Generalized Scharfetter–Gummel schemes for electro-thermal transport in degenerate semiconductors using the Kelvin formula for the Seebeck coefficient

Markus Kantner
Weierstrass Institute for Applied Analysis and Stochastics (WIAS), Mohrenstr. 39, 10117 Berlin, Germany

Abstract
Many challenges faced in today’s semiconductor devices are related to self-heating phenomena. The optimization of device designs can be assisted by numerical simulations using the non-isothermal drift-diffusion system, where the magnitude of the thermoelectric cross effects is controlled by the Seebeck coefficient. We show that the model equations take a remarkably simple form when assuming the so-called Kelvin formula for the Seebeck coefficient. The corresponding heat generation rate involves exactly the three classically known self-heating effects, namely Joule, recombination and Thomson–Peltier heating, without any further (transient) contributions. Moreover, the thermal driving force in the electrical current density expressions can be entirely absorbed in the diffusion coefficient via a generalized Einstein relation. The efficient numerical simulation relies on an accurate and robust discretization technique for the fluxes (finite volume Scharfetter–Gummel method), which allows to cope with the typically stiff solutions of the semiconductor device equations. We derive two non-isothermal generalizations of the Scharfetter–Gummel scheme for degenerate semiconductors (Fermi–Dirac statistics) obeying the Kelvin formula. The approaches differ in the treatment of degeneration effects: The first is based on an approximation of the discrete generalized Einstein relation implying a specifically modified thermal voltage, whereas the second scheme follows the conventionally used approach employing a modified electric field. We present a detailed analysis and comparison of both schemes, indicating a superior performance of the modified thermal voltage scheme.

Keywords: finite volume Scharfetter–Gummel method, semiconductor device simulation, electro-thermal transport, non-isothermal drift-diffusion system, degenerate semiconductors (Fermi–Dirac statistics), Seebeck coefficient

1. Introduction

Thermal effects cause many challenges in a broad variety of semiconductor devices. Thermal instabilities limit the safe-operating area of high power devices and modules in electrical energy technology [1, 2], electro-thermal feedback loops lead to catastrophic snapback phenomena in organic light-emitting diodes [3, 4] and self-heating effects decisively limit the achievable output power of semiconductor lasers [5–8]. The numerical simulation of semiconductor devices showing strong self-heating and thermoelectric effects requires a thermodynamically consistent modeling approach, that describes the coupled charge carrier and heat transport processes. In the context of semiconductor device simulation, the non-isothermal drift-diffusion system [9–14] has become the standard model for the self-consistent description of electro-thermal transport phenomena. This is a system of four partial differential equations, which couples the semiconductor device equations [15, 16] to a (lattice) heat flow equation for the temperature distribution in the device. On the step from the isothermal to the non-isothermal drift-diffusion system, additional thermoelectric transport coefficients must be included in the theory. The magnitude of the thermoelectric cross-effects is governed by the Seebeck coefficient (also thermopower), which quantifies the thermoelectric voltage induced by a temperature gradient (Seebeck effect) [17, 18]. The reciprocal phenomenon of the Seebeck effect is the Peltier effect, which describes the current-induced heating or cooling at material junctions. As a consequence of Onsager’s reciprocal relations, the
Seebeck and Peltier coefficients are not independent such that only the Seebeck coefficient must be specified [19]. Over the decades, several definitions have been proposed for the Seebeck coefficient [20–23]; recent publications list at least five coexisting different (approximate) formulas [24, 25]. In the context of semiconductor device simulation, the Seebeck coefficients are typically derived from the Boltzmann transport equation in relaxation time approximation [26–28] or defined according to the adage of the Seebeck coefficient being the “(specific) entropy per carrier” [13, 14, 18, 29]. These approaches are often focused on non-degenerate semiconductors, where the carriers follow the classical Maxwell–Boltzmann statistics. This approximation breaks down in heavily doped semiconductors, where the electron-hole plasma becomes degenerate and Fermi–Dirac statistics must be considered to properly take into account the Pauli exclusion principle. Degeneration effects are important in many semiconductor devices such as semiconductor lasers, light emitting diodes or transistors. Moreover, heavily doped semiconductors are considered as “good” thermoelectric materials, i.e., materials with high thermoelectric figure of merit [17, 18], for thermoelectric generators, which can generate electricity from waste heat [30, 31].

In this paper, we will consider an alternative model for the Seebeck coefficient, which is the so-called Kelvin formula for the thermopower [32]. The Kelvin formula recently gained interest in theoretical condensed matter physics and has been shown to yield a good approximation of the Seebeck coefficient for many materials (including semiconductors, metals and high temperature superconductors) at reasonably high temperatures [32–42]. The Kelvin formula relates the Seebeck coefficient to the derivative of the entropy density with respect to the carrier density and therefore involves only equilibrium properties of the electron-hole plasma, where degeneration effects are easily included. To our knowledge, the Kelvin formula has not been considered in the context of semiconductor device simulation so far. In Sec. 2, we show that the Kelvin formula yields a remarkably simple form of the non-isothermal drift-diffusion system, which shows two exceptional features:

1. The heat generation rate involves exactly the three classically known self-heating effects (Joule, Thomson–Peltier and recombination heating) without any further (transient) contributions.

2. The thermal driving force in the current density expressions can be entirely absorbed in a (nonlinear) diffusion coefficient via a generalized Einstein relation. Hence, the $\nabla T$ term is eliminated in the drift-diffusion form.

The second part of this paper (Sec. 3) deals with the discretization of the electrical current density expressions, which are required in (non-isothermal) semiconductor device simulation tools. The robust and accurate discretization of the drift-diffusion fluxes in semiconductors with exponentially varying carrier densities is a non-trivial problem, that requires a special purpose discretization technique. The problem has been solved by Scharfetter and Gummel for the case of non-degenerate semiconductors under isothermal conditions [43]. Since then, several adaptations of the method have been developed to account for more general situations (non-isothermal conditions [44–51], degeneration effects [52–57]). The Kelvin formula for the Seebeck coefficients allows for a straightforward generalization of the Scharfetter–Gummel approach to the non-isothermal case. We take up two different approaches to incorporate degeneration effects into the non-isothermal Scharfetter–Gummel formula and give an extensive numerical and analytical comparison of both methods. This includes an investigation of limiting cases and structure preserving properties of the discrete formulas (Sec. 3.3), a comparison with the numerically exact solution of the underlying two-point boundary value problem (Sec. 3.4) and a comparison of analytical error bounds (Sec. 3.5). Finally, in Sec. 3.6, we present a numerical convergence analysis of both schemes based on numerical simulations of a one-dimensional p-n-diode.

2. The non-isothermal drift-diffusion system using the Kelvin formula for the Seebeck coefficient

In this section we briefly review the non-isothermal drift-diffusion system, which provides a self-consistent description of the coupled electro-thermal transport processes in semiconductor devices. The model has been extensively studied by several authors from the perspective of physical kinetics or phenomenological non-equilibrium thermody-
The carrier densities \( n \) and \( p \) are the densities of electrons and holes, respectively, \( C \) is the built-in doping profile, \( q \) is the elementary charge and \( \varepsilon \) is the (absolute) dielectric constant of the material. The transport and recombination dynamics of the electron-hole plasma are modeled by the continuity equations (2)–(3), where \( \mathbf{j}_n/p \) are the electrical current densities and \( R \) is the (net-)recombination rate. The latter includes several radiative and non-radiative recombination processes (Shockley–Read–Hall recombination, Auger recombination, spontaneous emission etc.) \[15\]. The carrier densities \( n, p \) are connected with the electrostatic potential \( \phi \) via the state equations:

\[ n = N_c(T) \, \mathcal{F}_n \left( \frac{\mu_c + q\phi - E_{c}(T)}{k_B T} \right), \]
\[ p = N_v(T) \, \mathcal{F}_v \left( \frac{E_v(T) - q\phi - \mu_v}{k_B T} \right), \]

where \( k_B \) is Boltzmann’s constant, \( T \) is the absolute temperature, \( N_{c/v} \) is the effective density of states and \( E_{c/v} \) is the reference energy level (typically the band edge energy) of the conduction band or valence band, respectively. The function \( \mathcal{F} \) describes the occupation of the electronic states under quasi-equilibrium conditions, which is controlled by the quasi-Fermi energies \( \mu_{c/v} \) of the respective bands. The quasi-Fermi energies are connected with the quasi-Fermi potentials \( \phi_{n/p} \) via:

\[ \mu_c = -q\phi_n, \]
\[ \mu_v = -q\phi_p. \]

In non-degenerate semiconductors (Maxwell–Boltzmann statistics), \( \mathcal{F} \) is the exponential function \( \mathcal{F}(\eta) = \exp(\eta) \). Taking the degeneration of the electron-hole plasma due to Pauli-blocking into account (Fermi–Dirac statistics), \( \mathcal{F} \) is typically given by the Fermi–Dirac integral

\[ \mathcal{F}(\eta) = F_{c/v}(\eta) = \frac{1}{\Gamma(\nu+1)} \int_0^{\infty} d\xi \frac{\xi^{\nu}}{\exp(\xi - \eta) + 1}, \]

where the index \( \nu \) depends on the dimensionality of the structure. Isotropic, bulk materials with parabolic energy bands are described by \( \nu = 1/2 \); for two-dimensional materials (quantum wells) the index \( \nu = 0 \) applies. See Fig. 1 (a).
for a plot of the Fermi–Dirac integrals for different \( \nu \) as a function of the reduced Fermi energy. The function \( \mathcal{F} \) may also include non-parabolicity effects, see Appendix A. In the case of organic semiconductors, \( \mathcal{F} \) is often taken as the Gauss–Fermi integral \([58, 59]\) or a hypergeometric function \([60, 61]\).

The heat transport equation (4) describes the spatio-temporal dynamics of the temperature distribution in the device. Here, \( c_V \) is the (volumetric) heat capacity, \( \kappa \) is the thermal conductivity and \( H \) is the heat generation rate. The non-isothermal drift-diffusion model assumes a local thermal equilibrium between the lattice and the carriers, i.e., \( T = T_L = T_n = T_p \). The system (1)–(4) must be supplemented with initial conditions and boundary conditions (i.e., for electrical contacts, semiconductor-insulator interfaces, heat sinks etc). We refer to Refs. \([15, 62]\) for a survey on commonly used boundary condition models.

