Weierstraß-Institut
für Angewandte Analysis und Stochastik

im Forschungsverbund Berlin e.V.

Preprint ISSN 0946 – 8633

Modeling of Drift-Diffusion Systems

Holger Stephan[1]

submitted: 6 April 2006

[1] Weierstrass Institute
for Applied Analysis and Stochastics
E-Mail: stephan@wias-berlin.de

No. 1120
Berlin 2006

2000 *Mathematics Subject Classification.* 35K55, 80A20, 82D37, 82C31, 35B50, 35G25.

Key words and phrases. drift-diffusion systems, energy models, free energy, Lyapunov functions, positive solutions, inverse Hessian, cross diffusion.

Edited by
Weierstraß-Institut für Angewandte Analysis und Stochastik (WIAS)
Mohrenstraße 39
10117 Berlin
Germany

Fax: + 49 30 2044975
E-Mail: preprint@wias-berlin.de
World Wide Web: http://www.wias-berlin.de/

Contents

1 Introduction **2**

2 A short report on microscopic modeling **3**

3 Drift-diffusion problems **7**

3.1 The structure of the free energy 7

3.2 Examples of the free energy 9

3.3 One-species drift-diffusion problem 10

4 Many-species drift-diffusion problems **14**

4.1 Many diffusive non-interacting species 19

4.2 Many non-interacting species and pseudo-species 20

5 Drift-diffusion systems with a temperature field **24**

5.1 Thermodynamic properties and equilibrium state 25

5.2 Evolution equations and dissipation rate 27

6 Conclusions **33**

Abstract

We derive drift-diffusion systems describing transport processes starting from free energy and equilibrium solutions by a unique method. We include several statistics, heterostructures and cross diffusion. The resulting systems of nonlinear partial differential equations conserve mass and positivity, and have a Lyapunov function (free energy). Using the inverse Hessian as mobility, non-degenerate diffusivity matrices turn out to be diagonal, or—in the case of cross diffusion—even constant.

1 Introduction

A common problem of importance in applied mathematics is the description of the transport of particles under several influences in a spatial domain. Of particular interest is usually the time evolution of the concentrations of each type of particle, which we can describe by drift-diffusion systems ([1, 10, 14]). Typical problems of this kind are the transport of electrons and holes in semiconductor devices and the transport of impurities during the production of such devices. These problems are characterized by global electrostatic interaction of charged particles, anisotropic materials, heterostructures, spatial change of temperature and several statistics such as Boltzmann or Fermi-Dirac statistics ([1, 6, 3]).

Several processes can influence the time evolution of the concentrations studied. This paper will take the following processes into consideration.

Diffusion. Diffusion is the spatial redistribution of particles forced by their own kinetic energy. In this regard, diffusion is a typical phenomenological effect. We will describe it by means of effective driving forces, although these driving forces cannot be understood in a microscopic sense.

Drift. Drift is the transport of particles forced by a given outer field or a field generated by the particles in a self-consistent manner (e.g., an electrostatic field for charged particles).

Heat transport. It is clear that all processes depend on temperature. In many applications, it is necessary to consider a space dependent temperature field.

Let $\Omega \subset \mathbb{R}^d$ be a spatial domain and $x \in \Omega$. We denote time by $t \geq 0$ and the concentrations of m species X_1, ..., X_m by $u_1(x,t)$, ..., $u_m(x,t)$. Our goal is the derivation of a system of type $\dot{u} = \mathbf{E}(u)$ that describes the evolution of a vector of concentrations $u = (u_1, ..., u_m)$.

We know that a more exact description of the above mentioned processes has to be a microscopic one. Concentrations are an idealized point of view of an ensemble of classical point particles (or pseudo particles). We have to consider n_j particles of species $\mathbf{X}_j$ moving in some random medium. Unfortunately, this underlying microscopic structure is only partly known. Moreover, even if the microscopic model were well known, there would be no hope to derive macroscopic equations from microscopic ones, because of the complexity of the problems. This is regrettable, because from an analytical point of view, the microscopic equations are much simpler and have nice properties, like the existence of stationary solutions, time-decreasing Lyapunov functions (second law of thermodynamics), and positivity of the solutions. Therefore, we do not try to derive macroscopic equations starting from microscopic models. But in some simple cases—if there is no self-consistent interaction and many-particle effects can be neglected—the concentrations can be understood as probability densities. In these situations, microscopic and macroscopic equations coincide and should have the same properties. This is an important tool for the validation of the macroscopic equations.

We are going to derive evolution systems with the following general basic properties:

Conservation of positivity: $u_j(x,0) \geq 0 \Longrightarrow u_j(x,t) \geq 0$, $\forall\ t > 0$.

Existence of equilibrium concentrations $u_j^*(x)$, at least in the simplest situations.

Conservation of mass: $\int_\Omega u_j(x,0)dx = 1 \Longrightarrow \int_\Omega u_j(x,t)dx = 1$, $\forall\ t > 0$.

Second law of thermodynamics: There is a Lyapunov function $\Phi(u)$ called free energy and some concentration u^* with $\Phi(u) \geq \Phi(u^*)$, $\mathbf{E}(u^*) = 0$, and $\frac{d}{dt}\Phi(u) \leq 0$.

There is ample evidence that the difficulties of the mathematical analysis and numerical calculations of the evolution systems closely depend on these properties (see, e.g., [7, 8]). The derivation of evolution systems in a thermodynamically consistent manner was initiated in [1].

This paper is organized as follows: In section 2, we recall some basic properties of microscopic stochastic models and compare them with the simplest macroscopic problems. In section 3, we define the structure of the free energy and derive drift-diffusion equations for single species. In section 4, we derive drift-diffusion systems for many species, considering in detail the case when the many-species free energy is a sum of one-species and pseudo-species free energies. In section 5, we derive systems of equations for general drift-diffusion problems including a temperature field $T(x,t)$.

Notations. Indices counting species run from 1 to m and are denoted by i or j. Indices of the components of d-dimensional objects are denoted by α or β. If a is a $d \times d$ matrix, a^* its transpose and f and g are d-Vectors, then

$$\langle af, g\rangle = \langle f, a^* g\rangle = \sum_{\alpha,\beta=1}^{d} \int_\Omega a_{\alpha\beta}(x) f_\alpha(x) g_\beta(x) dx\ .$$

Moreover, we denote the scalar product of scalar valued functions by $\langle\cdot,\cdot\rangle$.

A function $f(x) \equiv 1$ is indicated by $\mathbb{1}$. We will often use operators $\mathbf{D}$ with kernel $\mathbb{1}$. A typical example is $\mathbf{D} = \mathbf{G}\nabla = \mathbf{G}$ grad. Here, $\mathbf{G}$ is a $d \times d$ matrix. Thus for a fixed f, we have $\mathbf{D}f : \mathbb{R} \to \mathbb{R}^d$. For the adjoint $\mathbf{D}^*$, we have $\mathbf{D}^* f : \mathbb{R}^d \to \mathbb{R}$, so in our example, $\mathbf{D}^* = -\nabla \mathbf{G}^* = -\mathrm{div}\ \mathbf{G}$.

We will use a dot to denote the derivative with respect to time, i.e. $\dot{g} = \frac{\partial}{\partial t} g(t)$ or $\dot{g} = \frac{d}{dt} g(t)$.

We formulate the results in the form of theorems. Since we are interested in modeling, the proofs are only formal.

2 A short report on microscopic modeling

Let us recall the main properties of stochastic microscopic models. For further information the reader is referred to [19, 21, 11]. Let $z \in \mathcal{Z}$ be the state and $\mathcal{Z}$ the state space of the interesting physical system. The state z is the set of parameters, necessary to distinguish two systems. Let $W(z,t)$ be the probability density function of the state z at time t. The problem is solved completely if this quantity is known for all times.

It is well known that $W(z,t)$ can be calculated for a given initial density $W_0(z) = W(z,0)$ by solving the Kolmogorov–Chapman equation

$$\dot{W}(t) = \mathbf{A}W(t) \ , \tag{1}$$

where $\mathbf{A}$ is the generator of a Markov semigroup. In case that $\mathcal{Z}$ is a manifold in $\mathbb{R}^n$, $\mathbf{A}$ has the form

$$\begin{aligned}(\mathbf{A}W)(z) &= \sum_{i,j=1}^{n} \frac{\partial^2}{\partial z_i \partial z_j}\Big(b_{ij}(z)W(z)\Big) - \sum_{i=1}^{n} \frac{\partial}{\partial z_i}\Big(a_i(z)W(z)\Big) + \\ &+ \fint_{\mathcal{Z}} \Big(Q(z',z)W(z') - Q(z,z')W(z)\Big)dz' \ . \end{aligned} \tag{2}$$

Here, $a_i(z)$, $b_{ij}(z)$, and $Q(z,z') \geq 0$ are suitable coefficient functions, and $\mathbf{B} = (b_{ij}(x))$ is a positive semidefinite matrix. The integral is to be understood as a principle value integral.

In general, equation (1) is to be understood in a weak sense, and W is the density of a probability measure that may not exist. In this paper, we do not consider regularity problems and will assume that the underlying measures have densities with respect to the Lebesgue measure. Looking at the problem from a physical point of view, we will assume that all quantities are well defined, and all equations have a solution. Assume further that the solutions of the instationary equations are unique and globally defined.

The microscopic model has the following important properties.

Theorem 1 *Let $W(z,t)$ be the probability density of some Markov process. Then $W(z,t)$ is the unique solution of equation (1) and the following properties hold:*

1. *The operator $\mathbf{A}$ has the structure (2) and is linear.*
2. *The equation (1) conserves positivity: $W_0(z) \geq 0 \Longrightarrow W(z,t) \geq 0, \ \forall \ t > 0$.*
3. *The equation (1) conserves mass: $\int_{\mathcal{Z}} W(z,t)dz = \int_{\mathcal{Z}} W_0(z)dz = 1, \ \forall \ t > 0$.*
4. *If $W(z,t)$ converges in time to W^*, then $\mathbf{A}W^* = 0$.*
5. *The stationary equation $\mathbf{A}W^* = 0$ has at least one nonnegative solution.*
6. *Any convex function $F : \mathbb{R} \longrightarrow \overline{\mathbb{R}}$ not depending on $\mathbf{A}$ generates a Lyapunov function of the form*

$$\Phi[W] = \Phi(t) = \int_{\mathcal{Z}} F\left(\frac{W(z,t)}{W^*(z)}\right) W^*(z)dz \ , \tag{3}$$

with the property

$$\frac{d}{dt}\Phi(t) = -\int\limits_{Z} \sum_{i,j=1}^{n} F''(\theta)\,\theta_i\theta_j b_{ij}(z) W^*(z,t)dz - \\ - \iint\limits_{ZZ} \Big(F(\theta) - F(\theta') - F'(\theta')(\theta-\theta')\Big) W^*(z) Q(z,z')dz\,dz' \le 0 \; .$$

Here, we have set

$$\theta_i = \frac{\partial}{\partial z_i} \log \frac{W(z,t)}{W^*(z)}, \;\; \theta = \frac{W(z,t)}{W^*(z)}, \;\; \theta' = \frac{W(z',t)}{W^*(z')} \; .$$

For the proof, see [19, 21] and the references there.

The microscopic description can serve as a starting point to derive macroscopic equations. This is a very hard problem, and we are not going to derive macroscopic equations in a rigorous manner. We are rather looking for microscopic equations that can be compared with the macroscopic ones in simple situations.

Let us assume that the state of one particle is z. Considering $n_1 + ... + n_m = n$ particles of m species $\mathbf{X}_1$, ..., $\mathbf{X}_m$, we have to investigate a microscopic equation for a probability density $W(z_{11}, ..., z_{1n_1}, ..., z_{mn_m}, t)$, a function depending on n parameters z_i and time t. To begin with, we will assume that n_j particles of species $\mathbf{X}_j$ can be described by a one-particle probability density $W(z_1, ..., z_m, t)$. We will later try to improve this simplification by using self consisting interactions (mean fields), nonlinear material coefficients and the correct statistics, which will describe how to count n_j particles for a given macroscopic problem. More precisely, this means that in order to describe a many-particle problem by a one-particle equation, we have to choose a specific function of type (3) from all possible convex functions.

