There is a protein, in which the structural element, E_1 appears for n times. The measured unfolding data (○) of the protein are presented in Figure 4A.

$$\underbrace{E_i - E_i - E_i - \cdots\cdots - E_i - E_i}_{n} \quad \mathbf{1}$$

$$\underbrace{E_i - E_j - E_j - \cdots\cdots - E_i - E_i}_{n} \quad \mathbf{2}$$

$$\underbrace{E_i - E_i - E_j - \cdots\cdots - E_k - E_j}_{n} \quad \mathbf{3}$$

Scheme 1. The proteins consisted of structural elements

Denaturant-induced unfolding of the protein is found to be a two-state process, just as that of the structural element, E_1.

There is a protein, in which the structural element, E_1 appears for n_1 times and E_2 does for n_2 times. The sum of n_1 and n_2 is n. The measured unfolding data (○) of the protein are presented in Figure 4B. Denaturant-induced unfolding of the protein is found to be a two-step process with accumulation of an intermediate state at around 3.5~5.5 M denaturant concentration. Therefore, denaturant-induced denaturation may be considered as a three-state transition and the mechanism for unfolding of the protein may be represented as:

$$F \Leftrightarrow I \Leftrightarrow U \tag{7}$$

where F, I and U are the folding, intermediate and unfolding states of the protein, respectively. The first transition which corresponded to the transformation of F state to the I state started at around at 1.5 M denaturant concentration and completed at 3.5 M denaturant with a midpoint occurring at 2.57 M denaturant. In fact the transition corresponds to the unfolding of E_1 in the protein.

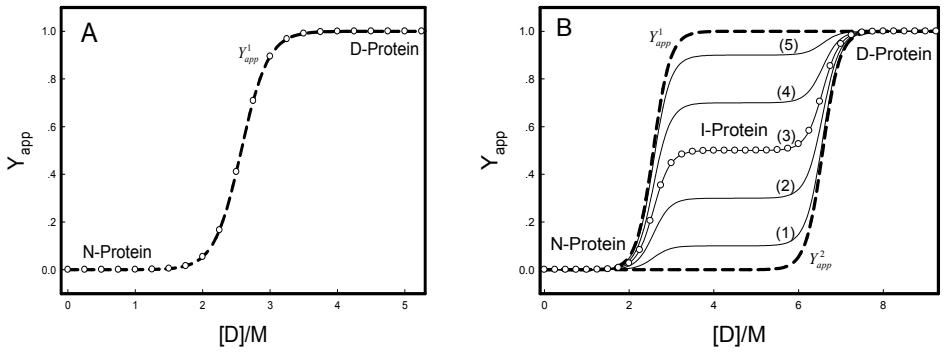


Fig. 4. The measured apparent fraction of unfolded proteins (○) at different denaturant concentrations. A: The protein consisted of same structural elements (E_1). B: a protein consisted of different structural elements (E_1 and E_2). Y^1_{app} and Y^2_{app} are the unfolding curves of E_1 and E_2, respectively. Based on eq(8) curves (1)~(5) are obtained from f_1 as 0.1, 0.3, 0.5, 0.7, and 0.9, respectively.

The second transition, which corresponded to the unfolding of the I state, started at around 5.5 M denaturant concentration and finally sloped off to the U state at 7.7 M denaturant with a midpoint occurring at 5.57 M denaturant. The transition corresponds to the unfolding of E_2 in the protein. The intermediate state of the protein is a partial unfolding state, in which E_1 exists in completely unfolding form and E_2 does in completely folding form. The three-state unfolding process can be considered to be composed of two sequential two-state unfolding processes. Then the measured unfolding data (○) of the protein in Figure 4B can be fitted by using eq(8).

$$Y_{app} = f_1 \bullet Y_{app}^1 + f_2 \bullet Y_{app}^2 \tag{8}$$

where f_1 and f_2 are molar fractions of structural elements E_1 and E_2 in the protein.

$$f_1 = \frac{n_1}{n} \qquad f_2 = \frac{n_2}{n} \qquad n = n_1 + n_2 \tag{9}$$

From eq(8) it is known that f_1 or f_2, Y_{app}^1 and Y_{app}^2 are the important factors to determine the three-state unfolding curve of protein. When f_2 is zero, the protein is consisted of structural elements, E_1, which appear n time in the protein. The three-state unfolding curve changes into two-state unfolding curve, Y_{app}^1. When f_1 is zero, the protein is consisted of structural elements, E_2, which appear n time in the protein. The three-state unfolding curve changes into two-state unfolding curve, Y_{app}^2.

Y_{app} has the scale as f_1 and f_2. It can be seen that the measured stable intermediate state of the protein appears at Y_{app} = 0.5 from Figure 4B. So we could determine the f_1 or f_2 value to be 0.5.

Protein is consisted of structural elements, E_1, E_2, E_3, ·····, E_{n-1}, which appears n_1, n_2, n_3, ·····, n_{n-1} times in the protein. The unfolding curves corresponding to E_1, E_2, E_3, ·····, E_{n-1} are Y_{app}^1, Y_{app}^2, Y_{app}^3, ·····, and Y_{app}^{n-1}, respectively, and $m_i \neq m_j$, $[D]_{1/2}^i \neq [D]_{1/2}^j$ (i and j change from 1 to n-1). The unfolding curve of the protein would be a n-state unfolding curve and can be fitted using eq(10).

$$\begin{aligned} Y_{app} &= f_1 \bullet Y_{app}^1 + f_2 \bullet Y_{app}^2 + f_3 \bullet Y_{app}^3 + \cdots\cdots + f_{n-1} \bullet Y_{app}^{n-1} \\ &= \sum_{i=1}^{n-1} f_i \bullet Y_{app}^i \end{aligned} \tag{10}$$

$$f_i = \frac{n_i}{n} \qquad \sum_{j=1}^{n-1} n_j = n$$

and

$$\sum_{i=1}^{n-1} f_i = 1 \tag{11}$$

So the four-state, five-state unfolding curves can be fitted by using eq(10).

3. The unfolding free energy represented by the curve (Y_{app}^1) and the (Y_{app}^2) are $\Delta G_1^0(H_2O)$ and $\Delta G_2^0(H_2O)$, respectively, for structural elements E_1 and E_2. Y_{app} is the unfolding curve of

protein which is consisted of structural elements, E_1, E_2. Then the three-state unfolding free energy, $\Delta G^0_{total}(H_2O)$ represented by the curve (Y_{app}) can be calculated from $\Delta G^0_1(H_2O)$ and $\Delta G^0_2(H_2O)$ using the same molar fraction and eq(12).

$$\Delta G^0_{\text{total}}(H_2O) = n \bullet (f_1 \bullet \Delta G^0_1(H_2O) + f_2 \bullet \Delta G^0_2(H_2O)) \tag{12}$$

$$\frac{\Delta G^0_{\text{total}}(H_2O)}{n} = (f_1 \bullet \Delta G^0_1(H_2O) + f_2 \bullet \Delta G^0_2(H_2O))$$

If $\Delta G^0_2(H_2O)$ is greater than $\Delta G^0_1(H_2O)$, $\Delta G^0_{total}(H_2O)/n$ will change from $\Delta G^0_1(H_2O)$ to $\Delta G^0_2(H_2O)$. From eq(11) f_1 and f_2 can change from 0 to 1. The variation of $\Delta G^0_{total}(H_2O)/n$ with f_2 is shown in Figure 5.

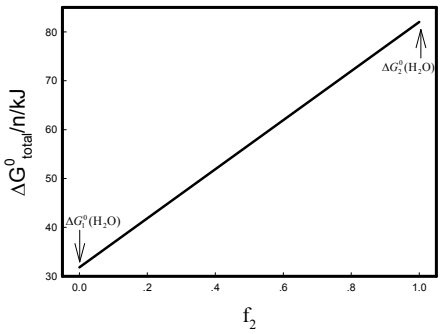


Fig. 5. The plot of $\Delta G^0_{total}(H_2O)/n$ against f_2 at 25°C. $\Delta G^0_1(H_2O)$ = 31.82 kJ·mole^{-1}, $\Delta G^0_2(H_2O)$ = 82.01 kJ·mole^{-1}.

Protein is consisted of structural elements, E_1, E_2, E_3, ·····, E_{n-1}, which appears n_1, n_2, n_3, ·····, n_{n-1} times in the protein. The unfolding free energies corresponding to E_1, E_2, E_3, ·····, E_{n-1} are $\Delta G^0_1(H_2O)$, $\Delta G^0_2(H_2O)$, $\Delta G^0_3(H_2O)$, ·····, and $\Delta G^0_{n-1}(H_2O)$, respectively, and $\Delta G^0_i(H_2O) \neq \Delta G^0_j(H_2O)$, (i and j change from 1 to n-1). The unfolding free energy, $\Delta G^0_{total}(H_2O)$ of the protein can be calculated using eq(13).

$$\begin{aligned}\Delta G^0_{\text{total}}(H_2O) &= n \bullet (f_1 \bullet \Delta G^0_1(H_2O) + f_2 \bullet \Delta G^0_2(H_2O) + f_3 \bullet \Delta G^0_3(H_2O) + \\ &\quad + f_{n-1} \bullet \Delta G^0_{\text{n-1}}(H_2O)) \\ &= n \bullet \sum_{\text{i=1}}^{\text{n-1}} f_i \bullet \Delta G^0_{\text{i}}(H_2O)\end{aligned} \tag{13}$$

$$\frac{\Delta G^0_{\text{total}}(H_2O)}{n} = \sum_{\text{i=1}}^{\text{n-1}} f_i \bullet \Delta G^0_{\text{i}}(H_2O)$$

When n is 4 the four-state unfolding free energy, $\Delta G^0_{total}(H_2O)/n$ can be calculated from $\Delta G^0_1(H_2O)$, $\Delta G^0_2(H_2O)$, and $\Delta G^0_3(H_2O)$. The plot of $\Delta G^0_{total}(H_2O)$ against f_1 and f_2 is shown in Figure 6.

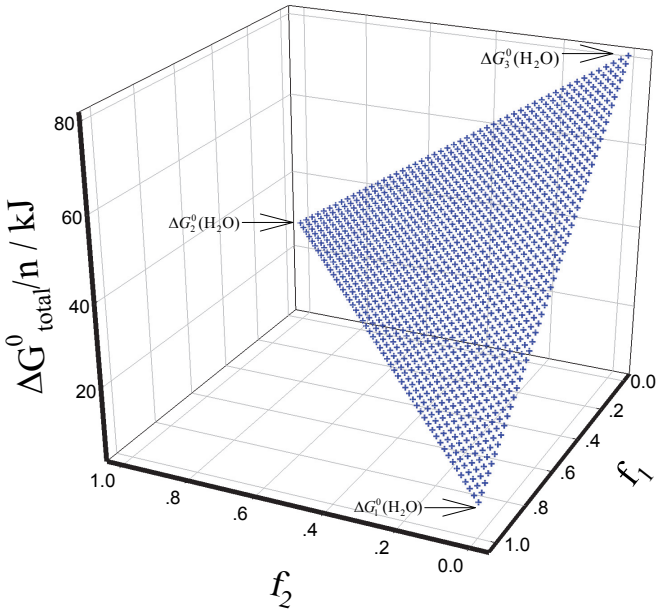


Fig. 6. The plot of $\Delta G^0_{total}(H_2O)$ /n against the f_1 and f_2 at 25°C. $\Delta G^0_1(H_2O)$ = 10.56 kJ·mole^{-1}, $\Delta G^0_2(H_2O)$ = 28.39 kJ·mole^{-1}, and $\Delta G^0_3(H_2O)$ = 81.40 kJ·mole^{-1}.

It can be seen that $\Delta G^0_{total}(H_2O)$ /n is a point located at the plane consisted of $\Delta G^0_1(H_2O)$, $\Delta G^0_2(H_2O)$, and $\Delta G^0_3(H_2O)$ in the space of $\Delta G^0_{total}(H_2O)$ /n, f_1, and f_2.

Although the unfolding curve of a protein, which is composed of same structural elements, i.e. E_i appears n times in the protein, is same with that of E_i, the $\Delta G^0_{protein}(H_2O)$ of the protein is different with the $\Delta G^0_i(H_2O)$ of E_i.

According to thermodynamics the $\Delta G^0_{Protein}(H_2O)$ of the protein should be n times of the $\Delta G^0_i(H_2O)$ of E_i (see eq(14)).

$$\Delta G^0_{Protein}(H_2O) = n \cdot \Delta G^0_i(H_2O) \tag{14}$$

3. Typical multistate unfolding curves

Three-state unfolding curves

From eq(8) three-state unfolding curve is dependence on Y^1_{app} (m_1, $[D]^1_{1/2}$)and the Y^2_{app} (m_2, $[D]^1_{1/2}$). Some typical three-state unfolding curves are given in Figure 7. When $m_1 = m_2$, and $[D]^1_{1/2} \neq [D]^1_{1/2}$, the three-state unfolding curve is shown in Figure 7A. When $m_1 < m2$, and $[D]^1_{1/2} \neq [D]^1_{1/2}$, the three-state unfolding curve is shown in Figure 7B. And when $m_1 > m_2$, and $[D]^1_{1/2} \neq [D]^1_{1/2}$, the three-state unfolding curve is shown in Figure 7C.

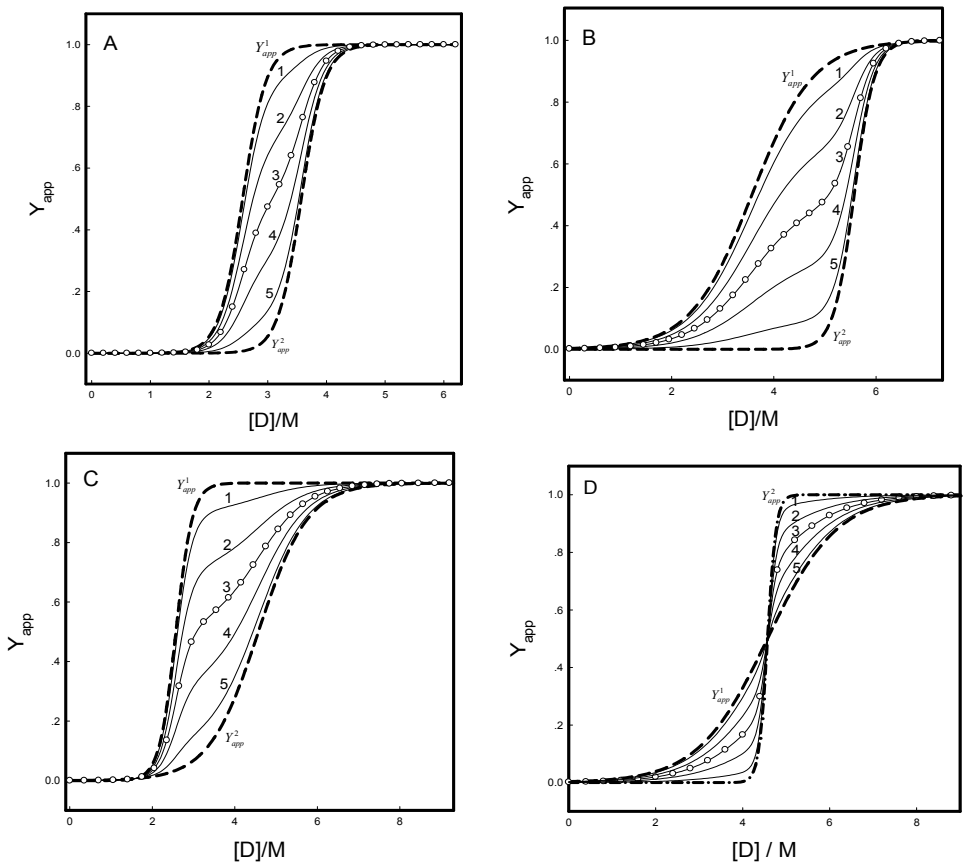


Fig. 7. Typical three-state unfolding curves. (○) represents measured value and Y^1_{app} , Y^2_{app} are the unfolding curves of structural elements E_1 and E_2, respectively. Solid lines (1), (2), (3), (4) and (5) are the fitted based eq(8) for f_1 are o.1, 0.3, 0.5, 0.7, and 0.9, respectively. A: $m_1 = m_2$= -24.72 kJ·M^{-1}·mole^{-1}, $[D]^1_{1/2}$ = 2.57M, and $[D]^2_{1/2}$ = 3.57M; B: m_1= -4.13 kJ·M^{-1}·mole^{-1}, m_2 = -24.72 kJ·M^{-1}·mole^{-1}, $[D]^1_{1/2}$ = 3.57M, and $[D]^2_{1/2}$ = 5.57M; C: m_1 = -24.72 kJ·M^{-1}·mole^{-1}, m_2 = -4.13 kJ·M^{-1}·mole^{-1}, $[D]^1_{1/2}$ = 2.57M, and $[D]^2_{1/2}$ = 4.57M ;D: m_1 = -2.48 kJ·M^{-1}·mole^{-1}, m_2 = -24.72 kJ·M^{-1}·mole^{-1}, $[D]^1_{1/2} = [D]^2_{1/2}$ = 4.57M

Of course, if the difference of $[D]^1_{1/2}$ and $[D]^1_{1/2}$ is greatly, stable intermediate state could be easily observed. When $m_1 < m_2$, and $[D]^1_{1/2} = [D]^1_{1/2}$, the three-state unfolding curve is shown in Figure 7D. The unfolding curves in Figure 7D are often considered as two-state unfolding.

Four-state unfolding curves

When n is 4 the unfolding curve appears three transitions from eq(10). The four-state unfolding curve is a function of Y^1_{app} , Y^2_{app} , Y^3_{app} , and f_i. Some typical four-state unfolding curves are shown in Figure 8.

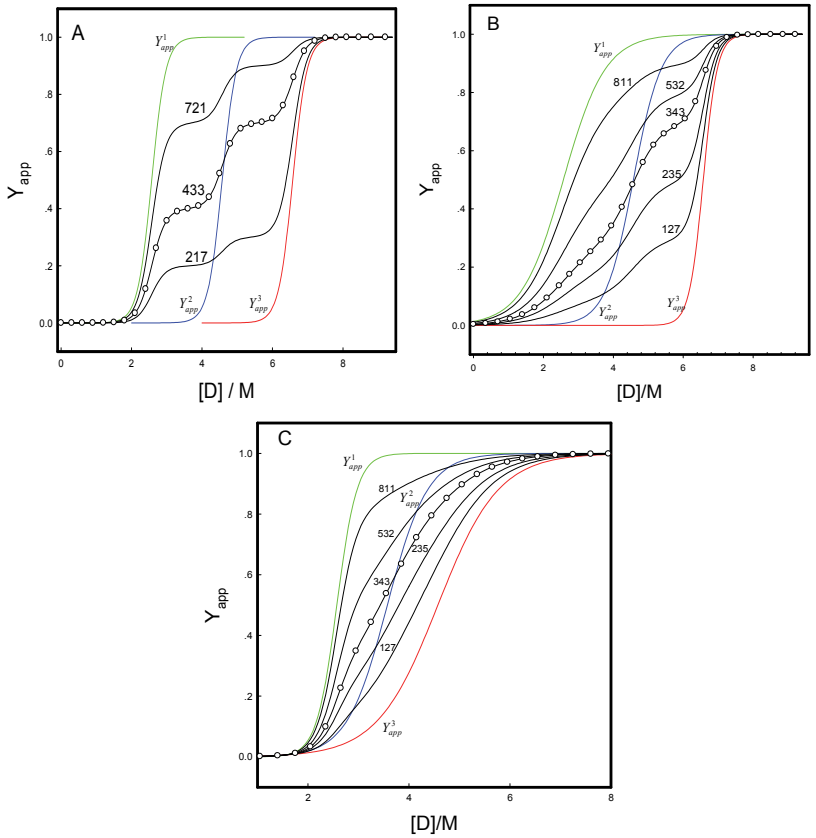


Fig. 8. Typical four-state unfolding curves. (○) represents measured value and Y^1_{app} (green line), Y^2_{app} (blue line), and Y^3_{app} (red line), are the unfolding curves of structural elements E_1, E_2 and E_3, respectively. Solid lines labeled as lst, representing f_1 as l tenth, f_2 as s tenth, and f_3 as t tenth, are the fitted based eq(10) when n is 4. A: $m_1 = m_2 = m_3 = -12.35$ kJ·M^{-1}·mole^{-1}, $[D]^1_{1/2} = 2.57$M, $[D]^2_{1/2} = 4.57$M, and $[D]^3_{1/2} = 6.57$M; B: $m_1 = -4.13$ kJ·M^{-1}·mole^{-1}, $m_2 = -6.19$ kJ·M^{-1}·mole^{-1}, and $m_3 = -12.35$kJ·M^{-1}·mole^{-1}, $[D]^1_{1/2} = 2.57$M, $[D]^2_{1/2} = 4.57$M, and $[D]^3_{1/2} = 6.57$M; C: $m_1 = -12.35$ kJ·M^{-1}·mole^{-1}, $m_2 = -6.19$ kJ·M^{-1}·mole^{-1}, and $m_3 = -4.13$kJ·M^{-1}·mole^{-1}, $[D]^1_{1/2} = 2.57$M, $[D]^2_{1/2} = 3.57$M, and $[D]^3_{1/2} = 4.57$M.

4. Structural element

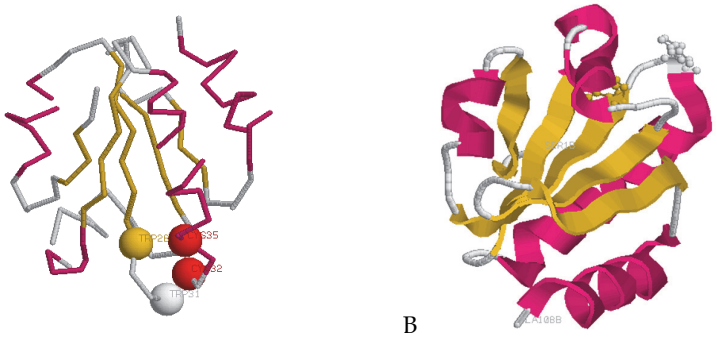
Structural element is a smallest structural unit in native protein, of which the unfolding obeys two-state model.

Native proteins are only marginally entities under physiological conditions. There are various no covalent bonds to which protein is subject—electrostatic interactions (both attractive and repulsive), hydrogen bonding (both intramolecular and to water), and

hydrophobic force, over an entire protein molecule. The low conformational stabilities of native proteins make them easily susceptible to denaturation by altering the balance of the weak nonbonding forces that maintain the native conformation. Native protein structure unfolds in a highly cooperative manner: Any partial unfolding of the structure destabilizes the remaining structure, which must simultaneously collapse to the random coil.

It has been shown that the unfolding or folding of small globular proteins occurs via a two-state process, whereas the unfolding or folding of larger proteins is complex and often involves the formation of intermediate [9]. Thioredoxin is characterized by an active site containing two cysteine residues separated by two other residues. The reversible oxidation of these cysteine residues to the disulfide form serves as a redox couple for a number of biological reactions. It was first characterized in yeast for its role in the reduction of methionine sulfoxide and inorganic sulphate[10, 11]. The backbone of oxidized form and ribbon model of reduced form for Escherichia coli thioredoxin are shown in Figure 9. The protein consists of single polypeptide chain of 108 amino acid residues. Either the oxidized or the reduced forms the protein is a compact molecule with 90% of its residues in α-helices, β-strands or reverse turns, which a core of twisted β-strand is flanked on either side by α-helices. The conversion between the oxidized and the reduced forms is accompanied by a change in conformation that is reflected in the fluorescence properties of Trp28.

The equilibrium unfolding of the oxidized form of Escherichia coli thioredoxin at pH 7 was studied by Santoro and Bolen [1] using guanidine hydrochloride (GdnHCl) and urea as denaturant at 25°C by monitoring the changes in ellipticity at 222nm.



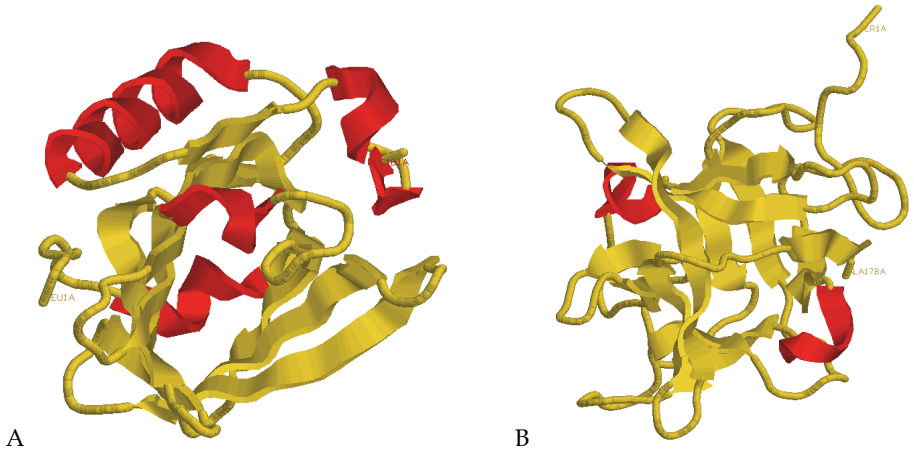
A B

Fig. 9. Backbone of oxidized form (A) and Ribbon model of reduced form (B) for E coli thioredoxin. The figures were generated using the Protein Data Bank files 1SRX and 2TRX, respectively. Trp residues are indicated as ball and stick, α- helices (pink) and β- strands (yellow) are shown.