The electrical current densities are driven by the gradients of the quasi-Fermi potentials and the temperature

\[
\mathbf{j}_n = -\sigma_n (\nabla \varphi_n + \mathbf{P}_n \nabla T), \quad \mathbf{j}_p = -\sigma_p (\nabla \varphi_p + \mathbf{P}_p \nabla T),
\]

(8)

where \( \sigma_n = qM_n \) and \( \sigma_p = qM_p \) are the electrical conductivities (with carrier mobilities \( M_{n/p} \) and \( P_{n/p} \) are the Seebeck coefficients. Finally, we consider a (net-)recombination rate of the form \([63]\)

\[
R = R (\phi, \varphi_n, \varphi_p, T) = \left( 1 - \exp \left( -\frac{\mu_n - \mu_p}{k_B T} \right) \right) \sum_{\alpha} r_\alpha (\phi, \varphi_n, \varphi_p, T),
\]

(9)

which combines several radiative and non-radiative recombination processes labeled by \( \alpha \) (e.g., Shockley–Read–Hall recombination, spontaneous emission, Auger recombination etc.). The functions \( r_\alpha = r_\alpha (\phi, \varphi_n, \varphi_p, T) \geq 0 \) are inherently non-negative and specific for the respective processes. We refer to Refs. \([15, 62]\) for commonly considered recombination rate models.

### 2.1. Kelvin formula for the Seebeck coefficient

In this paper, we consider the so-called Kelvin formula for the Seebeck coefficient \([32]\)

\[
P_n = -\frac{1}{q} \frac{\partial s(n, p, T)}{\partial n}, \quad P_p = \frac{1}{q} \frac{\partial s(n, p, T)}{\partial p},
\]

(10)

which relates the thermoelectric powers to the derivatives of the entropy density \( s = s(n, p, T) \) with respect to the carrier densities. The expression for the entropy density is easily derived from the free energy density \( f(n, p, T) \) of the system, which is a proper thermodynamic potential if the set of unknowns is chosen as \( (n, p, T) \) (“natural variables”). The expressions for the quasi-Fermi energies and the entropy density then follow as

\[
\frac{\partial f(n, p, T)}{\partial n} = +\mu_n (n, p, T), \quad \frac{\partial f(n, p, T)}{\partial p} = -\mu_p (n, p, T), \quad \frac{\partial f(n, p, T)}{\partial T} = -s(n, p, T).
\]

(11)

Taking the second derivatives, this yields the Maxwell relations

\[
\frac{\partial \mu_n(n, p, T)}{\partial T} = -\frac{\partial s(n, p, T)}{\partial n}, \quad \frac{\partial \mu_p(n, p, T)}{\partial T} = +\frac{\partial s(n, p, T)}{\partial p},
\]

(12)

which allow for an alternative representation of Eq. (10). The free energy density includes contributions from the quasi-free electron-hole plasma (ideal Fermi gas), the lattice vibrations (ideal Bose gas) and the electrostatic (Coulomb) interaction energy. Throughout this paper, we assume a free energy density of the form \([13]\)

\[
f(n, p, T) = f_{e-b}(n, p, T) + f_L (T) + f_{\text{Coul}} (p - n).
\]

(13)

The free energy density of the (non-interacting) electron-hole plasma reads \([13, 63]\)

\[
f_{e-b}(n, p, T) = k_B T \mathcal{F}^{-1} \left( \frac{n}{N_c(T)} \right) n - k_B T \mathcal{F}^{-1} \left( \frac{n}{N_c(T)} \right) + E_c(T) n + k_B T \mathcal{F}^{-1} \left( \frac{p}{N_v(T)} \right) p - k_B T \mathcal{F}^{-1} \left( \frac{p}{N_v(T)} \right) - E_v(T) p.
\]

(14)
Seebeck coefficient 

\[
\begin{align*}
E_f(T)/k_B &= E'_f(T)/k_B \\
\end{align*}
\]

\( \text{holes} \)

\( \text{electrons} \)

Fermi–Dirac

Maxwell–Boltzmann

Fig. 2. (a) Seebeck coefficient according to the Kelvin formula (16) as a function of the reduced Fermi energy \( \eta \) for power law type effective density of states \( N_{c/v} \propto T^{3/2} \) and \( \mathscr{F}(\eta) = F_{c/v}(\eta) \) in units of \( k_B/q \). The formula (16) takes degeneration effects (Fermi–Dirac statistics, solid lines) of the electron-hole plasma into account, which causes a deviation from the non-degenerate result (Maxwell–Boltzmann statistics, dashed lines) at \( \eta \geq 1 \). The temperature dependency of the band gap energy yields an offset of \( E'_f(T) = \chi + \frac{1}{2} E'_c(T) \) for electrons (red lines) and \( E'_c = \chi - \frac{1}{2} E'_c(T) \) for holes (blue lines). The plot is for \( \chi = -0.2 \) and \( k_B E'_c(T) = -5 \). (b) Seebeck coefficient for \( n \)-type GaAs. Solid lines are computed according to the Kelvin formula (16a) using Fermi–Dirac statistics, dashed lines indicate the corresponding non-degenerate limit. The respective ionized donor densities \( N_{D_i} \) are given in the plot in units of \( \text{cm}^{-3} \). The temperature dependency of the band gap energy \( E_g(T) \) is modeled by the Varshni model (18) with data from Ref. [62] and the effective mass \( m_e^*(T) = 0.067 - 1.2 \times 10^{-3} \text{K}^{-1} T^3 m_0 \), where \( m_0 \) is the free electron mass. The fitting parameter is set to \( \chi = -0.2 \). Experimental data: \( \nabla \) Carlson et al. [65], \( \bigtriangleup \) Amith et al. [66], \( \heartsuit \) Edmond et al. [67], \( \bigdiamond \) Homm et al. [69] and \( \bullet \) Emel’yanenko et al. [70] (data from Ref. [68]).

where \( \mathscr{F}^{-1} \) is the inverse of the function \( \mathscr{F} \) in the state equations (5) and \( \mathcal{G} \) denotes its antiderivative: \( \mathcal{G}'(\eta) = \mathcal{F}(\eta) \).

Note that Eq. (14) implies

\[
\frac{\partial f_{e-h}}{\partial n} = k_B T \mathcal{F}^{-1}\left( \frac{n}{N_c(T)} \right) + E_c(T), \quad \frac{\partial f_{e-h}}{\partial p} = k_B T \mathcal{F}^{-1}\left( \frac{p}{N_v(T)} \right) - E_v(T). \tag{15a}
\]

The lattice contribution \( f_l(T) \) yields the dominant contribution to the heat capacity \( c_V \). It can be derived from, e.g., the Debye model for the free phonon gas [64]. The Coulomb interaction energy \( f_{\text{Coul}} \) must be modeled such that the state equations (5) follow consistently from solving the defining relations for the quasi-Fermi energies (11) for the carrier densities. In order to supplement the “missing” electrostatic contributions in Eq. (15a), we specify the derivatives of \( f_{\text{Coul}} \) with respect to the carrier densities:

\[
\frac{\partial f_{\text{Coul}}}{\partial n} = -q \phi, \quad \frac{\partial f_{\text{Coul}}}{\partial p} = +q \phi. \tag{15b}
\]

We refer to Alblins et al. [13] for a rigorous mathematical treatment of the Coulomb interaction energy.

The Seebeck coefficients (10) are evaluated using Eqs. (11)–(15). Since \( f_{\text{Coul}} \) is independent of the temperature and \( f_l \) does not depend on the carrier densities, the evaluation of Eq. (10) requires only the Maxwell relations (12) and the derivatives of Eqs. (15a) with respect to the temperature. One obtains

\[
\begin{align*}
P_n(n, T) &= -\frac{k_B}{q} \left( \frac{T N_c'(T)}{N_c(T)} + \mathcal{F}^{-1}\left( \frac{n}{N_c(T)} \right) - \frac{1}{k_B} E_c'(T) \right), \tag{16a} \\
P_p(p, T) &= +\frac{k_B}{q} \left( \frac{T N_v'(T)}{N_v(T)} + \mathcal{F}^{-1}\left( \frac{p}{N_v(T)} \right) + \frac{1}{k_B} E_v'(T) \right). \tag{16b}
\end{align*}
\]

where the prime denotes the derivatives \( N_{c/v}'(T) = \partial_T N_{c/v}(T) \) and \( E_{c/v}'(T) = \partial_T E_{c/v}(T) \). For power law type temperature dependency \( N_{c/v}(T) \propto T^\theta \) (e.g., \( \theta = 3/2 \)), the factor in the first term reduces to a constant \( T N_{c/v}'(T) / N_{c/v}(T) = \theta \).
For temperature-dependent effective masses, the term is more complicated. The function
\[ g(x) = x \frac{dF^{-1}(x)}{dx} \]  
(17a)
quantifies the degeneration of the Fermi gas. For non-degenerate carrier statistics (Maxwell–Boltzmann statistics), Eq. (17a) reduces to exactly \( g \equiv 1 \). For degenerate carrier statistics one obtains \( g > 1 \), which implies a nonlinear enhancement of the diffusion current (see Sec. 2.4). For later use, we also introduce the function
\[ g_{\eta}(\eta) \equiv g(F(\eta)) = \frac{F(\eta)}{F'(\eta)}, \]  
(17b)
which is plotted in Fig. 1 (b). The last terms in Eq. (16) describe the contributions of the temperature dependency of the band edge energies to the Seebeck coefficients. The two terms are not independent, as they are required to satisfy \( E_g'(T) = E'_s(T) - E'_c(T) \), where \( E_s(T) \) is the energy band gap. A plot of the Seebeck coefficients (16) as functions of the reduced Fermi energy \( \eta \) is shown in Fig. 2 (a) for \( F(\eta) = F_{1/2}(\eta) \) and \( N_{c,v} \propto T^{3/2} \). The plot illustrates schematically the impact of the temperature derivatives of the band edge energies and the role of degeneration effects.

In the following, several consequences of the Kelvin formula for the Seebeck coefficients are described, which are very appealing for numerical semiconductor device simulation as they greatly simplify the model equations. Before going into details, we emphasize that the Kelvin formula is of course merely a convenient approximation and by no means exact. More accurate and microscopically better justified approaches to calculate the Seebeck coefficients are based on advanced kinetic models such as the semi-classical Boltzmann transport equation beyond the relaxation time approximation (retaining the full form of the collision operator \([71, 72]\)) or fully quantum mechanical methods \([36, 37, 39, 42]\).