Inspecting classical particles, we recall that the state of one particle is given by its space coordinate x and its momentum or velocity v, i.e., $z = (x, v) \in \mathbb{R}^{2d}$. Here, the evolution of the velocity is determined by some forces (Newton's law), and the space coordinate is varying according to the velocity. Supposing that the velocity is already relaxed to an equilibrium distribution, we can assume that the state of one particle is given by the space coordinate $x \in \Omega \subset \mathbb{R}^d$. For the general Brownian motion, [20] explicitly gives an equation for the $(v, x) \to (x)$ reduced problem. These results show that there are good reasons to reduce the problems if the temperature is constant. Besides, they demonstrate how the reduced equation implicitly depends on the equilibrium distribution of the velocity.

To derive models with a non-constant temperature field, we also have to take into account the velocity, which models the energy transfer between the particles and the medium. It is not even clear how to define a constant equilibrium quantity like temperature in the non-equilibrium case. Therefore, we model the dependence on temperature in section 5 in a purely phenomenological way.

The reduction leads to a probability density function $W(x_1, ..., x_m, t)$ with $x_i \in \Omega$. $W(x_1, ..., x_m)$ is the probability density function finding the particles of type i at point x_i. The equation for the evolution of W is equation (1)–(2), where $z = (x_1, ..., x_m)$ is a $m \times d$–dimensional parameter. The microscopic Lyapunov functions are

$$\Phi[W] = \int_{\mathcal{Z}} F\left(\frac{W(x_1, ..., x_m)}{W^*(x_1, ..., x_m)}\right) W^*(x_1, ..., x_m) dx_1 \cdots dx_m \ , \tag{4}$$

with any convex function $F : \mathbb{R} \longrightarrow \mathbb{R}$ and a stationary solution W^* satisfying $\mathbf{A}W^* = 0$.

As usual, the macroscopic concentration $u_i(x)$ are assumed to be one-particle probability densities and are defined as

$$u_i(x) = \int_{\Omega^{m-1}} W(x_1, ...x_{i-1}, x, x_{i+1}, ..., x_m) dx_1 \cdots dx_{i-1} dx_{i+1} \cdots dx_m \ , i = 1, ..., m. \tag{5}$$

From the norm condition $\int_\Omega W(x)dx = 1$ it follows the mass conservation of each species:

$$\int_\Omega u_i(x)dx = 1 \ , i = 1, ..., m \ . \tag{6}$$

Assuming that $u = (u_1, ..., u_m)$ is the solution of some equation $\dot{u} = \mathbf{E}(u)$, the linearity of $\mathbf{E}$ follows from the linearity of $\mathbf{A}$. The structure of this operator can be described.

Theorem 2 *Let $u = (u_1, ..., u_m)$ be the solution of some linear evolution equation $\dot{u} = \mathbf{E}(u)$, which conserves mass (6) and positivity. Then $u = (u_1, ..., u_m)$ is the solution of a system*

$$\dot{u}_j = \mathbf{A}_j u_j + \sum_{j \neq i=1}^{m} \mathbf{C}_{ij} u_i + \sum_{j \neq i=1}^{m} e_{ij}(u_i - u_j), \ j = 1, ..., m \ , \tag{7}$$

where the operators $\mathbf{A}_j$ are of type (2) and the operators $\mathbf{C}_{ij}$ are bounded integral operators of the type

$$(\mathbf{C}_{ij} f)(x) = \int_\Omega c_{ij}(x, y) f(y) dy - f(x) \int_\Omega c_{ij}(y, x) dy \ ,$$

with functions $c_{ij}(x, y) \geq 0$ that are connected with the nonnegative constants e_{ij} by

$$e_{ij} \geq \int_\Omega c_{ij}(y, x) dy \ .$$

The proof will be presented in [21].

Neglecting nonlinear effects, interaction of particles of the same kind (many-particle effects of one species) and other characteristics of the particles (e.g., velocity), a general evolution system for the concentrations of m species has to be of this structure. Let us point out that this system has a diagonal main part. This holds not only for linear systems, but also for nonlinear non-degenerate ones. If a nonlinear system $\dot{u} = \mathbf{E}(u)$ contains unbounded off-diagonal operators, they have to degenerate.

3 Drift-diffusion problems

In this section, we are going to derive an evolution equation for the drift and diffusion of one species. For this purpose, we start from a given free energy Φ and a given reference concentration u^*. In the case of no self-consistent interaction this should be the equilibrium solution.

3.1 The structure of the free energy

As usual, the free energy is a function of u and u^*,

$$\Phi(u) = \int_\Omega F(u(x), u^*(x))dx \ ,$$

with some function F that describes the local dependence of the free energy. It consists of a chemical part $\Phi_C(u)$ and a self-consistent interaction part $\Phi_I(u)$. As usual, we assume

$$\Phi(u) = \Phi_C(u) + \Phi_I(u) \ .$$

This ansatz is common [5, 25], but seems to be correct only in the Boltzmann case.

We assume that $\Phi_I(u)$ describes self-consistent interaction generated by a nonlocal interaction field described by a linear symmetric operator $\mathbf{K}$,

$$\Phi_I(u) = \frac{1}{2}\int_\Omega u(x)(\mathbf{K}u)(x)dx \ ,$$

where

$$(\mathbf{K}u)(x) = \int_\Omega k(x,y)u(y)dx, \ k(x,y) = k(y,x) \ .$$

As an example, this field can be understood as the electrostatic field. This field is generated by fixed charges in the domain and on the boundary, as well as by the diffusing particles them self. Neglecting dynamical effects, we assume that this field is instantaneously given by position and type of the particles. In this case, $\mathbf{K}$ is the solution operator of the Poisson equation including boundary conditions.

The chemical part $\Phi_C(u)$ contains the statistics of the problem, describing how many-particle effects have to be considered using a one-particle picture. $\Phi_C(u)$ has to be of type (3) with a special convex function. In general (for heterostructures), $u(x)$ and $u^*(x)$ can be nonsmooth. If the particles can pile up on interfaces or in points, the concentration can be a measure and $u(x)$ is not defined. Looking for free energies for general situations, the definition of the free energy has to take into account this behaviour.

We collect several definitions of the free energy depending on the considered problem (see [19]):

$$\begin{aligned}\Phi(u) &= \sup_{g\in\mathcal{C}}\left(\int_\Omega g(x)p(dx) - \int_\Omega G\big(g(x)\big)p^*(dx)\right) = &(8)\\ &= \int_\Omega F\left(\frac{p}{p^{*\mathrm{RN}}}(x)\right)p^*(dx) = &(9)\\ &= \int_\Omega F\left(\frac{u(x)}{u^*(x)}\right)u^*(x)dx = &(10)\\ &= \int_\Omega \tilde{F}(u(x))dx &(11)\end{aligned}$$

Here, $F = G^*$ is the convex conjugate of $G, \frac{p}{p^{*\mathrm{RN}}}$ is the Radon-Nikodym derivative of p with respect to p^*, and $\mathcal{C}$ is a suitable set of test functions.

In the most general case, if the concentrations are Radon measures, (8) is the correct definition of the free energy. If p is Radon-Nikodym differentiable with respect to p^*, we can consider (9) instead of (8). If, moreover, p and p^* are differentiable with respect to the Lebesgue measure with densities u and u^* resp., we can consider (10). If, finally, the equilibrium density is smooth enough, it is not necessary to point out the dependence on u^*, and (11) can be taken with some convex $\tilde{F}$.

As already mentioned, not caring about regularity, we will assume that all measures are differentiable with respect to the Lebesgue measure. Thus, we will use expression (10).

Since F is convex, the free energy is bounded from below by Jensen's inequality.

In the case of m species we have $u = (u_1, ..., u_m)$, and in general the free energy is of the form

$$\Phi(u) = \int_\Omega F\big(u_1(x), ..., u_m(x)\big)dx\ .$$

Here, $F : \mathbb{R}^m \to \mathbb{R}$ is a convex function of m variables. The gradient

$$\partial F(u) = \begin{pmatrix} \frac{\partial F(u)}{\partial u_1} \\ \vdots \\ \frac{\partial F(u)}{\partial u_m}\end{pmatrix}$$

is a monotone operator $\partial F : \mathbb{R}^m \longrightarrow \mathbb{R}^m$. The solution of the equation $\partial F(u) = \lambda$ is uniquely determined by its inverse $u = \partial F^{-1}(\lambda)$. The second derivative $\partial^2 F : \mathbb{R}^m \to \mathbb{R}^m \times \mathbb{R}^m$ is for fixed u a $m \times m$ matrix—the Hessian. Its inverse $\left[\partial^2 F(u)\right]^{-1}$ is the inverse Hessian. $\partial^2 F(u)$ and its inverse are positive semidefinite matrices if F is strict convex.

If $F(u)$ is the sum of convex functions of one species $F(u) = \sum_{i=1}^m F_i(u_i)$, then $\partial^2 F(u) = \mathrm{diag}(F_1''(u_1), ..., F_m''(u_m))$ is diagonal.

For a unique notation, we set for one species u

$$\partial F(u) = F'(u),\ \partial^2 F(u) = F''(u),\ \left[\partial^2 F(u)\right]^{-1} = \frac{1}{F''(u)}\ .$$

3.2 Examples of the free energy

Here, we examine some examples for the chemical part of the free energy. This is the logarithm of the number of possibilities to arrange the particles. This number has to take into account the distinguishability of particles. Describing a many-particle problem simplified by one-particle equations, this number tells us how to improve our one-particle picture for many particles. Having $n_1 + ... + n_m = n$ particles, one extremal situation is that all particles are distinguishable. In this case we have

$$\Phi_C(n_1, ..., n_m) = \log\big(n_1! \cdot n_2! \cdots n_m!\big) = \sum_{i=1}^{m} \log n_i!\ . \tag{12}$$

Assuming a large number of particles, we can use Stirling's formula and obtain

$$\Phi_C(n_1, ..., n_m) = \sum_{i=1}^{m} \log\left(\frac{n_i}{e}\right)^{n_i} = \sum_{i=1}^{m} n_i(\log n_i - 1)\ .$$

Identifying the particle number with the concentration, we get

$$\Phi_C(u_1, ..., u_m) = \sum_{i=1}^{m} F_B(u_i) = \sum_{i=1}^{m} u_i(\log u_i - 1)\ ,$$

the Boltzmann statistics, describing the free energy of many distinguishable (i.e., classical) particles. Sometimes $F_B(u_i) = u_i \log u_i$ is used, neglecting a non-interesting linear shift in the free energy density.

If we study quantum particles, some particles may be indistinguishable. This leads to a function $F(u) \leq F_B(u)$ for the statistics (the number of possibilities to arrange the particles decreases). It is natural to have $F(u) \approx F_B(u)$ for small u (few particles are more distinguishable than many). A typical family of statistic functions is for $\alpha \geq 0$

$$F_\alpha(u) = uI_\alpha^{-1}(u) - I_{\alpha+1}\big(I_\alpha^{-1}(u)\big)\ ,$$

where

$$I_\alpha(\lambda) = \frac{1}{\Gamma(\alpha+1)} \int_0^\infty \frac{\xi^\alpha}{1 + e^{\xi-\lambda}} d\xi$$

is the Fermi-integral (or polylog-function). We have the main properties $I'_\alpha(\lambda) = I_{\alpha-1}(\lambda)$, $I_\alpha(\lambda) \geq 0$ for $\alpha \geq -2$. I_α is monotone for $\alpha \geq -1$ and convex for $\alpha \geq 0$.

In the following table, we collect explicit representations and asymptotics for $F_\alpha(u)$ and related functions.

function	$u \to 0$	$u \to \infty$
$F_\alpha(u) = uI_\alpha^{-1}(u) - I_{\alpha+1}\big(I_\alpha^{-1}(u)\big)$	$u(\log u - 1)$	$\Gamma(\alpha+1)^{\frac{1}{\alpha+1}}\frac{\alpha+1}{\alpha+2}u^{\frac{\alpha+2}{\alpha+1}}$
$\partial F_\alpha(u) = I_\alpha^{-1}(u)$	$\log u$	$\Gamma(\alpha+1)^{\frac{1}{\alpha+1}}u^{\frac{1}{\alpha+1}}$
$\partial F_\alpha^{-1}(\lambda) = I_\alpha(\lambda) \geq 0$	e^λ	$\frac{\lambda^{\alpha+1}}{\Gamma(\alpha+1)}$
$\partial^2 F_\alpha(u) = \frac{1}{I_{\alpha-1}(I_\alpha^{-1}(u))}$	$\frac{1}{u}$	$\Gamma(\alpha+1)^{\frac{1}{\alpha+1}}\frac{1}{\alpha+1}u^{-\frac{\alpha}{\alpha+1}}$
$[\partial^2 F_\alpha(u)]^{-1} = I_{\alpha-1}\big(I_\alpha^{-1}(u)\big)$	u	$(\alpha+1)\Gamma(\alpha+1)^{\frac{-1}{\alpha+1}}u^{\frac{\alpha}{\alpha+1}}$

We have the special cases $F_\infty(u) = F_B(u)$ and $F_0(u) = \frac{1}{2}u^2$. $F_B(u)$ and related functions behave for any u like $F_\alpha(u)$ for $u \to 0$.