The GdnHCl- or urea-induced unfolding of thioredoxin is a two-state process. The unfolding free energy change is 7.8±0.2 kcal/mole (32.64±0.84 kJ/mole) in the absence of (GdnHCl) or 8.6±0.9 kcal/mole (35.98±3.77 kJ/mole) in the absence of urea. Thermal unfolding measurements give $\Delta G^0(H_2O)$ being 8.1±0.1 kcal/mole (33.89±0.42 kJ/mole) by using differential scanning calorimetry. Within error, $\Delta G^0(H_2O)$ values obtained from GdnHCl- and urea-induced, and thermal unfolding are in agreement, meaning that GdnHCl, urea, and thermal unfolding involve the same native form to unfolding form equilibrium in oxidized form of Escherichia coli thioredoxin.

The free energy required to denature a protein is ~ 0.4 kJ/mole of amino acid residues so that 100-residue protein is typically stable by only around 40 kJ/mole [12]. Since the measured $\Delta G^0(H_2O)$ of thioredoxin is close to 40 kJ/mole, the compact protein molecule can be considered as a structural element.

Goat β-lactoglobulin consists of single polypeptide chain of 101 amino acid residues. The protein contains α- helices (red) and β- sheets (yellow) as shown in Figure 10. The measured $\Delta G^0(H_2O)$ is 11.7±0.8 kcal/mole (48.95±3.35 kJ/mole) at 25°C by measuring optical rotation [13]. β-lactoglobulin can be considered as a structural element. Although α-chymotrypsin consists of single polypeptide chain of 178 amino acid residues the $\Delta G^0(H_2O)$ is only 8.3±0.4 kcal/mole (34.73±1.67 kJ/mole) at 25°C by measuring optical rotation [13]. If the $\Delta G^0(H_2O)$ of a 100-residue protein were 40 kJ/mole, the $\Delta G^0(H_2O)$ of α-chymotrypsin would have been around 70 kJ/mole at 25°C. It means that α-chymotrypsin contains same structural elements at least two.



A B

Fig. 10. Ribbon models of β-lactoglobulin (A) and α-chymotrypsin (B). The figures were generated using the Protein Data Bank files 3NPO and 1XG6, respectively. α- helices (red) and β- strands (yellow) are shown.

Human serum albumin (HSA) has been used as a model protein for protein folding and ligand-binding studies over many decades [5, 14-19]. The protein is a single chain protein with 585 amino acids, with a molecular weight of ~67,000 Da. Serum albumin homologs with very similar properties are found in other mammals. The protein contains only α-helices. The structure of this protein has been determined by X-ray crystallography of high resolution (Figure 11); it includes three homologous domains (I-III) that assemble a heart-shaped molecule. Each domain is formed by two subdomains that possess common structural motifs. Only one cysteine residue located at position 34 (in domain I) with a free sulfhydryl group. HSA plays a special role in transporting metabolites and drugs throughout the vascular system and also in maintaining the pH and osmotic pressure of plasma. Interestingly, the structure and dynamics of HAS are known to be influenced by a number of factors, like pH, temperature, and binding of different ligands.

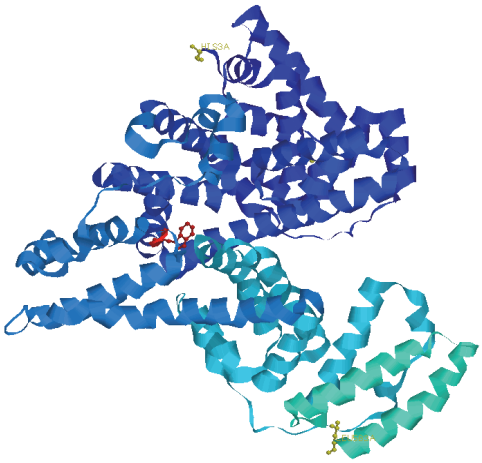


Fig. 11. Structure of human serum albumin (PDB: 1BJ5). Trp residueat 214 is indicated as ball and stick in red.

Denaturation of HSA has been studied by several workers using different methods like absorption difference at 287nm, fluorescence at 340nm upon excitation at 282nm, tryptophan fluorescence upon excitation at 295nm, ellipticity at 222nm, 1-anilinonaphthalene-8-sulfonate fluorescence, differential calorimetry, etc. Denaturation of HSA has been suggested to follow a single-step transition as well as two- or multiple-step transitions depending upon the reaction conditions and probes used. For example, urea denaturation of HSA has been shown to follow a two-step, three-state transition involving one intermediate when studied by fluorescence and ellipticity measurements but became single-step, two-state transition when studied by UV difference spectroscopy. Similarly, guanidine hydrochloride (GdnHCl) denaturation of serum albumin has been reported to follow both single-step and two-step transitions [16]. Meanwhile the measured $\Delta G^0(H_2O)$ has big difference. The measured $\Delta G^0(H_2O)$ is 5.98 kcal/mole (25.02 kJ/mole) and 6.00 kcal/mole (25.10 kJ/mole) at 25°C, 1M KCl pH 7.0 by measuring circular dichroism and fluorescence, respectively [5]. In the presence of 2,2,2-trifluoroethanol the measured $\Delta G^0(H_2O)$ is 8.04 kcal/mole (33.64 kJ/mole) and 8.06 kcal/mole (33.72 kJ/mole) [15]. 2,2,2-trifluoroethanol makes the $\Delta G^0(H_2O)$ increase around 2 kcal/mole. Halim et al [16] used bromophenol blue as a probe the $\Delta G^0(H_2O)$ of bovine serum albumin is 4.04 and 4.60 kcal/mole, respectively, from urea and GdnHCl two-state denaturation curves at pH 8.0, 25°C. However, the $\Delta G^0(H_2O)$ of HSA is 146.7 kJ/mole from sodium dodecyl sulphate two-state denaturation curves at pH 7.0, 35°C [17]. It means that HAS is a multi-element protein. When the structural elements are same at the experimental conditions the unfolding of HAS is a two-state process, and when the structural elements are different at another experimental conditions the unfolding behaves as a three-state process.

Urea-induced unfolding of bovine serum albumin and one of its fragments containing domain II+III has been studied by difference spectral and fluorescence emission measurements. The unfolding-refolding curves of both the proteins showed the presence of at least one stable intermediate when the transition was monitored at 288 nm. The presence

of the intermediate was not detectable at 293 nm where only tryptophan contributed towards the protein absorption. However, both the proteins did show the presence of intermediate when the denaturation was monitored fluorometrically. Since domain III of the albumin is devoid of tryptophan, it is concluded that the formation of intermediate in the unfolding-refolding transition of serum albumin involves (i) unfolding of domain III, (ii) minor structural transformations in domain II, and/or (iii) the separation of the sub-domains of domain III from each other [15-17].

5. The factors affected structural element

Acquisition of native globular conformation (3D structure) of a protein is governed by its structural elements and biological environment surrounding it. The stability of protein's 3D structure is endorsed by various intramolecular forces, such as hydrogen bonds, van der Waals, electrostatic and hydrophobic interactions involving various amino acid side chains as well as their milieu in the native structure. So protein structure is affected by factors such as experiment conditions (extreme temperature, pH and pressure) and presence of destabilizing agents (salt, alkali, denaturant and surfactant), and of stabilizing agents (metal ions, anions and small organic molecules) [15-19]. In fact they are the factors which structural element is affected. A few examples are given to show the effect of stabilizing agents on structural element.

Anticoagulation factor I (ACF I) isolated from the venom of *Agkistrodon acutus* is an activated coagulation factor X-binding protein in a Ca^{2+}-dependent fashion with marked anticoagulant activity. The protein is a single chain protein with 129 amino acids, in which there are 7 Trp. The crystal structure measurement shows that the protein has both α-helices and β- sheets and tryptophan residues locate at α-helices, β-sheets and loops. Metal ions-induced stabilization and unfolding of ACF I was studied in guanidine hydrochloride solutions by Xu et al [6] following the fluorescence and circular dichrosim. Metal ions can increase the structural stability of ACF I against guanidine hydrochloride denaturation and change its unfolding behavior. Reproduced unfolding curves of apo-ACF I and holo-ACF I from ref [6] are shown in Figure 12.

		$[D]_{1/2}$/M		-m/(kcal/M/mole)		
Holo-ACF I	F↔Ia	Ia↔Ib	Ib↔U	F↔Ia	Ia↔Ib	Ib↔U
Unfolding	0.66±0.02	1.60±0.04	2.52±0.05	3.76±0.09	3.63±0.04	2.59±0.02
Refolding	0.67±0.03	1.59±0.02	2.55±0.06	3.70±0.05	3.79±0.10	2.77±0.08

	ΔG^0/(kcal/mole)			
Holo-ACF I	F↔Ia	Ia↔Ib	Ib↔U	total
Unfolding	2.51±0.04	2.21±0.02	1.29±0.03	6.01±0.09
Refolding	2.60±0.07	2.16±0.03	1.31±0.02	6.07±0.12

Table 1. Thermodynamic parameters for unfolding and refolding of ACF I by GdnHCl at 25°C, monitored by measurement of fluorescence at 340nm

The guanidine hydrochloride induced unfolding of apo-ACF I is a two-state process with no detectable intermediate state, whereas the unfolding of holo-ACF I follows a three-step transition, with intermediate state Ia and intermediate state Ib. The thermodynamic parameters for unfolding and refolding of ACF I are shown in Table 1 [6]. The unfolding

data can be fitted by using eq (3) for two-state process of apo-ACF I and eq (10) for four-state process of holo-ACF I, respectively.

Using the new method it can be thought that ACF I is consisted of multi-structural elements, and the structural elements are same in apo-ACF I in guanidine hydrochloride. The stability of structural element can be described by parameters $[D]_{1/2}$ (1.22M) and m (-4.06 kcal/M/mole), i.e. $\Delta G^0(H_2O)$ is 4.96 kcal/mole. After binding Ca^{2+} the structural elements are changed. At least three-kind structural elements appear in holo-ACF I. Their stability is 4.22 kcal/mole, 7.07 kcal/mole, and 5.95 kcal/mole, for structural element 1, 2, and 3, respectively.

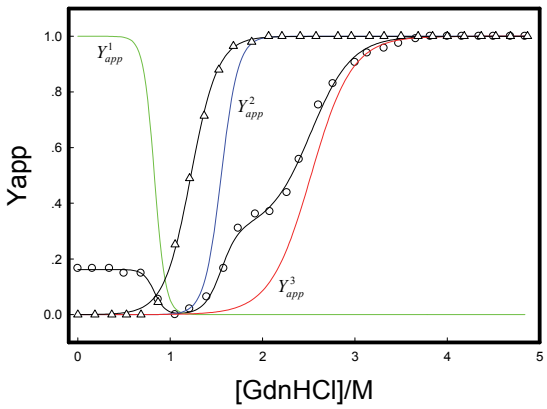


Fig. 12. Guanidine hydrochloride (GdnHCl)–induced unfolding of anticoagulation factor I (ACF I) in 0.02 M Tris-HCl buffer (pH 7.6, 25°C). Unfolding transitions of apo-ACF I(Δ), and holo-ACF I(○) monitored by measurement of fluorescence at 340 nm after exciting at 295 nm. The data were reproduced from ref [6]. Black lines are the fitting from eq (3) corresponding to two-state process and eq (10) corresponding to four-state process, respectively. The unfolding transition of structural element is shown as Y^1_{app} (green), Y^2_{app} (blue), and Y^3_{app} (red), respectively.

Parameters of $[D]_{1/2}$ and m values are 0.85M and -4.96 kcal/M/mole, 1.56M and -4.53 kcal/M/mole, 2.56M and -2.32 kcal/M/mole, for structural element 1, 2, and 3, respectively. The unfolding of holo-ACF I Y_{app} can be fitted from eq (10).

$$Y_{app} = 0.14 \times Y^1_{app} + 0.26 \times Y^2_{app} + 0.60 \times Y^3_{app} \tag{15}$$

where Y^1_{app}, Y^2_{app}, and Y^3_{app} are unfolding transition of structural element 1, 2, and 3, respectively.

The $\Delta G^0_{total}(H_2O)$ /n of holo-ACF I is 6.00 kcal/mole from eq (13).

$$\begin{aligned} \frac{\Delta G^0_{total}(H_2O)}{n} &= 0.14 \times \Delta G^0_1(H_2O) + 0.26 \times \Delta G^0_2(H_2O) + 0.60 \times \Delta G^0_3(H_2O) \\ &= 0.14 \times 4.22 + 0.26 \times 7.07 + 0.60 \times 5.95 = 6.00 \end{aligned} \tag{16}$$

If using the stability of structural element in apo-ACF I as standard the binding of Ca^{2+} makes structural element 1 be destabilized and 2 and 3 be stabilized. Of the various forces, hydrogen bonding and hydrophobic effect have been shown to make significant contribution to the protein stability. Being secondary contributor to protein's stability through salt bridges, charged residues have also been found important in manipulating protein stability [20]. It is that Ca^{2+} binding breaks the balance of charged residues makes the charged residues in apo-ACF I re-distribute and form new structural elements in holo-ACF I.

Second example is succinylation-induced conformational destabilization of lysozyme as studied by guanidine hydrochloride denaturation [20].Figure 13 shows the crystal structure of lysozyme which is made up of a single polypeptide chain of 129 amino acid residues arranged in the form of two domains and stabilized by four disulphide bonds. All six lysine residues are distributed in the molecular surface and all six tryptophan residues are located at α-helices and loops. The unfolding of the protein obeys two-state process, meaning that the protein is consisted of same structural element(s). Using fluorescence data upon excitation at 280nm value of $\Delta G^0_{total}(H_2O)$ /n is 8.56 kcal/mole and 'm' is -2,03kcal/M/mole for native lysozyme. When excitation at 295nm $\Delta G^0_{total}(H_2O)$ /n is 7.96 kcal/mole and 'm' is -1.93 kcal/M/mole. Using 100-fold molar excess of succinic anhydride, about 99% of lysine residues of lysozyme were modified. Though the succinylated lysozyme obeys two-state process yet a pronounced decrease in $\Delta G^0_{total}(H_2O)$ /n value is noticed with succinylated lysozyme preparation compared to native lysozyme. $\Delta G^0_{total}(H_2O)$ /n value is 4.40 kcal/mole and 4.67 kcal/mole excitation at 280nm and 295nm, respectively. The 'm' values are also changed from -2.03 kcal/M/mole and -1.93 kcal/M/mole (for native lysozyme) to -1.54 kcal/M/mole and -1.59 kcal/M/mole for succinylated lysozyme, respectively. The conformational destabilization of the modified protein can be attributed to the effect of succinylate on structural element(s), i.e. it is that the positive charges on lysine residues are neutralized by negative charges on succinylate, makes the structural element(s) be destabilized. Experimental measurements show that the net charge of both native and succinylated lysozyme is +8 and -9, respectively at pH 7.0.

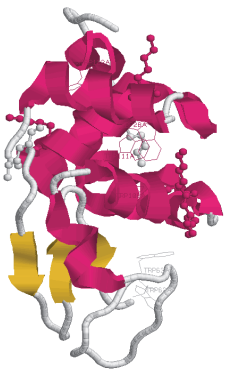


Fig. 13. Ribbon model of lysozyme produced with a Protein Data Bank file 1VDQ. Trp residues are indicated as wireframe and Lys residues are shown as ball and stick.

Third example is HAS, whose crystal is shown in Figure 11. In the presence of different concentrations of KCl urea-induced denaturation of HAS displays different pathways. When the concentration of KCl is 1.0M HAS undergoes a single-step transition with no intermediate. At pH 7.0, 25°C the variation of ΔG as a function of urea concentration is presented in Figure 14A (curve 2). According to eq (5) $\Delta G^0_{total}(H_2O)$ /n and 'm' values can be obtained to be 5.98±0.26 kcal/mole and –(1.03±0.04) kcal/M/mole, respectively. In the absence of KCl the unfolding of HAS is a two-step transition process and the urea-induced denaturation curve is shown in Figure 14B, using the data of Figure 1B or 1A in ref [5]. Muzammil et al [5] considered that first transition corresponded to the formation of intermediate state whereas second transition corresponded to the unfolding of the intermediate. Assuming both the transitions follow two-state mechanism the variations of ΔG as a function of urea concentration are presented in Figure 14A (curve 1 and 3) at pH 7.0, 25°C. According to eq (5) $\Delta G^0_1(H_2O)$ and $\Delta G^0_2(H_2O)$ are 3.66±0.35 kcal/mole and 5.38±0.24 kcal/mole, respectively. The 'm' values are –(0.95±0.10) kcal/M/mole and –(0.80±0.04) kcal/M/mole, respectively. From eq (6) the $[D]_{1/2}$, of which corresponds second transition from intermediate state to unfolding state, should be 6.73M.

However, the authors [5] reported $\Delta G^0_2(H_2O)$ to be 1.4 kcal/mole. Using 'm' value as -0.80 kcal/M/mole the calculated $[D]_{1/2}$ is only 1.75M from eq (6). The difference of $[D]_{1/2}$ is 4.98 M, at which the second transition from intermediate state to unfolding state starts. The authors [5] got $\Delta G^0_2(H_2O)$ from extrapolation of ΔG values upto the starting of second transition. Since $\Delta G^0_{total}(H_2O)$ /n being a thermodynamic property does not depend on the path free energy change in the absence of salt should be obtained by summing the $\Delta G^0_1(H_2O)$ and $\Delta G^0_2(H_2O)$. The $\Delta G^0_{total}(H_2O)$ /n (4.89 kcal/mole) in the absence of KCL is smaller than that (5.98 kcal/mole) in the presence of 1M KCl. The authors attributed the binding of Cl^- to the domain III to stabilize the domain [5].

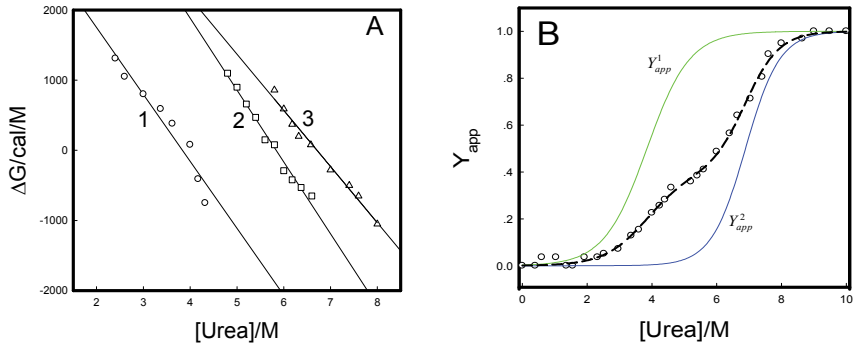


Fig. 14. Urea-induced unfolding of HAS at pH 7.0, 25°C in the presence as well as absence of KCl. A: Dependence of free energy change on urea concentration for the transitions (1): F↔I; (2): F↔U; and (3): I↔U. B: Normalized curve for the transition F↔U in the absence of KCl. Data (○) obtained from the Figure 2 in ref [5]. Medium dash is the fitting from eq (8), Y^1_{app} (green) and Y^2_{app} (blue) are unfolding transition of structural element 1 and 2, respectively.

In my views the calculation of $\Delta G_2^0(H_2O)$ is unreasonable. $\Delta G^0(H_2O)$ is a parameter to describe the stability of a protein. In two-state transition 'm' and $[D]_{1/2}$ values are also two parameters to describe the stability of the protein. When 'm' is same the $\Delta G^0(H_2O)$ is proportional to $[D]_{1/2}$. The protein whose unfolding curve appears at higher denaturant concentrations is more stable than that whose unfolding curve appears at lower denaturant concentrations For the unfolding curve of HAS (Figure 14B) in the absence of KCl, two transitions are observed. The first transition starts at around 2M urea and completes at 4.5M urea, and the second transition starts at around 5.2M urea and finally slopes off at 8.4M urea. The free energy change corresponded to second transition is greater than that corresponded to the first transition if 'm' is nearly same. The reported $\Delta G_2^0(H_2O)$ (=1.40 kcal/mole) < $\Delta G_1^0(H_2O)$ (=3.40 kcal/mole) [5]. Using the new method HAS is consisted of multi-structural elements. In the absence of KCl HAS contains at least two-kind structural elements, the unfolding shows three-state process. A least squares analysis of the data in Figure 14B according to eq (8) yielded a fitting curve, Y_{app} .

$$Y_{app} = 0.40 \times Y_{app}^1 + 0.60 \times Y_{app}^2 \tag{17}$$

where Y_{app}^1 and Y_{app}^2 are unfolding transition of structural element 1 and 2, respectively. According eq (12) the $\Delta G_{total}^0(H_2O)$ /n express as

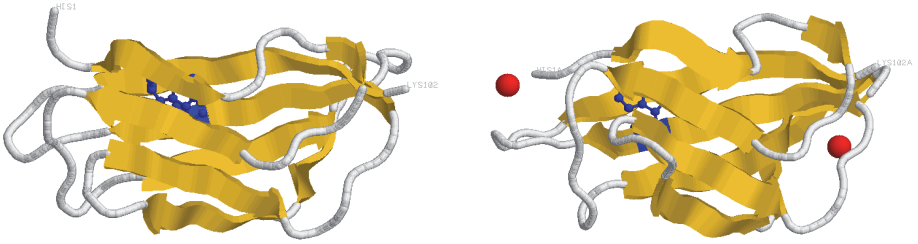
$$\begin{aligned} \frac{\Delta G_{total}^0(H_2O)}{n} &= 0.40 \times \Delta G_1^0(H_2O) + 0.60 \times \Delta G_2^0(H_2O) \\ &= 0.40 \times 3.66 + 0.60 \times 5.38 = 4.69 \end{aligned} \tag{18}$$

The $\Delta G_{total}^0(H_2O)$ /n (4.69 kcal/mole) in the absence of KCL is smaller than that (5.98 kcal/mole) in the presence of 1M KCl. The binding of Cl^- to the domain III to stabilize the domain

6. The unfolding of CopC

The protein CopC of *Pseudomonas syringae* pathovar tomato (CopC) is one of four proteins, CopA, B, C, and D, coded on the copper resistance operon (cop). CopC belonging to periplasmic protein, consists of 102 amino acids. In solution it adopts a fold essentially constituted by nine β sheets forming a barrel motif. The copper (Ⅰ) site located in C-terminal is constituted by His-48 and two or three of the four Met residues (40, 43, 46, 51) [$Cu^{I}(His)(Met)_x$ (x = 2 or 3)]. On the other hand, the protein is known to bind copper (Ⅱ) in an N-terminal position that is found consistent with a coordination arrangement including His-1, Glu-27, Asp-89, and His-91 [$Cu^{II}(His)_2(Asp)(Glu)(OH_2)$]. They represent novel coordination environments for copper in proteins and the two copper binding sites are about 30 Å apart. All of these have been confirmed by the method of nuclear magnetic resonance (NMR) and extended X-ray absorption fine structure spectroscopy (EXAFS) [21, 22]. CopC has high affinity for both copper ions although the Cu(II) site has the highest affinity (~10^{13} M^{-1}) [23, 24]. The binding sites of CopC can also occupied by other transition metal ions. However the affinity is lower than copper ions. For example, Hg(II) ions can bind to the both sites [25] and Ag(I) only binds to the Cu(I) site [26]. Figure 15 shows the NMR structure of apoCopC (A) and crystal structure of Cu^{2+}-CopC-Cu^{+} (B) drawn with the

Protein Data Bank files 1M42 and 2C9Q, respectively. A β-sandwich structure comprised of β-strands arranged in a Greek Key topology exists in apoCopC and Cu^{2+}-CopC-Cu^{+}. Single tryptophan residue 83, sandwiched between the two β sheets, locates at a rather hydrophobic microenvironment and has numerous contacts with residues in strand β2 and strand β7 [27].



A B

Fig. 15. NMR structure of apoCopC (1M42) (A) and crystal structure of Cu^{2+}-CopC-Cu^{+} and (2C9Q) (B) revealing the b-sandwich fold, at which Trp (blue) locates and Cu(II), Cu(I) (red).