2.2. Comparison with experimental data

Several empirical models for the temperature dependency of the band gap energy have been proposed in the literature \([73]\), including the commonly accepted Varshni model
\[ E_g(T) = E_{g,0} - \frac{\alpha T^2}{\beta + T}, \]  
(18)
where \( E_{g,0} \), \( \alpha \) and \( \beta \) are material specific constants \([74]\). In order to specify \( E'_{c,v}(T) \) from Eq. (18), we introduce a parameter \( \chi \) such that \( E'_c(T) = \left( \chi + \frac{1}{2} \right) E'_g(T) \) and \( E'_v(T) = \left( \chi - \frac{1}{2} \right) E'_g(T) \). In applications, \( \chi \) can be used as a fitting parameter. It shall be noted that the terms involving \( E'_{c,v}(T) \) in Eq. (16) are non-negligible and yield a significant contribution to the Seebeck coefficients at elevated temperatures. Indeed, some room temperature values of \( k_B T E'_g(300 \text{ K}) \) for important semiconductors are \(-2.95 \text{ (Si)}, -4.47 \text{ (Ge)}, -5.32 \text{ (GaAs)} \) \([62]\), which are on the same order of magnitude as the first term \( T N_{c,v}'(T) \beta N_{c,v}(T) \approx 1.5 \) in Eq. (16).

In Fig. 2 (b), the Kelvin formula is plotted along with experimental data for n-GaAs. We observe a good quantitative agreement of the formula (16a) with the experimental data in both the weak and the heavy doping regime for temperatures above 150 K. At high carrier densities (\( N_D^0 \geq 9 \times 10^{17} \text{ cm}^{-3} \)) the conduction band electrons become degenerate (see the deviation of the solid from the dashed lines), where the experimental values nicely follow the degenerate formula (16a). See the caption for details. At low temperatures (\( T < 150 \text{ K}, \text{not shown} \)), the Seebeck coefficient is increasingly dominated by the phonon drag effect \([69]\), which is not considered in the present model.

2.3. Heat generation rate

A commonly accepted form of the self-consistent heat generation rate \( H \) was derived by Wachutka \([9]\):
\[
H = \frac{1}{\sigma_n} \left\| \mathbf{j}_n \right\|^2 + \frac{1}{\sigma_p} \left\| \mathbf{j}_p \right\|^2 - T \mathbf{j}_n \cdot \nabla P_n - T \mathbf{j}_p \cdot \nabla P_p + q \left( T \frac{\partial \varphi_n(n, p, T)}{\partial T} - \varphi_n - T \frac{\partial \varphi_p(n, p, T)}{\partial T} + \varphi_p \right) R - T \left( \frac{\partial \varphi_p(n, p, T)}{\partial T} + P_p \right) \nabla \cdot \mathbf{j}_p - T \left( \frac{\partial \varphi_p(n, p, T)}{\partial T} + P_n \right) \nabla \cdot \mathbf{j}_n. \]  
(19)
Thus, the thermally driven part of the electrical current density in the diffusion equation for the present model is derived in Appendix D. The dissipated heat is closely related with the electrical power injected through the contacts. The global power balance equation for the present model is derived in Appendix D.

2.4. Electrical current densities in drift-diffusion form

In this section we recast the electrical current density expressions from the thermodynamic form (8) to the drift-diffusion form. As we will see below, the Kelvin formula for the Seebeck coefficient allows to entirely absorb the thermally driven part of the electrical current density in the diffusion coefficient via a generalized Einstein relation. Thus, the \( \nabla T \) term can be eliminated in the drift-diffusion form, which significantly simplifies the current density...
expression. Our derivation is based on rewriting the gradient of the quasi-Fermi potential using the free energy density \((13)\) and further thermodynamic relations stated above. In the following, we sketch the essential steps for the electron current density, the corresponding expression for the holes follows analogously. We obtain

\[
-q\nabla \varphi_n \stackrel{\text{Eq. (11)}}{=} \nabla \frac{\partial f}{\partial n} \stackrel{\text{Eq. (13)}}{=} \nabla \left( \frac{\partial f_{\text{Coul}}}{\partial n} + \frac{\partial f_{e-h}}{\partial n} \right) \stackrel{\text{Eq. (15b)}}{=} -q\nabla \phi + \frac{\partial^2 f_{e-h}}{\partial n^2} \nabla n + \frac{\partial^2 f_{e-h}}{\partial n \partial p} \nabla p + \frac{\partial^2 f_{e-h}}{\partial n \partial T} \nabla T,
\]

where we have separated the contributions from the Coulomb interaction energy \(f_{\text{Coul}}\) (leading to drift in the electric field) and the quasi-free electron-hole plasma (yielding Hessian matrix elements of the ideal Fermi gas’ free energy density \(f_{e-h}\)). The electrons and holes are decoupled in the non-interacting Fermi-gas such that

\[
\frac{\partial^2 f_{e-h}}{\partial n \partial p} = 0.
\]

Moreover, since (i) the Coulomb interaction energy is independent of the temperature and therefore does not contribute to the system’s entropy and (ii) the lattice contribution \(f_l\) is independent of the carrier densities, it holds

\[
\frac{\partial^2 f_{e-h}}{\partial n \partial T} = -\frac{\partial s}{\partial n},
\]

where \(s\) is the entropy density of the full system (see the last formula in Eq. (11)). Thus, we arrive at

\[
\nabla \varphi_n = \nabla \phi - \frac{1}{q} \frac{\partial^2 f_{e-h}}{\partial n^2} \nabla n + \frac{1}{q} \frac{\partial s}{\partial n} \nabla T,
\]

which must be substituted in Eq. (8) to obtain

\[
J_n = -\sigma_n \left( \nabla \phi - \frac{1}{q} \frac{\partial^2 f_{e-h}}{\partial n^2} \nabla n + \left[ \frac{1}{q} \frac{\partial s}{\partial n} + P_n \right] \nabla T \right) = -\sigma_n \nabla \phi + \sigma_n \frac{1}{q} \frac{\partial^2 f_{e-h}}{\partial n^2} \nabla n.
\]

In the last step, we have used the Kelvin formula (10) for the Seebeck coefficient. The temperature gradient term vanishes exactly, since reversing the order of the derivatives in the Hessian of the free energy density immediately yields the definition (10) and cancels with the Seebeck term in Eq. (8). The same result can be obtained by simply inverting the carrier density state equation (5) and using the explicit expression (16). With the electrical conductivities \(\sigma_n = qM_n n, \sigma_p = qM_p p\) and

\[
\frac{\partial^2 f_{e-h}}{\partial n^2} = \frac{k_B T}{n} g \left( \frac{n}{N_c(T)} \right), \quad \frac{\partial^2 f_{e-h}}{\partial p^2} = \frac{k_B T}{p} g \left( \frac{p}{N_v(T)} \right),
\]

(from Eq. (15a)), we finally arrive at the drift-diffusion form:

\[
J_n = -qM_n n \nabla \phi + qD_n (n, T) \nabla n, \quad J_p = -qM_p p \nabla \phi - qD_p (p, T) \nabla p.
\]

The diffusion coefficients are given by the generalized Einstein relation [26, 77]

\[
D_n (n, T) = \frac{k_B T M_n}{q} g \left( \frac{n}{N_c(T)} \right), \quad D_p (p, T) = \frac{k_B T M_p}{q} g \left( \frac{p}{N_v(T)} \right).
\]

Here the degeneracy factor \(g\) describes an effective enhancement of the diffusion current that depends nonlinearly on the carrier densities, which results from the increased average thermal energy of the carriers in the case of Fermi–Dirac statistics (see above). The diffusion enhancement due to carrier degeneracy has been found to be important in, e.g., semiconductor laser diodes [78], quantum-photonic devices operated at cryogenic temperatures [79, 80] as well as organic field-effect transistors [81] and light emitting diodes [58]. We emphasize that the drift-diffusion form (22) of the current densities is fully equivalent to the thermodynamic form (8). Thus, even though the \(\nabla T\) term is eliminated, the thermoelectric cross-coupling via the Seebeck effect is fully taken into account via the temperature dependency of the diffusion coefficient. A generalization to the case of hot carrier transport (with multiple temperatures) is described in Appendix B. For Seebeck coefficients that deviate from the Kelvin formula, additional thermodiffusion terms emerge.
3. Non-isothermal generalization of the Scharfetter–Gummel scheme for degenerate semiconductors

The typically exponentially varying carrier densities in semiconductor devices lead to numerical instabilities when using a standard finite difference discretization. In particular, the naive discretization approach results in spurious oscillations and may cause unphysical results such as negative carrier densities [82, 83]. A robust discretization scheme for the drift-diffusion current density was introduced by Scharfetter and Gummel [43], who explicitly solved the current density expressions as a separate differential equation along the edge between two adjacent nodes of the mesh. The resulting discretized current density expressions feature exponential terms that reflect the characteristics of the doping profile and allow for numerically stable calculations. Over the last decades, several generalizations of the Scharfetter–Gummel method have been proposed for either degenerate semiconductors [52–57] or non-isothermal carrier transport with included thermoelectric cross effects [44–51].

In this section, we derive two different generalizations of the Scharfetter–Gummel scheme for degenerate semiconductors obeying the Kelvin formula for the Seebeck coefficient. Both schemes differ in the treatment of degeneration effects and are obtained by extending the approaches previously developed in Refs. [54, 56] and [52]. First, we outline the finite volume method in Sec. 3.1 and then introduce the non-isothermal Scharfetter–Gummel schemes in Sec. 3.2.

We study important limiting cases and structure preserving properties of the discretizations (Sec. 3.3), give a detailed comparison with the numerically exact solution of the underlying two-point boundary value problem (Sec. 3.4) and derive analytical error bounds (Sec. 3.5). Finally, we present a numerical convergence analysis by means of numerical simulations of a one-dimensional p–n-diode in Sec. 3.6.