Other widely used statistic functions are the so-called Tsallis-entropies (see [23])

$$F_q(u) = \frac{u - u^q}{1-q} \ .$$

The main interesting functions can be calculated explicitly, but some important properties for applications, like $[\partial^2 F_q(u)]^{-1} \approx u$ for $u \to 0$ and $\partial F_q^{-1}(\lambda) \geq 0$, are not fulfilled.

If we consider m species and looking for separated concentrations u_i for every species, we implicitly assume that particles from different species are distinguishable. This means that the number of possibilities to distribute particles of different kind is the product of one-particle numbers. Let $F_i(u_i)$ be the free energy of the i-th species, then we get from (12)

$$\Phi_C(u_1, ..., u_m) = \sum_{i=1}^m F_i(u_i) \ . \tag{13}$$

Therefore, a general free energy for many species like (3), which is not of type (13), is difficult to understand as a function of separated concentrations u_i.

3.3 One-species drift-diffusion problem

In this case, we simply assume that the state is the space coordinate $z = x$, and the probability density function can be understood as a concentration $u(x,t) = W(z,t)$. In the case in which interaction and nonlinearities are absent, the equation for u is an equation of type (1) with some operator $\mathbf{A}$ of type (2). Now the problem is to derive an equation for u, knowing a stationary solution u^* for the no-interaction case ($\mathbf{A}u^* = 0$), some free energy of the form

$$\Phi(u) = \int_\Omega F\left(\frac{u(x,t)}{u^*(x)}\right) u^*(x)dx + \frac{1}{2}\int_\Omega (\mathbf{K}u)(x)u(x)dx \tag{14}$$

with a convex function F, a symmetric operator $\mathbf{K}$, and a conservation law

$$\int_\Omega u(x,t)dx = \int_\Omega u^*(x)dx = 1 \ . \tag{15}$$

We calculate the stationary state $\bar{u}(x)$ by Lagrange's method, varying the functional $\Phi(u)$ under the constraint (mass conservation). We consider the Lagrangian

$$L(u) = \Phi(u) - \lambda \left(\int_\Omega u(x,t)dx - 1 \right) \ .$$

The stationary state $\bar{u}(x)$ is the solution to the Euler–Lagrange equation

$$\frac{\delta L(u)}{\delta u} = \partial F \left(\frac{\bar{u}(x)}{u^*(x)} \right) + (\mathbf{K}u)(x) - \lambda = 0 \ .$$

Now we introduce an interacting potential ψ as

$$(\mathbf{K}u)(x) = \psi(x) \ .$$

As an example, this potential can be understood as the electrostatic potential. In this case, $\mathbf{K}$ is the solution operator of the Poisson equation including boundary conditions. It follows

$$\begin{aligned} \bar{u}(x) &= \partial F^{-1}\big(\lambda - \psi(x)\big)u^*(x) \ , \\ (\mathbf{K}\bar{u})(x) &= \psi(x) \ . \end{aligned}$$

In the case of no interacting $\mathbf{K} = \mathbf{O}$, $\psi = 0$, we get, from (15),

$$\partial F^{-1}(\lambda) = 1$$

and therefore $\lambda = \partial F(1)$. It follows $\bar{u}(x) = u^*(x)$, as expected.

We are now going to derive an evolution equation for the concentration $u(x,t)$ in such a way that the desired equation can be generalized to more complicated situations such as nonlinear material laws and heterostructures. The driving force for a diffusion equation is often assumed to be the gradient of the concentration ∇u. In general, this is not suitable for the following reasons:

- It is well known from thermodynamics (see, e.g., [26]) that driving forces should be gradients of some potentials. Potentials are intensive variables whereas the concentration is an extensive variable.
- The driving force has to vanish in the equilibrium case, i.e., it has to be a gradient of a function that is constant in the equilibrium. Therefore, the gradient of the concentration is the correct driving force if and only if the equilibrium concentration is constant. This is, of course, a very special case, which is not fulfilled, for example, by heterostructures.

- From an analytical point of view, the driving force has to be a smooth function. Zero is such a smooth function, and we can study the problem with a smooth driving force at least near the equilibrium. In heterostructures, it is typical that the concentration is not smooth even in equilibrium.

The only intensive variable, which is surely constant in the equilibrium, is the Lagrange multiplier. Therefore, we assume that the Lagrange multiplier λ is the chemical potential, depending on x and t, and

$$\begin{aligned} u(x,t) &= \partial F^{-1}(\lambda(x,t) - \psi(x,t))u^*(x) \;, \\ (\mathbf{K}u)(x,t) &= \psi(x,t) \end{aligned} \tag{16}$$

is the state equation, connecting the potential and the concentration.

Let $\mathbf{D}$ be a linear operator with a kernel, containing the stationary potential, i.e., $\mathbf{D}\mathbb{1} = 0$. We assume that $\mathbf{D}\lambda$ is the driving force of the diffusion vanishing in the stationary state $\lambda = \text{const}$.

Now we postulate an evolution equation for u of the form

$$\frac{\partial}{\partial t} u = -\mathbf{D}^* \boldsymbol{\sigma}\, \mathbf{D}\, \lambda \;, \tag{17}$$

$$u = u^* \partial F^{-1}(\lambda - \psi) \;, \tag{18}$$

$$\mathbf{K}u = \psi \;. \tag{19}$$

with some suitable operator $\boldsymbol{\sigma}$ which is called the conductivity in the semiconductor literature.

We know that in the simple case of one diffusing species the equations (1)–(2) and (17) describe the same physical problem. As usual, $\mathbf{D}$ is taken as a gradient $\mathbf{D} = \mathbf{G}\nabla$ with a $d \times d$ matrix $\mathbf{G}(x)$ (implicitly, this means Maxwell distributed velocities with constant temperature; see [20]).

So far, the operator $\boldsymbol{\sigma}$ is arbitrary. Taking $\mathbf{D}$ proportional to the gradient, we choose $\boldsymbol{\sigma}$ in the form

$$\boldsymbol{\sigma} = \boldsymbol{\mu}\, u^* \left[\partial^2 F(u/u^*)\right]^{-1} \tag{20}$$

with some suitable $d \times d$ matrix $\boldsymbol{\mu}$. This choice of $\boldsymbol{\sigma}$ is not common. Often (see, e.g., [17, 25, 5]), $\boldsymbol{\sigma}$ is taken proportional to u. This is right for Boltzmann statistics. If we want to choose $\boldsymbol{\sigma}$ in such a way that equation (17) is of type (1)–(2) in the case of no interaction, then (20) is the only possible way. Moreover, the many-species case shows that the mobility is the product of some diagonal diffusion matrix and the inverse Hessian. This is the Einstein relation. Microscopic considerations [12] give further arguments for this choice.

Finally, we get the system

$$\frac{\partial}{\partial t} u = \nabla \mathbf{G}^* \boldsymbol{\mu} u^* \left[\partial^2 F(u/u^*)\right]^{-1} \mathbf{G}\, \nabla\, \lambda \;, \tag{21}$$

$$u = u^* \partial F^{-1}(\lambda - \psi) \;, \tag{22}$$

$$\mathbf{K}u = \psi \;. \tag{23}$$

Theorem 3 *The system (21)–(23) conserves mass and positivity, and has a time-decreasing free energy* $\Phi(t)$*.*

Proof: The conservation of mass follows for arbitrary positive semidefinite $\boldsymbol{\sigma}$:

$$\frac{d}{dt}\langle \mathbb{1}, u(t)\rangle = \langle \mathbb{1}, \dot{u}(t)\rangle = -\langle \mathbb{1}, \mathbf{D}^* \boldsymbol{\sigma}\, \mathbf{D}\,\lambda\rangle = -\langle \mathbf{D}\mathbb{1}, \boldsymbol{\sigma}\, \mathbf{D}\,\lambda\rangle = 0\ .$$

Moreover, for such $\boldsymbol{\sigma}$, we have

$$\frac{d}{dt}\Phi\big(u(t)\big) = \left\langle \partial F\left(\frac{u(t)}{u^*}\right), \dot{u}(t)\right\rangle = -\left\langle \lambda, \mathbf{D}^* \boldsymbol{\sigma}\, \mathbf{D}\,\lambda\right\rangle = -\left\langle \boldsymbol{\sigma}\, \mathbf{D}\,\lambda, \mathbf{D}\,\lambda\right\rangle \le 0\,.$$

The positivity of the solution follows from the property of the Boltzmann or general Fermi–Dirac statistics $\partial F^{-1} \ge 0$. ■

Assuming no interaction $\mathbf{K} = \mathbf{O}$ and inserting the state equation, we get

$$\frac{\partial}{\partial t}u(x,t) \quad = \quad \nabla \mathbf{G}^* \boldsymbol{\mu} u^* \big[\partial^2 F(u/u^*)\big]^{-1} \mathbf{G}\,\nabla\partial F\left(\frac{u}{u^*}\right)\ . \tag{24}$$

This is the right equation in the case that u^* is not smooth like in the case of heterostructures.

Assuming for a moment sufficient smoothness of u and u^*, we would get from equation (24)

$$\begin{aligned}
\frac{\partial}{\partial t}u(x,t) &= \nabla \mathbf{G}^* \boldsymbol{\mu} u^* \big[\partial^2 F(u/u^*)\big]^{-1} \mathbf{G}\,\partial^2 F(u/u^*)\nabla\frac{u}{u^*} = \\
&= \nabla \mathbf{G}^* \boldsymbol{\mu} u^* \mathbf{G}\left(\frac{1}{u^*}\nabla u - \frac{u}{u^*}\frac{1}{u^*}\nabla u^*\right) = \\
&= \nabla \mathbf{G}^* \boldsymbol{\mu}\, \mathbf{G}\left(\nabla u - u\,\nabla \log u^*\right) = \\
&= \nabla \mathbf{G}^* \boldsymbol{\mu}\, \mathbf{G}\left(\nabla u + u\,\nabla V\right)\ ,
\end{aligned} \tag{25}$$

where we introduced a function $V(x)$ describing a given external potential or a non-homogeneity of the space. This is the Fokker–Planck equation, a special case of (1)–(2) with $Q = 0$, $\mathbf{B} = \mathbf{G}^* \boldsymbol{\mu}\, \mathbf{G}$, and

$$a_i(x) = \sum_j \frac{\partial}{\partial x_j}\, b_{ij}(x) + \frac{1}{u^*(x)}\frac{\partial}{\partial x_j}\, u^*(x)\ .$$

The equilibrium concentration is the well-known Gibbs equilibrium $u^*(x) = \frac{1}{Z}e^{-V(x)}$ for sufficiently fast growing potentials such that the statistical sum $Z = \int e^{-V(x)}dx$ is well defined. Equation (25) is a typical drift-diffusion equation, considered for instance in [17, 25, 5].

Let us remark that starting from a given free energy

$$\Phi(u) = \int_\Omega F\left(u(x)\right)dx + \int_\Omega V(x)u(x)dx + \frac{1}{2}\int_\Omega (\mathbf{K}u)(x)u(x)dx, \tag{26}$$

we get equation (25) only in the Boltzmann case. Indeed, it is not clear why equation (25) should describe the linear, non-interaction case only for Boltzmann statistics. A free energy of type (10) generated by any convex function F is a Lyapunov function for equation (25) (see, [19, 24, 2]). Thus, ansatz (14) seems to be better than (26).

Equation (24) can be generalized by allowing nonlinear matrices $\boldsymbol{\mu}$ (depending on the concentration u, the driving force $\nabla\lambda$, or other quantities). The mass conservation and the time decrease of the free energy is fulfilled as well, whereas, in general, the positivity of the solution only follows for positive $(\partial F)^{-1}$.