The fluorescence properties of tryptophan residues in protein can be considered as three discrete spectral classes: One is the buried in nonpolar regions of the protein, the fluorescence maximum position appears at 310~330 nm and spectral band width is around 48~49 nm: two is located at the surface of protein and completely exposed to water, maximum peak is in 350~353 nm and half bandwidth is in the range of 59~61 nm; the latter is in limited contact with water which is probably immobilized by bonding at the macromolecular surface, the fluorescence maximum peak is located at 340~342 nm and the half bandwidth is about 53~55 nm.

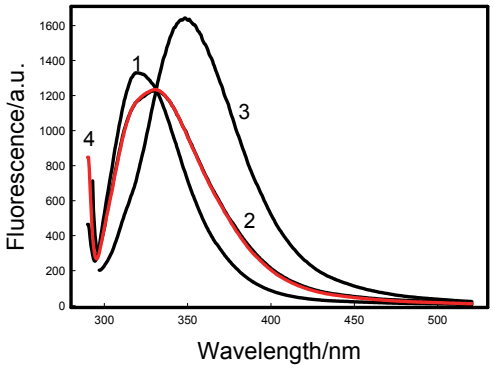


Fig. 16. Fluorescence spectra of apoCopC in the absence or presence of urea, in 20 mM PBS and 0.1 M NaCl, pH 6.0. The concentration of urea / M is 1: 0; 2: 5; 3: 9.4, respectively. 4 is the fitted curve according to two-state mechanism.

Figure 16 is the spectra of apoCopC in different experimental conditions. In the presence of 9.4 M urea the fluorescence peak of apoCopC locates at 353 nm and the half bandwidth reaches to 60 nm (Figure 16 spectrum 3), meaning that apoCopC exists in unfolding state. In

the absence of urea the fluorescence peak of native apoCopC appears at 320 nm and the half bandwidth is 48 nm (Figure 16 spectrum 1), meaning that the tryptophan residue 83 is buried in nonpolar regions of apoCopC. In the presence of 5.0 M urea the fluorescence peak appears at 331 nm and the half bandwidth reaches to 66 nm (Figure 16 spectrum 2).

Obviously spectrum 2 is the summing of spectra produced by two-class tryptophan residue in protein(s). With the increasing of urea concentration an iso-fluorescence point keeps at 332nm from spectrum 1 to spectrum 3, meaning that there are two species of apoCopC in the presence of 5.0 M urea. Curve 4 (red line) in Figure 16 is fitted from spectrum 1 and 2, in which the contribution of spectrum 1 is 0.73 and that of spectrum 2 is 0.27. It can be seen that the fitted curve is nearly coincided with the measured spectrum. In the presence of 5.0 M urea apoCopC exists in native and completed unfolding forms. The fluorescence property of tryptophan residue 83 in apoCopC is an index to show the protein in either native or unfolding states. It can be deduced that the unfolding of apoCopC would obeys two-state mechanism.

When the tryptophan residue 83 is mutated as leucine residue the protein is nearly immeasurable by fluorescence. Meanwhile the mutagenesis make the ability of binding to Cu(II), the thermal stability of apoCopC be decreased [28]. The important role of the tryptophan residue 83 is not only in using as a fluorescence probe but also in keeping the β-sandwich hydrophobic structure. The thermodynamic stability apoCopC can be investigated at pH 7.4 by monitoring the fluorescence intensity at around 405 nm [25, 29, 30]. As known Cu^{2+}-CopC and Hg^{2+}-CopC-Hg^{2+} are prepared by adding Cu(II) or Hg(II). The unfolding of apoCopC, Cu^{2+}-CopC and Hg^{2+}-CopC-Hg^{2+} are reversible. According to two-state model, the free energy change can be measured, respectively, and the order of stability is Cu^{2+}-CopC >> Hg^{2+}-CopC-Hg^{2+} ~ apoCopC. The binding of Cu(II) makes the stability of CopC be increased greatly. Although Hg(II) can also binding the Cu(II) site the contribution to the stability can be neglected. The reason is the weak binding of Hg(II) to the sites of apoCopC.

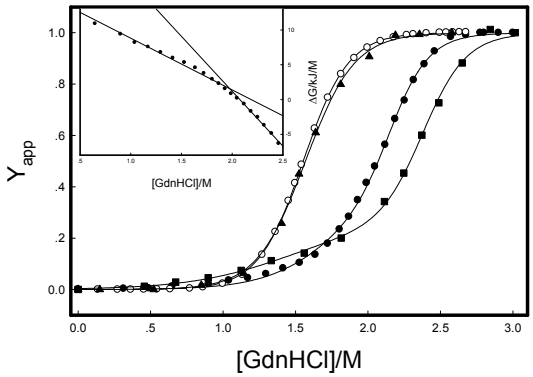


Fig. 17. GdnHCl-induced unfolding of CopC by monitoring the fluorescence 400/320 nm ratio for apoCopC (○), Cu^{2+}-CopC (●), CopC-Ag^{+} (▲), and Cu^{2+}-CopC-Ag^{+} (■) forms at pH 7.4, 25°C. The solid lines are fitting curves in terms of two-state mechanism for apoCopC (○) and CopC-Ag^{+} (▲), of three-state for Cu^{2+}-CopC (●) and Cu^{2+}-CopC-Ag^{+} (■) from eq (2) and eq (8), respectively. Inset: Dependence of free energy change of Cu^{2+}-CopC (●) on GdnHCl concentration.

Using guanidine hydrochloride as chemical denaturant the unfolding profiles are presented in Figure 17 by monitoring the changes of fluorescence ratio at 400nm and 320nm. ApoCopC unfolds in two-state equilibrium reaction. It means that apoCopC is consisted of a kind of structural element(s). The $\Delta G^0(H_2O)/n$, $[D]_{1/2}$ and 'm' values are listed in Table 2. When used urea as denaturant the obtained parameters are also in Table 2. Compared the data in Table 2 the $\Delta G^0(H_2O)/n$ values agree between guanidine hydrochloride and urea as chemical denaturant. The unfolding curve of CopC-Ag^+ roughly is a repeat of that of apoCopC. Like Hg(II), the binding of Ag(I) to the Cu(I) site of apoCopC does not also make the stability of the protein be increased. In another word, the structural element(s) in apoCopC don't affected by the binding of Hg(II) or Ag(I).

The corresponding GdnHCl-induced unfolding with Cu^{2+}-CopC reveals some difference. If the unfolding were considered as two-state process the GdnHCl-concentration at transition midpoint would have been 2.06 M which is higher than 1.55 M, and the dependence of ΔG on GdnHCl concentrations would have been as inset in Figure 17. Indeed the binding of Cu(II) makes the transition midpoint of apoCopC from 1.55 M shift to 2.06 M and stabilize the protein. However, there are evidently two slopes in dependence of ΔG on GdnHCl concentrations (Inset of Figure 17). This indicates that the equilibrium-unfolding reaction is three-state process, involving an intermediate species populated at 1.4 to 1.7 M guanidine hydrochloride [7].

Protein	Denaturant		-m kJ/M/mole	$[D]_{1/2}$ M	$\Delta G^0(H_2O)/n$ kJ/mole
apoCopC	GdnHCl		16.16±0.22	1.55	24.98±0.31
	Urea		4.32±0.18	5.61	24.24±1.04
CopC-Ag^+	GdnHCl		15.74±0.27	1.55	24.33±0.41
	$-m_{FI}$ kJ/M/mole	$-m_{IU}$ kJ/M/mole	$\Delta G^0_{FI}(H_2O)$ kJ/mole	$\Delta G^0_{IU}(H_2O)$ kJ/mole	$\Delta G^0_{total}(H_2O)/n$ kJ/mole
Cu^{2+}-CopC	8.83±0.31	19.86±0.18	12.83±0.31	42.52±0.30	36.82
Cu^{2+}-CopC-Ag^+	7.67±0.25	18.930±0.23	10.54±0.28	45.08±0.43	36.71
Cu^{2+}-Y79W-W83F	9.78±0.36	21.27±0.31	12.33±0.39	44.22±0.66	37.80
		Urea – induced			
Cu^{2+}-CopC	2.21±0.03	4.35±0.07	9.71±0.13	36.70±0.54	29.36

Table 2. Thermodynamic parameters for various forms of CopC from chemical denaturant-induced unfolding at pH 7.4, 25°C

According to the new model the structural element(s) of apoCopC is changed by the binding of Cu(II), from one-kind structural element(s) changing to two-kind structural elements, which one is more stable than another. Naturally the equilibrium-unfolding reaction of the protein displays three-state behavior. The data (●) measured guanidine hydrochloride experiment with Cu^{2+}-CopC are fitted well in eq (19).

$$Y_{app} = 0.19 \times Y^1_{app}(m_1, [D]^1_{1/2}) + 0.81 \times Y^2_{app}(m_2, [D]^2_{1/2}) \quad (19)$$

where Y_{app}^{1} and Y_{app}^{2} are the unfolding transitions, m_1,$[D]_{1/2}^{1}$ and m_2,$[D]_{1/2}^{2}$ are the parameters described the stabilities of structural elements 1 and 2, respectively. A least-squares curve fitting analysis is used to calculate the $\Delta G_1^0(H_2O)$, $\Delta G_2^0(H_2O)$, m_1, and m_2. The $\Delta G_1^0(H_2O)$ and $\Delta G_2^0(H_2O)$ correspond to $\Delta G_{FI}^0(H_2O)$ and $\Delta G_{IU}^0(H_2O)$, and m_1 and m_2 are m_{FI} and m_{IU} respectively, in Table 2. From eqs (8) and (20) free energy change of Cu^{2+}-CopC can be calculated to be 36.82 kJ/mole.

$$\begin{aligned}\frac{\Delta G_{total}^0(H_2O)}{n} &= 0.19\times\Delta G_1^0(H_2O)(m_1,[D]_{1/2}^{1}) + 0.81\times\Delta G_2^0(H_2O)(m_2,[D]_{1/2}^{2}) \\ &= 0.19\times 12.83 + 0.81\times 42.52 = 36.82\end{aligned} \tag{20}$$

Cu^{2+}-CopC is consisted of two-kind structural elements. Thus the unfolding profile appears to involve two sequential transitions. The first transition corresponds to the unfolding of structural element one, whose free energy change is 12.83 kJ/mole and the guanidine hydrochloride concentration at the transition midpoint is 1.53 M, close to 1.55 M of apoCopC or apoCopC-Ag^{+}. The second transition corresponds to the unfolding of another, whose free energy change is 42.52 kJ/mole and the guanidine hydrochloride concentration at the transition midpoint is 2.12 M. The probability of one appeared in the protein is 0.19 and another is 0.81. This indicates that the binding of Cu(II) makes the secondary structure of 81% integrative apo-CopC be stabilized. The stabilized energy $\Delta\Delta G^0_{Cu(II)}$ is 11.84 kJ/mole. Since Cu(II) binds at N-terminal of apoCopC the first transition corresponds to the partial unfolding of apoCopC and it is the part the C-terminal of apoCopC.

Using urea as chemical denaturant the unfolding of Cu^{2+}-CopC shows three-state process. The free energy change is calculated to be 29.16 kJ/mole by using the new model. The stabilized energy $\Delta\Delta G^0_{Cu(II)}$ is only 4.18 kJ/mole. At pH 7.4 apoCopC is a protein charged positively since the pI of the protein is 8.27. After binding of Cu(II) the positive charges of the protein is increased and anion ions would make Cu^{2+}-CopC be stable. As an ionic denaturant, guanidine hydrochloride makes Cu^{2+}-CopC be stable.

In a similar way the unfolding of Cu^{2+}-CopC-Ag^{+} is studied by measuring the fluorescence ratio at 320 nm and 400 nm. Two transitions appear in the unfolding profile (Figure 17). Corresponding to the unfolding of C-terminal the first transition midpoint is shifted from 1.53 M to lower guanidine hydrochloride concentration (1.42 M), while the second transition midpoint is shifted from 2.12 M to higher guanidine hydrochloride concentration (2.38 M). Comparing with Cu^{2+}-CopC the binding of Ag(I) makes structural element one be unstable and structural element two be stable. This means that the structural element one and two are affected by the binding of Ag(I). Although the structural elements are marked as structural element one and two in Cu^{2+}-CopC-Ag^{+}, they are different with that in Cu^{2+}-CopC.

The data (■) of Cu^{2+}-CopC-Ag^{+} are fitted well in eq (21).

$$Y_{app} = 0.24\times Y_{app}^{1}(m_1,[D]_{1/2}^{1}) + 0.76\times Y_{app}^{2}(m_2,[D]_{1/2}^{2}) \tag{21}$$

It is noticed that Y_{app}^{1} and Y_{app}^{2} in eq (21) are different with that in eq (19). From eqs (8) free energy change of Cu^{2+}-CopC-Ag^{+} can be calculated to be 36.71 kJ/mole. As Ag(I) binding to apoCopC the binding of Ag(I) to Cu^{2+}-CopC doesn't change free energy change of the protein and affected only the structural elements and the probabilities of structural elements appeared in the protein. Comparing eq (21) with eq (19), the contribution of structural element one is increased and two is decreased.

Hydrogen bonding and hydrophobic forces have major influence to the structure of native protein. Some molecules are often used in researching stability of native proteins since their perturbation on the hydrogen bonding and hydrophobic forces [31, 32]. Apo-Y79W-W83F is a mutant of apoCopC, in which tyrosine residue 79 is mutated as tryptophan and tryptophan residue is did as phenylalanine. As shown in Figure 18A the fluorescence spectrum of apoY79W-W83F is dominated by tryptophan residue at position 79 and centered at 328 nm. Comparing the fluorescence spectrum of apoCopC the maximum peak is shifted toward to red around 8 nm and the tryptophan residue is yet a typical for buried tryptophyl side chains in apolar environment. The protein binds one equivalent of Cu(II) with high affinity and the fluorescence at 328 nm is quenched[33].

Since the hydrogen bonding and hydrophobic forces in apoCopC are changed, the GdnHCl-induced unfolding profile of apo-Y79W-W83F shows an obvious two transitions. According to the new model apo-Y79W-W83F is consisted of two-kind structural elements, the unfolding curve can be fitted from Y_{app}^{1} and Y_{app}^{2} from eq (8), and free energy change can be calculated using eq (12). Although free energy change corresponding to apo-Y79W-W83F unfolding is nearly same with that of apoCopC (unpublished data) the structural elements in apo-Y79W-W83F are different with that of apoCopC.

Figure 18B shows GdnHCl-induced unfolding of Cu^{2+}-Y79W-W83F (◯). Analyzing the unfolding data according to two-state unfolding model, the free energy changes plotted as a function of denaturant concentration are an obviously inflection point, suggesting that GdnHCl-induced unfolding of Cu^{2+}-Y79W-W83F does not proceed via a simple two-state process but involves one intermediate state. Using eq (8) the unfolding data of Cu^{2+}-Y79W-W83F in Figure 18B can be fitted well. Y_{app}^{1} and Y_{app}^{2} are the unfolding curves of structural elements 1 and 2, respectively. The obtained thermodynamic parameters are listed Table 2.

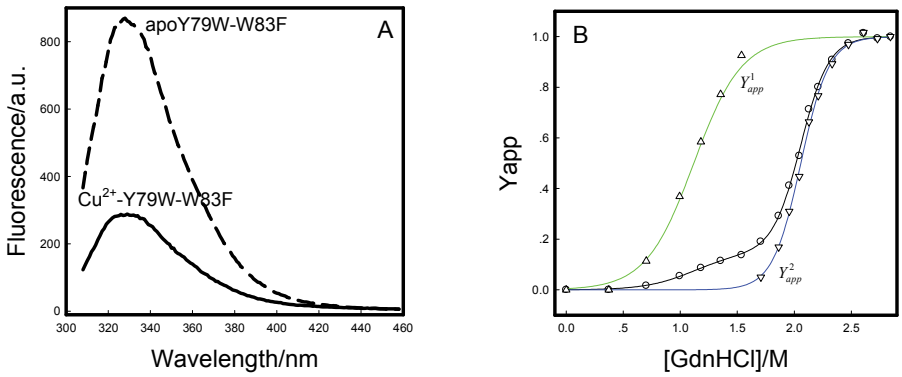


Fig. 18. Fluorescence spectra of Y79W-W83F (A) and GdnHCl-induced unfolding of Cu^{2+}-Y79W-W83F by monitoring the fluorescence 400/320 nm ratio (○) (B). Lines are the fitting from two-state and three-state eqs. (Δ) and (∇) are the normalized data for first and second transitions, respectively.

Tryptophan residue 83 locates at N-terminal and tyrosine residue 79 locates at C-terminal of apoCopC. After binding Cu(II) the positions of the residues have no large change (Figure

15). Two transitions are observed in GdnHCl-induced unfolding profiles of Cu^{2+}-CopC-Ag^{+}, Cu^{2+}-CopC, and Cu^{2+}-Y79W-W83F. The first transition corresponds to the unfolding of C-terminal and second does to that of N-terminal. It can be seen that the mutation from tryptophan to phenylalanine residue at position of 83 makes the N-terminal stability be increased, while the mutation from tyrosine to tryptophan residue makes the C-terminal stability be decreased a little. This indicates that the aromatic ring stacking at position 83 plays an important role in keeping the secondary native structure of apoCopC.

7. Summary

A new model to calculate the free energy change of proteins unfolding is presented. In the new proposed model proteins are consisted of different structural elements; the unfolding of structural element obeys two-state mechanism and the free energy change of the element can be obtained by using linear extrapolation method; if the protein is consisted of same structural elements unfolding of the protein displays a two-state process, while the unfolding of the protein shows a multi-state behavior; A least-square fitting is used to analyze the contribution of structural element to the protein and free energy change of the protein is the probable summing of all structural element. The model can be used to analyze and compare the stability of proteins, which are similar in size and structure.

8. References

[1] Santoro M M and Bolen D W, A test of the linear extrapolation of unfolding free energy changes over an extended denaturant concentration range, Biochemistry, 1992, 31, 4901-4907.

[2] Bolen D W and Santoro M M, Unfolding free energy changes determined by the linear extrapolation method. 2. Incorporation of ΔG0N-U values in a thermodynamic cycle, Biochemistry, 1988, 27, 8069-8074.

[3] Santoro M M and Bolen D W, Unfolding free energy changes determined by the linear extrapolation method. 1. Unfolding of phenylmethanesulfonyl α-chymotrypsin using different denaturant, Biochemistry, 1988, 27, 8063-8068.

[4] Matthews C R and Crisanti M M, Urea-induced unfolding of the α subunit of tryptophan synthase: evidence for a multistate process, Biochemistry, 1981, 20, 784-792.

[5] Muzammil S, Kumar Y and Tayyab S, Anion-induced stabilization of human serum albumin prevents the formation of intermediate during urea denaturation, Proteins: Structure, Function, and Genetics, 2000, 40, 29-38.

[6] Xu X L, Liu Q L, Yu H M and Xie Y S, Ca(II)- and Tb(III)-induced stabilization and refolding of anticoagulation factor I from the venom of Agkistrodon acutus, Protein Science, 2002, 11, 944-956.

[7] Hussain F, Sedlak E and Wittung-Stafshede P, Role of copper in folding and stability of cupredoxin-like copper-carrier protein CopC, Arch Biochem Biophys, 2007, 467, 58-66.

[8] Wang X Y, Zhang Z R and Perrett S, Characterization of the activity and folding of the glutathione transferase from Escherichia coli and the roles of residues Cys10 and His106, Biochem J, 2009, 417, 55-64.

[9] Englander S W, Protein folding intermediates and pathways studied by hydrogen exchange. Annu Rev Biophys Biomol Struct, 2000, 29, 213-239.

[10] Soderberg B O, Three-dimensional structure of Escherichia coli thioredoxin-S2 to 2.8 A resolution. Proc Natl Acad Sci USA, 1975, 72, 2305-2309.
[11] Katti S K, LeMaster D M, Crystal structure of thioredoxin from Escherichia coli at 1.68A resolution. J Mol Biol, 1990, 212, 167-184.
[12] Voet D, Voet J G, Biochemistry, Chapter 7: Three-dimensional structures of proteins, Second Ed. 1995, p141.
[13] Greene Jr R F, Pace C N, Urea and guanidine hydrocholide denaturation of ribonuclease, lysozyme, α-chymotrypsin, and β-lactoglobulin. J Biol Chem, 1974, 249, 5388-5393.
[14] Wallevik K, Reversible denaturation of human serum albumin by pH, Temperature, and Guanidine Hydrochloride Followed by Optical Rotation. J Biol Chem, 1973, 248, 2650-2655.
[15] Kumar Y, Muzammil S, and Tayyab S, Influence of Fluoro, Chloro and Alkyl Alcohols on the Folding Pathway of Human Serum Albumin, J Biochem, 2005, 138, 335-341.
[16] Halim A A A, Kadir H A, and Tayyab S, Bromophenol blue binding as a probe to study urea and guanidine hydrochloride denaturation of bovine serum albumin. J Biochem. 2008, 144, 33–38.
[17] Mostafa R T, Moghaddamnia S H, Ranjbar B, Amani M, Marashi S A, Conformational study of human serum albumin in pre-denaturation temperatures by differential scanning calorimetry, circular dichroism and UV spectroscopy. J Biochem Mol Biol, 2006, 39, 530-536.
[18] Yang B S, Yang P, Song L H, The binding of Gd(III) to human serum albumin, Kexue Tongbao, 1984, 29, 1502-1505.
[19] Yang B S, Yang P, The action of rare earth ions with human serum albumin. Chin J Biochem Biophys, 1989, 21, 302-305.
[20] Ong H N, Arumugam B, Tayyab S, Succinylation-induced conformational destabilization of lysozyme as studied by guanidine hydrochloride denaturation. J Biochem, 2009, 146, 895–904.
[21] Arnesano F, Banci L, Bertini I, et al. Solution structure of CopC: a cupredoxin-like protein involved in copper homeostasis. Structure, 2002, 10, 1337–1347.
[22] Arnesano F, Banci L, Bertini I, et al. A strategy for the NMR characterization of type II copper(II) proteins: the case of the copper trafficking protein CopC from Pseudomonas syringae. J Am Chem Soc, 2003, 125, 7200–7208.
[23] Koay M, Zhang L Y, Yang B S, Maher M J, Xiao Z G, Wedd A G, CopC Protein from Pseudomonas syringae: Intermolecular Transfer of Copper from Both the Copper(I) and Copper(II) Sites. Inorg Chem, 2005, 44, 5203-5205.
[24] Pang E G, Zhao Y Q, Yang B S, Fluorescence study on the interaction between apoCopC and cupric. Chin Sci Bullet, 2005, 50, 2302-2305.
[25] Zheng X Y, Pang E G, Zhao Y Q, Yang B S. Spectral studies on the interaction between mercuric ion and apoCopC. Chin J Chem, 2007, 25: 630–634.
[26] Song Z, Zheng X Y, Pang E G, Yang B S, Spectral studies on the interaction of Vitamin B6 with different CopC. Chem J Chin Univ, 2011, 32 .
[27] Zheng X Y, Pang E G, Zhao Y Q, Jiao Y, Yang B S, Investigation on the Inclusion Behavior of ApoCopC with Vitamin B6. Supramol Chem, 2008, 20, 553–557.
[28] Li H Q, Zheng X Y, Pang E G, Zhao Y Q, Yang B S, The effect of Trp83 mutant on the properties of CopC. Spectrochim Acta A, 2008, 70(2)，384-388.

[29] Zheng X Y, Pang E G, Li H Q, Zhao Y Q, Yang B S, The role of cupric in maintaining the structure of CopC. Chin Sci Bullet，2007, 52, 743-747.
[30] Li H Q, Zhao Y Q, Yang B S, Apo-CopC and CopC-Cu(II) unfolding characteristics in GuHCl solution. Chin J Chem, 2009, 27, 1762-1766.
[31] Li H Q, Zhao Y Q, Zheng X Y, Yang B S，Fluorescence spectra study the perturbations of CopC native fold by 2-p-toluidinynaphthalene-6-sulfonate. Spectrochim Acta A, 2009, 72(1), 56–60.
[32] Li H Q, Zheng X Y, Zhao Y Q, Yang B S, Conformation disturbance of CopC by α-naphthylamine. Chem Res Chin Univ, 2009, 25(3), 269-272.
[33] Zheng X Y, Yang B S, An improved method for measuring the stability of a three-state unfolding protein. Chinese Sci Bullet, 2010, 55, 4120-4124.