3.1. Finite volume discretization

We assume a boundary conforming Delaunay triangulation [84] of the point set \( \mathbf{R} = \{ \mathbf{r}_k \}_{k=1 \ldots N_{\text{nodes}}} \), \( \mathbf{r}_k \in \Omega \), where \( \Omega \subset \mathbb{R}^d \) is the computational domain with dimensionality \( d \in \{1, 2, 3\} \). The dual mesh is given by the Voronoï cells

\[
\Omega_K = \{ \mathbf{r} \in \Omega : \| \mathbf{r} - \mathbf{r}_K \| \leq \| \mathbf{r} - \mathbf{r}_L \| \text{ for all } \mathbf{r}_L \in \mathbf{R} \text{ with } \mathbf{r}_L \neq \mathbf{r}_K \},
\]

which provides a non-overlapping tessellation \( \Omega = \bigcup_K \Omega_K \) of the domain. This represents an admissible mesh in the sense of Ref. [85]. The finite volume discretization of the system (1)–(4) is obtained by integration over the cell \( \Omega_K \) and usage of the divergence theorem [83, 85]. The discrete (stationary) non-isothermal drift-diffusion system reads

\[
- \sum_{L \in \mathcal{N}(K)} s_{K,L} \Delta \left( \phi_{L, r_k} - \phi_{K, r_k} \right) = \frac{q}{|\Omega_K|} (C_K + p_K - n_K) \quad (24a)
\]

\[
- \sum_{L \in \mathcal{N}(K)} s_{K,L} J_{n,K,L} = -\frac{q}{|\Omega_K|} |\mathbf{R}_K| \quad (24b)
\]

\[
+ \sum_{L \in \mathcal{N}(K)} s_{K,L} J_{p,K,L} = -\frac{q}{|\Omega_K|} |\mathbf{R}_K| \quad (24c)
\]

\[
- \sum_{L \in \mathcal{N}(K)} s_{K,L} (T_L - T_K) = \frac{1}{2} \sum_{L \in \mathcal{N}(K)} s_{K,L} \left( H_{J,K,L} + H_{T-P,K,L} + |\Omega_K| H_{R,K} \right) \quad (24d)
\]

with the flux projections

\[
J_{n,K,L} = (\mathbf{r}_L - \mathbf{r}_K) \cdot \mathbf{j}_n \quad J_{p,K,L} = (\mathbf{r}_L - \mathbf{r}_K) \cdot \mathbf{j}_p \quad (25)
\]

on the edge \( \overline{KL} := \{ x \mathbf{r}_L + (1 - x) \mathbf{r}_K : x \in \mathbb{R}, 0 \leq x \leq 1 \} \). The geometric factors in Eq. (24) are the volume \( |\Omega_K| \) of the \( K \)-th Voronoï cell and the edge factor

\[
s_{K,L} = \frac{|\partial \Omega_K \cap \partial \Omega_L|}{\| \mathbf{R}_L - \mathbf{r}_K \|}. \quad (26)
\]

The symbol \( \mathcal{N}(K) \) denotes the set of nodes adjacent to \( K \). For the sake of simplicity we restrict ourselves to the case of a homogeneous material. This limitation is not important for the flux discretization, as the discrete fluxes appear only along possible heterointerfaces (edges of the primary simplex grid, see Fig. (3)) but never across. In the case of heterostructures, the currents along material interfaces are weighted by the respective edge factors. Moreover, boundary terms on \( \partial \Omega \cap \Omega_K \neq \emptyset \) are omitted in Eq. (24), which are treated in the standard way as
The problem can be solved exactly by separation of variables. One obtains
\[ x_1 \in \Omega_L \text{ and } x_2 \in \Omega_L \]
where \( x \in \Omega_L \). The electrostatic field \( E \) is assumed to be constant on the edge. For the electron current density, this yields the two-point boundary value problem (BVP)
\[ J_n = R(\phi; \varphi_n, \varphi_p, T_k). \]
Similarly, the discrete self-heating terms are \( H_{f,p} = \frac{1}{L} \int_{\Omega} \frac{\partial}{\partial x} (\varphi_p - \varphi_n) \] \( T \) is obtained by integrating the current density expressions along the edge.

\[ H_{j,n,k,l} = -J_{n,k,l} \left( \varphi_{n,k} + P_{n,k,l} (T_L - T_K) \right) - J_{p,k,l} \left( \varphi_{p,k} + P_{p,k,l} (T_L - T_K) \right), \]
\[ H_{f,p,k,l} = -T_{k,l} J_{n,k,l} \left( P_{n,k} - P_{n,k} \right) - T_{k,l} J_{p,k,l} \left( P_{p,k} - P_{p,k} \right), \]
\[ H_R = q \left( \varphi_{p,k} - \varphi_{n,k} + T \right) \left( P_{p,k} - P_{n,k} \right) \]
\[ R_k. \]

The finite volume discretization of the Joule and Thomson–Peltier heating terms is not straightforward [86–88]. Details on the derivation of Eqs. (27a)–(27b) are provided in Appendix C. The discretization of the edge current densities \( J_{n,k,l} \), the edge-averaged temperature \( T_{k,l} \), and the Seebeck coefficients \( P_{n,p,k,l} \) along the edge \( KL \) are subject to the following sections.

3.2. Discretization of the current density expression

The discretization of \( J_{n,k,l} \) is obtained by integrating the current density expressions (22) along the edge \( KL \) between two adjacent nodes of the mesh. Since the Kelvin formula implies a remarkably simple form of the electrical current densities in drift-diffusion form, where the thermal driving force is eliminated exactly (see Sec. 2.4), this allows for a straightforward adaptation of the Scharfetter–Gummel schemes developed for the isothermal case. We assume the electrostatic field \( E = -\nabla \phi \) and the temperature gradient \( \nabla T \) to be constant along the edge \( KL \), such that
\[ \phi (x) = x \phi_L + (1 - x) \phi_K, \quad T (x) = x T_L + (1 - x) T_k, \]
where \( x \in [0, 1] \) parametrizes the coordinate on the edge \( r (x) = x r_L + (1 - x) r_K \). Tactically, these assumptions have already been used above in Eqs. (24a) and (24d). Moreover, also the mobilities \( M_{n,p} \) and the fluxes \( J_{n,p,k,l} \) are assumed to be constant on the edge. For the electron current density, this yields the two-point boundary value problem (BVP)
\[ k_B T (x) g \left( \frac{n (x)}{N_c (T (x))} \right) \frac{dn}{dx} = q (\phi_L - \phi_K) n (x) + \frac{J_{n,k,l}}{M_n}, \quad n (0) = n_K, \quad n (1) = n_L. \]

on \( x \in [0, 1] \). The problem for the holes current density is analogous.

In the non-degenerate case (Maxwell–Boltzmann statistics) the degeneracy factor is exactly \( g = 1 \), such that the problem can be solved exactly by separation of variables. One obtains
\[ \int_{n_K}^{n_L} \frac{dn}{J_{n,k,l}} + n \frac{q (\phi_L - \phi_K)}{k_B T (x)} = \int_0^1 dx. \]
where the integral on the right hand side yields the (inverse) logarithmic mean temperature

\[
\int_0^1 \frac{dx}{T(x)} = \int_0^1 \frac{dx}{xT_L + (1-x)T_K} = \frac{1}{\Lambda(T_L, T_K)} \equiv \frac{1}{T_{K,L}}.
\]

and \(\Lambda(x, y) = \frac{x - y}{\log (x/y)}\), (29)

where \(\Lambda(x, y)\) is the logarithmic mean. Solving for the flux yields the non-isothermal Scharfetter–Gummel scheme

\[
J_{n,K,L}^{\text{deg}} = M_c k_B T_{K,L} \left( n_L B \left( X_{n,K,L}^{\text{deg}} \right) - n_K B \left( -X_{n,K,L}^{\text{deg}} \right) \right), \quad X_{n,K,L}^{\text{deg}} = \frac{q(\phi_L - \phi_K)}{k_B T_{K,L}},
\]

(30)

where \(B(x) = x/(x^2 - 1)\) is the Bernoulli function. The Bernoulli function is closely related to the logarithmic mean: \(B(x) = 1/\Lambda(x, 1)\). At isothermal conditions \(T_K \equiv T_L\), Eq. (30) reduces to the original Scharfetter–Gummel scheme [43].

In the case of Fermi–Dirac statistics \((g \neq 1)\), no closed-form solution exists such that approximate solutions of the BVP (28) are required. As the degeneracy factor \(g\) depends on both the carrier density and temperature, the problem is not even separable

3.2.1. Modified drift scheme

Following Refs. [54, 56], we solve the BVP (28) by freezing the degeneracy factor \(g(n/N_c(T)) \rightarrow g_{n,K,L}\) to a carefully chosen average. The resulting problem has the same structure as in the non-degenerate case (see above), but with a modified thermal voltage \(k_B T_{K,L}/q \rightarrow k_B T_{K,L}/g_{n,K,L}/q\) along the edge, which takes the temperature variation and the degeneration of the electron gas into account. This yields the modified drift scheme

\[
J_{n,K,L}^g = M_c k_B T_{K,L} g_{n,K,L} \left( n_L B \left( X_{n,K,L}^g \right) - n_K B \left( -X_{n,K,L}^g \right) \right), \quad X_{n,K,L}^g = \frac{q(\phi_L - \phi_K)}{k_B T_{K,L} g_{n,K,L}},
\]

(31a)

where \(T_{K,L}\) is the logarithmic mean temperature (29). In order to ensure the consistency with the thermodynamic equilibrium and boundedness \(g_{n,K} \leq g_{n,K,L} \leq g_{n,L}\) (for \(\eta_{n,K} \leq \eta_{n,L}\) or with \(K \leftrightarrow L\) else), the edge-averaged degeneracy factor is taken as [54, 56]

\[
g_{n,K,L} = \frac{\eta_{n,L} - \eta_{n,K}}{\log (\mathcal{F}(\eta_{n,L})) - \mathcal{F}(\eta_{n,K})} = \frac{\mathcal{F}^{-1}(n_L/N_c(T_L)) - \mathcal{F}^{-1}(n_K/N_c(T_K))}{\log (n_L/N_c(T_L)) - \log (n_K/N_c(T_K))}.
\]

(31b)

In the limit of \(\eta_{n,L} = \eta_{n,K}\), it approaches the common nodal value

\[
\lim_{\eta_{n,L} \rightarrow \eta_{n,K}} g_{n,K,L} = \mathcal{F}(\eta_{n,L}) \equiv g_{\eta}(\eta_{n,L}).
\]

For constant temperature \(T_L = T_K = T\), the scheme reduces to the modified Scharfetter–Gummel scheme discussed in Refs. [54, 56]. It can thus be regarded as a non-isothermal generalization of this approach. In the non-degenerate limit \(g = 1\), it reduces to the non-isothermal, non-degenerate Scharfetter–Gummel scheme (30).