If the material is anisotropic then $\boldsymbol{\mu}$ can describe this anisotropy. As an example, $\boldsymbol{\mu}$ can be a diagonal matrix with unequal diagonal elements, describing the different behaviour in various directions (see [16] for examples in semiconductor physics). The matrix $\mathbf{G}$ describes the orientation of the material with respect to the coordinate system. Thus, $\mathbf{G}$ is an orthogonal matrix, and $\mathbf{G}^*$ is its inverse.

4 Many-species drift-diffusion problems

Now we consider the case of m interacting species. The main problem of the modeling of many-species drift-diffusion systems is the right structure of the free energy. It is not obvious how to generalize the free energy (8)–(10) to the many-species case. If the chemical part of the free energy is a sum of one-species free energies, it is natural to set

$$\Phi_C(u) = \Phi_C(u_1, ..., u_m) = \sum_{i=1}^{m} \int_\Omega F_i\left(\frac{u_i(x)}{u_i^*(x)}\right) u_i^*(x)dx \ . \tag{27}$$

This is the right form of the free energy (13), taking into account stationary solutions u_i^*. The only interaction of particles of different type is the self-consistent interaction via a global field. We consider this case in the following subsection.

Cross interaction can also occur without self-consistent interaction. A possible way to analyze this case is the following. We start with some free energy of type (27), but presume that the particles are not independent of each other. Therefore, we are interested only in some of them. Or, in other words, cross interaction can be understood as the result of some pseudo-species. We will examine this idea in the second subsection.

In general, the dependence of the free energy on the equilibrium concentrations u_j^* is so far unknown. A derivation from the microscopic free energy (4), using (5), is possible only when making further assumptions such as Boltzmann statistics and factorization of the m-particle probability density, and ultimately leads to (27).

Therefore, to formally sketch the derivation of an evolution system, we start with a free energy modeling cross interaction

$$\Phi(u) = \int_\Omega F(u_1(x), ..., u_m(x))dx + \frac{1}{2}\langle u, \mathbf{K}u\rangle + \langle u, \mathbf{K}_0 f\rangle \ , \tag{28}$$

with no further structure given, taking into account, however, a given external field, generated by some given function $f(x)$ via an operator $\mathbf{K}_0$.

To illustrate the structure of the terms in (28), we consider the case of electrostatic interaction, assuming that the particles are charged (particle $\mathbf{X}_j$ with charge number q_j), and $f(x)$ is a density of given and fixed charges. This is a typical situation in semiconductor physics.

Let $k(x,y)$ be Green's function of Poisson's equation including boundary conditions, $\mathbf{K}_1$ the integral operator with kernel $k(x,y)$, and $\psi_0(x)$ the solution of Poisson's equation

$$-\nabla\big(\epsilon\nabla\psi_0(x)\big) = f(x) + \sum_{i=1}^{m} q_i u_i(x) \ .$$

Thus,

$$\psi_0(x) = \int_\Omega k(x,y)\left(f(y) + \sum_{i=1}^{m} q_i u_i(y)\right) dy = \mathbf{K}_1(f + \sum_{i=1}^{m} q_i u_i) \ .$$

We set

$$\psi(x) = \begin{pmatrix} q_1 \\ \vdots \\ q_m \end{pmatrix} \psi_0(x) = (\mathbf{K}u)(x) + (\mathbf{K}_0 f)(x) \ , \tag{29}$$

with

$$(\mathbf{K}_0 f)(x) = \int_\Omega k(x,y)\begin{pmatrix} q_1 \\ \vdots \\ q_m \end{pmatrix} f(y)dy = \begin{pmatrix} q_1 \\ \vdots \\ q_m \end{pmatrix} \mathbf{K}_1 f \ ,$$

and

$$(\mathbf{K}u)(x) = \int_\Omega k(x,y)\begin{pmatrix} q_1^2 & \cdots & q_1 q_m \\ \vdots & \ddots & \vdots \\ q_m q_1 & \cdots & q_m^2 \end{pmatrix}\begin{pmatrix} u_1(y) \\ \vdots \\ u_m(y) \end{pmatrix} dy = \mathbf{K}_0 \sum_{i=1}^{m} q_i u_i \ .$$

Considering cross interaction without reactions, we assume that the particles can interact but cannot change their type. Thus, we assume conservation laws for each species

$$\int_\Omega u_i(x,t)dx = \int_\Omega u_i(x,0)dx = 1, \ i = 1, ..., m \ . \tag{30}$$

Varying the Lagrangian

$$L(u,\lambda) \ = \ \Phi(u) - \sum_{i=1}^{m} \lambda_i \left(\int_\Omega u_i(x,t)dx - 1\right) \ ,$$

we get the Euler–Lagrange equation

$$\begin{aligned}\lambda &= \partial F(u) + \mathbf{K}u + \mathbf{K}_0 f = \\ &= \partial F(u) + \psi \ ,\end{aligned}$$

where $\lambda = (\lambda_1, ..., \lambda_m)$ is a vector of Lagrange multipliers, and $\psi = \mathbf{K}u + \mathbf{K}_0 f$. The stationary solution is the solution to the system

$$u^* = (\partial F)^{-1}(\lambda - \psi) \ , \tag{31}$$

$$\mathbf{K}_1^{-1}\psi_0 = \sum_{i=1}^{m} q_i u_i^* + f \ . \tag{32}$$

This seems to be the right formulation because, in general, the operators $\mathbf{K}$ and $\mathbf{K}_0$ are not given. In the case of electrostatic interaction, equation (32) is Poisson's equation, with a well-known operator $\mathbf{K}_1^{-1}$, the same for all problems, whereas the solution operators $\mathbf{K}$ and $\mathbf{K}_0$ depend on the domain and on the boundary conditions. To calculate the stationary solution, we have to solve the system (31)–(32) simultaneously (and to observe (30)).

As in the one-species case, we assume that in the non-equilibrium case the Lagrange multiplier $\lambda = (\lambda_1(x,t), ..., \lambda_m(x,t))$ is a space-, time- and species-dependent quantity—the chemical potential—and

$$u(x,t) = (\partial F)^{-1}\big(\lambda(x,t) - \psi(x,t)\big) \tag{33}$$

is the state equation, where $\psi(x,t)$ is defined by (29) and depends only on a scalar function ψ_0. In the case of electrostatic interaction, $\lambda - \psi$ is called the vector of chemical potentials.

Let $\mathbf{D}_i$ be linear operators with a kernel, containing the stationary potential, i.e., $\mathbf{D}_i \mathbb{1} = 0$. We assume that $\mathbf{D}_i \lambda_i$ are the driving forces of the drift-diffusion vanishing in the stationary state $\lambda_i = \text{const}$.

Now we postulate a system of evolution equations for u of the form

$$\frac{\partial}{\partial t} u_j(x,t) = \mathbf{D}_j^* J_j, \ j = 1, ..., m \ , \tag{34}$$

where the flows J_i are connected with the driving forces by

$$\begin{pmatrix} J_1 \\ \vdots \\ J_m \end{pmatrix} = - \begin{pmatrix} \sigma_{11} & \cdots & \sigma_{1m} \\ \vdots & \ddots & \vdots \\ \sigma_{m1} & \cdots & \sigma_{mm} \end{pmatrix} \begin{pmatrix} \mathbf{D}_1 \lambda_1(x,t) \\ \vdots \\ \mathbf{D}_m \lambda_m(x,t) \end{pmatrix} \ , \tag{35}$$

with some suitable operator $\boldsymbol{\sigma} = (\sigma_{ij})$. Taking into account Onsager's principle, we assume that the $d \times d$ matrices σ_{ij} form a symmetric $m \times m$ matrix, i.e., $\sigma_{ij} = \sigma_{ji}$.

As in the one-species case, we have mass conservation

$$\frac{d}{dt} \int_\Omega u_i(x,t) dx = \langle \mathbb{1}, \dot{u}_i \rangle = \langle \mathbb{1}, \mathbf{D}_i^* J_i \rangle = \langle \mathbf{D}_i \mathbb{1}, J_i \rangle = 0 \ , \ i = 1, ..., m \ .$$

To investigate the time dependence of the free energy (42), we calculate

$$
\begin{aligned}
\frac{d}{dt}\Phi\big(u(t)\big) &= \sum_{i=1}^{m}\int_{\Omega}\lambda_i(x,t)(\mathbf{D}_i^*J_i)(x,t)dx = \\
&= \sum_{i=1}^{m}\langle\lambda_i,\mathbf{D}_i^*J_i\rangle = -\sum_{i,j=1}^{m}\langle\lambda_i,\mathbf{D}_i^*\sigma_{ij}\mathbf{D}_j\lambda_j\rangle = -\sum_{i,j=1}^{m}\langle\sigma_{ij}\mathbf{D}_j\lambda_j,\mathbf{D}_i\lambda_i\rangle\ .
\end{aligned}
$$

Thus, the free energy decreases if the dissipation rate

$$
R(t) = \sum_{i,j=1}^{m}\langle\lambda_i,\mathbf{D}_i^*\sigma_{ij}\mathbf{D}_j\lambda_j\rangle \tag{36}
$$

is nonnegative. This means for m arbitrary d-dimensional vectors y_i

$$
R = \sum_{i,j=1}^{m}\langle\sigma_{ij}y_j,y_i\rangle = \sum_{i,j=1}^{m}\sum_{\alpha,\beta=1}^{d}\sigma_{ij}^{\alpha\beta}y_j^{\alpha}y_i^{\beta} \geq 0\ .
$$

For some special choices of the matrices σ_{ij}, the positivity of the dissipation rate can be shown. We consider three of them:

1) Let σ_{ij} be positive semidefinite matrices, then

$$
\begin{aligned}
R(t) &= \sum_{i,j=1}^{m}\langle\sigma_{ij}\mathbf{D}_j\lambda_j,\mathbf{D}_i\lambda_i\rangle = \\
&= \sum_{i=1}^{m}\langle\sigma_{ii}\mathbf{D}_i\lambda_i,\mathbf{D}_i\lambda_i\rangle + \sum_{i<j=1}^{m}\Big\langle\sigma_{ij}(\mathbf{D}_i\lambda_i+\mathbf{D}_j\lambda_j),(\mathbf{D}_i\lambda_i+\mathbf{D}_j\lambda_j)\Big\rangle \geq 0\ .
\end{aligned}
$$

2) Let $\boldsymbol{\sigma}$ be a matrix of the type

$$
\boldsymbol{\sigma} = \begin{pmatrix}
a_1+a_{12}+...+a_{1m} & -a_{12} & ... & a_{1m} \\
-a_{12} & a_2+a_{12}+...+a_{2m} & ... & -a_{2m} \\
\vdots & \vdots & \ddots & \vdots \\
-a_{1m} & -a_{2m} & ... & a_m+a_{1m}+...+a_{m,m-1}
\end{pmatrix}, \tag{37}
$$

where a_i and a_{ij} are positive semidefinite $d\times d$ matrices. Then,

$$
\begin{aligned}
R(t) &= \sum_{i,j=1}^{m}\langle\sigma_{ij}\mathbf{D}_j\lambda_j,\mathbf{D}_i\lambda_i\rangle = \\
&= \sum_{i=1}^{m}\langle a_i\mathbf{D}_i\lambda_i,\mathbf{D}_i\lambda_i\rangle + \sum_{i,j=1}^{m}\langle a_{ij}\mathbf{D}_i\lambda_i,\mathbf{D}_i\lambda_i\rangle - \sum_{i,j=1}^{m}\langle a_{ij}\mathbf{D}_i\lambda_i,\mathbf{D}_j\lambda_j\rangle = \\
&= \sum_{i=1}^{m}\Big\langle a_i\mathbf{D}_i\lambda_i,\mathbf{D}_i\lambda_i\Big\rangle + \sum_{i<j=1}^{m}\Big\langle a_{ij}(\mathbf{D}_i\lambda_i-\mathbf{D}_j\lambda_j),(\mathbf{D}_i\lambda_i-\mathbf{D}_j\lambda_j)\Big\rangle \geq 0\ .
\end{aligned}
$$

3) Let $\sigma_{ij}^{\alpha\beta} = \mu^{\alpha\beta} H_{ij}$, where $\mathbf{H} = (H_{ij})$ is a positive semidefinite $m \times m$ matrix, and $\boldsymbol{\mu} = (\mu^{\alpha\beta})$ is a $d \times d$ matrix with positive semidefinite entries. Then we have

$$R = \sum_{i,j=1}^{m} \sum_{\alpha,\beta=1}^{d} \mu^{\alpha\beta} H_{ij} y_j^\alpha y_i^\beta = \sum_{\alpha,\beta=1}^{d} \mu^{\alpha\beta} \sum_{i,j=1}^{m} H_{ij} y_j^\alpha y_i^\beta = \sum_{\alpha,\beta=1}^{d} \mu^{\alpha\beta} \gamma^{\alpha\beta} \geq 0 \; .$$