15

Phase Diagram and Waterlike Anomalies in Core-Softened Shoulder-Dumbbell Complex Fluids

Paulo A. Netz[1], Guilherme K. Gonzatti[1], Marcia C. Barbosa[2], Juliana Z. Paukowski[2], Cristina Gavazzoni[2] and Alan Barros de Oliveira[3]
[1]*Instituto de Química, Universidade Federal do Rio Grande do Sul, Porto Alegre*
[2]*Instituto de Física, Universidade Federal do Rio Grande do Sul, Porto Alegre*
[3]*Departamento de Física, Universidade Federal de Ouro Preto, Ouro Preto*
Brazil

1. Introduction

Water is not only by far the most important liquid in nature but also the one with the most puzzling behavior. Its density decreases upon cooling below 4 °C and the compressibility and specific heat increase anomalously by cooling. Water has also unusually high boiling, freezing and critical points and also a very high viscosity. In certain range of pressures, the diffusivity of water molecules increases with increasing pressure. Nevertheless, many very simple molecular water models proved to be able to reproduce several water anomalies (Angell et al., 2000; Errington & Debenedetti, 2001; Guillot, 2002; Kumar et al., 2005; Lynden-Bell, 2010; Netz, Starr, Barbosa & Stanley, 2004; Netz et al., 2001; Pi et al., 2009; Stanley et al., 2008). Consequently, very complex thermodynamic, dynamic or macroscopic features in principle could be captured in simple intermolecular model potentials.
In addition, some other complex materials also exhibit anomalous behavior. This is the case of Se_xTe_{1-x} (Thurn & Ruska, 1976), and $Ge_{15}Te_{85}$ (Tsuchiya, 1991) which have density anomaly. Liquid sulfur displays a sharp minimum in the density (Sauer & Borst, 1967), related to a polymerization transition (Kennedy & Wheeler, 1983). Waterlike anomalies were also found in simulations for silica (Angell et al., 2000; Angell & Kanno, 1976; Poole et al., 1997; Sharma, Chakraborty & Chakravarty, 2006; Shell et al., 2002), silicon (Sastry & Angell, 2003) and BeF_2 (Agarwal et al., 2007; Angell et al., 2000; Hemmer et al., 2001). Despite being sucessfull in the description of many aspects of water and other anomalous materials behavior, however, none of these potentials could indeed reproduce all desired properties.
The understanding of the phase behavior of fluids based solely on the information about the intermolecular interactions remains as one of the biggest challenges in the statistical thermodynamics. The scientific community has already recognized that such big challenge demands a smart strategy: beginning with the simplest models.
It is recognized that the origin of the anomalies is related to the competition between open low-density and closed high-density structures, which depend on the thermodynamic state of the liquid (Krekelberg et al., 2008). In simple isotropic models, this competition is described by two preferred interparticle distances. This simple recipe gives raise to a myriad of models and

approaches (Almarza et al., 2009; Balladares & Barbosa, 2004; Buldyrev et al., 2002; Buldyrev & Stanley, 2003; de Oliveira & Barbosa, 2005; de Oliveira et al., 2007; de Oliveira, Franzese, Netz & Barbosa, 2008; de Oliveira, Netz & Barbosa, 2008; de Oliveira et al., 2009; 2006a;b; Fomin et al., 2008; Franzese et al., 2001; 2002; Gibson & Wilding, 2006; Hemmer & Stell, 1970; Henriques & Barbosa, 2005; Henriques et al., 2005; Jagla, 1998; Kurita & Tanaka, 2004; Netz, Raymundi, Camera & Barbosa, 2004; Pretti & Buzano, 2004; Scala et al., 2000; Skibinsky et al., 2004; Wilding & Magee, 2002; Xu et al., 2005).

Among these models, the core-softened shoulder potential (Cho et al., 1996; de Oliveira et al., 2006a;b; Netz, Raymundi, Camera & Barbosa, 2004) reproduces qualitatively water's structural, density and diffusion anomalies. This potential is built summing a Lennard-Jones potential and a displaced Gaussian term, as follows:

$$U^*(r) = 4\left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right] + a\exp\left[-\frac{1}{c^2}\left(\frac{r-r_0}{\sigma}\right)^2\right], \quad (1)$$

Depending of the choice of a, c and r_0, a whole family of potentials can be built, which shapes ranging from double well to a ramp-like shoulder (Netz, Raymundi, Camera & Barbosa, 2004). Molecular dynamics simulations using double well intermolecular potentials constructed with this function show anomalous behavior in the stable region of the phase diagram if the outer minimum is deeper than the inner minimum. In the case of a deeper inner minimum, anomalous behavior is also present but inside an unstable region. Another choice of parameters, yielding a ramp-like shoulder profile is $a = 5.0$, $r_0/\sigma = 0.7$ and $c = 1.0$, which result in stable fluid phases in a large portion of the phase diagram.

Using this potential, we found that the hierarchy of anomalies, i.e. the loci in the pressure-temperature phase diagram of the state points displaying anomalous behavior (de Oliveira et al., 2006a;b), is the same as in water (Errington & Debenedetti, 2001; Netz et al., 2001): the structurally anomalous region in the pressure-temperature phase diagram enclosing the region of dynamic (diffusivity) anomalies, that in its turn is enclosing region of the thermodynamic (density) anomalies. In principle this shoulder potential can represent in an effective and orientation-averaged way the interaction between water pentamers (Krekelberg et al., 2008) characterized by the presence of two structures – one open and one closed – as discussed above. Similarly, the thermodynamic and dynamic anomalies result from the competition between the two length scales associated with the open and closed structures. The open structure is favored by low pressures and the closed structure is favored by high pressures, but only becomes accessible at sufficiently high temperatures.

Even though core-softened potentials have been mainly used for modeling water (de Oliveira & Barbosa, 2005; de Oliveira et al., 2009; 2006a;b; Franzese, 2007; Gibson & Wilding, 2006; Xu et al., 2006; Yan et al., 2006; 2005), many other materials present the so called water-like anomalous behaviour. In this sense, it is reasonable to use core-softened potentials as the building blocks of a broader class of materials which we can classify as anomalous fluids. One interesting aspect in this investigation is the role of multiple length scales and of anisotropy. It has been shown that the presence of many length scales in a spherical symmetric potential leads to only one region in the pressure-temperature phase-diagram with thermodynamic and dynamic anomalies (Netz et al., 2006) because actually only the smaller and the larger scales compete.

Obviously anisotropic systems are not only more complicated but the addition of other degrees of freedom result in a richer phase-diagram. For instance diatomic particles interacting through a Lennard-Jones potential (Kriebel & Winkelmann, 1996; Sumi et al.,

2004) exhibit a solid phase that occupies higher pressures and temperatures in the pressure-temperature phase diagram. In the case of the solid phase, the diatomic particles exhibit two close-packed arrangements instead of one as observed in the monoatomic Lennard-Jones (Vega et al., 2003). Therefore, it is expected that the dimeric system interacting through two length scales potential might have a phase-diagram with a larger solid phase region in the pressure-temperature phase diagram than the one occupied by the solid phase in the monomeric system.

In the recently proposed model of dimeric molecules linked as rigid dumbbells, interacting with the shoulder potential described above (de Oliveira et al., 2010) the anisotropy due to the dumbbell leads indeed to a much larger solid phase and to the appearance of a liquid crystal phase. This system has a peculiar phase behavior: by isothermal increase of the pressure at moderate temperatures, the system, beginning in the fluid phase, becomes solid, at high pressures it becomes fluid again and by further increase of the pressure it becomes liquid-crystal-like before becoming a solid.

The regions of thermodynamic, dynamic and structural anomalies display the same hierarchy as in the monomeric case. In general, the range of pressures and temperatures occupied by these anomalous regions are larger than those observed in the monomeric case. The differences are more pronounced for the regions of diffusion and structural anomalies and not so large for the thermodynamic anomalies.

These results showed that the dimeric particles interacting with the core softened shoulder potential have a different behavior when compared with the monomeric particles with the same interaction potential. This raises the following question: how to quantify the influence of the anisotropy, i.e. what is the role of the strength of anisotropy, the interparticle separation λ, on the phase and anomalous behavior?

We consider a set of $N/2$ dimeric molecules formed each by two spherical particles of diameter σ, linked rigidly in pairs with the distance of λ between their centers of mass, as depicted in Fig. 1. Each particle within a dimer interacts with all particles belonging to other dimers with the intermolecular continuous shoulder potential (de Oliveira et al., 2006b) given by equation 1 with the parameters: $a = 5.0$, $r_0/\sigma = 0.7$, $c = 1.0$ and comparing several values of λ.

In the following section we describe some details of the simulation methodology, the choice of state points and the methods of analysis and quantification of the anomalous behavior. In the next section, detailed results from simulations using λ = 0.50 are compared with the previously simulated system with λ = 0.20. The results pointed out to further simulations with other values of λ, which are also discussed in this section. Finally, in the conclusions section, a survey of all we have learnt from this system is given, as well as we point out some perspectives.

2. Methods

We performed molecular dynamics simulations in the canonical ensemble using $N = 500$ particles (250 dimers) in a cubic box with periodic boundary conditions, interacting with the intermolecular potential described above. The cutoff radius was set to 5.5 length units. Pressure, temperature, density, and diffusion are calculated in dimensionless units, as detailed elsewhere (de Oliveira et al., 2010). We first compared the previous (de Oliveira et al., 2010) results, obtained with the choice of $\lambda/\sigma = 0.20$ with a new choice $\lambda/\sigma = 0.50$, in a broad range of temperatures ($0.10 \leq T^* \leq 3.00$) and densities ($0.10 \leq \rho^* \leq 0.50$). In order to follow the trend of the anomalies on lambda, several additional sets of simulations in a rather small

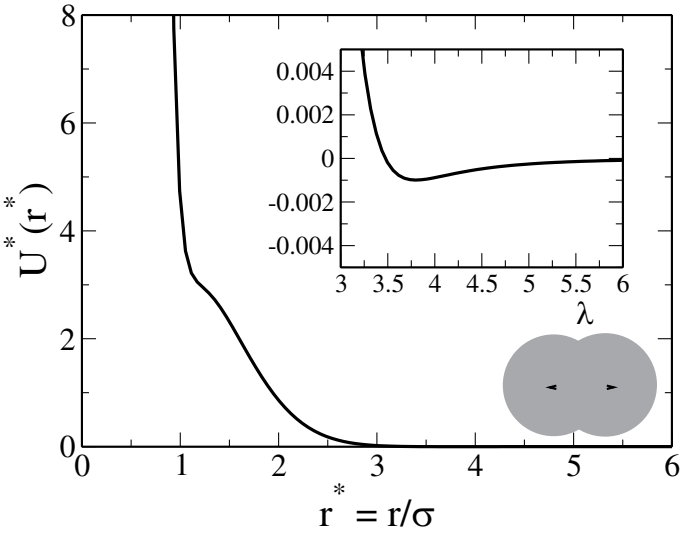


Fig. 1. Effective potential versus distance, in reduced units.

region of the phase diagram ($0.60 \leq T^* \leq 0.94$ and $0.20 \leq \rho^* \leq 0.28$) were carried out, for λ/σ = 0.10, 0.30, 0.40 and 0.70.

Thermodynamic and dynamic properties were calculated over 700 000 steps after previous 200 000 equilibration steps. The time step was 0.001 in reduced units, the time constant of the Berendsen thermostat (Berendsen et al., 1984) was 0.1 in reduced units. The internal bonds between the particles in each dimer remain fixed using the SHAKE (Ryckaert et al., 1977) algorithm, with a tolerance of 10^{-12} and maximum of 100 interactions for each bond.

The stability of the systems was checked analyzing the dependence of pressure on density and also by visual analysis of the final structure, searching for cavitation. The structure of the system was characterized using the intermolecular radial distribution function, $g(r)$ (RDF), which does not take into account the correlation between atoms belonging to the same molecule. The diffusion coefficient was calculated using the slope of the least square fit to the linear part of the mean square displacement, $< r^2(t) >$ (MSD), averaged over different time origins. Both the $g(r)$ and $< r^2(t) >$ where computed taking the origin as the center of mass of a dimer. The phase boundary between solid and fluid phase was mapped by analysis of the change of the pattern of the mean squared displacement and radial distribution function.

For the quantification of the structural anomaly we use the translational order parameter t (Errington & Debenedetti, 2001), given by

$$t \equiv \frac{1}{\xi_c} \int_0^{\xi_c} |g(\xi) - 1| d\xi, \tag{2}$$

which decreases upon increasing density in the structurally anomalous region. Here $\xi \equiv r\rho^{1/3}$ is the distance r in units of the mean interparticle separation, computed by the center of mass of the dimers, $\rho^{-1/3}$, ξ_c is the cutoff distance set to half of the simulation box times $\rho^{-1/3}$. $g(\xi)$ is the radial distribution function as a function of the (reduced) distance ξ from a reference particle. For an ideal gas $g = 1$ and $t = 0$. In the crystal phase $g \neq 1$ over long distances and t is large.

In a previous work we showed that (de Oliveira et al., 2010), depending on the chosen temperature and density, the system could be in a fluid phase metastable with respect to the solid phase. In order to check this case, for $\lambda/\sigma = 0.50$, two sets of simulations were carried out, one with a ordered crystalline initial configuration and other with a liquid configuration obtained from previous equilibrium simulations. No noticeable differences were detected in the results of both sets.

3. Results

We carried out a detailed set of simulations of rigid dumbbells interacting with the potential 1 with $\lambda/\sigma = 0.50$. In order to locate the anomalous regions for this system, the following methodology was adopted.

First of all, it is needed to locate the phase boundary, i.e. the loci in the phase diagram separating the solid and fluid phases. This was determined by the change of pattern in the mean square displacement and radial distribution function. Fig. 2 shows that, for a chosen fixed density ($\rho = 0.22$ in this figure), the system becomes fluid for temperatures above $T = 0.39$, as can be seen by the abrupt change of the slope of the mean square diplacement and the change of a solid-like radial distribution function below this temperature to a fluid-like g(r) above that. This transition was independent of the starting configuration: simulations starting from the ordered crystalline initial configuration and from a previoulsy equilibrated liquid configuration showed the same behavior. For other densities the location of the phase boundary was done in the same way.

Second, in order to map the region of thermodynamic anomalies, one has to locate the temperature of maximum density, TMD. The TMD was determined plotting the pressure against the temperature along isochores and locating the minima, as shown in Fig. 3. The isochores located between $2.0 < P^* < 3.5$ have minimum at a certain temperature and pressure corresponding to a temperature of maximum density (TMD) at a given P^*.

Third, the region of diffusion anomalies is defined as the region where the mobility behaves anomaloulsy, increasing D with increasing density. Indeed, for $T^* > 0.65$, the plot of the diffusion coefficient D against the density ρ along isotherms shows the same anomalous behavior (see Figure 3) already known for the waterlike core-softened shoulder monomeric (de Oliveira et al., 2006a;b) and dimeric (de Oliveira et al., 2010) systems. For low densities $\rho < \rho_{D\min}$ D decreses with increasing ρ (normal behavior); for intermediate densities $\rho_{D\min} < \rho < \rho_{D\max}$, D increases with increasing ρ (anomalous behavior) and for high densities $\rho > \rho_{D\max}$, the behavior turns out to be the same of the normal fluids. For each temperature there is a $\rho_{D\min}$ and a $\rho_{D\max}$ and consequently a $P_{D\min}(T)$ and a $P_{D\max}(T)$. The line of $P_{D\max}(T)$ in the P-T phase diagram illustrated in Fig. 5 is similar to the diffusivity maxima observed in experiments for water as well as in simulations of water (Errington & Debenedetti, 2001; Mittal et al., 2006; Netz et al., 2001), silica, (Poole et al., 1997; Sharma, Mudi & Chakravarty, 2006; Shell et al., 2002), other isotropic potentials (Xu et al., 2006; Yan et al., 2006; 2005) and for the potential Fig. 1 in the case in which the particles are monomeric (de Oliveira et al., 2006b). The line of $P_{D\min}(T)$ is also shown in Fig. 5. The region of diffusion anomaly in the P^*-T^* phase diagram is bounded by $P_{D\min}(T)$ and $P_{D\max}(T)$.

Finally, the location of the structurally anomalous region demands mapping the extrema in the translational order parameter t. Fig. 4 shows the behavior of t along isotherms against density. It is expected, for normally behaved fluids, the increase of translational order parameter with increasing compression (as in the red and blue portions of the isotherms). The region where t decreases with increasing compression (delimited by the maximum and minimum) is shown

in black and corresponds to the structurally anomalous region. For low temperatures, the behavior is strikingly anomalous, stronger than in the monomeric case. It is also shown the typical pattern in the radial distribution functions, corresponding to the three regions seen in the figure. For low densities, the compression leads to an increase on the first peak of the radial distribution function, thus increasing t. For intermediate densities, a new, short range peak appears. The intensity of this inner peak increases and the former first peak (now termed the second peak) decreases in intensity. This competition is the main characteristic of the structurally anomalous region. Finally, a normal behavior is recovered at very high densities, because the inner peak dominates and grows leading to the increase in t.

Fig. 5 shows the pressure-temperature phase diagram of the shoulder-dumbbell system with the parameters described above and with the choice of $\lambda/\sigma = 0.5$. For comparison, it is also shown the corresponding phase diagram obtained previously (de Oliveira et al., 2010) with $\lambda/\sigma = 0.2$. For both cases, two solid regions are clearly seen, as well as fluid phases with different densities.

The boundaries of the anomalous regions depend clearly on λ. The solid-fluid phase boundary and the TMD are shifted towards lower temperatures and slightly higher pressures with the increase of lambda. The high-temperature limit of the boundary of the region of diffusion anomalies becomes also shifted towards lower temperatures, whereas the lower-temperature part of this regions becomes broader. The structurally anomalous region shrinks with increasing λ.

In order to confirm the influence of λ on the phase diagram and regions of anomalous behavior, we carried out several simulations restricted to the small region in the phase diagram between the densities $\rho = 0.20$ and $\rho = 0.28$ (which correspond roughly to $2 < P^* < 4$) and temperatures T = 0.60 and T = 0.94. Figure 6 shows the solid-fluid phase boundary for the systems with λ/σ = 0.10, 0.20, 0.40, 0.50 and 0.70. The increase of the interparticle separation shifts the solid-fluid phase boundary to lower temperatures.

Figure 7 show the radial distribution functions for a chosen state point (density 0.20 and temperature 0.60) as a function of λ. The smaller values of λ tend to increase the solid-like nature of the system.

Comparing the results of the simulations corresponding to the several values of λ with the results of the monomeric shoulder-dumbbell simulations, we see an unexpected non-monotonic behavior. The effect of the introduction of a rather small anisotropy due to the dimeric nature of the particle (small values of the interparticle separation λ) leads to the increase of the size of the regions of anomalies (de Oliveira et al., 2010). Nevertheless, the increase of λ shrinks those regions.

For very high densities, however, the boundary between solid and liquid phases is not sharp and a broad metastable region exhibiting liquid-crystal-like behavior is found. This state can be recognized by the comparison of the mean square displacement and radial distribution function for several temperatures with a fixed density. In this case there is no clear-cut change of pattern as the state point moves from solid to liquid (as in Figure 2 and it is seen, instead, a seemingly contradictory liquid-like mean square displacement and solid-like radial distribution function (see Fig 8).

This liquid-crystal-like behavior was also found in the shoulder-dumbbell system with $\lambda/\sigma = 0.20$ (de Oliveira et al., 2010). In this state the particles have a crystal-like ordering, but diffuse as a one-dimensional string.

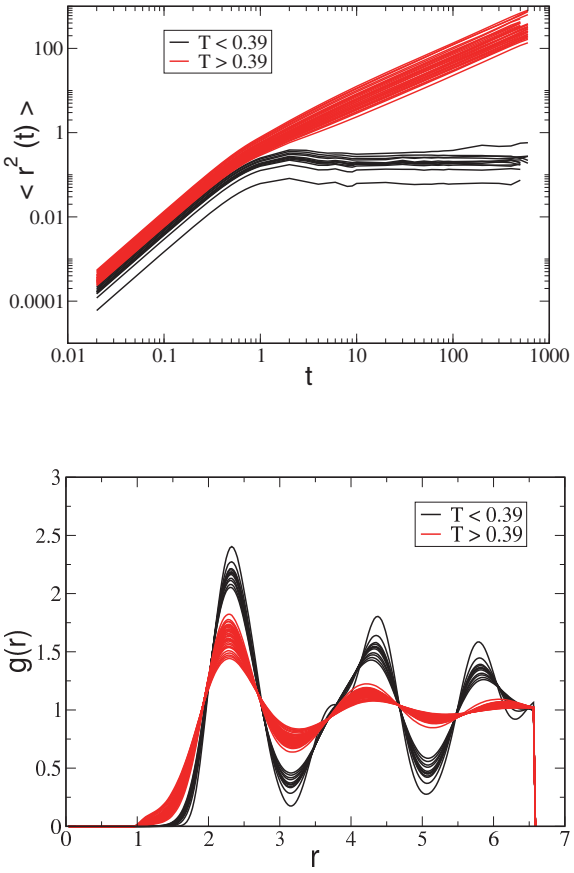


Fig. 2. Mapping the phase boundary between solid and fluid by the change of profile in the mean square displacement (a) and radial distribution function (b) for rigid dumbbells with interatomic separation $\lambda/\sigma = 0.50$ and density 0.22.

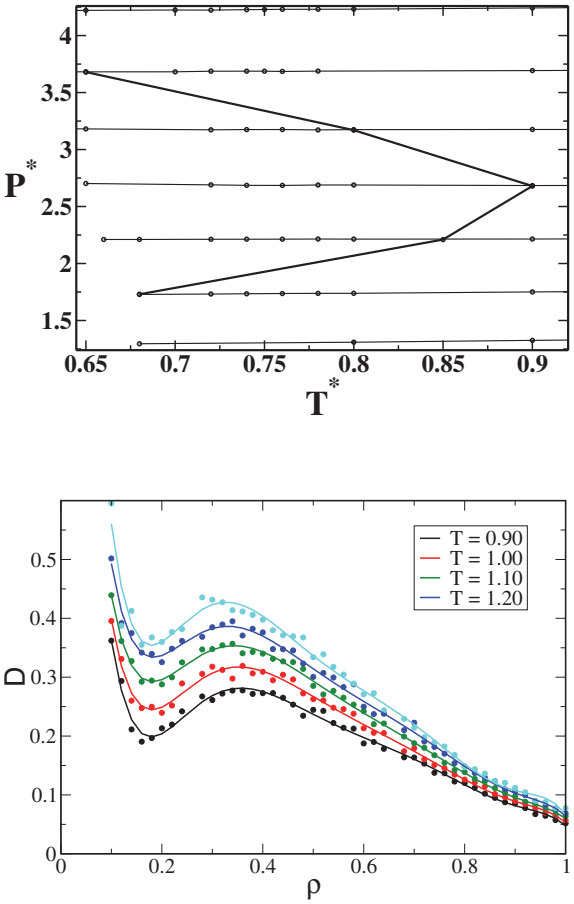


Fig. 3. Locating the TMD by the minimum of pressure along isochors (a) and the region of diffusion anomaly (b) for rigid dumbbells with interatomic separation $\lambda/\sigma = 0.50$.

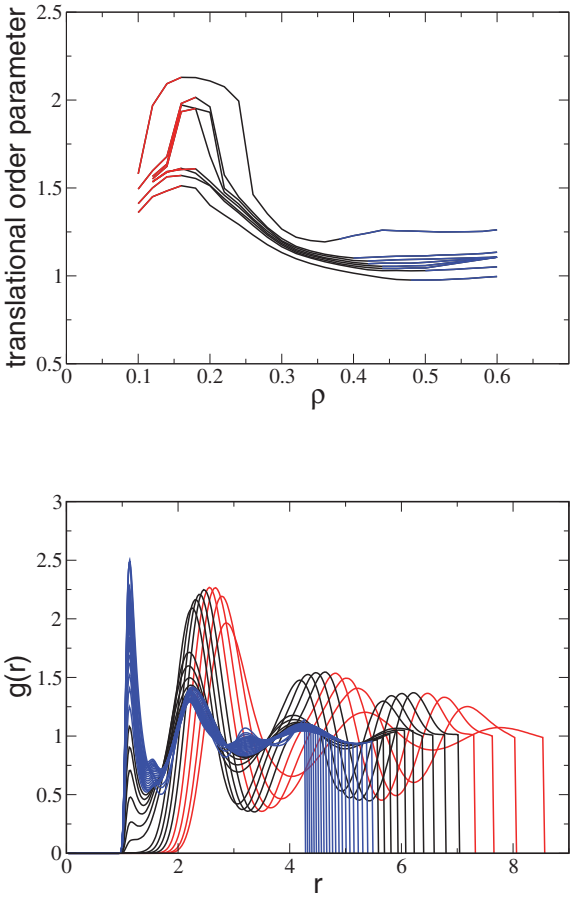


Fig. 4. (a) Translational order parameter for rigid dumbbells with interatomic separation $\lambda/\sigma = 0.50$ as a function of density for temperatures T = 0.30, 0.40, 0.42, 0.44, 0.46, 0.48, 0.50 and 0.60 (from top to bottom); (b) radial distribution functions (g (r)) for a fixed temperature T = 0.30 and several densities. The colors correspond to the densities where the translational order parameter increases (red), decreases (black) or increases (blue) under increasing density, as in (a).