3.2.2. Modified drift scheme

The traditional approach for the inclusion of degeneration effects in the Scharfetter–Gummel scheme, that is widely used in commercial software packages, is based on introducing the correction factors [52, 90, 91]

\[
\gamma(\eta) = \frac{\mathcal{F}(\eta)}{\exp(\eta)}
\]

(32)

and rearranging the current density expression with nonlinear diffusion (22) (involving the generalized Einstein relation (23)) into a form with linear diffusion and a modified drift term:

\[
\mathbf{j}_n = \sigma_n \mathbf{E}_n + M_c k_B T \nabla n + \frac{k_B}{q} \sigma_n \rho_n(T, \eta_n) \nabla T, \quad \mathbf{E}_n = -\nabla \left( \phi + \frac{k_B T}{q} \log(\gamma(\eta_n)) \right).
\]

(33)
Here, the degeneration of the electron gas induces a thermodiffusion term with the coefficient
\[
\rho_n(T, \eta_n) = \log(\gamma(\eta_n)) - \frac{TN'_c(T)}{N_c(T)} \gamma'(\eta_n) / \gamma(\eta_n) = \log(\gamma(\eta_n)) + \frac{TN'_c(T)}{N_c(T)} (g_\eta(\eta_n) - 1),
\]
that vanishes exactly in the non-degenerate limit \( \gamma(\eta) \equiv 1 \equiv g_\eta(\eta) \). Hence, the function \( \rho_n(T, \eta_n) \) quantifies the difference between the degenerate and the non-degenerate Seebeck-coefficient, see Fig. 2 (a). On the step from Eq. (22) to (33), we have used the relation
\[
g_\eta(\eta) = (1 + \gamma'(\eta) / \gamma(\eta))^{-1}.
\]
A plot of the correction factor (32) is given in Fig. 1 (c).

The current density expression (33) is discretized by projecting the current on the edge \( KL \), assuming the effective electric field \( E_0 \) to be a constant along the edge, and freezing \( \rho_n(T, \eta_n) \to \rho_{n,K,L} \) to a constant average value. Here, different averages can be taken for \( \rho_{n,K,L} \), see Fig. 4 (c). The influence of this choice will be discussed below in Sec. 3.4. Along the same lines as above, one arrives at the modified drift scheme
\[
J^\gamma_{n,K,L} = M_n k_B T_{K,L} \left( n_L B(X^\gamma_{n,K,L}) - n_K B(-X^\gamma_{n,K,L}) \right),
\]
with
\[
X^\gamma_{n,K,L} = q(\phi_L - \phi_K)/k_B T_{K,L} + \frac{T_L \log(\gamma(\eta_{n,L})) - T_K \log(\gamma(\eta_{n,K})) - \rho_{n,K,L} \log(T_L/T_K)}{T_{K,L}}.
\]
Again, \( T_{K,L} \) is the logarithmic mean temperature (29). The corresponding non-degenerate limit (30) is easily recovered by \( \gamma(\eta_{n,L/K}) \to 1 \) and \( \rho_{n,K,L} \to 0 \).

3.3. Limiting cases and structure preserving properties

In the following, we investigate some important limiting cases and structure preserving properties of the generalized Scharfetter–Gummel schemes (31) and (35). This includes an analysis of the consistency of the discrete expressions with fundamental thermodynamical principles (thermodynamic equilibrium, second law of thermodynamics). To this end, it is convenient to rewrite both expressions using the identity \( B(-x) = \exp(x) B(x) \) and the logarithmic mean \( \Lambda \) (see Eq. (29)) as
\[
J^\gamma_{n,K,L} = -\sigma^\gamma_{n,K,L} (\varphi_{n,L} - \varphi_{n,K} + P^\gamma_{n,K,L} (T_L - T_K)) \quad \text{with} \quad \sigma^\gamma_{n,K,L} = qM_n \frac{\Lambda(n_L \exp(-\frac{1}{2} X^\gamma_{n,K,L}), n_K \exp(\frac{1}{2} X^\gamma_{n,K,L}))}{\sinhc(\frac{1}{2} X^\gamma_{n,K,L})},
\]
and
\[
J^\gamma_{n,K,L} = -\sigma^\gamma_{n,K,L} (\varphi_{n,L} - \varphi_{n,K} + P^\gamma_{n,K,L} (T_L - T_K)) \quad \text{with} \quad \sigma^\gamma_{n,K,L} = qM_n \frac{\Lambda(n_L \exp(-\frac{1}{2} X^\gamma_{n,K,L}), n_K \exp(\frac{1}{2} X^\gamma_{n,K,L}))}{\sinhc(\frac{1}{2} X^\gamma_{n,K,L})},
\]
where \( \sinhc(x) = \sin(x)/x \). This representation directly corresponds to the continuous current density expression in the thermodynamic form (8), where the conductivity along the edge \( \sigma^\gamma_{n,K,L} \) is determined by a “tilted” logarithmic average of the nodal carrier densities. Both expressions (36)–(37) have a common structure, but differ in the discrete conductivity \( \sigma^\gamma_{n,K,L} \neq \sigma^\gamma_{n,K,L} \) (due to \( X^\gamma_{n,K,L} \neq X^\gamma_{n,K,L} \)) and the discrete Seebeck coefficients \( P^\gamma_{n,K,L} \neq P^\gamma_{n,K,L} \) along the edge, which are implicitly prescribed by the Scharfetter–Gummel discretization procedure. The latter read
\[
P^\gamma_{n,K,L} = -\frac{k_B}{q} \log \left( \frac{N_c(T_L)}{N_c(T_K)} \right) \frac{g_{n,K,L}}{\log(T_L/T_K)} - \frac{(T_L - T_{K,L}) \eta_{n,L} - (T_K - T_{K,L}) \eta_{n,K}}{T_L - T_K} \frac{1}{k_B} \frac{E_c(T_L) - E_c(T_K)}{T_L - T_K}
\]
and
\[
P^\gamma_{n,K,L} = -\frac{k_B}{q} \log \left( \frac{N_c(T_L)}{N_c(T_K)} \right) \frac{1}{\log(T_L/T_K)} + \rho_{n,K,L} - \frac{T_L \log(\gamma(\eta_{n,L})) - T_K \log(\gamma(\eta_{n,K})) - T_{K,L} \log(\gamma(\eta_{n,L})/\gamma(\eta_{n,K}))}{T_L - T_K}
\]
\[
- \frac{(T_L - T_{K,L}) \eta_{n,L} - (T_K - T_{K,L}) \eta_{n,K}}{T_L - T_K} \frac{1}{k_B} \frac{E_c(T_L) - E_c(T_K)}{T_L - T_K}.
\]
The discrete Seebeck coefficients (38) enter the discrete Joule heat term (27a) and thus the discrete entropy production rate (see Sec. 3.3.5 below). Out of the thermodynamic equilibrium, the discrete Seebeck coefficients (38) determine the point of compensating (discrete) chemical and thermal current flow such that \( J_{n,K,L}^{\text{upw}} = 0 \). In other words, there is a non-equilibrium configuration with \( T_K \neq T_L \) and \( \varphi_{n,K} \neq \varphi_{n,L} \), where the discrete Seebeck coefficient \( p_{n,K,L}^{\varphi} = \frac{\varphi_{n,K} - \varphi_{n,L}}{T_L - T_K} \) equals the (negative) ratio of both discrete driving forces such that the discrete current density is zero. This compensation point is in general slightly different between both schemes (see inset of Fig. 4 (c)). In the limit of a small temperature gradient and a small difference in the reduced Fermi energy along the edge, both discrete Seebeck coefficients approach in leading order the continuous expression (16a)

\[
p_{n,K,L}^{\varphi} = -\frac{k_B}{q} \left[ \frac{N'(\bar{T})}{N(\bar{T})} \delta \eta - \frac{1}{k_B} E'_i(\bar{T}) \right] + O(\delta \eta^2) + O(\delta \eta \delta \Theta) + O(\delta \Theta^2),
\]

where \( \delta \Theta = (T_L - T_K)/\bar{T} \), \( \bar{T} = \frac{1}{2} (T_L + T_K) \). \( \delta \eta = \eta_{n,L} - \eta_{n,K} \) and \( \bar{\eta} = \frac{1}{2} (\eta_{n,L} + \eta_{n,K}) \).

3.3.1. Thermodynamic equilibrium

In the thermodynamic equilibrium (thermodynamic equilibrium \( T_K = T_L \) and chemical equilibrium \( \varphi_{n,K} = \varphi_{n,L} \)), both the discrete current densities (31) and (35) are exactly zero. This is easily seen from Eqs. (36)–(37), where the discrete driving force \( \varphi_{n,L} - \varphi_{n,K} + p_{n,K,L}^{\varphi} (T_L - T_K) \) vanishes under thermodynamic equilibrium conditions.

3.3.2. Strong electric field (drift-dominated limit)

Due to the asymptotics of the Bernoulli function \( B(x \to \infty) = 0 \) and \( B(x \to -\infty) \sim -x \), the modified thermal voltage scheme (31) approaches the first-order upwind scheme

\[
J_{n,K,L}^{\varphi} (\delta \Phi_{K,L} \to \pm \infty) \sim -q M_n \left( \frac{n_L + n_K}{2} + \frac{n_K - n_L}{2} \text{sign} (\delta \Phi_{K,L}) \right) \delta \Phi_{K,L} = J_{n,K,L}^{\text{upw}}
\]

in the limit of a strong electrostatic potential gradient \( \delta \Phi_{K,L} = \phi_L - \phi_K \to \pm \infty \). The upwind scheme is a stable, first-order accurate discretization for advection-dominated problems, where the coefficient is evaluated in the “donor” cell of the flow [92]. Hence, this asymptotic feature of the original Scharfetter–Gummel scheme, which is important for the robustness of the discretization as it avoid spurious oscillations, is preserved in the degenerate and non-isothermal case. The modified drift scheme (35) approaches the upwind scheme as well

\[
J_{n,K,L}^{\varphi} (\delta \Phi_{K,L} \to \pm \infty) \sim -q M_n \left( \frac{n_L + n_K}{2} + \frac{n_K - n_L}{2} \text{sign} (\delta \Phi_{K,L}) \right) \left( \delta \Phi_{K,L} + \frac{k_B}{q} \left( \log \left( \frac{\gamma (\eta_{n,L}) \eta_{n,K}^{\text{rd}}}{\gamma (\eta_{n,K}) \eta_{n,L}^{\text{rd}}} \right) - (T_L - T_K) \rho_{n,K,L} \right) \right),
\]

however, in the case of strong degeneration the convergence is significantly slowed down if the nodal correction factors \( \gamma (\eta_{n,K}) \neq \gamma (\eta_{n,L}) \) and temperatures \( T_L \neq T_K \) are very different. This is shown in Fig. 4 (c), where the modified drift scheme shows a constant offset from the numerically exact solution of the BVP (28) for \( \delta \Phi_{K,L} \to -\infty \).