As typical special case we choose $\mathbf{D}_i = \mathbf{G}_i \nabla$. $\boldsymbol{\sigma}$ is an $m \times m$ matrix of $d \times d$ matrices. The inverse Hessian is an $m \times m$ matrix. Thus, it is natural to set

$$\sigma_{ij} = \mu_{ij} \big[\partial^2 F(u)\big]_{ij}^{-1}, \; \mu_{ij} = \mu_{ji} \; ,$$

where μ_{ij} are suitable $d \times d$ matrices, and $\big[\partial^2 F(u)\big]_{ij}^{-1}$ are scalars. This leads to the system

$$\frac{\partial}{\partial t} u_j(x,t) \;\; = \;\; -\nabla \mathbf{G}_j^* J_j, \; j = 1, ..., m \; , \tag{38}$$

with the flows

$$J_j = -\sum_{i=1}^{m} \mu_{ji} \big[\partial^2 F(u)\big]_{ji}^{-1} \mathbf{G}_i \nabla \lambda_i \; . \tag{39}$$

The matrix $\boldsymbol{\mu}$ describes material laws and, therefore, we assume that $\boldsymbol{\mu}$ is not degenerated, i.e., $\mu_{ij} \geq \varepsilon > 0$. To get more information about the structure of $\boldsymbol{\mu}$, we assume sufficient smoothness and get from the state equation (33)

$$\lambda_i = \big[\partial F(u)\big]_i + q_i \psi_0 = \frac{\partial F(u)}{\partial u_i} + q_i \psi_0$$

and, therefore,

$$\nabla \lambda_i = \sum_{k=1}^{m} \frac{\partial^2 F(u)}{\partial u_i \partial u_k} \nabla u_k + q_i \nabla \psi_0 \; .$$

Putting all together, we get (using that $\big[\partial^2 F(u)\big]_{ik}$ are scalars)

$$\begin{aligned} \frac{\partial}{\partial t} u_j(x,t) \;\; &= \;\; \nabla \mathbf{G}_j^* \sum_{k=1}^{m} \sum_{i=1}^{m} \mu_{ji} \big[\partial^2 F(u)\big]_{ji}^{-1} \big[\partial^2 F(u)\big]_{ik} \mathbf{G}_i \nabla u_k + && (40) \\ &+ \;\; \nabla \mathbf{G}_j^* \sum_{i=1}^{m} \mu_{ji} \big[\partial^2 F(u)\big]_{ji}^{-1} \mathbf{G}_i \nabla \psi_i, \; j = 1, ..., m \; . && (41) \end{aligned}$$

We compare this expression for no interaction $\psi_i = 0$ and spatial isotropy $\mathbf{G}_i = \mathbf{I}$, with the linear system (7). This shows that

$$\sum_{i=1}^{m} \mu_{ji} \big[\partial^2 F(u)\big]_{ji}^{-1} \big[\partial^2 F(u)\big]_{ik}$$

has to be a diagonal matrix. This is the case if $\boldsymbol{\mu}$ and $\big[\partial^2 F(u)\big]_{ik}$ are diagonal matrices (see the following subsection), or μ_{ji} do not depend on j. Setting $\mu_j = \mu_{ji}$, we get

$$\sum_{i=1}^{m} \mu_j \big[\partial^2 F(u)\big]^{-1}_{ji} \big[\partial^2 F(u)\big]_{ik} = \mu_j \delta_{jk} \ .$$

From $\mu_{ji} = \mu_{ij}$ follows that $\boldsymbol{\mu}$ has to be a constant diagonal matrix with positive entries μ. This leads to

$$\sigma_{ij} = \mu \big[\partial^2 F(u)\big]^{-1}_{ij} \ .$$

Since (41) is an unbounded operator, too, if $\big[\partial^2 F(u)\big]^{-1}_{ji}$ is not diagonal, it has to degenerate. for $u \to 0$.

Finally, we get the following

Theorem 4 *An evolution system, describing drift and diffusion of a vector $u = (u_1, ..., u_m)$ of concentrations, according to a general free energy (28), has to have the structure*

$$\begin{aligned}
\frac{\partial}{\partial t} u_j(x,t) &= \nabla \mathbf{G}^*_j \mu \sum_{i=1}^{m} \big[\partial^2 F(u)\big]^{-1}_{ji} \mathbf{G}_i \nabla \lambda_i, \ j = 1, ..., m \ , \\
\psi &= \mathbf{K} u + \mathbf{K}_0 f \ , \\
u &= (\partial F)^{-1} \big(\lambda - \psi\big) \ ,
\end{aligned}$$

where the inverse Hessian has the asymptotic $\big[\partial^2 F(u)\big]^{-1}_{ji} \sim u_i$ for $u_i \to 0$, and μ is a $d \times d$ matrix with positive entries.

Let us remark that in the case of charged particles, the global interaction can be written as (32) with u_i instead of u^*_i.

The matrix $\boldsymbol{\mu}$ can contain anisotropic effects. We have to remark that $\boldsymbol{\mu}$ is the same for each species. This means it is not possible to consider different anisotropic effects for different species. Microscopic considerations like in [12] lead to the same results. The diagonal matrix $\boldsymbol{\mu}$ can contain different diagonal elements only if the Hessian is diagonal. We consider this case in the following subsection.

4.1 Many diffusive non-interacting species

As a special case, let us assume that the free energy of m species is the sum of m one-species free energies

$$\Phi(u) = \sum_{i=1}^{m} \int_\Omega F_i \left(\frac{u_i(x,t)}{u^*_i(x)} \right) u^*_i(x) dx + \frac{1}{2} \langle u, \mathbf{K} u \rangle + \langle u, \mathbf{K}_0 f \rangle \ , \tag{42}$$

with some reference concentrations u_i^* (equilibrium concentrations) given. The F_i are assumed to be convex functions. We have $\lambda_i = \partial F_i(1)$ in the equilibrium and non-interacting case.

Introducing the interaction potential ψ and similar to the one species case, we get m state equations

$$\begin{aligned} u_i(x,t) &= \partial F_i^{-1}\big(\lambda_i(x,t) + q_i\psi_0(x,t)\big)u_i^*(x) \iff \\ \lambda_i(x,t) &= \partial F_i\left(\frac{u_i^*(x)}{u_i(x,t)}\right) - q_i\psi_0(x,t),\ i = 1,...,m\ , \end{aligned}$$

connecting the potential and the concentration.

The Hessian

$$\partial^2 F(u) = \mathrm{diag}\left(\partial^2 F_1\left(\frac{u_1}{u_1^*}\right),...,\partial^2 F_m\left(\frac{u_m}{u_m^*}\right)\right)$$

is a diagonal matrix and so is its inverse. Setting $\mathbf{D}_i = \mathbf{G}_i\nabla$ and

$$\boldsymbol{\sigma} = \mathrm{diag}\left(\mu_1 u_1^*\left[\partial^2 F_1\left(\frac{u_1}{u_1^*}\right)\right]^{-1},...,\mu_m u_m^*\left[\partial^2 F_m\left(\frac{u_m}{u_m^*}\right)\right]^{-1}\right)\ , \tag{43}$$

we get the following result.

Theorem 5 *An evolution system, describing drift and diffusion of a vector $u = (u_1,...,u_m)$ of concentrations, according to a general free energy (42), has to have the structure*

$$\begin{aligned} \frac{\partial}{\partial t}u_j(x,t) &= \nabla\mathbf{G}_j^*\mu_j u_j^*\left[\partial^2 F_j\left(\frac{u_j}{u_j^*}\right)\right]^{-1}\mathbf{G}_j\nabla\lambda_j,\ j = 1,...,m\ , \\ \psi &= \mathbf{K}u + \mathbf{K}_0 f\ , \\ u &= (\partial F)^{-1}\big(\lambda - \psi\big)\ , \end{aligned}$$

where μ_j are positive semidefinite $d \times d$ matrices.

Let us remark that, if no global self-consistent interaction is present, these are m uncoupled equations.

4.2 Many non-interacting species and pseudo-species

We have assumed so far that the chemical part of the free energy $\Phi_C(u)$ is a sum of terms, each of which is the free energy of one kind of particle. This, of course, does not describe cross interaction, i.e., interaction between particles of several types. Such interaction can be a drift, forced by a self-consistent mean field, reactions, and also diffusion cross effects. In this case, the free energy can be generated by a general convex function of m variables. A special case of such effects is the situation in which one can understand the interaction as the result of some pseudo-species that one has to take into consideration but which one does not want to model.

This is the case if the free energy contains no cross effects, but the species are not independent of each other. Let us consider $m+1$ species $u = (u_0, u_1, ..., u_m)$ with the free energy

$$\Phi_C(u) = \frac{1}{\alpha_0}\int_\Omega F_0\left(\frac{u_0(x,t)}{u_0^*(x)}\right)u_0^*(x)dx + \sum_{i=1}^m \frac{1}{\alpha_i}\int_\Omega F_i\left(\frac{u_i(x,t)}{u_i^*(x)}\right)u_i^*(x)dx\ ,$$

and the 0-th species is a function of the others $u_0 = H(u_1, ..., u_m)$. We will consider the case that this dependence is linear, i.e.

$$u_0 = \beta - \alpha_0 \sum_{i=1}^m \frac{1}{\alpha_i} u_i\ , \tag{44}$$

with coefficients $\alpha_i > 0$ and $\beta > 0$. We can consider two different cases, depending on the sign of α_0.

For $\alpha_0 > 0$, we have, from (44),

$$\frac{\beta}{\alpha_0} = \sum_{i=0}^m \frac{1}{\alpha_i} u_i\ .$$

This occurs for particles occupying $\frac{\beta}{\alpha_0}$ places and α_i particles of kind u_i occupying one place. If we know the concentration of m species, the concentration of the 0-th species is given. Here, u_0 can be the concentration of a real particle or the concentration of vacancies. This is the typical situation one encounters in phase separation problems (see, e.g., [14]).

For $\alpha_0 < 0$, we get, from (44),

$$u_0 = \beta + |\alpha_0| \sum_{i=1}^m \frac{1}{\alpha_i} u_i\ .$$

The particles in this setting do not behave independently of each other, but the dependence can be understood as the formation of some pseudo-particle. This pseudo-particle has to be taken care of, but we are not interested in its evolution. A case in point is the modeling of electron-hole scattering in semiconductor theory (see, e.g., [22]).

For simplicity, we only explore the case with Boltzmann statistics $F_i(y) = F(y) = y\log y$. The general case is similar, but we cannot describe it explicitly, and therefore, the real structure of the object is not transparent. Moreover, the author is not aware of papers dealing with cross interaction in the general case.

Thus, let us consider the free energy

$$\begin{aligned}\Phi_C(u) &= \int_\Omega F(u_1, ..., u_m)dx = \\ &= \frac{1}{\alpha_0}\int_\Omega u_0(x,t)\log\frac{u_0(x,t)}{u_0^*(x)}dx + \sum_{i=1}^m \frac{1}{\alpha_i}\int_\Omega u_i(x,t)\log\frac{u_i(x,t)}{u_i^*(x)}dx \quad (45)\end{aligned}$$

and the dependence (44). We have to investigate whether this function is convex, and what is the structure of the evolution system.