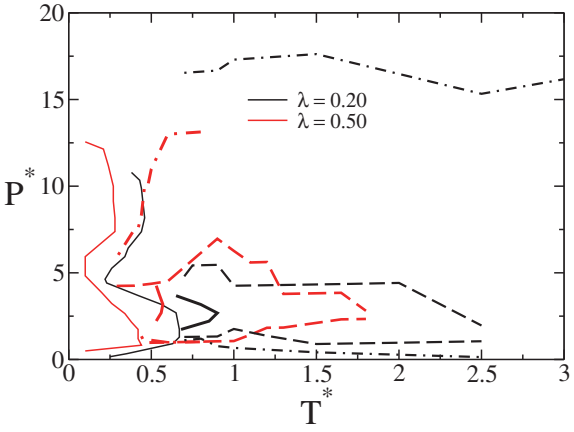


Fig. 5. Pressure versus temperature phase diagram of the potential illustrated in Figure 1 for for rigid dumbbells with interatomic separation $\lambda/\sigma = 0.20$ and $\lambda/\sigma = 0.50$. The thin solid lines represent the boundaries between the fluid and the solid phases, the bold solid line, the TMD's lines, the dashed lines represent the D extrema lines and the dotted-dashed lines the boundary of the structurally anomalous region.

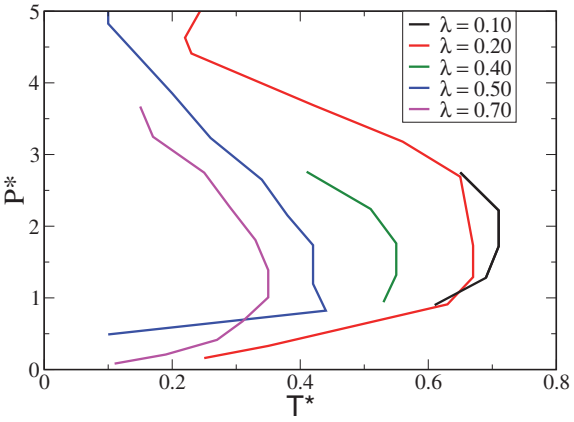


Fig. 6. Phase boundary for systems with λ/σ = 0.10, 0.20, 0.40, 0.50 and 0.70.

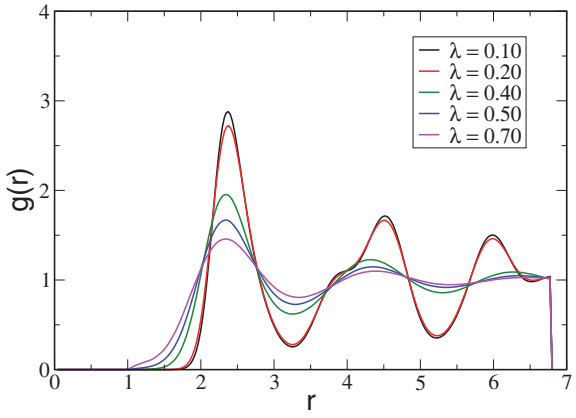


Fig. 7. Radial distribution functions for a chosen state point (density 0.20 and temperature 0.60) showing the influence of the interparticle separation λ on the structure.

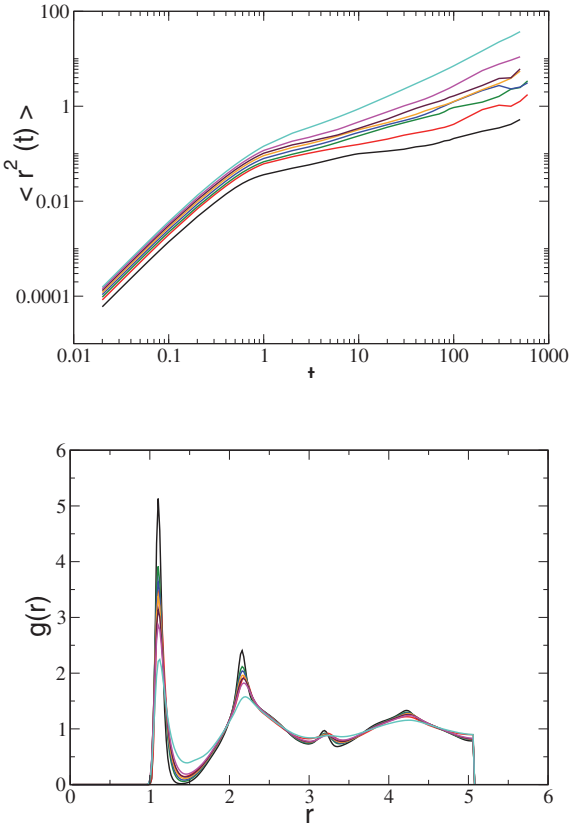


Fig. 8. Mean square displacement (a) and radial distribution function (b) for rigid dumbbells with interatomic separation $\lambda/\sigma = 0.50$ and density 0.48. The curves show $< r^2 >$ or $g(r)$ for temperatures T = 0.10 (black), 0.12, 0.14, 0.16, 0.18, 0.20, 0.22, 0.24 and 0.26 (turquoise).

4. Conclusions

In this work we have simulated a system consisting of rigid dimers in which each monomer interacts with monomers from the other dimers through a core-softened shoulder potential. In order to check how the anisotropy induced by the dimeric structure affects the presence of density, diffusion and structural anomalies, we have obtained the pressure temperature phase diagram, the diffusion constant and the structural order parameter of the system for different λ, distance between the bonded particles in each dimer, values.

The understanding of this effect is a fundamental step towards modeling more complex systems such as polymers in which anomalies would be present.

We found that this system has thermodynamic, dynamic and structural anomalies similar to the anomalies present in the monomeric case, but the size of the anomalous region in the pressure temperature phase diagram is dependent on the value of λ.

In particular we found that in the range $0.1 < \lambda < 0.7$ the pressures and the temperatures of the melting line decrease with increasing λ. This result is consistent with the idea that the anisotropy enhanced by large values of λ made it difficult for the system to reach the proper solid configuration. Any small thermal motion disrupts the structure that finds it easier to be in a solid phase. The same argument also explains why the region in the pressure temperature phase diagram occupied by the density, diffusion and structural anomalous regions is smaller for large λ values than for smaller values. As the anisotropy becomes larger, the competition between the two length scales that is responsible for the anomalies becomes less important for the organization of the particles than the dimer scale. As a result the anomalous regions do exist only for a narrow range of pressures and temperatures where the dimeric scale does not prevail.

5. References

Agarwal, M., Sharma, R. & Chakravarty, C. (2007). Ionic melts with waterlike anomalies: Thermodynamic properties of liquid BeF_2, *J. Chem. Phys.* 127: 164502.

Almarza, N. G., Capitan, J. A., Cuesta, J. A. & Lomba, E. (2009). Phase diagram of a two-dimensional lattice gas model of a ramp system, *J. Chem. Phys* 131: 124506.

Angell, C. A., Bressel, R. D., Hemmatti, M., Sare, E. J. & Tucker, J. C. (2000). Water and its anomalies in perspective: tetrahedral liquids with and without liquid-liquid phase transitions., *Phys. Chem. Chem. Phys.* 2: 1559.

Angell, C. A. & Kanno, H. (1976). Density maxima in high-pressure supercooled water and liquid silicon dioxide, *Science* 193: 1121.

Balladares, A. & Barbosa, M. C. (2004). Density anomaly in core-softened lattice gas, *J. Phys.: Cond. Matter* 16: 8811.

Berendsen, H. J. C., Postuma, J. P. M., van Gunsteren, W. F., DiNola, A. & Haak, J. R. (1984). Molecular dynamics with coupling to an external bath, *J. Chem. Phys.* 81: 3684–3690.

Buldyrev, S. V., Franzese, G., Giovambattista, N., Malescio, G., Sadr-Lahijany, M. R., Scala, A., Skibinsky, A. & Stanley, H. E. (2002). Models for a liquid-liquid phase transition, *Physica A* 304: 23.

Buldyrev, S. V. & Stanley, H. E. (2003). A system with multiple liquid-liquid critical points, *Physica A* 330: 124.

Cho, C. H., Singh, S. & Robinson, G. W. (1996). Liquid water and biological systems: the most important problem in science that hardly anyone wants to see solved, *Faraday Discuss.* 103: 19.

de Oliveira, A. B. & Barbosa, M. C. (2005). Density anomaly in a competing interactions lattice gas model, *J. Phys.: Cond. Matter* 17: 399.

de Oliveira, A. B., Barbosa, M. C. & Netz, P. A. (2007). Interplay between structure and density anomaly for an isotropic core-softened ramp-like potential, *Physica A* 386: 744.

de Oliveira, A. B., Franzese, G., Netz, P. A. & Barbosa, M. C. (2008). Waterlike hierarchy of anomalies in a continuous spherical shouldered potential, *J. Chem. Phys* 128: 064901.

de Oliveira, A. B., Netz, P. A. & Barbosa, M. C. (2008). Which mechanism underlies the water-like anomalies in core-softened potentials?, *Euro. Phys. J. B* 64: 481.

de Oliveira, A. B., Netz, P. A. & Barbosa, M. C. (2009). An ubiquitous mechanism for water-like anomalies, *Europhys. Lett.* 85: 36001.

de Oliveira, A. B., Netz, P. A., Colla, T. & Barbosa, M. C. (2006a). Structural anomalies for a three dimensional isotropic core-softened potential, *J. Chem. Phys.* 125: 124503.

de Oliveira, A. B., Netz, P. A., Colla, T. & Barbosa, M. C. (2006b). Thermodynamic and dynamic anomalies for a three-dimensional isotropic core-softened potential, *J. Chem. Phys.* 124: 084505.

de Oliveira, A. B., Neves, E. B., C., G., Paukowski, J. Z., Netz, P. A. & Barbosa, M. C. (2010). Liquid crystal phase and waterlike anomalies in a core-softened shoulder-dumbbells system, *J. Chem. Phys* 132: 164505.

Errington, J. R. & Debenedetti, P. D. (2001). Relationship between structural order and the anomalies of liquid water, *Nature (London)* 409: 318.

Fomin, D. Y., Frenkel, D., Gribova, N. V. & Ryzhov, V. N. (2008). Quasibinary amorphous phase in a three-dimensional system of particles with repulsive-shoulder interactions, *J. Chem. Phys.* 127(064512).

Franzese, G. (2007). Differences between discontinuous and continuous soft-core attractive potentials: The appearance of density anomaly, *J. Mol. Liq.* 136(3): 267.

Franzese, G., Malescio, G., Skibinsky, A., Buldyrev, S. V. & Stanley, H. E. (2001). Generic mechanism for generating a liquid-liquid phase transition, *Nature (London)* 409: 692.

Franzese, G., Malescio, G., Skibinsky, A., Buldyrev, S. V. & Stanley, H. E. (2002). Metastable liquid-liquid phase transition in a single-component system with only one crystal phase and no density anomaly, *Phys. Rev. E* 66: 051206.

Gibson, H. M. & Wilding, N. B. (2006). Metastable liquid-liquid coexistence and density anomalies in a core-softened fluid, *Phys. Rev. E* 73: 061507.

Guillot, B. (2002). A reappraisal of what we have learnt during three decades of computer simulations on water, *J. Mol. Liq.* 101: 219.

Hemmer, P. C. & Stell, G. (1970). Fluids with several phase transitions, *Phys. Rev. Lett.* 24: 1284.

Hemmer, P. C., Velasco, E., Medeiros, L., Navascués, G. & Stell, G. (2001). Solid-solid transitions induced by repulsive interactions, *J. Chem. Phys.* 114: 2268.

Henriques, V. B. & Barbosa, M. C. (2005). Liquid polymorphism and density anomaly in a lattice gas model, *Phys. Rev. E* 71: 031504.

Henriques, V. B., Guissoni, N., Barbosa, M. A., Thielo, M. & Barbosa, M. C. (2005). Liquid polyamorphism and double criticality in a lattice gas model, *Mol. Phys.* 103: 3001.

Jagla, E. A. (1998). Phase behavior of a system of particles with core collapse, *Phys. Rev. E* 58: 1478.

Kennedy, S. J. & Wheeler, J. C. (1983). On the density anomaly in sulfur at the polymerization transition, *J. Chem. Phys.* 78: 1523.

Krekelberg, W. P., Mittal, J., Ganesan, V. & Truskett, T, M. (2008). Structural anomalies of fluids: Origins in second and higher coordination shells, *Phys. Rev. E* 77: 041201.

Kriebel, C. & Winkelmann, J. (1996). Polarizable dipolar two-center Lennard-Jones fluids: Computer simulations and equation of state, *J. Chem. Phys.* 105: 9316.

Kumar, P., Buldyrev, S. V., Sciortino, F., Zaccarelli, E. & Stanley, H. E. (2005). Static and dynamic anomalies in a repulsive spherical ramp liquid: Theory and simulation, *Phys. Rev. E* 72: 021501.

Kurita, R. & Tanaka, H. (2004). Critical-like phenomena associated with liquid-liquid transition in a molecular liquid, *Science* 306: 845.

Lynden-Bell, R. M. (2010). Towards understanding water: simulation of modified water models, *Journal of Physics, Condensed Matter* 22: 284107.

Mittal, J., Errington, J. R. & Truskett, T. M. (2006). Quantitative link between single-particle dynamics and static structure of supercooled liquids, *J. Phys. Chem. B* 110: 18147.

Netz, P. A., Buldyrev, S., Barbosa, M. C. & Stanley, H. E. (2006). Thermodynamic and dynamic anomalies for dumbbell molecules interacting with a repulsive ramplike potential, *Physical Review E* 73: 061504.

Netz, P. A., Raymundi, J. F., Camera, A. S. & Barbosa, M. C. (2004). Dynamic anomalies of fluids with isotropic doubled-ranged potential, *Physica A* 342: 48.

Netz, P. A., Starr, F., Barbosa, M. C. & Stanley, H. E. (2004). Computer simulation of dynamical anomalies in stretched water, *Brazilian Journal of Physics* 34: 24.

Netz, P. A., Starr, F. W., Stanley, H. E. & Barbosa, M. C. (2001). Static and dynamic properties of stretched water, *J. Chem. Phys.* 115: 344.

Pi, H. L., Aragones, J. L., Vega, C., Noya, E. G., Abascal, J. L. F., Gonzalez, M. A. & McBride, C. (2009). Anomalies in water as obtained from computer simulations of the tip4p/2005 model: density maxima, and density, isothermal compressibility and heat capacity minima, *Molecular Physics* 107: 365.

Poole, P. H., Hemmati, M. & Angell, C. A. (1997). Comparison of thermodynamic properties of simulated liquid silica and water, *Phys. Rev. Lett.* 79: 2281.

Pretti, M. & Buzano, C. (2004). Thermodynamic anomalies in a lattice model of water, *J. Chem. Phys* 121: 11856.

Ryckaert, J. P., Ciccotti, G. & Berendsen, H. J. C. (1977). Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes, *J. Comput. Phys.* 23: 327.

Sastry, S. & Angell, C. A. (2003). Liquid-liquid phase transition in supercooled silicon, *Nature Mater.* 2: 739.

Sauer, G. E. & Borst, L. B. (1967). Lambda transition in liquid sulfur, *Science* 158: 1567.

Scala, A., Sadr-Lahijany, M. R., Giovambattista, N., Buldyrev, S. V. & Stanley, H. E. (2000). Applications of the Stell-Hemmer potential to understanding second critical points in real systems, *J. Stat. Phys.* 100: 97.

Sharma, R., Chakraborty, S. N. & Chakravarty, C. (2006). Entropy, diffusivity, and structural order in liquids with waterlike anomalies, *J. Chem. Phys.* 125: 204501.

Sharma, R., Mudi, A. & Chakravarty, C. (2006). Diffusional anomaly and network dynamics in liquid silica, *J. Chem. Phys.* 125: 044705.

Shell, M. S., Debenedetti, P. G. & Panagiotopoulos, A. Z. (2002). Molecular structural order and anomalies in liquid silica, *Phys. Rev. E* 66: 011202.

Skibinsky, A., Buldyrev, S. V., Franzese, G., Malescio, G. & Stanley, H. E. (2004). Liquid-liquid phase transitions for soft-core attractive potentials, *Phys. Rev. E* 69: 061206.

Stanley, H. E., Kumar, P., Franzese, G., Xu, L., Yan, Z., Mazza, M. G., Buldyrev, S. V., Chen, S. H. & Mallamace, F. (2008). Liquid polyamorphism: Possible relation to the anomalous behaviour of water, *European Physical Journal* 161: 1.

Sumi, T., Shirahama, H. & Sekino, H. (2004). A density-functional study for the liquid-vapor coexistence curve of nitrogen fluid, *J. Chem. Phys.* 121: 1014.

Thurn, H. & Ruska, J. (1976). Change of bonding system in liquid SexTe1.1 alloys as shown by density measurements, *J. Non-Cryst. Solids* 22: 331.

Tsuchiya, T. (1991). The anomalous negative thermal expansion and the compressibility maximum of molten Ge-Te alloys, *J. Phys. Soc. Jpn.* 60: 227.

Vega, C., McBride, C., de Mibuel, E., J., B. F. & Galindo, A. (2003). The phase diagram of the two center lennard-jones model as obtained from computer simulation and wertheim's thermodynamic perturbation theory, *J. Chem. Phys.* 118: 10696.

Wilding, N. B. & Magee, J. E. (2002). Phase behavior and thermodynamic anomalies of core-softened fluids, *Phys. Rev. E* 66: 031509.

Xu, L., Buldyrev, S., Angell, C. A. & Stanley, H. E. (2006). Thermodynamics and dynamics of the two-scale spherically symmetric jagla ramp model of anomalous liquids, *Phys. Rev. E* 74: 031108.

Xu, L., Kumar, P., Buldyrev, S. V., Chen, S.-H., Poole, P., Sciortino, F. & Stanley, H. E. (2005). Relation between the widom line and the dynamic crossover in systems with a liquid-liquid phase transition, *Proc. Natl. Acad. Sci. U.S.A.* 102: 16558.

Yan, Z., Buldyrev, S. V., Giovambattista, N., Debenedetti, P. G. & Stanley, H. E. (2006). Family of tunable spherically symmetric potentials that span the range from hard spheres to waterlike behavior, *Phys. Rev. E* 73: 051204.

Yan, Z., Buldyrev, S. V., Giovambattista, N. & Stanley, H. E. (2005). Structural order for one-scale and two-scale potentials, *Phys. Rev. Lett.* 95: 130604.

16

Effect of Magnetic and Mechanical Fields on Phase Liquid Crystalline Transitions in Solutions of Cellulose Derivatives

S. A. Vshivkov
Ural State University
Russia

1. Introduction

Phase diagrams of multicomponent systems provide full information on the thermodynamic compatibility of components in wide concentration and temperature ranges. One of the first phase diagrams of polymer–solvent systems was published by Papkov et al., (1937), Rogovin et al. (1937), Kargin et al. (1939). In 1941, the works of Tager & Kargin (1941) devoted to the thermodynamics of polymer solutions were published. From the end of the 1940s, systematic research into the thermodynamic properties and construction of phase diagrams of polymer solutions have been performed at the Laboratory of Colloid Chemistry (at the Polymer Chair with 1958), Ural State University. Over the course of sixty years, phase diagrams have been constructed for hundreds of polymer systems with amorphous and crystalline phase separations. Many of these data were included in textbooks, monographs, and reviews: Tager (2007), Papkov (1981), Nesterov & Lipatov (1987), Vshivkov (1991), Chalykh et al. (1998), Malkin & Kulichikhin (1996), Vshivkov et al. (1998), Vshivkov & Rusinova (1998, 2001), Rusinova & Vshivkov (1997), Klenin (1995). The phase liquid crystalline transitions of the cellulose derivatives solutions are studied at the polymer chair of Ural State University last ten years.

Academician Kargin was the first who described the ability of polymers to produce mesophases. In 1941, he wrote "interactions between big molecules is rather strong even when the interaction between individual units is weak. As a result, this can lead to the orientation of big molecules in one common direction". In 1956 Robinson (1956, 1958) has discovered, that poly(γ-benzyl-L-glutamate) (PBG) can form the liquid crystals in concentrated solutions in chloroform, methylene chloride, trichlorethane, dioxane, *m*-cresol. Flory (1956) has suggested the phase diagram for a polymer – solvent system with the liquid crystalline transition. At a later date such diagrams were built for the systems: PBG – DMF (Wee & Miller, 1971), polycarbobenzoxyline – DMF (Miller et al., 1974, 1978), poly-*p*-benzamide – DMA (Papkov et al., 1974), poly-p-benzamide – DMA, LiCl (Iovleva et al., 1976), poly(*p*-phenyleneterephthalamide) – H_2SO_4 (Papkov & Kulichikhin, 1977, Andreeva et al., 1981), PBG – ethylene chloride, PBG – benzyl alcohol (Sasaki et al., 1983), poly(*p*-phenyleneterephthalamide) – H_2SO_4 – water (Nakajima et al., 1978), PBG – *m*-cresol (Kiss & Porter, 1977), copolymer on the bases of p-phenylenediamine + terephthalic acid and 4,4"-diphenyldicarboxylic acid) – H_2SO_4 Lukashova et al. , 1978), PBG – DMF and PBG –

dicloracetic acid (Konevets et al., 1985), poly(*p*-benzamide) – DMA – LiCl (Salaris et al., 1976), polyhexylisocyanate and poly(50 % butyl + 50 % p-anisole-3-propyl)isocyanat in tetrachlorethane (Aharoni & Walsh, 1979), copolymer on the bases of *p*-phenylenterephyhalaamide + benzimidazole) – H_2SO_4 (Iovleva & Banduryan, 2010).
The influence of a magnetic field on the liquid crystal structure was studied by Meuer (1968), de Gennes (1968). The authors have considered such a field distortion of the cholesteric structure and find that the transition to complete nematic order occurs at a critical field strength given by

$$Hc = \frac{\pi^2}{2}\left(\frac{K_{22}}{\Delta\chi_m}\right)^{1/2}\frac{1}{p_0} \quad (1)$$

Where p_0 is the zero field pitch, $\Delta\chi_m$, the diamagnetic anisotropy of the liquid crystal, K_{22} , the twist elastic constant. As the field increases the pitch is predicted to oncrease slowly at first and then diverge logarithmically as the critical field is approached. The theory has been verified for lyotropic liquid crystals of PBG in a number of different solvents (Chandrasekhar, 1977, Iizuka, 1973, DuPre & Duke, 1974, 1975, DuPre et al., 1976, 1977, Patel & DuPre, 1979). Molecules of liquid crystals orient themselves in the magnetic field so that their long chains are oriented parallel to the magnetic field lines (Miller, 1978). This orientation is associated with the molecular anisotropy of macromolecules rather than the existence of permanent magnetic moments.
The LC state in solutions and melts of cellulose derivatives was characterized later on 1970–1980s (Kulichikhin & Golova, 1985, Meeten & Navard, 1982, Bhadani & Gray, 1983, Navard & Haudin, 1981, Yunusov et al., 1982, Iovleva, 1989, Vshivkov et al., 2006, 2007, Vshivkov & Rusinova, 2007, 2008). Molecules of cellulose and cellulose derivatives are characterized by a rigid helical conformation and, hence, they are capable of ordering and formation of cholesteric liquid crystals in concentrated solutions. Investigation of the LC state in polymer solutions is of evident practical importance because, owing to the ability to orient under the action of external fields, such solutions are used for the preparation of high-modulus fibers. To control the above processes, the knowledge of phase diagrams for the systems under processing is crucial. However, data on the phase diagrams of such systems in the applied magnetic and mechanical fields are not numerous (Vshivkov & Rusinova, 2007, 2008). The goal of this work is to study phase liquid crystalline transitions of cellulose derivative solutions in magnetic and in a shear stress fields for the following systems: CEC – DMA, CEC – DMF, CEC – mixture of trifluoroacetic acid with methylene chloride, HPC – ethanol, HPC – acetic acid, HPC – DMA, HPC – DMF, HPC –water and poly(γ-benzyl-L-glutamate) (PBG) – DMF.