3.3.3. No electric field (diffusive limit)

In the case of a vanishing electrostatic potential gradient \( \delta \Phi_{K,L} = \phi_L - \phi_K = 0 \) the schemes take the form

\[
\lim_{\delta \Phi_{K,L} \to 0} J_{n,K,L}^{\varphi} = M_n k_B T_{K,L} g_{n,K,L} (n_L - n_K),
\]

\[
\lim_{\delta \Phi_{K,L} \to 0} J_{n,K,L}^{\text{upw}} = M_n k_B T_{K,L} \frac{1}{N(\bar{T})} \left( \frac{T_L^{\varphi_{n,K,L}}}{\gamma (\eta_{n,L}) \eta_{n,K}^{\text{rd}}} - \frac{T_K^{\varphi_{n,K,L}}}{\gamma (\eta_{n,K}) \eta_{n,L}^{\text{rd}}} \right) n_L - \frac{\rho_{n,K,L}}{\gamma (\eta_{n,K}) \eta_{n,L}^{\text{rd}}}.\]

The modified thermal voltage scheme (31) approaches the central finite difference discretization (40), which is a stable discretization for diffusion-dominated transport problems [92]. With the edge-averaged degeneracy factor \( g_{n,K,L} \), Eq. (40) nicely reflects the structure of the diffusive part of the continuous current density expression (22) involving the generalized Einstein relation (23). For the modified drift scheme (35), the limiting expression is a weighted finite difference discretization. Due to the different treatment of the degeneracy of the electron gas via the correction factors (32), it does not yield a discrete analogue of the generalized Einstein relation.
3.3.5. Non-negativity of the discrete dissipation rate

The continuous entropy production rate (dissipation rate) per volume (see Eq. (D.4))

\[ s_{\text{tot}} = \frac{1}{T} (\mu_e - \mu_v) R + \frac{\kappa}{T^2} \| \nabla T \|^2 + \frac{1}{T} H_J, \]

has contributions from carrier recombination, heat flux and Joule heating. Given the finite volume discretization described above, this is straightforwardly achieved for the contributions from the carrier recombination and the heat flux, however, it is less obvious for the Joule heating term. In fact, the non-negativity of the discrete Joule heating term is non-trivial and can be violated in general when using a naive discretization approach as in Ref. [94].

We show that the discrete dissipation rate is evidently non-negative for both generalized Scharfetter–Gummel schemes (31) and (35). This follows immediately from their consistency with the thermodynamic equilibrium (see Sec. 3.3.1) in conjunction with the discrete form (27a) of the heating term. Substituting Eq. (36) in (27a), one obtains

\[ H_J = - (\nabla \varphi_p + P_v \nabla T) \cdot \left( \varphi_n - \varphi_n \right), \]

which is zero only in the (discrete) thermodynamic equilibrium and positive else. The discrete conductivity \( \sigma_{n,K,L} \) is positive by construction, see Eq. (36). Analogous expressions are obtained for the holes’ current contribution and the modified drift scheme (35). In conclusion, the consistency of the discrete system (24) with the second law of thermodynamics relies on using the respective Seebeck coefficients (38).

3.4. Comparison with numerically exact solution

We investigate the accuracy of the schemes (31) and (35) by comparing them with the numerically exact solution of the BVP (28). In the isothermal case, this has been carried out before in a similar way by Farrell et al. [95] (for different Scharfetter–Gummel schemes), which inspired the investigation of highly accurate Scharfetter–Gummel type discretizations based on the direct numerical integration of the arising integral equation using quadrature rules in Ref. [96]. In the present non-isothermal case, the problem is more complicated because of the spatially varying temperature distribution along the edge. It is convenient to recast the problem (28) into the form

\[ \frac{dy(x)}{dx} = \frac{T(\tilde{T})}{T(x)} \left( \frac{\phi_n(T)}{N_n(T)} \right) J \frac{\nabla \tilde{T}(x)}{\nabla y(x)} - \delta \Theta \frac{\nabla \tilde{T}(x)}{\nabla y(x)} \cdot \left( \frac{\phi_p(T)}{N_p(T)} \right) F(y), \]

with the notations

\[ T(x) \equiv \left( 1 + \frac{1}{2} x \right) \delta \Theta \tilde{T}, \quad \tilde{T} = \frac{1}{2} (T_L + T_K), \quad \delta T = T_L - T_K, \quad \delta \Theta = \frac{1}{2} (\eta_{n,L} + \eta_{n,K}), \quad \delta \eta_n = \eta_{n,L} - \eta_{n,K} \]

and the non-dimensionalized quantities

\[ J = \frac{J_{n,K,L}}{M_e k_B T N_n(T)}, \quad \delta \Phi = \frac{q(\phi_L - \phi_K)}{k_B T}, \quad \delta \Theta = \frac{\delta T}{T}. \]
The exact current $J_{\text{exact}} = J_{\text{exact}}(\delta\Phi, \delta\eta_n, \delta\Theta, \tilde{\eta}_n, \tilde{T})$ is a function of five parameters that satisfies the BVP (41). We solve the BVP (41) numerically using the shooting method [97], where we combine a 4th order Runge–Kutta method with Brent’s root finding algorithm [98]. The problem is invariant under the simultaneous transformation

$$\delta\Phi \to -\delta\Phi, \quad \delta\eta_n \to -\delta\eta_n, \quad \delta\Theta \to -\delta\Theta, \quad x \to 1 - x, \quad J \to -J$$

(i.e., the sign of the current changes when changing the nodes $K \leftrightarrow L$), such that we can restrict our analysis to $\delta\Theta \geq 0$, when exploring the accuracy of the discrete current in the $(\delta\Phi, \delta\eta_n)$-plane. The comparison is carried out for $F(q) = F_{1/2}(q)$ and $N_c = 2 \left( m^* k_B T / (2\pi h^2) \right)^{1/2}$. The Fermi–Dirac integrals $F_{1/2}(q)$ and $F_{-1/2}(q)$ are evaluated using MacLeod’s algorithm [99].

Figure 4 shows the numerically exact current $J_{\text{exact}}$ along with the approximations $J_x$ (modified thermal voltage scheme (31)) and $J_f$ (modified drift scheme (35)) as a function of the normalized electric field $\delta\Phi$ along the edge.

For weak degeneracy, both schemes agree with the numerically exact solution – even in the case of non-isothermal conditions. This is shown in Fig. 4(a) for $\tilde{\eta}_n = -5$ and $\delta\Theta = 0.1$. At strong electric fields $\delta\Phi \to \pm\infty$, all schemes approach the upwind scheme $J_{\text{upw}}$ (grey dashed line, cf. Eq. (39)). The schemes (31) and (35) differ in the treatment of degeneration effects, which becomes apparent for increased $\tilde{\eta}_n$. Figure 4(b) shows the results for $\tilde{\eta}_n = 2$ at isothermal conditions $\delta\Theta = 0$. The modified thermal voltage scheme $J_x$ (red line) yields an acceptable deviation from the exact result $J_{\text{exact}}$ (black line) over the whole range of $\delta\Phi$. The error vanishes at strong electric fields where both $J_x$ and $J_{\text{exact}}$ converge to the upwind scheme $J_{\text{upw}}$. The modified drift scheme (purple line), however, shows a significant error at large (negative) $\delta\Phi$, where it overestimates the current density significantly (about 33% relative error at $\delta\Phi = -3$).

This behavior results from the different treatment of the degeneration effects, that degrades the convergence of the modified drift scheme in the case of strong degeneration (see Sec. 3.3.2). The plot highlights two other important exceptional points (pure diffusion and zero current), where both schemes show a similar accuracy. In the presence of an additional temperature gradient along the edge, see Fig. 4(c), the approximation error of both schemes increases. The upper inset shows that the choice of the average of $\rho_{nK,L}$ (see Eq. (34)) has only a minor impact on the modified drift scheme. The lower inset zooms on the region where the currents become zero. Here, all schemes provide a satisfying accuracy, but none of them is exact, i.e., they yield a small spurious discrete current and intersect with the exact solution only in the vicinity of the exact zero current point. We observe that the modified drift schemes show a slightly better performance in this case, i.e., the Seebeck coefficient (38b) appears to be slightly better than (38a).

The normalized absolute errors $(J - J_{\text{exact}})/J$ (with $J = M_n k_B T N_c(T)$) of the two schemes (31) and (35) are shown in the $(\delta\Phi, \delta\eta_n)$-plane in Fig. 5 under isothermal ($\delta\Theta = 0$, top row (a)–(d)) and non-isothermal ($\delta\Theta = 1/6$, bottom row (e)–(h)) conditions and for different levels of degeneration (weak $\tilde{\eta}_n = 0$ and strong $\tilde{\eta}_n = 5$). In the limit of very fine meshes $(\delta\Phi, \delta\eta_n, \delta\Theta) \to (0, 0, 0)$, both schemes coincide and the deviation from the exact current approaches...
zero. One observes that the size of the white regions with a normalized absolute error below 0.01 (in the following denoted as “low error domain”) are generally larger for the modified thermal voltage scheme than for the modified drift scheme. Thus, the modified thermal voltage scheme is expected to yield a higher accuracy on sufficiently fine meshes. This will be evidenced by the numerical simulation of a p-n-diode in Sec. 3.6. The plots in Fig. 5 feature two additional lines, that refer to special limiting cases where both schemes yield a very high accuracy. The case of a pure drift current (i.e., no diffusion $n_L = n_K$) is indicated by a red line; the zero current line (blue) refers to the curve in the $(\delta \Phi, \delta \eta_n)$-plane where the exact current vanishes ($J_{\text{exact}} = 0$ in the BVP (41)). In the isothermal case ($\delta \Theta = 0$), the latter corresponds to the thermodynamic equilibrium, in the non-isothermal case ($\delta \Theta \neq 0$) it refers to the situation of compensating chemical and thermal driving forces. Both schemes are exact in the case of a pure drift current, i.e., they asymptotically approach the upwind scheme (39), which is important for the robustness of the discretization in order to avoid spurious oscillations. The modified thermal voltage scheme shows a high accuracy also for slight deviations from the pure drift line, even in the case of strong degeneration, see Fig. 5 (b, f). In contrast, the modified drift scheme yields significant errors (much higher than 0.01) already for tiny deviations from the pure drift line in the strongly degenerate case, see Fig. 5 (d, h). This behavior has already been observed above in Fig. 4 (b, c) and was predicted analytically in Sec. 3.3.2. Note that in the non-isothermal case, the temperature gradient shifts the pure drift line from $\delta \eta_n \text{pure drift}|_{\delta \Theta = 0} = 0$ (in Fig. 5 (a-d)) to $\delta \eta_n \text{pure drift} \approx -\delta \Theta g_\omega(\bar{\eta}_n) T N_c(T)/N_c(T)$ (see Fig. 5 (e-h)). A prominent feature of the modified drift scheme is the additional intersection with the exact solution (see also Fig. 4 (c) at $\delta \Theta \approx 3.5$), which leads to additional “fingers” of the low error domain, see Fig. 5 (c, d, g, h), that are not associated with any special limiting case. The same feature has been observed for the so-called inverse activity scheme described in Ref. [95]. Finally, we study the consistency of the discretization schemes with the zero current line. In the isothermal case, both schemes are exact and therefore consistent with the thermodynamic equilibrium, see Fig. 5 (a-d). In the strongly degenerate case, however, the zero current line is only partially located within the low error domain, since the schemes intersect with the exact solution only in the vicinity of the zero current line and not exactly on it (see also the inset of Fig. 4 (c)). Nevertheless, the zero current lines of the discrete schemes, which are plotted as dashed orange lines in Fig. 5 (e-h), nicely overlap with the exact zero current line (blue). Thus, the spurious non-zero currents are very low and the little discrepancy in this limiting case is of minor importance.
3.5. Analytical error estimate