The i-th component of the state equation $\partial F(u) = \lambda$ reads (For simplicity, we consider the no-interaction case. In general, λ has to be replaced by $\lambda - \psi$):

$$\frac{\partial F}{\partial u_i} = \frac{1}{\alpha_i} \log \frac{u_i}{u_i^*} - \frac{1}{\alpha_i} \log \frac{u_0}{u_0^*} = \lambda_i \ ,$$

or, using (44) and setting $k_i = \frac{u_i^*}{u_0^*} e^{\alpha_i \lambda_i}$,

$$u_i = k_i u_0 = k_i \beta - k_i \alpha_0 \sum_{i=1}^{m} \frac{1}{\alpha_i} u_i \ .$$

This is equivalent to the system ($\mathbf{I}$ is the identity matrix)

$$\left[\mathbf{I} + \alpha_0 \begin{pmatrix} k_1 \\ \vdots \\ k_m \end{pmatrix} \left(\frac{1}{\alpha_1}, ..., \frac{1}{\alpha_m} \right) \right] \begin{pmatrix} u_1 \\ \vdots \\ u_m \end{pmatrix} = \beta \begin{pmatrix} k_1 \\ \vdots \\ k_m \end{pmatrix} \ .$$

This is the state equation, connecting u with λ (via k_i). This linear system has a matrix that is the sum of the identity and a rank-1 matrix and therefore can be easily solved. One gets

$$u_i = \frac{\beta u_i^* e^{\alpha_i \lambda_i}}{u_0^* + \alpha_0 \sum\limits_{i=1}^{m} \frac{1}{\alpha_i} u_i^* e^{\alpha_i \lambda_i}} \ .$$

Theorem 6 *The free energy density*

$$F(u_1, ..., u_m) dx = \frac{1}{\alpha_0} u_0(x,t) \log \frac{u_0(x,t)}{u_0^*(x)} + \sum_{i=1}^{m} \frac{1}{\alpha_i} u_i(x,t) \log \frac{u_i(x,t)}{u_i^*(x)}$$

with (44) is convex for any $\alpha_0 \neq 0$.

Proof: For the Hessian we get

$$\partial^2 F(u) = \begin{pmatrix} \frac{1}{\alpha_1 u_1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \frac{1}{\alpha_m u_m} \end{pmatrix} + \frac{\alpha_0}{u_0} \begin{pmatrix} \frac{1}{\alpha_1} \\ \vdots \\ \frac{1}{\alpha_m} \end{pmatrix} \left(\frac{1}{\alpha_1}, ..., \frac{1}{\alpha_m} \right) \ .$$

Let $x = (x_1, ..., x_m) \in \mathbb{R}^m$ be arbitrary. We have

$$\begin{aligned}
\langle \partial^2 F(u)x, x \rangle &= \sum_{i=1}^m \frac{x_i^2}{\alpha_i u_i} + \frac{\alpha_0}{u_0} \left(\sum_{i=1}^m \frac{x_i}{\alpha_i} \right)^2 = \\
&= \frac{1}{u_0} \left(u_0 \sum_{i=1}^m \frac{x_i^2}{\alpha_i u_i} + \alpha_0 \left(\sum_{i=1}^m \frac{x_i}{\alpha_i} \right)^2 \right) = \qquad (46) \\
&= \frac{1}{u_0} \left(\beta \sum_{i=1}^m \frac{x_i^2}{\alpha_i u_i} - \alpha_0 \sum_{i,j=1}^m \frac{x_i^2}{\alpha_i \alpha_j} \frac{u_j}{u_i} + \alpha_0 \sum_{i,j=1}^m \frac{x_i x_j}{\alpha_i \alpha_j} \right) = \\
&= \frac{1}{u_0} \left(\beta \sum_{i=1}^m \frac{x_i^2}{\alpha_i u_i} - \alpha_0 \sum_{i<j=1}^m \frac{1}{\alpha_i \alpha_j} \left(x_i \sqrt{\frac{u_j}{u_i}} - x_j \sqrt{\frac{u_i}{u_j}} \right)^2 \right) . \qquad (47)
\end{aligned}$$

This bilinear form is positive for $\alpha_0 > 0$ because of (46) and for $\alpha_0 < 0$ because of (47). Hence, $F(u)$ is convex for any $\alpha_0 \neq 0$. ■

Rewriting the evolution system, we have to calculate $\left[\partial^2 F(u)\right]^{-1}$. We have

$$\det \partial^2 F(u) = \frac{\beta}{\alpha_1 \cdots \alpha_m} \frac{1}{u_0 u_1 \cdots u_m} \geq 0 .$$

Solving for fixed u the equation $\partial^2 F(u)x = y$, we get the inverse Hessian

$$\left[\partial^2 F(u)\right]^{-1} = \begin{pmatrix} \alpha_1 u_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \alpha_m u_m \end{pmatrix} - \frac{\alpha_0}{\beta} \begin{pmatrix} u_1 \\ \vdots \\ u_m \end{pmatrix} (u_1, ..., u_m) . \qquad (48)$$

Theorem 7 *An evolution system, describing drift and diffusion of a vector $u = (u_0, u_1, ..., u_m)$ of concentrations, depending on each other via (44) and according to a free energy with a chemical part (45), has to have the structure*

$$\begin{aligned}
\frac{\partial}{\partial t} u_j(x,t) &= \nabla \mathbf{G}_j^* \mu \alpha_j u_j \mathbf{G}_j \nabla \lambda_j + \nabla \mathbf{G}_j^* \mu \frac{\alpha_0}{\beta} u_j \sum_{i=1}^m u_i \mathbf{G}_i \nabla \lambda_i \ , \ j = 1, ..., m \ , \\
\psi &= \mathbf{K}u + \mathbf{K}_0 f \ , \\
u &= (\partial F)^{-1} (\lambda - \psi) \ ,
\end{aligned}$$

where μ is a $d \times d$ matrix with positive entries.

Proof: The proof is obtained by putting the inverse Hessian (48) into the system (38)–(39). ■

To illustrate this system, we consider two examples, setting for simplicity $\mathbf{G}_i = \mathbf{I}$.

Example 1): Setting $\alpha_i = 1$, $\beta = 1$, and $\boldsymbol{\mu} = \mathbf{I}$, we have for $m = 2$ a system modeling phase separation:

$$\frac{\partial}{\partial t} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \nabla \begin{pmatrix} u_1(1-u_1) & -u_1 u_2 \\ -u_1 u_2 & u_2(1-u_2) \end{pmatrix} \begin{pmatrix} \nabla \lambda_1 \\ \nabla \lambda_2 \end{pmatrix} .$$

This system is considered in [14].

Example 2): Setting $\alpha_0 = -\alpha_1 \cdots \alpha_m$, $\beta = 1$, and $\boldsymbol{\mu} = \frac{1}{u_0}\mathbf{I}$, we get for $m = 2$ a system modeling electron-hole scattering:

$$\frac{\partial}{\partial t}\begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \nabla \frac{1}{1+\alpha_2 u_1 + \alpha_1 u_2} \begin{pmatrix} \alpha_1 u_1(1+\alpha_2 u_1) & \alpha_1 u_1 \alpha_2 u_2 \\ \alpha_1 u_1 \alpha_2 u_2 & \alpha_2 u_2 (1+\alpha_1 u_2) \end{pmatrix} \begin{pmatrix} \nabla\lambda_1 \\ \nabla\lambda_2 \end{pmatrix} .$$

This system was derived in [18, 15, 22]. Our derivation shows that electron-hole scattering can be understood as an independent diffusion of electrons, holes, and some pseudo-particle. Analytical properties of this system were considered in [4].

5 Drift-diffusion systems with a temperature field

All material properties of drift–diffusion systems depend on temperature. Thus, quantities such as conductivities depend on temperature in parameter-like way. All processes that lead to drift or diffusion are connected to the transfer of energy and, in turn, with local heat production and change of temperature. As usual, we assume that the production of heat is small or dissipates out of the interesting domain sufficiently fast. In many typical situations, however, this is not true. Moreover, the transport of heat—especially in heterostructures—is so slow that whilst other processes can be assumed already in an equilibrium, the heat transfer cannot. It is therefore necessary to consider a spatial change of temperature—a temperature field.

Temperature is an equilibrium value, so this point of view implicitly assumes that we can regard small spatial regions as being in equilibrium to define the temperature.

This understanding of temperature contradicts the methods presented in the first sections for the derivation of evolution equations for non-equilibrium situations. As a consequence, it appears to be impossible to derive an energy model by varying a global free energy functional. We therefore derive an evolution system in a purely phenomenological way, assuming that temperature is a local parameter, and that we can vary a suitable functional (this will be the entropy) in a local way. This precludes the derivation of global drift terms, generated by a non-local operator $\mathbf{K}$. To still be able to study drift, we allow global drift terms only when the inverse $\mathbf{K}_1^{-1}$ can be understood in a local way. This is the case, for instance, if $\mathbf{K}$ describes electrostatic interaction with homogeneous Neumann boundary conditions.

A better derivation of models involving temperature has to consider the true state of the (classical) particles—spatial coordinates and momentum or velocity. The change in velocity can be a measure of the heat production, understanding heat as the loss of kinetic energy—the velocity variation of a probability distribution in phase space.

The derivation of evolution equations for concentrations while taking into account the probability distribution in phase space is a hard problem, in general. In [20], this is done in a rigorous way for general Brownian motion. The results there show that even in simple cases the type of the resulting evolution equations can change in time.

Although the derivation of the equations is a purely phenomenological one, the presented equations were successfully used in numerical applications [3]. Therefore, the methods appear to be useful, regardless of the difficulties of their rigorous foundation.

5.1 Thermodynamic properties and equilibrium state

We consider global self-consistent electrostatic interaction, assuming that the particles $\mathbf{X}_i$ are charged with charge q_i. Let f be a given and fixed distribution of charges, then the electrostatic interaction is described by Poisson's equation

$$\mathbf{K}^{-1}\psi_0 = -\nabla \cdot (\epsilon \nabla \psi_0) = f + q = f + \sum_{i=1}^{m} q_i u_i \ ,$$

where ϵ is the dielectric permittivity. This equation can be understood locally.

Assuming the temperature to be a local equilibrium quantity, as mentioned we have to consider global (capital letters) and local (small letters) quantities:

$$\begin{aligned} S(t) &= \int_\Omega s(x,t)\ dx \quad \text{(entropy)}, \\ E(t) &= \int_\Omega u_0(x,t)\ dx \quad \text{(total energy)}, \\ \Phi(t) &= \int_\Omega h(x,t)\ dx \quad \text{(free energy)}. \end{aligned}$$

From thermodynamics we use the following local relations between entropy and energies:

$$s = -h' \ , \tag{49}$$

$$u_0 = h + Ts \ . \tag{50}$$

Here and from now on, $g'(T)$ means the derivative of some function $g(T)$ with respect to the temperature as a parameter. ∂_i denotes the derivative with respect to u_i as a parameter.

The equations (49) and (50) are equilibrium relations, but we assume that they hold in every point.

Let $c(T)$ be a suitable function describing the heat capacity (e.g., $c(T) = c_L T(\log T - 1)$, where c_L is the heat capacity of a crystal lattice). We postulate the free energy in the form

$$\begin{aligned} \Phi(u,T) &= -\int_\Omega c(T)dx + \sum_{i=1}^{m} \int_\Omega F_i(u_i,T)dx + \\ &+ \frac{1}{2}\int_\Omega u(x)(\mathbf{K}u)(x)dx + \int_\Omega u(x)(\mathbf{K}_0 f)(x)dx \end{aligned}$$

(here $\mathbf{K}$ and $\mathbf{K}_0$ are the same operators as in the previous section). This definition uses extensive (concentrations) and intensive (temperature) values as independent variables of the free energy. This seems to be inconsistent. For technical reasons, we set $F_i(u_i, T) = T u_i f_i(u_i, T)$ and assume that $F_i(\cdot, T)$ is convex. Our definition assumes that the temperature is not a thermodynamic value but a parameter (with unknown nature) depending on x and t. For a more correct definition, a right understanding of the temperature is necessary. For the same reason, F_i is not assumed to be convex as a function of T.