2. Results and discussion

2.1 Experimental

Cyanoethyl cellulose (CEC) sample with a degree of substitution of 2.6 and M_w= 1.9x105 and hydroxypropyl cellulose (HPC) samples with a degree of substitution of 3.4 and M_w=9.5x10^4 (HCP-1), 1.4x10^5 (HPC-2), and 1.15x10^6 (HPC-3) were used. According to X-ray studies, degree of crystallinity of HPC samples did not exceed 15%. The degree of crystallinity for the CEC samples was 35%. All X-ray measurements were performed on a DRON-13

diffractometer (Cu Kα-irradiation). The CEC and PBG sample with M_η = 2.4× 10^5 were synthesized at the laboratory of the Institute of Macromolecular Compounds, Russian Academy of Sciences.

Dimethylformamide (DMF), dimethylacetamide (DMA), acetic acid (reagent grade), twice-distilled water, ethanol and a 1 : 1 (by mass) mixture of trifluoroacetic acid with methylene chloride were used as solvents. The polymer solutions were prepared in sealed ampoules for several weeks at 298 (water), 340 (in ethanol), 350 K (DMF, DMAA, and TFAA–methylene chloride mixed solvent).

Phase-transition temperatures T_{ph} were estimated by the cloud-point method. The solution temperatures were varied at a rate of 12 K/min. The structure of solutions was examined with the help of an "Olympus BX-51" polarization microscope. A polarization photoelectric setup was used to determine the type of phase transition in solutions: a sealed ampoule containing the transparent polymer solution was placed in the gap between the crossed polaroids and the temperature of the ampoule was decreased. The polarized light of an LGN-015 He–Ne laser was transmitted through the polaroids in the direction normal to the ampoule containing the solution. When the solution was transparent (isotropic) the intensity of the transmitted light was zero. As the system became turbid upon variation in temperature or increase in the concentration of solution, the transmitted light intensity increased. This indicated formation of the anisotropic phase, that is, the LC phase transition. Experiments in the magnetic field were performed using a setup generating a constant magnetic field with an intensity of up to 15 000 Oe (Fig. 1).

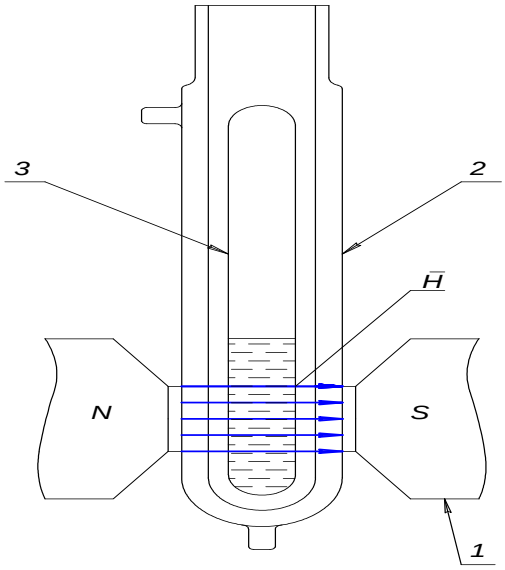


Fig. 1. Schematic presentation of the magnetic facility: (1) electromagnet tips, (2) the thermostating jacket and (3) the ampoule with the test solution. H is the magnetic field vector.

The sealed ampoule containing a transparent polymer solution was placed between the magnet poles. The magnetic field vector was directed normal to the solution layer (~5 mm thick) in the sealed ampoule. The temperature of solution was varied with the thermostating jacket, and the onset temperature of opalescence development was measured. This temperature was related to the appearance of the LC state. The coefficients of magnetic susceptibility χ were determined by means of a vibration magnetometer. The values of χ are presented in the table 1. The energy of the magnetic field E stored by the solution volume unit was calculated via equation $E=\chi H^2$, where H is the magnetic intensity.

System	$-\chi \cdot 10^7$
HPC-1	18,1
CEC	5,3
HPC-1 – ДМА (ω_2=0.5)	6,9
CEC - ДМА (ω_2=0.5)	6,7
HPC-1 – Water (ω_2=0.5)	3,4
HPC-1 – CH_3COOH (ω_2=0.3)	71,3
PBG	78,8
PBG – DMF (ω_2=0.4)	2,3

Table 1. The value of χ for the studied systems.

The phase transition temperature under dynamic conditions was measured using two methods: (1) a polymer solution that occurred in the isotropic state at elevated temperatures was placed in a gap between a glass rotor and a stator of the plastoviscometer. The shear rate was set constant, and the working unit was cooled at a rate of 12 K/h. A temperature corresponding to the onset of solution opalescence was taken as the phase transition temperature. (2) a polymer solution was placed in a metallic working unit of the rheometer HAAKE MARS. The shear stress was measured as a function of temperature, and viscosity η was calculated. The temperature corresponding to a sharp change in the run of the η – T curve was taken as the phase transition temperature T_{ph}.

2.2 The HPC, CEC and PBG solutions in organic solvents

2.2.1 Influence of the molecular weight of polymers on the phase liquid crystalline transitions

The boundary curves delimiting transparent isotropic and opalescent anisotropic solutions for HPC-1–DMA, HPC-2–DMAA, HPC-1–ethanol, HPC – acetic acid, HPC-2–ethanol, CEC–DMA, CEC–DMF, CEC–(methylene chloride/TFAA) and PBG – DMF systems are determined. Under conventional light, the concentrated solutions of HPC and CEC are opalescent. This is suggests formation of cholesteric liquid crystals.

Figures 2 a – c show boundary curves delimiting transparent isotropic and opalescent anisotropic solutions for solutions of the polymers with the different molecular weights.

It is seen, that as the molecular mass of the polymer increases, the boundary curve corresponding to the development of anisotropic LC phase in solutions is shifted to lower concentrations. This behavior agrees with the existing theoretical concepts (Flory, 1956). According to Flory, the critical concentration of a polymer, φ_2^*, above which the LC order arises, is related to the asymmetry of macromolecules x (the length-to-diameter aspect ratio) through the following relationship: $\varphi_2^* = \frac{8}{x}(1-\frac{2}{x})$. As the molecular mass of the polymer increases, the degree of anisometry of macromolecules x increases; as a result, φ_2^*

decreases. The boundary curves of the HPC-1 and HPC-2 in acetic acid coincide practically, because the molecular weights of these samples are not too different.

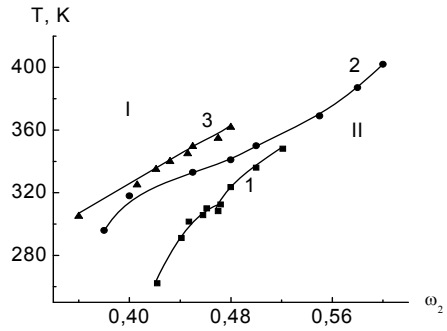


Fig. 2 a. Boundary curves for the systems: (1) HPC-1 – ethanol, (2) HPC-2 – ethanol, (3) HPC-3 – ethanol. I – isotropic solutions, II – anisotropic solutions.

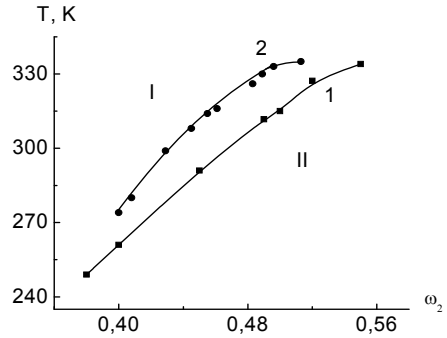


Fig. 2 b. Boundary curves for the systems: (1) HPC-1 – DMA, (2) HPC-3 – DMA. I – isotropic solutions, II – anisotropic solutions.

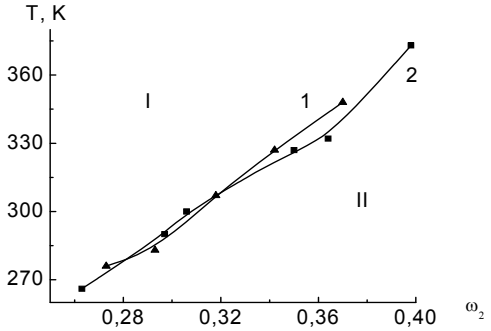


Fig. 2 c. Boundary curves for the systems: (1) HPC-1 – acetic acid, (2) HPC-2 – acetic acid. I – isotropic solutions, II – anisotropic solutions.

2.2.2 Influence of the solvent nature on the phase liquid crystalline transitions

Figures 2 d – 2 g show the boundary curves for the polymer solutions in the different solvents. It is seen from these figures and from the table 2, that as the solvent polarity is increased (solvent dipole moment μ is increased), the LC phase appears at higher concentrations and lower temperatures since with an increase in polarity, a solvent ruptures bonds between macromolecules to a higher extent. It is seen, that DMA and DMF are the best solvents of the cellulose derivatives. The cellulose and cellulose derivative molecules may form the hydrogen bonds between the chains. So the good solvent has to be also an electron donor. This ability is determined by the ionization potential φ. As φ is decreased, ω_2^* is increased, that is such solvent is better.

Solvent	μ dipole moment, D [61]	φ-ionization potential, eV [62]	ω_2^* (mass fraction) T=298 K			
			HPC-1	HPC-2	HPC-3	CEC
DMA	3,86	≤9,65	0,45		0,43	0,42
DMF	3,81	≤10,16				0,42
ethanol	1,69	10,25	0,44	0,38	<0,36	
acetic acid	1,74	10,35	0,3	0,305		
water	1,84	12,59	0,25	0,25	0,195	

Table 2. Physical parameters of the solvents and critical polymer concentration ω_2^*, above which the LC order arises. T = 298 K.

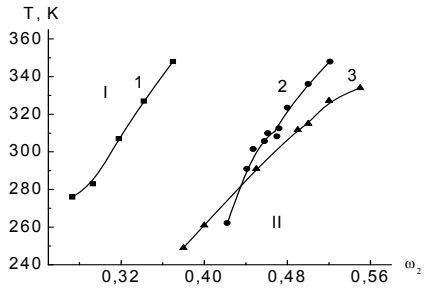


Fig. 2 d. Boundary curves for the HPC-1 solutions in (1) acetic acid, (2) ethanol, (3) DMA. I – isotropic solutions, II – anisotropic solutions.

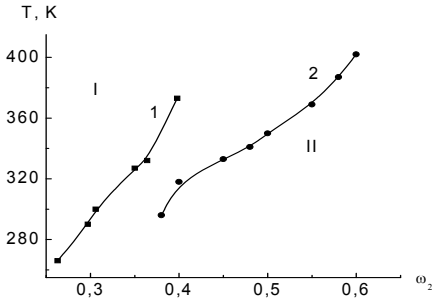


Fig. 2 e. Boundary curves for the HPC-2 solutions in (1) acetic acid, (2) ethanol. I – isotropic solutions, II – anisotropic solutions.

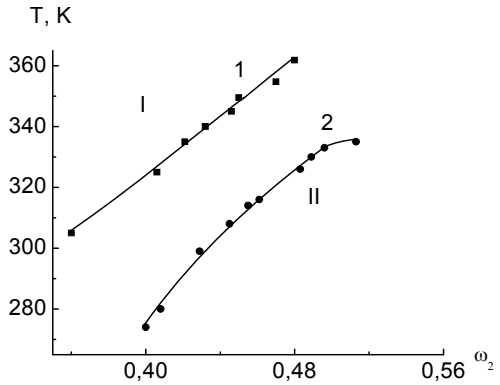


Fig. 2 f. Boundary curves for the HPC-3 solutions in: (1) ethanol, (2) DMA. I – isotropic solutions, II – anisotropic solutions.

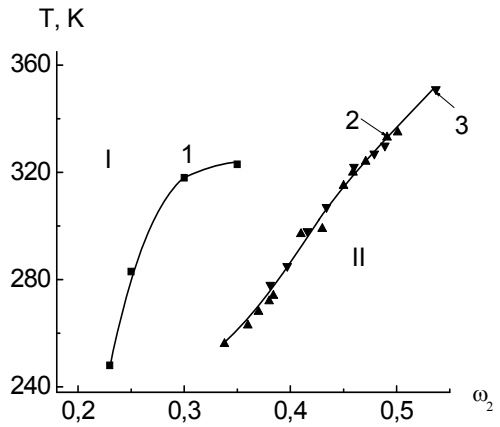


Fig. 2 g. Boundary curves for the CEC solutions in: (1) mixture of trifluoroacetic acid with methylene chloride (1 : 1), (2 ▲) DMF, (3 ▼) DMA. I – isotropic solutions, II – anisotropic solutions.

2.2.3 Influence of the shear stress field on the liquid crystalline phase transitions

Figures 3 a, b, c show the temperature dependences of viscosity for the solutions under study. The above dependences are described by curves with well-pronounced sharp maxima. This behavior is typical of the solutions with LC transitions (Kulichikhin & Golova, 1985, Vshivkov & Rusinova, 2008, Gray, 1962). According to Gray (1962), this profile of the temperature dependences of viscosity corresponds to the (isotropic liquid)–(nematic liquid crystal) phase transition. Therefore, upon cooling of HPC, CEC and PBG solutions under deformation conditions, no cholesteric crystals are formed: in other words, under dynamic conditions, a liquid crystal changes its type from cholesteric to nematic. The results obtained are in good agreement with the data of other authors (Volkova et al., 1986), who showed that the shear deformation of CEC solutions (c= 30%) in trifluoroacetic acid and a 2 : 1

TFAA–methylene chloride mixture results in the formation of similar textures that indicated the formation of an XRD-detectable nematic liquid crystal. Thus, the deformation of CES solutions leads to the change of an LC type from cholesteric to nematic. When the deformed solutions were studied by the method of polarization microscopy, the development of striped textures was observed (fig. 4). This fact is indicative of the formation of the domain supramolecular structure (Papkov & Kulichikhin, 1977, Aharoni & Walsh, 1979). Since, compared to cholesteric liquid crystals, nematic liquid crystals exist at higher temperatures, the temperature–concentration region corresponding to the existence of anisotropic solutions under the mechanical field should change.

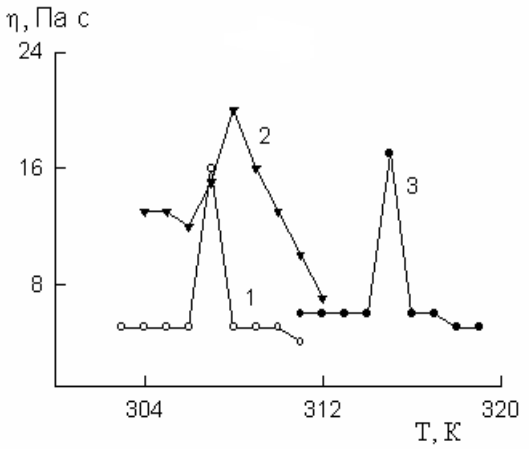


Fig. 3 a. Temperature dependences of viscosity for solutions (1, 3) HPC-3–DMA and (2) PBG–DMF (2); C = 42.9 (1), 19.4 (2), 44.5 % (3). Shear rate $\gamma = 8\ s^{-1}$.

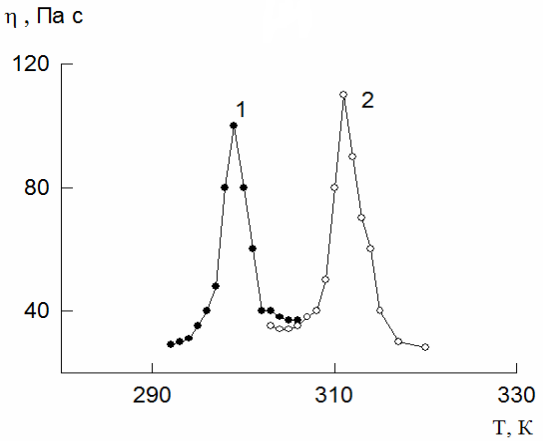


Fig. 3 b. Temperature dependences of viscosity for solutions: HPC-3– ethanol; c = (1) 36 and (2) 42.1%; Shear rate $\gamma = 8\ s^{-1}$.

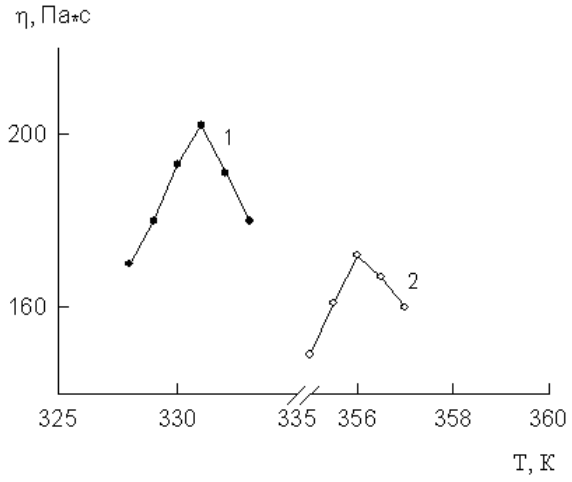


Fig. 3 c. Temperature dependences of viscosity for solutions: HPC-1–DMF: c = (1) 50 and (2) 54.9%. Shear rate γ = 8 s–1.

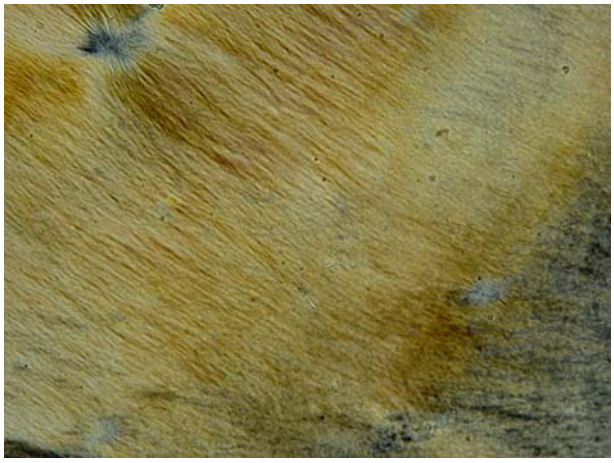


Fig. 4. Mocrograph of the CEC solution in DMA c=51.2 % after deformation. γ=60 s^{-1}. x120.

The phase transitions in the CEC–DMF and CEC–DMAA systems under static conditions and in a shear field are studied. The mechanical field leads to an extension of the temperature – concentration region of the existence of the LC phase (fig. 5), a phenomenon that is due to the change of orientation of CEC macromolecules in solutions.

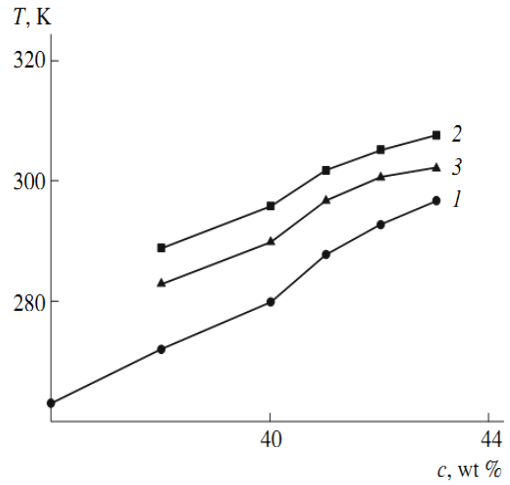


Fig. 5 a. Boundary curves of CEC – DMF system. γ: (1) 0, (2) 12 and (3) 60 s $^{-1}$.

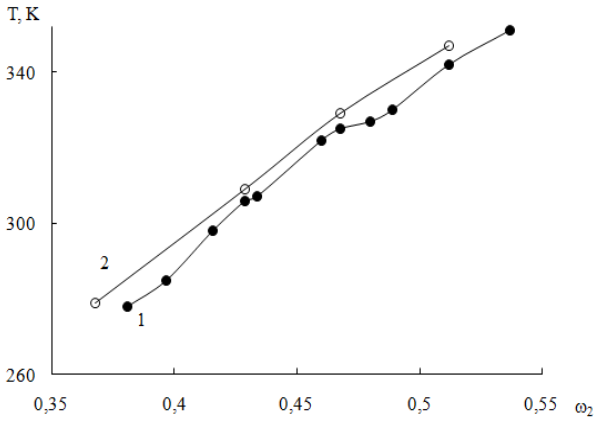


Fig. 5 b. Boundary curves of CEC – DMA system . γ=0 (1) и 12 s^{-1} (2).

The dependence of ΔT (ΔT is the difference of phase transition temperatures under dynamic and static conditions) on the shear rate is described by a curve with a maximum (fig. 6). The same behavior was reported for some polymer–solvent and polymer–polymer systems with crystalline phase separation (Vshivkov et al., 1998, Vshivkov & Rusinova, 2001).

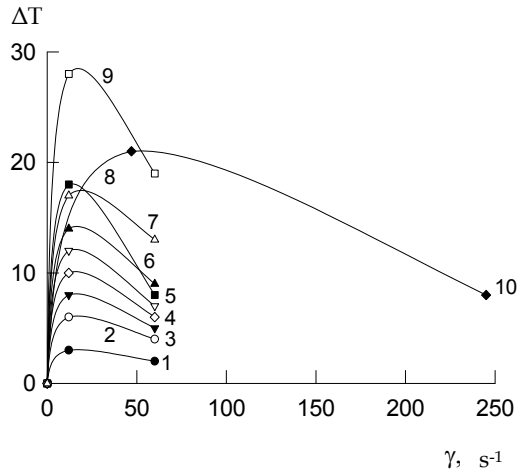


Fig. 6 a. ΔT vs. shear rate for the following solutions: HPC-2 – DMA (1 – 3), CEC – DMF (4, 6, 7), HPC-1 – DMA (5, 8, 9), PE ($M=2.3x10^5$) – p-xylene (10). c = 48.9 (1), 44.6 (2), 40.8 (3), 43.0 (4), 50.0 (5), 41.0 (6), 38.0 (7), 45.0 (8) and 40.0 % (9, 10).

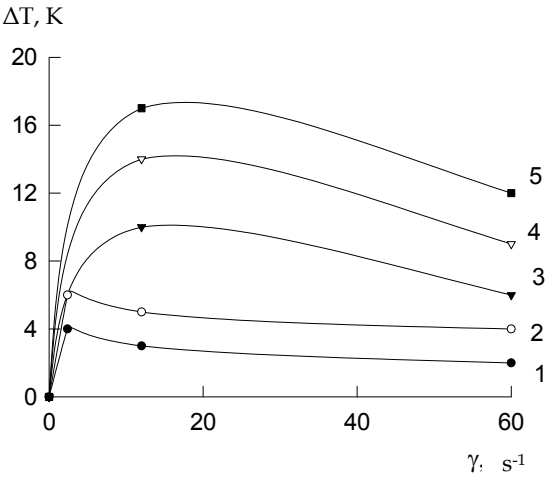


Fig. 6 b. ΔT vs. shear rate for the following solutions: (1, 2) CEC in DMA and (3–5) DMF; c = (1) 42.9, (2) 51.2, (3) 43.0, (4) 41.0, and (5) 38.0%.

This pattern of the curve was associated with two opposite processes in the system, namely, the orientation of macromolecules along the flow direction, which is favorable for phase transition, and the destruction of nuclei of the new phase by a mechanical field, a process that retards the formation of the LC phase. In the examined range of shear rates, the orientation processes dominate, thereby resulting in the elevation of the formation temperature (relative to static conditions) of the LC phase. as manifested in the elevation of the LC phase transition temperature. For comparison, Fig. 6 a shows the data for the PE – *p*-

xylene system with crystalline phase separation (Vshivkov et al., 1998). As follows from fig. 6 a , the orientational processes (increase in ΔT) during the LC transition are observed at lower shear rate (by approximately an order of magnitude).

2.2.4 Influence of the magnetic field on the liquid crystalline phase transitions

Application of the magnetic field raises the temperature of LC phase formation T_{ph} in HPC, CEC and PBG solutions; that is, it widens the temperature–concentration region of the existence of anisotropic solutions. Molecules of liquid crystals orient themselves in the magnetic field so that their long chains are oriented parallel to the magnetic field lines [46]. According to published data (Meuer (1968), de Gennes (1968), Chandrasekhar, 1977, DuPre & Duke, 1974, 1975, DuPre et al., 1976, 1977, Patel & DuPre, 1979), the cholesteric liquid crystal–nematic liquid crystal phase transition occurs in magnetic field. From a certain critical intensity, magnetic field causes untwisting of the cholesteric helix. Eventually, nematic liquid crystals are formed which occur at higher temperatures than cholesteric liquid crystals.

Polarization microscopy studies revealed a striped texture of HPC and CEC solutions treated in magnetic field (fig. 7), thus suggesting formation of large supramolecular structures—domains. A similar phenomenon was reported for other polymer–solvent systems (Papkov & Kulichikhin, 1977).