We compare both schemes (31) and (35) by deriving an upper error bound. We follow the approach developed by Farrell et al. [95] and extend it to the non-isothermal case. Using the identities

\[ B(x) - B(-x) = -x, \quad B(x) + B(-x) = x \coth \left( \frac{x}{2} \right), \]

we obtain the series expansion of the discrete currents (31) and (35) at \((\delta \Phi, \delta \eta_n, \delta \Theta) = (0, 0, 0)\) up to second order as

\[
J_\delta = -\mathcal{F}'(\eta_n) \delta \Phi + \mathcal{F}(\eta_n) \left( \delta \eta + \frac{\tilde{T} N'_c(T)}{N_c(T)} g_q(\eta_n) \delta \Theta \right) \frac{\tilde{X}_\delta}{2} \coth \left( \frac{\tilde{X}_\delta}{2} \right) + O(\delta^3),
\]

\[
J_\gamma = -\mathcal{F}(\eta_n) \delta \gamma + \mathcal{F}'(\eta_n) \left( \delta \gamma + \frac{\tilde{T} N'_c(T)}{N_c(T)} g_q(\eta_n) \delta \Theta \right) \frac{\tilde{X}_\gamma}{2} \coth \left( \frac{\tilde{X}_\gamma}{2} \right) + O(\delta^3),
\]

where

\[
\tilde{X}_\delta = \frac{1}{g_q(\eta_n)} \delta \Phi, \quad \tilde{X}_\gamma = \delta \Phi - \frac{g_q(\eta_n)}{g_\eta(\eta_n)} \left( \delta \eta + \frac{\tilde{T} N'_c(T)}{N_c(T)} g_q(\eta_n) \delta \Theta \right),
\]

and \(O(\delta^3) \equiv O(\delta \eta_n^3) + O(\delta \eta_n^2 \delta \Theta) + O(\delta \Theta^3) + O(\delta \Phi \delta \eta_n) + O(\delta \Phi \delta \Theta) + O(\delta \Phi \delta \eta_n \delta \Theta) + O(\delta \Phi \delta \Theta^2)\) denotes the third-order corrections. The second-order expansion of the modified drift scheme is independent of the kind of average used for \(\rho_{n,\text{ave}} = \rho_n(\tilde{\eta}_n, \tilde{T}) + O(\delta \eta_n^2) + O(\delta \eta_n \delta \Theta) + O(\delta \Theta^2)\), as only its zeroth-order contribution (where all means coincide) is relevant here. Using the inequality [95]

\[ 1 \leq x \coth(x) \leq 1 + |x|, \]

we arrive at the error estimates for the modified thermal voltage scheme (neglecting third-order terms)

\[
\left| J_\delta - J_1 \right| \leq \frac{1}{2} \mathcal{F}'(\tilde{\eta}_n) \left( |\delta \Phi \delta \eta_n| + \frac{\tilde{T} N'_c(T)}{N_c(T)} g_q(\tilde{\eta}_n)|\delta \Phi \delta \Theta| \right),
\]

(42)

and the modified drift scheme

\[
\left| J_\gamma - J_1 \right| \leq \frac{1}{2} \mathcal{F}'(\tilde{\eta}_n) \left( |\delta \Phi \delta \eta_n| + \frac{\tilde{T} N'_c(T)}{N_c(T)} g_q(\tilde{\eta}_n)|\delta \Phi \delta \Theta| + \frac{g_q(\tilde{\eta}_n)}{g_\eta(\tilde{\eta}_n)} \left| \delta \eta_n + \frac{\tilde{T} N'_c(T)}{N_c(T)} g_q(\tilde{\eta}_n) \delta \Theta \right|^2 \right).
\]

(43)

where \(J_1 = \mathcal{F}(\tilde{\eta}_n) \delta \Phi - \delta \Phi + \frac{\tilde{T} N'_c(T)}{N_c(T)} g_q(\tilde{\eta}_n) \delta \Theta\) is the first-order exact solution of the BVP (41). Both schemes converge to the exact result as their first-order terms agree with \(J_1\). The first two terms of Eq. (43) coincide with Eq. (42). The error bound for the modified drift scheme has an additional second-order contribution that becomes significant in the case of strong degeneration \(\tilde{\eta} \gg 1\) where \((g_q(\tilde{\eta}_n) - 1)/g_q(\tilde{\eta}_n) \to 1\). Therefore, the maximum error of the modified thermal voltage scheme is guaranteed to be smaller than that of the modified drift scheme in the case of degenerate carrier statistics. This analytical result is consistent with the numerical results shown in Figs. 4 and 5 and holds in both the isothermal and the non-isothermal case. For non-degenerate carrier statistics, both error estimates (42) and (43) coincide, since both schemes reduce to the non-degenerate scheme (30).

3.6. Numerical simulation of a p-n-diode

We consider a one-dimensional GaAs-based p-n-diode and compare the convergence of the total current density \((J_{tot} = J_n + J_p)\) under mesh refinement using both discretization schemes. The device consists of a 1\(\mu\)m n-doped section with \(C = N_d^* = 2 \times 10^{19} \text{cm}^{-3}\) followed by a 1\(\mu\)m long p-doped section with \(C = -N_a^* = -2 \times 10^{18} \text{cm}^{-3}\). We use Fermi–Dirac statistics \(\mathcal{F}(\eta) = F_{1/2}(\eta)\) and take Shockley–Read–Hall recombination, spontaneous emission and Auger recombination into account [15, 62]. The material parameters, mobility models (depending on temperature and doping density) and the temperature-dependent heat conductivity model are taken from Ref. [62]. The mobilities and thermal conductivity along the edges are taken as the harmonic average of the respective nodal values. We use Dirichlet boundary conditions on both ends of the diode, modeling ideal Ohmic contacts (charge neutrality at the
boundary) and ideal heat sinks with $T_{\text{contact}} = 300$ K. The simulations are carried out on equidistant grids with varying number of mesh points $N_{\text{nodes}}$ and mesh size $h = 2 \mu m/(N_{\text{nodes}} - 1)$. The nonlinear systems are solved using a Newton iteration method with a fully analytical Jacobian matrix [83].

Figure 6 (a) shows the current-voltage curves obtained by both discretization schemes on a coarse grid (13 nodes, $h \approx 1.7 \times 10^{-7}$ m) under isothermal conditions, i.e., without self-heating and Seebeck effect. For the evaluation of the error, we use a reference solution that was computed on a fine grid with 65535 nodes ($h \approx 3.1 \times 10^{-11}$ m), where the relative error between both schemes is about $9.6 \times 10^{-9}$. At 2 V the computed currents differ significantly from the reference result: The relative error is about 13% for the modified thermal voltage scheme and 15% for the modified drift scheme. The convergence of the computed current densities to the reference result under mesh refinement is shown in Fig. 6 (b). The modified drift scheme (Eq. (35), blue squares) shows a monotonous, quadratic convergence for decreasing $h$. The modified thermal voltage scheme (red circles) converges non-monotonously as it intersects with the reference solution at $h = 4.5$ nm. (c) Convergence of the absolute error of both schemes.

The numerical results for the non-isothermal case, where self-heating and the Seebeck effect are taken into account, are shown in Fig. 7. The results are qualitatively very similar to the isothermal case shown in Fig. 6; quantitatively the advantage of the modified thermal voltage scheme over the modified drift scheme is even greater. On a coarse grid, the total current is underestimated by both schemes with a relative error of about 8%, see Fig. 7 (a). For sufficiently fine meshes ($h < 0.5 \times 10^{-8}$ m), the error of the modified thermal voltage scheme is always more than one order of magnitude smaller than that of the modified drift scheme. Conversely, the modified thermal voltage scheme reaches the same accuracy as the modified drift scheme already on a coarse grid with less than half of the number of nodes. Thus, the modified thermal voltage scheme saves about one refinement step. The convergence of the absolute error is plotted in Fig. 6 (c), where the inset highlights the origin of the non-monotonous convergence behavior of the modified thermal voltage scheme.

4. Summary and conclusion

We discussed the non-isothermal drift-diffusion system for the simulation of electro-thermal transport processes in semiconductor devices. It was shown that the model equations take a remarkably simple form when assuming the Kelvin formula for the Seebeck coefficient. First, the heat generation rate involves exactly the three classically known self-heating effects (Joule heating, recombination heating, Thomson–Peltier effect) without any further transient contributions. Moreover, our modeling approach immediately yields the correct average kinetic energy of the carriers
in the recombination heating term, independently of any scattering parameter. Second, the Kelvin formula enables a simple representation of the electrical current densities in the drift-diffusion form, where the thermal driving force can be entirely absorbed in the (nonlinear) diffusion coefficient via the generalized Einstein relation. The Kelvin formula accounts for the degeneration of the electron-hole plasma (Fermi–Dirac statistics) and was shown to be in a good quantitative agreement with experimental data reported for n-GaAs.