We assume that $\Phi(u, T)$ can be represented as an integral of some locally defined function

$$\Phi(u, T) = \int_\Omega h(u, T) dx \ ,$$

where we define

$$h(u, T) = -c(T) + T \sum_{i=1}^{m} u_i f_i(u_i, T) + \frac{\epsilon}{2} \left|\nabla \psi_0\right|^2 \ .$$

Here, we used the local representation of $\mathbf{K}_1^{-1}$ from

$$\frac{1}{2}\langle \mathbf{K}u, u\rangle + \langle \mathbf{K}_0 f, u\rangle = \frac{1}{2}\langle \mathbf{K}_1^{-1}\psi_0, \psi_0\rangle \ .$$

Furthermore, we assume conservation of mass

$$\int_\Omega u_i(x, t) dx = \int_\Omega u_i(x, 0) dx = 1, \ \ i = 1, ..., m \ ,$$

and conservation of total energy

$$E(t) = E(0) \ .$$

Looking for the maximum of the entropy, we consider the Lagrangian

$$L(T, u) \ = \ S + \sum_{i=1}^{m} \lambda_i \left(\int_\Omega u_i(x) dx - 1 \right) + \theta(E - E_0) \ .$$

Dealing with local quantities, we assume that the variation of $L(T, u)$ with respect to some function is the partial derivative of the local function

$$l(T, u) = s + \sum_{i=1}^{m} \lambda_i u_i + \theta u_0 \ .$$

For the interaction term we use the following local property, coming from Poisson's equation

$$\frac{\partial}{\partial u_i} \frac{\epsilon}{2} \left|\nabla \psi_0\right|^2 = \frac{\partial}{\partial u_i} \psi_0 \left(f + \sum_{i=1}^{m} q_i u_i \right) = q_i \psi_0 \ .$$

To calculate s and u_0 we use equations (49) and (50). For the entropy density, we get

$$s = -h' = c'(T) - \sum_{i=1}^{m} u_i\Big(f_i(u_i,T) + Tf_i'(u_i,T)\Big) = c'(T) - \sum_{i=1}^{m} u_i P_i \ .$$

Here, we defined the specific entropy per particle

$$P_i = f_i(u_i,T) + Tf_i'(u_i,T) \ . \tag{51}$$

From equation (50), we get a representation for the total energy.

$$u_0 = h + Ts = c'(T)T - c(T) - T^2 \sum_{i=1}^{m} u_i f_i'(u_i,T) + \frac{\epsilon}{2}\big|\nabla\psi_0\big|^2 \ .$$

Varying the Lagrange functional, we get the Euler–Lagrange equations for the equilibrium quantities T^* and u_i^*

$$\frac{\delta L}{\delta T} = \frac{\partial}{\partial T} l = s' + \theta h' + \theta s + \theta T^* s' = s'\Big(1 + \theta T^*\Big) = 0 \ .$$

From this equation we get a condition for the equilibrium temperature T^*

$$T^* = -\frac{1}{\theta} \ . \tag{52}$$

As expected, the temperature has to be constant in the equilibrium.

For the concentrations we get, using (52),

$$\frac{\delta L}{\delta u_i} = \frac{\partial}{\partial u_i} l = \lambda_i - q_i \frac{\psi_0}{T^*} - f_i(u_i^*,T^*) - u_i^* \partial_i f_i(u_i^*,T^*) = 0 \ .$$

Hence,

$$\lambda_i - q_i \frac{\psi_0}{T^*} = f_i(u_i^*,T^*) + u_i^* \partial_i f_i(u_i^*,T^*) \tag{53}$$

is the state equation. The connection between u_i^* and λ_i is uniquely defined because the right-hand side of this equation is the monotone function $\partial F_i(\cdot,T^*)$.

5.2 Evolution equations and dissipation rate

We understand the Lagrange multipliers as thermodynamic potentials, being nonconstant in space, time, and for the species in the non-equilibrium case

$$\theta \to \theta(x,t), \ \lambda_1 \to \lambda_1(x,t), \ \ldots, \ \lambda_m \to \lambda_m(x,t) \ .$$

For formal reasons we set

$$\lambda_0(x,t) = \theta(x,t) = -\frac{1}{T} \ . \tag{54}$$

Now we understand equation

$$\lambda_i - q_i \frac{\psi_0}{T} = f_i(u_i, T) + u_i \partial_i f_i(u_i, T) \tag{55}$$

as the state equation, connecting potentials and concentrations in the non-equilibrium case, too.

Sometimes—especially in the semiconductor literature—electrochemical potentials or quasi-Fermi levels φ_i are introduced. In our notation, the potentials

$$\lambda_i(x,t) = \frac{\varphi_i(x,t)}{T(x,t)}$$

are the electrochemical potentials divided by the temperature. Using the inverse function of the free energy, this leads to the well-known definition of the connection between concentration and potential by the statistics

$$u_i = \partial F_i^{-1} \left(\frac{\varphi_i - q_i \psi_0}{T} \right) .$$

Now we are going to postulate evolution equations for the wanted functions. As usual, dealing with a temperature field, a heat equation is derived. It seems to be more natural to consider an energy transport equation instead of a heat equation because u_0 is an extensive value like the concentrations, whereas T is an intensive value. It it easy to derive a heat equation from an energy transport equation (see [3]).

We postulate the evolution system in the form

$$\begin{aligned} \dot{u}_0 - \nabla \cdot \left(\psi_0 \epsilon \nabla \dot{\psi}_0 \right) &= \mathbf{D}_0^* J_0 \ , \\ \dot{u}_1 &= \mathbf{D}_1^* J_1 \ , \\ &\ \ \ldots \\ \dot{u}_m &= \mathbf{D}_m^* J_m \ . \end{aligned}$$

This is the same system as in the constant-temperature case instead of the first equation. The electrostatic term describes the evolution of the energy induced by the electrostatic field. We define the flows setting

$$\begin{pmatrix} J_0 \\ J_1 \\ \vdots \\ J_m \end{pmatrix} = - \begin{pmatrix} \sigma_{00} & \sigma_{01} & \ldots & \sigma_{0m} \\ \sigma_{10} & \sigma_{11} & \ldots & \sigma_{1m} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{m0} & \sigma_{m1} & \ldots & \sigma_{mm} \end{pmatrix} \cdot \begin{pmatrix} \mathbf{D}_0 \lambda_0 \\ \mathbf{D}_1 \lambda_1 \\ \vdots \\ \mathbf{D}_m \lambda_m \end{pmatrix} .$$

Taking into account Onsager's principle, we assume $\sigma_{ij} = \sigma_{ji}$. Thus, we have the following representations for the flows

$$J_i = -\sum_{j=1}^{m} \sigma_{ij} \mathbf{D}_j \lambda_j + \sigma_{i0} \mathbf{D}_0 \frac{1}{T} \ , \tag{56}$$

$$J_0 = -\sum_{j=1}^{m} \sigma_{0j} \mathbf{D}_j \lambda_j + \sigma_{00} \mathbf{D}_0 \frac{1}{T} \ . \tag{57}$$

To prove the second law of thermodynamics, we have to show that the entropy increases in time. Let us assume that the evolution of the entropy is governed by some equation

$$\dot{s}(t) = \mathbf{D}_s^* J_s \ , \tag{58}$$

where J_s is the entropy flow, and $\mathbf{D}^*$ is some operator, which is so far unknown. We have to show

$$\frac{d}{dt}S(t) \quad = \quad \int_\Omega \dot{s}(t)dx = \langle 1, \mathbf{D}_s^* J_s \rangle \geq 0 \ . \tag{59}$$

To calculate this quantity, we have to find a connection of the time derivative $\dot{s}(t)$ to the time derivatives in the evolution equations.

For this purpose we derive some formulae, useful in the future. Differentiating the state equation with respect to the temperature, we obtain

$$\frac{\partial}{\partial T}\left(\lambda_i(x,t) - q_i \frac{\psi_0}{T}\right) = \frac{\partial}{\partial T}\Big(f_i(u_i,T) + u_i \partial_i f_i(u_i,T)\Big) \ ,$$

$$q_i \frac{\psi_0}{T^2} = f_i'(u_i,T) + u_i \partial_i f_i'(u_i,T) \ .$$

Hence,

$$q_i \psi_0 \dot{u}_i = T^2 \dot{u}_i (f_i'(u_i,T) + u_i \partial_i f_i'(u_i,T)) \ . \tag{60}$$

Differentiating by time, from Poisson's equation follows

$$-\nabla \cdot (\epsilon \nabla \dot{\psi}_0) = \sum_{i=1}^m q_i \dot{u}_i \ ,$$

$$-\sum_{i=1}^m q_i \dot{u}_i \psi_0 = \psi_0 \Big(\nabla \cdot (\epsilon \nabla \dot{\psi}_0)\Big) \ . \tag{61}$$

Using the relations (60) and (61), we calculate the time derivative of the total energy:

$$
\begin{aligned}
\dot{u}_0 &= \epsilon\nabla\psi_0\cdot\nabla\dot{\psi}_0 + c''(T)T\dot{T} - 2T\dot{T}\sum_{i=1}^{m}u_i f_i'(u_i,T) - T^2\sum_{i=1}^{m}\dot{u}_i f_i'(u_i,T) - \\
&- T^2\sum_{i=1}^{m}u_i\Big(f_i''(u_i,T)\dot{T} + \partial_i f_i'(u_i,T)\dot{u}_i\Big) = \\
&= \epsilon\nabla\psi_0\cdot\nabla\dot{\psi}_0 + c''(T)T\dot{T} - \dot{T}T\sum_{i=1}^{m}u_i(2f_i'(u_i,T) + Tf_i''(u_i,T)) - \\
&- T^2\sum_{i=1}^{m}\dot{u}_i\Big(f_i'(u_i,T) + u_i\partial_i f_i'(u_i,T)\Big) = \\
&\overset{(60)}{=} \epsilon\nabla\psi_0\cdot\nabla\dot{\psi}_0 + c''(T)T\dot{T} - \\
&- \dot{T}T\sum_{i=1}^{m}u_i(2f_i'(u_i,T) + Tf_i''(u_i,T)) - \sum_{i=1}^{m}q_i\dot{u}_i\psi_0 = \\
&\overset{(61)}{=} \epsilon\nabla\psi_0\cdot\nabla\dot{\psi}_0 + c''(T)T\dot{T} - \\
&- \dot{T}T\sum_{i=1}^{m}u_i(2f_i'(u_i,T) + Tf_i''(u_i,T)) + \psi_0\Big(\nabla\cdot(\epsilon\nabla\dot{\psi}_0)\Big) = \\
&= \nabla\cdot\Big(\psi_0\epsilon\nabla\dot{\psi}_0\Big) + c_w\dot{T}
\end{aligned}
$$

with

$$
c_w = T\Big(c''(T) - \sum_{i=1}^{m}u_i(2f_i'(u_i,T) + Tf_i''(u_i,T))\Big)\ . \tag{62}
$$

Now, using the relations (60), (61), and (62), we can calculate the time derivative of the entropy and get

$$
\begin{aligned}
\dot{s} \quad &= \quad c''(T)\dot{T} - \sum_{i=1}^{m} \dot{u}_i \Big(f_i(u_i,T) + T f_i'(u_i,T) \Big) - \\
&- \quad \sum_{i=1}^{m} u_i \Big(f_i'(u_i,T)\dot{T} + \partial_i f_i'(u_i,T)\dot{u}_i + \dot{T} f_i'(u_i,T) + T\dot{T} f_i''(u_i,T) + \\
&+ \quad T\partial_i f_i'(u_i,T)\dot{u}_i \Big) = \dot{T}\left(c''(T) - \sum_{i=1}^{m} u_i \Big(2f_i'(u_i,T) + T f_i''(u_i,T) \Big) \right) - \\
&- \quad \sum_{i=1}^{m} \dot{u}_i \Big(f_i(u_i,T) + u_i\partial_i f_i(u_i,T) + T f_i'(u_i,T) + T u_i \partial_i f_i'(u_i,T) \Big) = \\
&\overset{(62)}{=} \quad c_w(T)\frac{\dot{T}}{T} - \sum_{i=1}^{m} \dot{u}_i \Big(f_i(u_i,T) + u_i\partial_i f_i(u_i,T) + T f_i'(u_i,T) + T u_i \partial_i f_i'(u_i,T) \Big) = \\
&\overset{(55)}{=} \quad c_w(T)\frac{\dot{T}}{T} - \sum_{i=1}^{m} \dot{u}_i \left(-q_i \frac{\psi_0}{T} + \lambda_i + T f_i'(u_i,T) + T u_i \partial_i f_i'(u_i,T) \right) = \\
&\overset{(61)}{=} \quad c_w(T)\frac{\dot{T}}{T} - \sum_{i=1}^{m} \dot{u}_i \left(-q_i \frac{\psi_0}{T} + \lambda_i + q_i \frac{\psi_0}{T} \right) = c_w(T)\frac{\dot{T}}{T} - \sum_{i=1}^{m} \dot{u}_i \lambda_i \ .
\end{aligned}
$$

Thus, we obtain a connection between the time derivatives of u_i, s, and ψ_0:

$$
\dot{s} = \frac{1}{T}\dot{u}_0 - \sum_{i=1}^{m} \dot{u}_i \lambda_i - \frac{1}{T} \nabla \cdot \left(\psi_0 \epsilon \nabla \dot{\psi}_0 \right) \ . \tag{63}
$$

To prove the validity of the second law of thermodynamics (59), we use (63) and the definition (54) of λ_0 to obtain

$$
\begin{aligned}
\frac{d}{dt} S(t) \quad &= \quad \int_\Omega \frac{1}{T} \left(\dot{u}_0 - \nabla \cdot \left(\psi_0 \epsilon \nabla \dot{\psi}_0 \right) \right) dx - \sum_{i=1}^{m} \int_\Omega \dot{u}_i \lambda_i dx = \\
&= \quad \langle \frac{1}{T}, \mathbf{D}_0^* J_0 \rangle - \sum_{i=1}^{m} \langle \lambda_i, \mathbf{D}_i^* J_i \rangle = - \sum_{i=0}^{m} \langle \lambda_i, \mathbf{D}_i^* J_i \rangle = \\
&= \quad - \sum_{i=0}^{m} \langle \mathbf{D}_i \lambda_i, J_i \rangle = \sum_{i,j=0}^{m} \langle \sigma_{ij} \mathbf{D}_j \lambda_j, \mathbf{D}_i \lambda_i \rangle \ .
\end{aligned} \tag{64}
$$

To show that the dissipation rate

$$
R(t) \quad = \quad \sum_{i,j=0}^{m} \langle \sigma_{ij} \mathbf{D}_j \lambda_j, \mathbf{D}_i \lambda_i \rangle \tag{65}
$$

is nonnegative, we have to investigate the natural properties of the matrix $(\sigma_{ij})_{i,j=0}^m$. In the case of constant temperature, the positivity of the dissipation rate is a consequence of the positivity of the inverse Hessian. This follows from the convexity of the chemical part of the free energy. In the case with temperature field, we assumed that T is a parameter in the material laws, and it is not clear whether $F(u,T)$ is still a convex function. Thus, we can only use the connection of the matrix $(\sigma_{ij})_{i,j=1}^m$ with a positive defined inverse Hessian. Therefore, we have to use some further assumptions for the matrices σ_{0j} and σ_{00}.