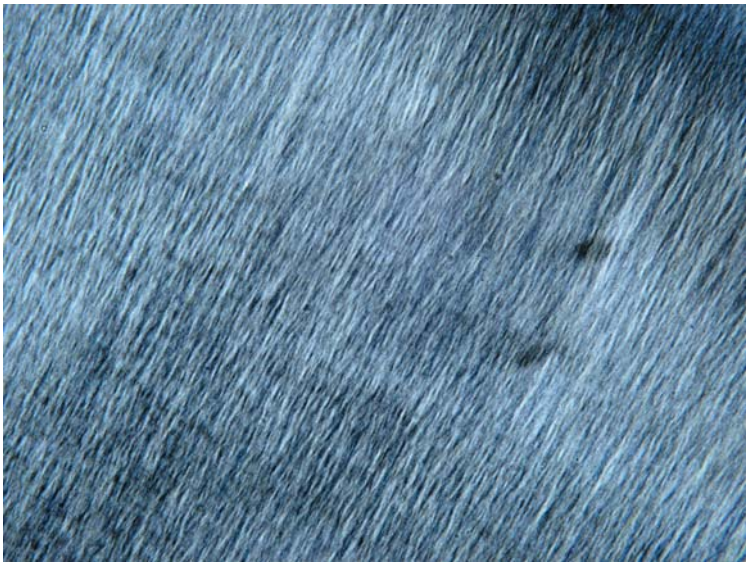


Fig. 7. Micrograph of the HPC-1 solution in DMA. c=52.0 %. x120. H=9 kOe.

It was discovered that after the magnetic field was switched off, an increased Tph was preserved in solutions for many hours. This is clearly seen from fig. 8, which demonstrates the time dependence of the time dependences of ΔT (ΔT is the difference in LC phase transition temperatures in the presence and absence of magnetic field) are determined. This fact provides evidence that structures induced by the magnetic field are preserved in solutions. Thus, the systems under study possess memory.

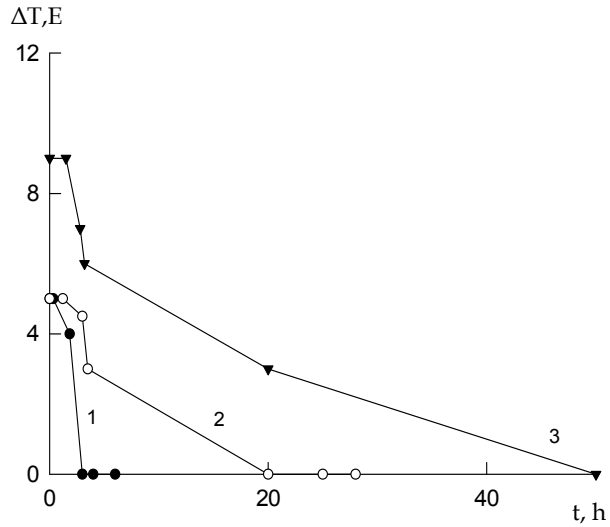


Fig. 8 a. Time dependence of ΔT for solutions: HCP-3 in DMAA (c = 51.3%) at (1) 370, (2) CEC in DMAA (c = 48.3%) at (2) 370 and (3) 298 K treated with magnetic field with intensity H = 7 kOe.

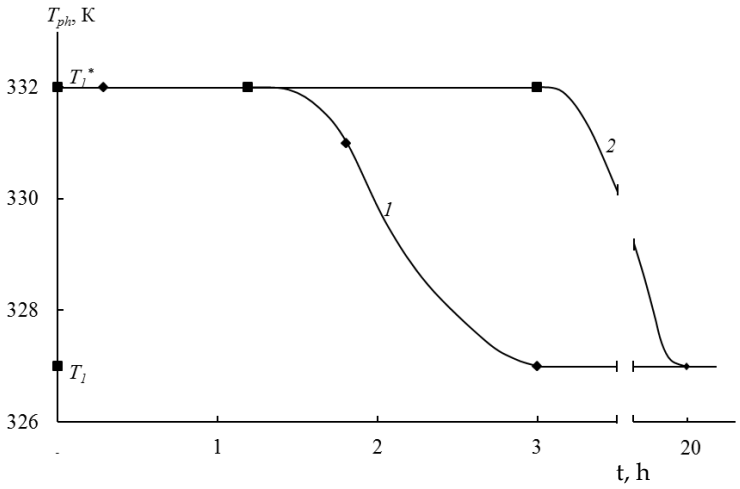


Fig. 8 b. Time dependence of T_{ph} for solution CEC in DMF (c = 48.3 %) at (1) 370 and (2) 298 K treated with magnetic field with intensity H = 7 kOe.

On the basis of the above data, the times of relaxation τ were calculated for the nematic liquid crystal – cholesteric liquid crystal reverse transition in solutions after switching off

the magnetic field. Calculations were performed according to the common exponential equation. The values of τ were found to be 18 h (HPC-3–DMAA); τ_1= 11 h at 298 K and τ_2= 8 h at 370 K (CEC–DMAA). The calculation results made it possible to estimate the order of the enthalpy of activation ΔH* for the nematic liquid crystal–cholesteric liquid crystal transition in solutions after switching off the magnetic field. The value of ΔH* is estimated via the equation $\ln(\tau_1/\tau_2) = (\Delta H^*/R)(1/T_1 - 1/T_2)$ as ~ 4 kJ/mol, in qualitative agreement with rather low enthalpies of LC phase transitions (Chandrasekhar, 1977, Plate, 1988).
Figures 9 a – 9 c show the boundary curves measured for the HPC-3-DMAA, HPC-1 – ethanol and CEC–DMAA systems at various magnetic field intensities H.

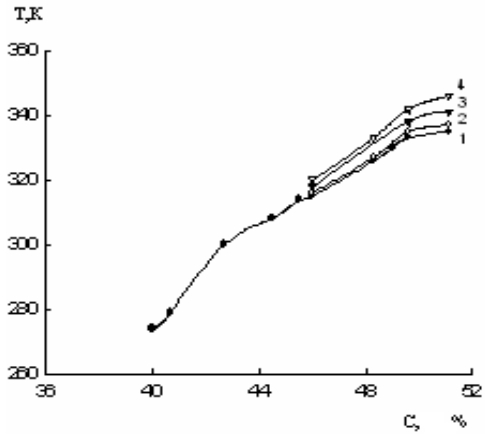


Fig. 9 a. Boundary curves for the HPC-3–DMAA system. H = (1) 0, (2) 3, (3) 5, and (4) 9 kOe.

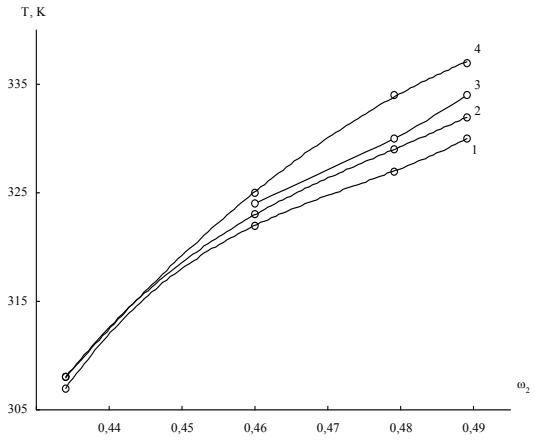


Fig. 9 b. Boundary curves for the CEC – DMAA system. H = (1) 0, (2) 3, (3) 5, and (4) 9 kOe.

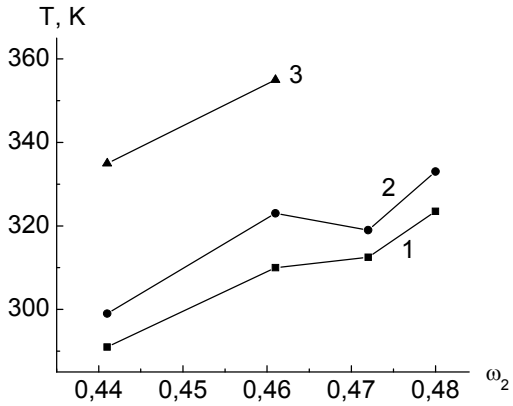


Fig. 9 c. Boundary curves for the HPC-1 – ethanol for system. H = (1) 0, (2) 3, (3) 5 kOe.

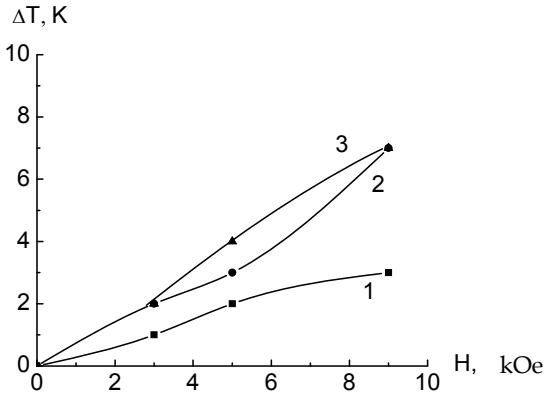


Fig. 10 a. ΔT vs magnetic field intensity for the CEC – DMA system. ω_2: (1) 0.46, (2) 0.48, (3) 0.49.

As is seen, with an increase in H, the temperature–concentration region of LC solutions widens. The higher the value of H, the more pronounced the shift of the boundary curves. A similar behavior was observed for solutions of CEC in DMF and PBG in DMF.

Figures 10 a – 10 e show the concentration dependences of ΔT for the cellulose ester – solvent systems (ΔT is the difference of phase transition temperatures in magnetic field and in its absence). It is seen, that as the magnetic field intensity H is increased, the ΔT value increases. It testifies about the macromolecule orientation increase.

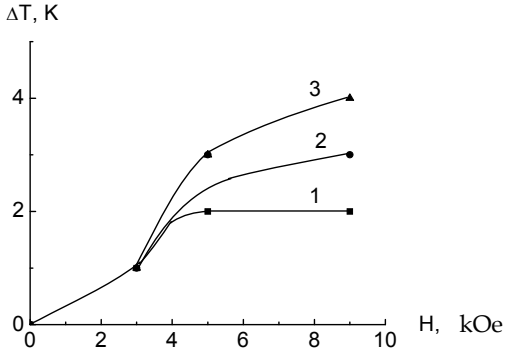


Fig. 10 b. ΔT vs magnetic field intensity for the CEC – DMF system. ω_2: (1) 0.47, (2) 0.49, (3) 0.50.

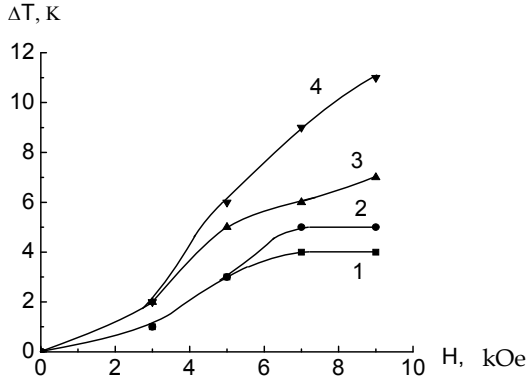


Fig. 10 c. ΔT vs magnetic field intensity for the HPC-3 – DMA system. ω_2: (1) 0.46, (2) 0.48, (3) 0.50, (4) 0.51.

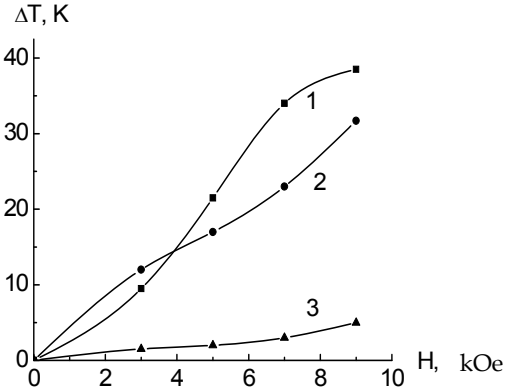


Fig. 10 d. ΔT vs magnetic field intensity for the HPC-1 – DMA system. ω_2: (1) 0.49, (2) 0.52, (3) 0.55.

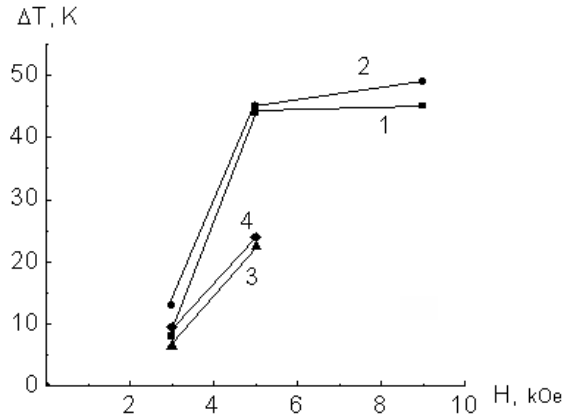


Fig. 10 e. ΔT vs magnetic field intensity for the HPC-1 – ethanol system. ω_2: (1) 0.44, (2) 0.46, (3) 0.47, (4) 0.48.

2.3 HPC solutions in water

Aqueous solutions of HPC belong to systems with strong electron-donor (hydrogen) bonds (Belousov & Panov, 1983). Because of the presence of two mobile protons and two unshared electron pairs at the oxygen atom, a water molecule may function both as an electron donor and an electron acceptor and form four hydrogen bonds with an energy of 20 kJ/mol. Therefore, a loose structure with a large free volume is formed in water. In the case of water, the fraction of nonspecific interaction is as low as 7%. Intermolecular interactions of HPC with water may be determined by both the hydrophilic hydration giving rise to hydrogen bonding between a polymer and a solvent and the hydrophobic hydration of water, which consists in densification of water structure around nonpolar methyl and methylene groups of HPC molecules during formation of solutions. Many studies were devoted to phase equilibrium in the HPC–water system, and the LCST values were reported in a number of papers (Vshivkov et al., 2007, Fischer et al., 1995, Kagemoto et al., 1970, Nystrom & Bergman, 1978, Werbowyi & Gray, 1976, 1979, 1980, Nishio et al., 2002, Fortin & Charlet, 1989, Ryotarou & Yoshiyuki, 2003, Guido, 1995, Furusawa & Tagawa, 1985, Suto et al., 1989, Lu & Schwartz, 2002, Bergman & Sundelof, 1977). Figure 11 a displays the phase diagram measured for the HPC-1–water system. This diagram is largely consistent with the phase diagrams described for this system. Four regions can be distinguished in the diagram: (I) the region of isotropic transparent solutions; (II) the region of anisotropic transparent solutions; (III) the region of heat-induced phase separation giving rise to formation of a white anisotropic precipitate; and (IV) the region of anisotropic solutions opalescent over the entire volume (the colorless solutions are observed, which is typical of cholesterol LC solutions. (1* refers to colorless solutions and 2* refers to blue solutions, which is typical of cholesterol LC solutions (Kapustin, 1978). According to Fisher et al. (1995), the crystal solvates are formed in solutions at an HPC concentration of ~ 80 % or above. The boundary curve 1 that characterized the heated induced phase transition has a binodal shape. It appears that the breakdown of hydrophilic and hydrophobic hydration of HPC initially leads to the amorphous phase separation of solutions and formation of two coexistent dilute and concentrated phases (the LCST is 298 K). Simultaneously, anisotropic crystal solvates precipitate in the concentrated phase.

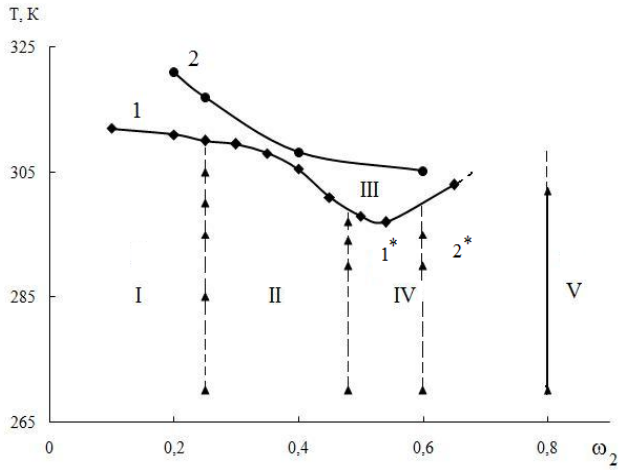


Fig. 11 a. Phase diagram for the HPC-1–water system. Shear rate γ = 0 (1) and 12 s^{-1} (2). See text for explanations.

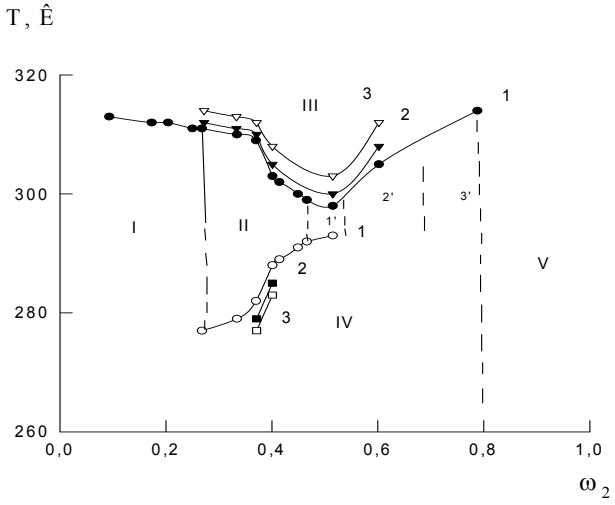


Fig. 11 b. Boundary curves for the HPC-2–water system at γ = (1) 0, (2) 12, and (3) 60 s–1. Color of the solutions: (1') red, (2') green, (3') violet. Comments are given in text.

Figures 11 b and 12 show the phase diagrams for the HPC-2–water and HPC-3–water systems measured under static conditions and in the shear field. The phase diagram of the HPC-2 – water system virtually coincides with that of the HPC-1–water system, since a difference in the molecular masses of these samples is insignificant; the LCST is 298 K. The diagram comprises the regions similar to those characteristic of the HPC-1–water system.

The distinctive feature is a change in the color of solutions with the weight fraction of the polymer $\omega_2 > 0.45$. The observed color transition from red to violet through green may be related to a reduction in the cholesteric helix pitch with an increase in the polymer concentration in solutions. Also note that, in contrast to HPC-1 solutions, in the case HPC-2 solutions, the curve delimiting regions II and IV shifts toward lower polymer concentrations at low temperatures. The molecular mass of the HPC-3 sample is almost an order of magnitude higher than that of the HPC-1 and HPC-2 samples, therefore we failed to prepare HPC-3 solutions in a wide concentration range because of its poor solubility. Three regions may be distinguished in the phase diagram: (I) the region of isotropic solutions; (II) the region of transparent anisotropic solutions; and (III) the region, where heating causes phase separation accompanied by formation of the while anisotropic precipitate. A comparison of the phase diagrams indicates that an increase in the molecular mass of the polymer leads to a shift in the boundary curve reflecting the formation of anisotropic LC solutions to lower concentrations. This fact is consistent with the current theoretical concepts [20].

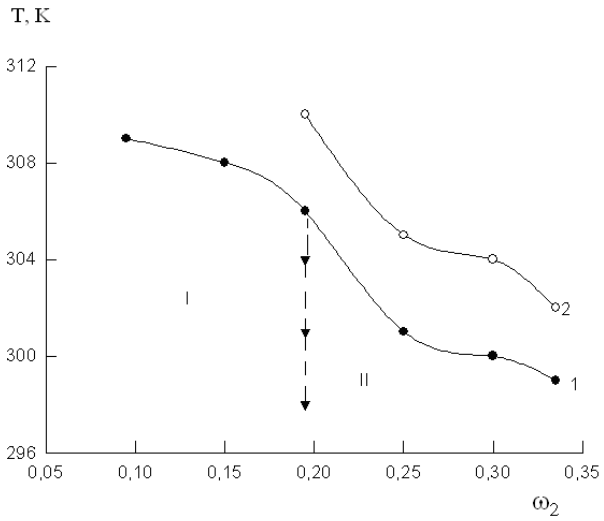


Fig. 12. Boundary curves for the HPC-3–water system at γ = (1) 0 and (2) 12 s–1. Comments are given in text.

2.3.1 Influence of the shear stress field on the liquid crystalline phase transitions

The deformation of these systems increases the temperatures of heat-induced separation and decreases the temperatures of transition from region II to region III. As the shear rate is increased, the absolute value of ΔT increases for both transitions and achieves 7 K (ΔT is a difference between phase transition temperatures under dynamic and static conditions). This phenomenon may be explained by the breakdown of nuclei of a new phase under the action of the mechanic field, as was observed for a number of polymer – solvent systems characterized by amorphous and crystalline phase separation (Vshivkov et al., 1998, Vshivkov & Rusinova, 2001).

2.3.2 Influence of the magnetic field on the liquid crystalline phase transitions

Application of magnetic field causes an increase in the phase transition temperature under heating, which is likely associated with a change in the orientation of macromolecules in solution (fig. 13).

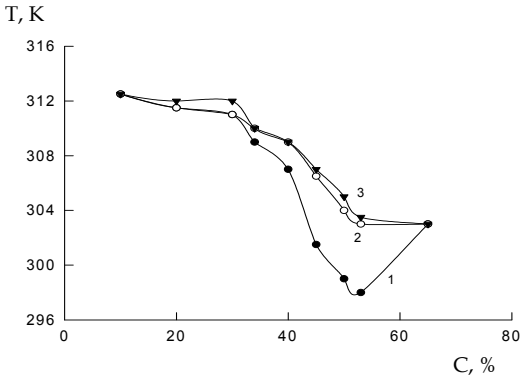


Fig. 13. Boundary curves for the HPC-1–water system. H = (1) 0, (2) 5, and (3) 9 kOe.

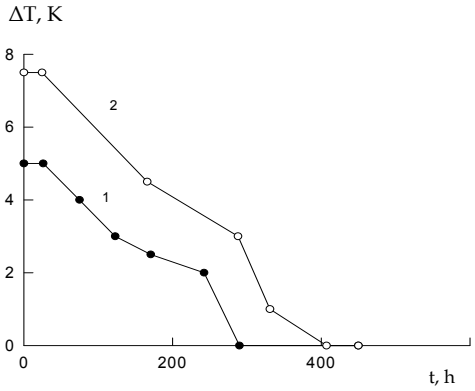


Fig. 14. Time dependence of ΔT for HPC-1 solutions in water. c = (1) 53.5 and (2) 49.6%. H= 5 Oe.

Like the HPC and CEC solutions in organic solvents, the HPC – water system possesses memory: an increased T_{ph} is preserved for many hours after the magnetic field is switched off (fig. 14). The calculated time of relaxation necessary to achieve the initial orientation of macromolecules is τ = 260 h (c = 53.5%) and 103 h (c = 49.6%). With an increase in the concentration of the polymer, the time of relaxation grows, since the viscosity of the system increases.

3. Conclusion

Phase diagrams have been constructed for the cellulose ester – water systems under static conditions, in the shear stress and magnetic fields. As the molecular mass of the polymer is

increased, the curves delimiting isotropic and anisotropic solutions shift to lower concentrations. The deformation causes the formation of domain structure, the changes in the type of the liquid crystal, and in the phase transition temperatures of solutions both upon heating and cooling. As the molecular mass of the polymer increases, the ability of macromolecules to orient under the shear stress field decreases. The concentration and dependence ΔT is described by the curve with maxima. The formation of domains in solutions was observed under the shear stress field.

The magnetic field also widens the temperature–concentration region of the existence of the LC phase. This effect is related to the cholesteric liquid crystal – nematic liquid crystal phase transition and the orientation of macromolecules in the direction parallel to the magnetic field lines. In this case, large supramolecular structures (domains) develop in solutions. The effect of magnetic field on the variation in LC transitions with the polymer concentration in solution shows an extremal pattern. Figures 15 and 16 demonstrate the concentration dependences of ΔT for HPC1– DMAA, HPC-3–DMAA, and HPC-1–water systems measured at various magnetic field intensities.

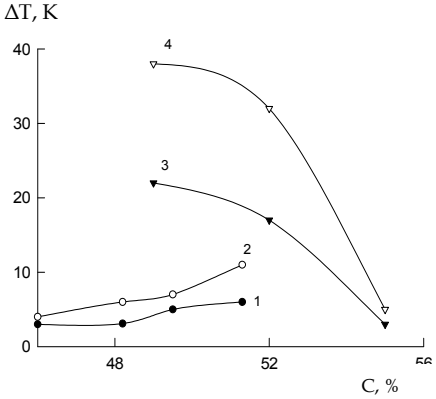


Fig. 15. Concentration dependence of ΔT for solutions of (1, 2) HPC-3 and (3, 4) HPC-1 in DMAA. H =(1, 3) 5 and (2, 4) 9 kOe.

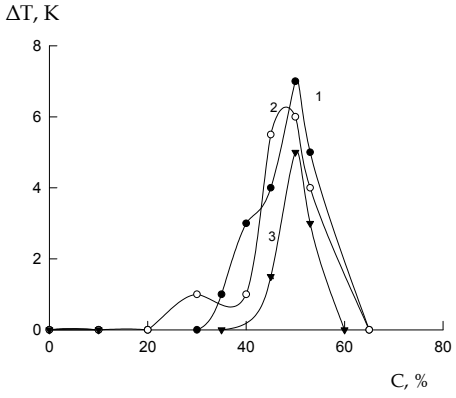


Fig. 16. Concentration dependence of ΔT for HPC-1 solutions in water. H = (1) 5, (2) 9, and (3) 13 kOe.

In analyzing the effect of the polymer concentration on magnetic field-induced changes in phase transition temperatures, two factors should be taken into account. First, as concentration increases, the number of macromolecules capable of orientation in the magnetic field grows; as a consequence, T_{ph} should increase. Second, a rise in the polymer concentration in solution facilitates densification of the fluctuation network of entanglements. This impedes the occurrence of orientation processes and weakens the effect of the magnetic field. On the whole, the concentration dependence of ΔT is apparently described by a curve with a maximum.

It should be noted that for solutions of an HPC-1 sample with a lower molecular mass, the value of ΔT is much higher. This observation indicates a more distinct orientation of smaller molecules in the magnetic field, in agreement with the data from (Kol"tsov et al., 1995).

Figures 17 and 18 plot ΔT as a function of lnE for HPC-1–DMAA, HPC-3–DMAA, and CEC – DMAA systems. It is seen that these dependences are described by straight lines. The analogous dependence is determined for the PBG – DMF system. With an increase in the magnetic field energy stored by solutions, the value of ΔT increases. The effect of field on the phase transitions shows a threshold character: a change in T_{ph} begins from a certain critical intensity of the field H_{cr}. These values are 2.3 and 2.0 kOe for the HPC – DMAA and CEC –DMAA systems, respectively, and 2.3 kOe for the PBG – DMF system. In order of magnitude, these values are consistent with H_{cr} necessary for the nematic liquid crystal - cholesteric liquid crystal phase transition (Papkov & Kulichikhin, 1977, Chandrasekhar, 1980). In this case, $\Delta T = K\ln(E/E_0)$ or $T_{ph}(H > H_{cr}) = T_{ph}(H = 0) + K\ln(E/E_0)$. Coefficient K depends on the molecular mass of the polymer and its concentration in solution.

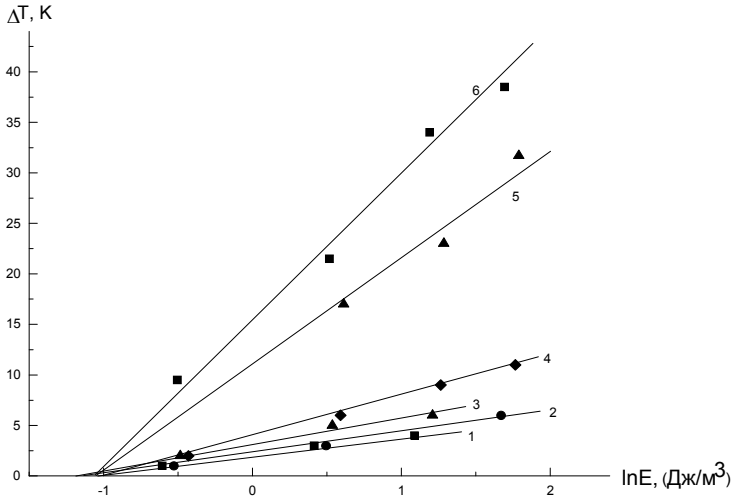


Fig. 17. Plot of ΔT vs. lnE for solutions of (1–4) HPC-3 and (5, 6) HPC-1 in DMAA c = (1) 46.1, (2) 48.3, (3) 49.6, (4) 51.3, (5) 52.0, and (6) 49.0%.

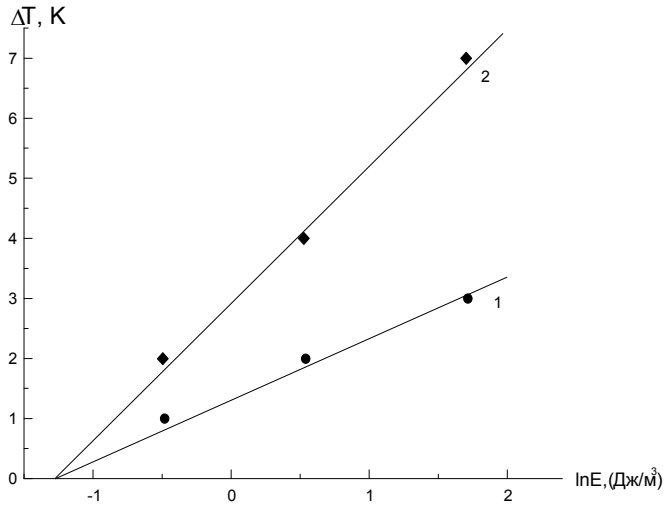


Fig. 18. Plot of ΔT vs. lnE for CEC solutions in DMAA. c = (1) 46.0 and (2) 48.8%.

Thus it is revealed a perfect analogy in influence magnetic and mechanical fields in their influence on phase liquid crystal transitions in solutions of rigid chain polymers.

4. Chart of symbols

- PBG – poly(γ-benzyl-L-glutamate, DMF – dimethylformamide, DMA – dimethylacetamide, Hc – a critical field strength,
- p_0 - zero field pitch, $\Delta\chi_m$ - diamagnetic anisotropy of the liquid crystal, K_{22} - the twist elastic constant, LC – liquid – crystalline, CEC – cyanoethyl cellulose, HPC – hydroxypropyl cellulose, T_{ph} - phase-transition temperature, X – coefficients of magnetic susceptibility, H – magnetic intensity, E – energy of the magnetic field stored by the solution volume unit, η – viscosity , φ_2^* – critical concentration of a polymer, , above which the LC order arises,
- x – asymmetry of macromolecules, φ – ionization potential, ΔT – the difference of phase transition temperatures under dynamic and static conditions , γ – shear rate, ΔT – the difference in LC phase transition temperatures in the presence and absence of magnetic field, τ – times of relaxation, ΔH* – enthalpy of activation for the nematic liquid crystal–cholesteric liquid crystal transition, ω_2 - mass portion of polymer,

5. Acknowledgment

This work was supported by the Ministry of Education and Science of the Russian Federation (project no. AVTsP 2.1.1/1535 "Development of the Scientific Potential of the Higher School") and by Federal Agency of Russian Federation (project no. NK-43 P(4), Federal Program "Scientific and Scientific – Pedagogic Personnel of Russia).

6. References

Aharoni, Sh.M. & Walsh, E.K. (1979). Rigid backbone polymers. III. Partitioning of isocyanate polymers between isotropic and anisotropic phases. *J. Polym. Sci. Polym. Lett. Ed.*, Vol. 17, pp. 321-327.

Andreeva, V. M., Konevets, V. I. & Tager, A. A. (1981). Research of the lightscattering of poly(*p*-phenilene-terephtalamide) solutions at the transition to a liquid crystal condition. *Vysokomol. Soedin.*, B, Vol. 23, No. 7 , pp. 538-541.

Belousov, V. P. & Panov, M. Yu. (1983). *Thermodynamics of Nonelectrolyte Aqueous Solutions*. Khimiya, Leningrad. [in Russian].

Bergman, R. & Sundelof, Z.O. (1977). Diffusion transport and thermodynamic properties in concentrated water solutions of hydroxypropyl cellulose at temperatures up to phase separation. *Eur. Polym. J.*, Vol. 13, No. 6, pp. 881-889.

Bhadani, S.N. & Gray, D.G. (1983). Cellulose-based liquid crystalline polymers; esters of (hydroxypropyl) cellulose. *Mol. Cryst. Liquid Cryst.*, Vol. 99, No. 1, pp. 29-38.

Chalykh, A. E., Gerasimov, V. K. & Mikhailov, Yu. M. (1998). *Phase Diagrams of Polymer Systems.* Yanus-K, Moscow. [in Russian].

Chandrasekhar, S. (1977). *Liquid crystals*. Cambridge Univ. Press, London, Mir, Moscow.

Chiba, R. & Nishio, Y. Electrooptical behavior of liquid-crystalline (hydroxypropyl) cellulose / inorganic salt aqueous solutions. *Macromolecules.* Vol. 36, No. 5, pp. 1706-1712.

DuPre, D.B. & Duke, R.W. (1975). Temperature, concentration, and molecular weight dependence of the twist elastic constant of cholesteric poly-γ-benzyl-L-glutamate. *J. Chem Phys., Vol. 63, No. 1., pp. 143-148.*

Duke, R.W., DuPre, D.B., Hines, W.A. & Samulski, E.T. (1976). Poly(γ-benzyl L-glutamate) helix-coil transition. Pretransition phenomena in the liquid crystal phase. *J. Am. Chem. Soc.,* Vol. 98, No. 10, pp. 3094-3101 .

DuPre, D.B., Duke, R.W., Hines, W.A. & Samulski, E.T. (1977). The Effect of trifluoroacetic acid on the viscoelastic properties of a polypeptide liquid crystal. *Mol. Cryst. Liq. Cryst.*, Vol. 40, No.2, pp. 247-259.

Fischer, H., Murray, M., Keller, A. & Odell, J. A. (1995). On the phase diagram of the system hydroxypropylcellulose-water. *J. Mater. Sci.*, Vol. 30, No. 18, pp.4623-4627.

Flory, P. J. (1956). Phase Equilibria in Solutions of Rod-Like Particles. Proc. R. Soc. Lond., A, V. 234, No. 1, pp. 73-89.

Fortin, S. & Charlet, G. (1989). Phase diagram of aqueous solutions of (hydroxypropyl) cellulose. *Macromolecules,* Vol. 22, No. 5, pp. 2286-2292.

Furusawa, K. & Tagawa, T. (1985). Adsorption behavior of water-soluble polymer with lower critical solution temperature. *Colloid Polym. Sci.*, Vol. 263, No. 5, pp. 353-360.

de Gennes, P.G. (1968). Calcul de la distorsion d"une structure cholesterique par un champ magnetique. *Solid State Comm.* Vol. 6, No. 3, pp. 163-165.

63. Gray, G. W. (1962) in Molecular Structure and the Properties of Liquid Crystals. Academic, New York,), p. 92.

Guido, S. (1995). Phase behavior of aqueous solutions of hydroxypropyl cellulose. *Macromolecules,* Vol. 28, No. 13, pp. 4530-4539.

Iizuka, E. (1973).The effect of magnetic fields on the structure of cholesteric liquid crystals of polypeptides. *Polymer J.* Vol. 4, No. 4, pp. 401-408.

Iovleva, M.M., Papkov, S.P., Milkova, L.P., Kalmykova, V.D., Volokhina, A.V.& Kudryavtsev, G.I. (1976). Temperature-concentration boundaries of a liquid crystal state of poly (*p*-benzamide). *Vysokomol. Soedin.*, Ser. B. V. 18, No.11, pp. 830- 832.

Iovleva, M.M. (1989). The schematic phase diagram for the cellulose – monohydrate methylmorpholene-N-oxide system. *Vysokomol. Soedin.*, Ser. A. Vol. 31, No.4, pp. 808- 812.

Iovleva, M.M., Banduryan, S.I. (2010).Phase equilibria in polymer-solution systems: development in the field of fibres. *Vysokomol. Soedin*., Vol. 52, No.11, pp. 2033- 2037.

Kapustin, A. P. *Experimental Methods for Investigation of Liquid Crystals*. Nauka, Moscow, 1978). [in Russian].

Kargin, V. A., Papkov, S. P. & Rogovin, Z. A. (1939). About solubility of high-molecular compounds. *Zh. Fiz. Khim.*, Vol. 13, No. 2, pp. 206-210.

Kiss, G. & Porter, R.S. (1977). Rheology of the concentrated solutions of poly-γ-benzyl-*L*-glutamate. *Am. Chem. Soc. Polym. Prepr.*, Vol. 18, No.1, pp. 185-186 .

Klenin, V. I. *Thermodynamics of Systems with Flexible Chain Polymers*. Saratov. Gos. Univ., Saratov, 1995). [in Russian].

Kol'tsov, A. I., Kapralov, V. M., Rot, Kh., Hachaturov, A.S., & Shabsels, B.M. (1995).Orientation of solutions of poly-γ-benzyl-*L*-glutamate fractions of various molecular weight in a magnetic field according to a deuteron magnetic resonance data. *Polymer Science*, B, Vol. 37, No.10, pp. 500-502.

Konevets, V. I., Andreeva, V. M., Tager, A. A., Ershova, I.A. & Kolesnikova E.N. (1985). Studying of the moderate-concentrated solution structure of some polyamides in the composition region, preceding the liquid crystal formation. *Vysokomol. Soedin.*, Ser. A. Vol. 27, No.5 , pp. 959- 967.

Kulichikhin, V. G. & Golova, L. K. (1985). Liquid – crystalline state of cellulose and cellulose derivatives. *Khim. Drev*., No. 3, pp. 9-27.

Lu, X., Hu, Z. & Schwartz, J. (2002). Phase transition behavior of hydroxypropylcellulose under interpolymer complexation with poly(acrylic acid). *Macromolecules,* Vol. 35, No. 7, pp. 9164-9168.

Lukashova, I.N., Volokhina, A.V. & Papkov, S. P. (1978). Phase transformations and viscosity properties of aromatic copolyamide solutions. *Vysokomol. Soedin*., Ser. B. Vol. 20, No.2, pp. 151-154.

Malkin, A. Ya. & Kulichikhin, S. G. (1996). Phase transitions in the polymer systems, caused by action of mechanical fields. *Polymer Science*, B, Vol. 38, No. 2, pp.362-374.

Meeten, G.H. & Navard, P. (1982). The cholesteric nature of cellulose triacetate solutions. *Polymer,* Vol. 23, No. 12, pp. 1727-1731.

Meuer, R.B. (1968). Distortion of a cholesteric structure by a magnetic field. *Appl. Phys. Lett.*, Vol. 14, No. 3, pp. 208-212..

Miller, W.G., Wu, C.C., Wee, E.L., Santee, G.L. & Rai, J.H. ed K.D. Goebel. (1974). Thermodynamics and dynamics of polypeptide liquid crystals. *Pure Appl. Chem*., Vol. 38, No. 1., pp. 37- 58.

Miller, W.G., Kou, L., Tohyama, K. & Voltaggio, V. (1978).Kinetic aspects of the formation of the ordered phase in stiff-chain helical polyamide acids. *J. Polym. Sci. Polym. Symp*. Vol. 65, pp. 91-106.

Miller, W.G. (1978). Stiff chain polymer lyotropic liquid crystals. *Ann. Rev. Phys.Chem.* Vol. 29, pp. 519- 535.

Nakajima, A., Hirai, T. & Hayashi, T. (1978). Phase relationship of rodlike polymer, poly(p-phenyleneterephthalamide), in sulfuric acid-water system. *Polymer Bulletin*, Vol. 1, No. 2, pp. 143-147.

Navard, P., Haudin, I.M. & Dayan, D.G. (1981). Cholesteric liquid crystalline phases based on (acetoxypropyl)cellulose. *Macromolecules*. Vol. 14, No. 6, pp. 715-719. (1981).

Nesterov, A. E. & Lipatov, Yu. S. (1987). *Phase state of polymer solutions and blends*. Naukova Dumka, Kiev. [in Russian].

Nikitskii, B.P. (1971). *Chemist"s Handbook*. Moscow. Chimiya.

Nishio, Y., Chiba, R., Miyashita, Y., Oshima, K., Miyajima, T., Kimura, N. & Suzuki, H. (2002). Salt addition effects on mesophase structure and optical properties of aqueous hydroxypropyl cellulose. *Polym. J. (Tokyo)*, Vol. 34, No. 9, pp. 149- 157.

Nystrom, B. & Bergman, R. (1978). Velocity sedimentation transport properties in dilute and concentrated solutions of hydroxypropyl cellulose in water at different temperatures up to phase separation. *Eur. Polym. J.* , Vol. 14, No. 6, pp. 431-437.

Papkov, S. P., Rogovin, Z. A.& Kargin, V. A. (1937). About solubility of cellulose and its ethers. *Zh. Fiz. Khim*., Vol. 10, No. 1, pp. 156-157.

Papkov, S. P. , Kargin, V. A. & Rogovin, Z. A. (1937). Solubility of cellulose and its ethers. II. About applicability of a rule of phases to cellulose ether solutions. *Zh. Fiz. Khim.* , Vol. 10, No. 4-5 , pp. 607-619.

Papkov, S.P., Kulichikhin, V.G., Kalmykova, V.D. & Malkin, A.Ya. (1974).
Rheological properties of anisotropic poly(para-benzamide) solutions. *J. Polym. Sci. : Polym. Phys. Ed*. Vol. 12, pp. 1753-1770.

Papkov, S. P. & Kulichikhin, V. G. (1977). *Liquid crystalline state of polymers*. Khimiya, Moscow. [in Russian].

Papkov, S. P. (1981). *Phase equilibrium in polymer–solvent system*. Khimiya, Moscow. [in Russian].

Patel, D. L. & DuPre, D.B. (1979). Poly (γ-benzyl-L-glutamate) liquid crystals in benzene. *Mol. Cryst. Liq. Cryst* . Vol. 53, No. 2, pp. 323 – 334.

Platé, N. A. (1988). Liquid crystal polymers. Khimiya, Moscow; (1993) Plenum, New York.

Robinson C. (1956). Liquid-crystalline structures in solutions of a polypeptide. *Trans. Faraday Soc*. Vol. 52. No.1, pp. 571- 592.

Robinson, C., Ward, J.C. & Buvers, R.B. (1958). Liquid-crystalline structures in polypeptide solutions. *Disc. Faraday Soc.* V. 25. No.1, pp. 29-42.

Rogovin, Z. A., Kargin, V. A. & Papkov, S. P. (1937). Solubility of cellulose and its ethers. III. Heterogeneity influence of acetylcellulose on its solubility. *Zh. Fiz. Khim*., Vol. 10, No.6 , pp. 793-797.

Rusinova, E. V. & Vshivkov, S. A. (1997). Phase transitions in polymer blends induced by mechanical fields. *Polymer Science*, A. Vol. 39, No. 10, pp. 1066-1073.

Rusinova, E. V.,Vshivkov, S. A. & Alekseeva, M.S. (2007). Phase transitions in liquid crystalline solutions of hydroxypropyl cellulose under deformation. *Polymer Science*, B, Vol. 49, No. 1-2, pp. 26-29.

Salaris, F., Valenti, B., Costa, G. & Ciferri, A. (1976). Phase equilibria of poly(*p*-benzamide) in *N*,*N*-dimethylacetamide/lithium chloride solutions. *Makromol. Chem*., Vol. 177, pp. 3073- 3076.

Sasaki S., Tokuma K. & Uematsu I. (1983). Phase behavior of poly(γ-benzyl L-glutamate) solutions in benzyl alcohol. *Polymer Bulletin*. V. 10, No. 6, pp. 539-546.

Suto, S., Nishibory, W., Kudo, K. & Karasawa, M. Lyotropic liquid crystalline solutions of hydroxypropyl cellulose in water: Effect of salts on the turbidity and viscometric behavior. *J. Appl. Polym. Sci.*, Vol. 37, No. 3, pp. 737-749.

Tager, A. & Kargin, V. A. (1941). Convertibility research of nitrocellulose, benzylcellulose and gelatin solutions. I. *Zh. Fiz. Khim.* Vol. 15, No. 9, pp. 1029- 1035.

Tager, A. & Kargin, V. A. (1941). Process of dissolution and swelling of cellulose ethers. *Zh. Fiz. Khim.* Vol. 15, No.9, pp. 1036-1054.

Tager. A. A. (1993). *Foundations of nonelectrolyte solutions science*. Ural. Gos. Univ., Yekaterinburg. [in Russian].

Tager, A. A. (2007). *Physical chemistry of polymers*. World of Science, Moscow. [in Russian].

Volkova, L. A., Kutsenko, L. I., Kulakova, O. M. & Mel'tser, Yu. A. (1986). Liquid crystalline state of the cyanethylcellulose concentrated solutions in trifluoroacetic acid and in its mixes with organic solvents. *Vysokomol. Soedin.*, B, Vol. 28, No. 1, pp. 27- 31.

Vshivkov, S. A. (1991). *Methods for investigation of phase equilibrium of polymer solutions*. Ural. Gos. Univ., Sverdlovsk. [in Russian].

Vshivkov, S. A., Kulichikhin, S. G. & Rusinova, E. V. (1998). Phase transitions in polymer solutions induced by mechanical fields. *Russian Chem. Rev.* Vol. 67, No. 3, pp. 233-243.

Vshivkov, S. A. & Rusinova, E. V. (2001). Phase Transitions in Polymer Systems Induced by Mechanical Field. Ural. Gos. Univ., Yekaterinburg. [in Russian].

Vshivkov, S.A., Rusinova, E.V., Kudrevatykh, N.V., Galyas, A.G., Alekseeva, M.S. & Kuznetsov, D.K. (2006). Phase transitions of hydroxypropylcellulose liquid-crystalline solutions in magnetic field. *Polymer Science.* A. Vol. 48, No 10, pp. 1115-1119.

Vshivkov, S.A., Adamova, L.V., Rusinova, E.V., Safronov, A.P., Dreval, V.E. & Galyas, A.G. (2007). Thermodynamics of liquid-crystalline solutions of hydroxypropyl cellulose in water and ethanol. *Polymer Sci.* A. Vol. 49, No. 5, pp. 578-583.

Vshivkov, S.A., Rusinova, E.V. & Kutsenko, L.I. (2007). Phase transitions in solutions of hydroxypropyl cellulose and cyanoethyl cellulose. *Polymer Sci.* B. Vol. 49, No. 5-6, pp.114-117.

Vshivkov, S.A. & Rusinova. E.V. (2007). Phase diagrams of a hydroxypropyl cellulose – water system under static conditions and in the shear field. *Polymer Sci*. B. Vol. 49, No. 7-8, pp. 209-212.

Vshivkov, S.A. Rusinova, E.V. Kutsenko, L.I. & Galyas, A.G. (2007). Phase transitions in liquid-crystalline cyanoethyl cellulose solutions in magnetic field. *Polymer Sci*. B. Vol. 49, No.7-8, pp. 200-202.

Vshivkov, S.A. Rusinova, E.V. & Kutsenko, L.I. (2007). Phase transitions of liquid crystalline cyanoethyl cellulose solutions under static conditions and on shear field. *Polymer Sci.* B. Vol. 49, No. 5-6, pp. 152-154.

Vshivkov, S.A. & Rusinova, E.V. (2007). Phase liquid-crystalline transitions in hydroxypropylcellulose-ethanol and hydroxypropylcellulose – acetic acid systems under deformation. *Polymer Sci.*, B. Vol. 49, No. 9-10, pp. 229-231.

Vshivkov, S.A. & Rusinova, E.V. (2008). Phase and structural transformations of liquid – crystalline polymer systems in the mechanical field. *Polymer Sci.*, A. Vol. 50, No. 2, pp. 135-141.

Vshivkov, S.A. & Rusinova, E.V. (2008). Effect of magnetic fields oh phase transitions in solutions of cellulose derivatives. *Polymer Sci.*, A. Vol. 50, No. 7, pp. 725-732.

Wee E. L. & Miller W. G. (1971). Liquid crystal-isotropic phase equilibriums in the system poly(gamma-benzyl-alpha-L-glutamate)-methylformamide. *J. Phys. Chem.*, V. 75, No. 10, pp. 1446-1452.

Werbowyj, R. S. & Gray, D. G. (1976). Liquid Crystalline Structure In Aqueous Hydroxypropyl Cellulose Solutions. *Mol. Cryst. Liq. Cryst.*, Vol. 34, No.4, pp. 97-103.

Werbowyj, R. S. & Gray, D. G. (1979). Cholesteric structure in concentrated hydroxypropyl cellulose solutions. *Am. Chem. Soc., Polym. Prepr. Div. Polym. Chem.*, Vol. 20, No. 1, pp. 102-105.

Werbowyj, R. S. & Gray, D. G. (1980). Ordered phase formation in concentrated hydroxpropylcellulose solutions. *Macromolecules*, *Vol. 13, No. 1, pp. 69-73.*

Yunusov, B.Yu., Khanchich, O.A., Dibrova, A.K., Primkulova, M.T. & Serkov, A.T. (1982). Phase transition at heating of liquid crystal solutions of acetates of cellulose in trifluoroacetic acid. *Vysokomol. Soedin.,* B. Vol. 24, No.6 , pp. 414-418.