Such an assumption we obtain investigating the entropy flow. Comparing equation (64) with (58), it is natural to understand

$$J_s = \frac{1}{T}J_0 + \sum_{i=1}^m \lambda_i J_i \tag{66}$$

as the entropy flow. On the other side, the entropy flow is the sum of entropy flows, generated by particle and heat transport. Therefore, the entropy flow has to be

$$J_s = \sum_{i=1}^m P_i J_i - \kappa_L T \mathbf{D}_0 \frac{1}{T} \ , \tag{67}$$

where κ_L is the heat conductivity of the material. The first term is the entropy flow, induced by the particles, where P_i is the specific entropy per particle introduced by (51). The last term is a generalization of the well-known temperature-induced entropy flow $-\kappa_L\frac{\nabla T}{T}$. If we have no anisotropy, i.e. $\mathbf{D}_0 = -\nabla$, the quantities coincide because of $(\nabla T)/T = -T\nabla(1/T)$.

Combining (66) and (67), and using the representations of the flows (56) and (57), we get

$$\begin{aligned}
0 &= \sum_{i=1}^m (P_i-\lambda_i)J_i - \frac{1}{T}J_0 - \kappa_L T\mathbf{D}_0\frac{1}{T} = \\
&= \sum_{i=1}^m (P_i-\lambda_i)J_i - \frac{1}{T}\left(\sum_{j=1}^m \sigma_{0j}\mathbf{D}_j\lambda_j - \sigma_{00}\mathbf{D}_0\frac{1}{T}\right) = -\kappa_L T\mathbf{D}_0\frac{1}{T} + \\
&+ \sum_{i=1}^m (P_i-\lambda_i)\left(\sum_{j=1}^m \sigma_{ij}\mathbf{D}_j\lambda_j - \sigma_{i0}\mathbf{D}_0\frac{1}{T}\right) - \frac{1}{T}\left(\sum_{j=1}^m \sigma_{0j}\mathbf{D}_j\lambda_j - \sigma_{00}\mathbf{D}_0\frac{1}{T}\right) = \\
&= \sum_{j=1}^m\left(\sum_{i=1}^m \sigma_{ij}(P_i-\lambda_i) - \frac{1}{T}\sigma_{0j}\right)\mathbf{D}_j\lambda_j + \qquad (68)\\
&+ \left(\sigma_{00}\frac{1}{T} - \sum_{i=1}^m (P_i-\lambda_i)\sigma_{i0} - \kappa_L T\right)\mathbf{D}_0\frac{1}{T} \ . \qquad (69)
\end{aligned}$$

This has to be an identity, not depending on the values $\mathbf{D}_j\lambda_j$ or $\mathbf{D}_0\frac{1}{T}$. Therefore, the matrix elements are no longer arbitrary, we have, from (68) for σ_{0i},

$$\sigma_{0j} \ = \ T\sum_{i=1}^m (P_i-\lambda_i)\sigma_{ij} \tag{70}$$

and hence, from (69) for σ_{00},

$$\sigma_{00} = T\sum_{i=1}^{m}(P_i-\lambda_i)\sigma_{i0}+\kappa_L T^2 = T^2\sum_{i,j=1}^{m}(P_i-\lambda_i)(P_j-\lambda_j)\sigma_{ij}+\kappa_L T^2 \ . \quad (71)$$

This is a transformation from the quantities $\sigma_{00}, \sigma_{01}, ..., \sigma_{0m}$ to other quantities $\kappa_L, P_1, ...P_m$. We will see that this transformation leads to a nonnegative dissipation rate. Using these expressions, we get for the dissipation rate (65)

$$\begin{aligned} R(t) &= \left\langle \kappa_L T\mathbf{D}_0\frac{1}{T}, T\mathbf{D}_0\frac{1}{T}\right\rangle + \\ &+ \sum_{i,j=1}^{m}\left\langle \sigma_{ij}\left[T(P_i-\lambda_i)\mathbf{D}_0\frac{1}{T}-\mathbf{D}_i\lambda_i\right], \left[T(P_j-\lambda_j)\mathbf{D}_0\frac{1}{T}-\mathbf{D}_j\lambda_j\right]\right\rangle \ . \end{aligned}$$

This quantity is nonnegative if the matrices σ_{ij}, $i, j = 1, ..., m$ are positive semidefinite, and $\kappa_L \geq 0$.

6 Conclusions

Here, we summarize some principles for the derivation of drift-diffusion systems.

- We start with a free energy, describing drift and diffusion of the particles. The chemical part of this free energies should be of type $\Phi_C(u) = \sum_i \int F_i(u_i/u_i^*)u_i^*$ with stationary state u_i^*. Modeling cross diffusion, some of the particles can be understood as pseudo-particles. It is not known how to write down the free energy in the case of general cross interaction and non-constant stationary states.
- The functions F_i have to be convex with the property $\partial F_i^{-1} \geq 0$.
- The stationary state u^* can be calculated by Lagrange's method, taking into account the constraints and solving the Euler–Lagrange equation.
- The Lagrange multiplier λ has to be understood as a thermodynamic potential and the Euler–Lagrange equation has to be understood as the state equation, connecting the intensive thermodynamic potential and the extensive concentration.
- The driving force is an operator with kernel $\mathbb{1}$ acting on the thermodynamic potential. Then, the stationary solution of the evolution system is the solution of the Euler–Lagrange equation.
- Microscopic considerations show that the operator connecting the driving forces should be taken as the product of a diagonal matrix (diffusivity) and the inverse Hessian of the free energy. Cross diffusion requires constant diagonal diffusivities.

- For energy models, it is not clear how to write down the free energy to use this method in a rigorous way. The presented derivation is a purely phenomenological one and far from accuracy. Nevertheless, this method is very successful in applications.

References

[1] G. Albinus, H. Gajewski, R. Hünlich, *Thermodynamic design of energy models of semiconductor devices*, Nonlinearity, **16** (2002), pp. 367–383.

[2] A. Arnold, P. Markowich, G. Toscani, A. Unterreiter, *On logarithmic Sobolev inequalities, Csiszar-Kullback inequalities, and the rate of convergence to equilibrium for Fokker-Planck type equations.* Preprint. Berlin, Techn. Univ., Fachbereich Mathematik (3), Preprint-Reihe Mathematik, 592, 1998.

[3] U. Bandelow, H. Gajewski, R. Hünlich, *Fabry-Perot lasers: Thermodynamics-based modeling*, in: Optoelectronic devices. Advanced simulation and analysis (2005), J. Piprek (ed.), Springer-Verlag, New York (2005).

[4] L. Chen and A. Jüngel, *Analysis of a parabolic cross-diffusion semiconductor model with electron-hole scattering*, 2005, to appear in J. Diff. Eqs.

[5] T. Frank, *Lyapunov and free energy functionals of generalized Fokker-Planck equations*, Physics Letters A 290: 93-100 (2001).

[6] H. Gajewski, *Analysis und Numerik des Ladungsträgertransports in Halbleitern*, GAMM-Mitteilungen, 16 (1993), pp. 35–57.

[7] H. Gajewski, K. Gärtner, *A dissipative discretization scheme for a nonlocal phase segregation model.* Z. Angew. Anal. Mech. 85, 815-822 (2005).

[8] H. Gajewski, J.A. Griepentrog, *A descent method for the free energy of multicomponent systems.* Discrete Contin. Dynam. Systems 15, 505-528 (2006).

[9] H. Gajewski, K. Zacharias, *On a nonlocal phase separation model*, J. Math. Anal. Appl. 286, 11-31 (2003).

[10] H. Gajewski, K. Zacharias, *Global behaviour of a reaction-diffusion system modelling chemotaxis*, Math. Nachr., **195** (1998), pp. 77–114.

[11] C. W. Gardiner. *Handbook of Stochastic Methods.* Springer Series in Synergetics. Volume 13. Second Edition. 1985.

[12] G. Giacomin, J.L. Lebowitz, R. Marra, *Macroscopic evolution of particle systems with short- and long-range interactions*, Nonlinearity 13 (2000), 2143-2162.

[13] A. Glitzky, R. Hünlich, *Global estimates and asymptotics for electro-reaction-diffusion systems in heterostructures*, Appl. Anal. 66 (1997) pp. 205-226.

[14] J.A. Griepentrog, *On the unique solvability of a phase separation problem for multicomponent systems*, Banach Center Publ., **66** (2004), pp. 153–164.

[15] D.E. Kane, R.M. Swanson, *Modeling of electron-hole scattering. in semiconductor devices*, IEEE Trans. Electron Devices,. ED-40, Nr. 8, 1496-1500 (1993).

[16] M. Lades, *Modeling and Simulation of Wide Bandgap Semiconductor Devices: 4H/6H-SiC*, Dissertation, TU–Munich, 2000.

[17] J. A. Carrillo, A. Jüngel, P.A. Markowich, G. Toscani, A.Unterreiter, *Entropy dissipation methods for degenerate parabolic problems and generalized Sobolev inequalities*, Monatshefte für Mathematik, 133, 1-82, 2001.

[18] T.T. Mnatsakanov, *Transport coefficients and Einstein relation in a High Density Plasma of Solids*, phys. stat. sol. (b) 143, pp.225-234 (1987).

[19] H. Stephan, *Lyapunov functions for positive linear evolution problems*, ZAMM, 85, No. 11, 766-777.

[20] H. Stephan, *A dimension-reduced description of general Brownian motion by non-autonomous diffusion-like equations*, Journal of Mathematical Physics, Analysis, Geometry (MAG). 2005, v.12, Nr. 2, p.187–202.

[21] H. Stephan, *Functional Analytic Properties of Microscopic Models*, in preparation.

[22] D. Reznik, W. Gerlach, *Generalized drift-diffusion model of bipolar transport in semiconductors.* Preprint, Fachbereich Elektrotechnik, TU Berlin, Germany, 1995.

[23] C. Tsallis, *Possible generalisation of Boltzmann–Gibbs statistics*, Journal of Statistical Physics, vol. 52, p. 479 (1988).

[24] C. Tzanakis, A.C. Grecos, *Classical Markovian Kinetic Equations: Explicit Form and H-Theorem.* arXiv:physics/9708031.

[25] C. Villani, *Trend to equilibrium for dissipative equations, functional inequalities and mass transportation*, in: Recent Advances in the Theory and Applications of Mass Transport, M. C. Carvalho and J. F. Rodrigues, editors, Contemporary Mathematics, vol. 353, Amer. Math. Soc., Providence, RI (2004), 95–109.

[26] G. Wachutka, *Rigorous thermodynamic treatment of heat generation and conduction in semiconductor device modelling*, IEEE Trans. CAD 9 (1990),1141-9.