We have derived two non-isothermal generalizations of the finite volume Scharfetter–Gummel scheme for the discretization of the current densities, which differ in their treatment of degeneration effects. The first approach is based on an approximation of the discrete generalized Einstein relation and implies a specific modification of the thermal voltage. The second scheme is based on including the degeneration effects into a modification of the electric field, which is similar to the conventional method that is widely used in commercial device simulation software packages [90, 91]. We presented a detailed analysis of both schemes by assessing their accuracy in comparison to the numerically exact solution of the underlying two-point boundary value problem. Moreover, we derived analytical error bounds and investigated important structure-preserving properties of the discretizations, including the consistency with the thermodynamic equilibrium, the non-negativity of the discrete dissipation rate (second law of thermodynamics on the discrete level) and their asymptotic behavior in the drift- and diffusion-dominated limits. Finally, we performed a numerical convergence study for a simple example case. Our results indicate a significantly higher accuracy and faster convergence of the modified thermal voltage scheme in comparison to the modified drift scheme. This result holds under both isothermal and non-isothermal conditions. The higher accuracy — about one order of magnitude for sufficiently fine grids in the present case study — of the modified thermal voltage scheme makes it a favorable discretization method for problems exhibiting stiff solutions (internal layers at p-n junctions, boundary layers at electrical contacts) or devices with a complicated multi-dimensional geometry, where the number of nodes required to reach the asymptotic accuracy regime is extremely large and routinely exceeds the available computational power.

In more general situations, where the Seebeck coefficient deviates from the Kelvin formula (e.g., due to the phonon drag effect), we suggest to combine the two discretization techniques by decomposing the Seebeck coefficient into a Kelvin formula part and an excess contribution: \( P_n = P_n^{\text{Kelvin}} + P_n^{\text{exc}} \). The first part can be absorbed in the generalized Einstein relation, which allows for the treatment described in Sec. 3.2.1 and inherits the improved accuracy of the modified thermal voltage scheme. The excess part \( P_n^{\text{exc}} \) must be averaged along the edge and plays a similar role as the \( \rho_{n,K,L} \) term in Sec. 3.2.2 (leading to an additive correction in the argument of the Bernoulli function).

**Acknowledgements**

This work was funded by the German Research Foundation (DFG) under Germany’s Excellence Strategy – EXC2046: MATH+ (project AA2-3). The author is grateful to Thomas Koprucki for carefully reading the manuscript and giving valuable comments.
Appendix A. Non-parabolic energy bands

The results presented in this paper hold for Fermi–Dirac statistics with arbitrary density of states, provided that the Kelvin formula for the Seebeck coefficient is applicable. In particular, this includes the case of non-parabolic energy bands. For example, we assume the scalar Kane-type model [28] for the conduction band energy dispersion

\[ E = E_c + \varepsilon_k, \]

\[ (1 + \alpha \varepsilon_k) \varepsilon_k = \frac{\hbar^2 k^2}{2m_c} \]

with the non-parabolicity parameter \( \alpha \geq 0 \). This corresponds to the density of states (in 3D)

\[ D_c(E) = \frac{1}{V} \sum_{k \beta} \delta(E - E(k)) \frac{1}{2\pi^2} \left( \frac{2m_k^{\*}}{\hbar^2} \right)^{3/2} \sqrt{(E - E_c)(1 + \alpha (E - E_c))(1 + 2\alpha (E - E_c))} \Theta(E - E_c), \]

which recovers with the parabolic case for \( \alpha = 0 \). Here, \( \Theta \) is the Heaviside step function. For weak non-parabolicity \( 0 \leq \alpha/\beta \ll 1 \), the corresponding electron density can be expanded in a series of Fermi–Dirac integrals (7)

\[ n = \int_{\mathbb{R}} dE D_c(E) \frac{1}{\exp(\beta [E - \phi + \mu_e]) + 1} = N_c(T) \mathcal{F}(\eta_n), \quad \mathcal{F}(\eta_n) = \sum_{n=0}^{\infty} \frac{1}{n!} \frac{\Gamma \left( \frac{n}{\beta} + 1 \right) \Gamma \left( \frac{n}{\beta} \right)}{\Gamma (n+1)} \left( \frac{\alpha}{\beta} \right)^n \eta_n^n, \]

with \( \eta_n = \beta (\mu_e + q\phi - E_c) \) and \( \beta = (k_B T)^{-1} \), which is clearly of the type (5).

The desired asymptotic behavior in the non-degenerate limit \( \mathcal{F}(\eta \ll -1) \sim \exp(\eta) \) is restored by rescaling

\[ N_c(T) \rightarrow \sqrt{\frac{\beta}{n_\alpha}} \exp \left( \frac{\beta}{2\alpha} K_2 \left( \frac{\beta}{2\alpha} \right) N_c(T) \right), \quad \mathcal{F}(\eta) \rightarrow \left( \sqrt{\frac{\beta}{n_\alpha}} \exp \left( \frac{\beta}{2\alpha} K_2 \left( \frac{\beta}{2\alpha} \right) \right) \right)^{-1} \mathcal{F}(\eta), \]

where \( K_2 \) is the modified Bessel function of second kind.

Appendix B. Generalization to the case of multiple temperatures

The results for the Kelvin formula can be generalized to the multi-temperature case, i.e., to systems describing hot carrier transport where the carrier ensembles have temperatures different from the lattice temperature \( T_n \neq T_L \neq T_p \). We postulate the free energy density of the electron-hole plasma

\[ f_{e-h}(n, p, T_n, T_p, T_L) = k_B T_L \left[ n \mathcal{F}^{-1} \left( \frac{n}{N_c(T_L, T_n)} \right) - N_c(T_n, T_L) \mathcal{F}^{-1} \left( \frac{n}{N_c(T_n, T_L)} \right) \right] + n E_c(T_L) \]

\[ + k_B T_p \left[ p \mathcal{F}^{-1} \left( \frac{p}{N_e(T_p, T_L)} \right) - N_e(T_p, T_L) \mathcal{F}^{-1} \left( \frac{p}{N_e(T_p, T_L)} \right) \right] - p E_e(T_L), \]

where the effective density of states \( N_{e/h} = N_{e/h}(T_{n/p}, T_L) \) is now a function of the carrier temperature and the lattice temperature. In the case of 3D bulk materials, this might be for example \( N_{e/h} = 2 \left( m_{e/h}^{\*}(T_L) k_B T_{n/p} / (2\pi \hbar^2) \right)^{3/2} \), where the effective masses are allowed to depend on the lattice temperature \( T_L \). Moreover, the band gap energy and the band edge energies \( E_{e/h} \) depend solely on \( T_L \). The free energy density of the phonon gas \( f_L = f_L(T_L) \) and the electrostatic interaction energy \( f_{\text{Coul}} = f_{\text{Coul}}(p - n) \) are the same as in Sec. 2.1.

The multi-temperature free energy density \( f(n, p, T_n, T_p, T_L) = f_{e-h}(n, p, T_n, T_p, T_L) + f_{\text{Coul}}(p - n) + f_L(T_L) \) is a thermodynamic potential, that yields expressions for the chemical potentials

\[ + \mu_e = \frac{\partial f}{\partial n} = k_B T_n \mathcal{F}^{-1} \left( \frac{n}{N_c(T_n, T_L)} \right) + E_c(T_L) - q\phi \quad \text{(B.2a)} \]

\[ - \mu_e = \frac{\partial f}{\partial p} = k_B T_p \mathcal{F}^{-1} \left( \frac{p}{N_e(T_p, T_L)} \right) - E_e(T_L) + q\phi \quad \text{(B.2b)} \]
and the entropy densities

\[ s_n = -\frac{\partial f}{\partial T_n}, \quad s_p = -\frac{\partial f}{\partial T_p}, \quad s_L = -\frac{\partial f}{\partial T_L}. \]

The state equations for the carrier densities follow from Eq. (B.2) as

\[ n = N_c(T_n, T_L) \mathcal{R} \left( \frac{\mu_c + q \phi - E_v(T_L)}{k_B T_n} \right), \quad p = N_c(T_p, T_L) \mathcal{R} \left( \frac{E_v(T_L) - q \phi - \mu_c}{k_B T_p} \right). \]

In analogy to Eq. (8), we postulate the generalized electrical current density expressions

\[ j_n = -\sigma_n (\nabla \varphi_n + P_{n,n} \nabla T_n + P_{L,n} \nabla T_L), \quad j_p = -\sigma_p \left( \nabla \varphi_p + P_{p,p} \nabla T_p + P_{L,p} \nabla T_L \right), \tag{B.3} \]

with the thermopowers given by the Kelvin formula

\[ P_{n,n} = -\frac{1}{q} \frac{\partial s_n}{\partial n}, \quad P_{L,n} = -\frac{1}{q} \frac{\partial s_L}{\partial n}, \quad P_{p,p} = +\frac{1}{q} \frac{\partial s_p}{\partial p}, \quad P_{L,p} = +\frac{1}{q} \frac{\partial s_L}{\partial p}. \tag{B.4} \]

Here, we do not describe the corresponding hot carrier transport system in full detail, but rather focus on the central property of the electrical current densities, which is the simple representation in the drift-diffusion form (without thermodiffusion). If the thermal driving forces can be absorbed in a generalized Einstein relation for the diffusion coefficient, this enables a generalized Scharfetter–Gummel discretization of type (31) \textit{(modified thermal voltage scheme)}. With the definitions given above, the drift-diffusion form of Eq. (B.3) reads (along the lines in Sec. 2.4)

\[ j_n = -\sigma_n \left( -\frac{1}{q} \nabla \mu_c + P_{n,n} \nabla T_n + P_{L,n} \nabla T_L \right) \approx -\sigma_n \left( -\frac{1}{q} \nabla \frac{\partial f}{\partial n} + P_{n,n} \nabla T_n + P_{L,n} \nabla T_L \right), \]

(holes analogously). As a result, the generalized Einstein relation – which now yields a nonlinear diffusion coefficient that depends on both the lattice and the carrier temperature – allows to absorb the thermal driving forces also in the multi-temperature case, if the respective Kelvin formulæ (B.4) for the Seebeck coefficients are assumed.

Appendix C. Discretization of the heat source term

This section explains the finite volume discretization of the heat source term (27) entering the discrete heat equation (24d) in more detail. First, the discrete recombination heating term is obtained by integration over the \( K \)-th Voronoï cell as

\[ H_{R,K} = \frac{1}{|\Omega_K|} \int_{\Omega_K} dV q (\varphi_p - \varphi_n + \Pi_p - \Pi_n) R \approx q (\varphi_{p,K} - \varphi_{n,K} + T_K (P_{p,K} - P_{n,K})) R_K, \]

using the local values of the quasi-Fermi potentials, temperature, Seebeck coefficients and \( R_K = R(\varphi_K, \varphi_{n,K}, \varphi_{p,K}, T_K) \).

The discretization of the Joule and Thomson–Peltier heating terms is more involved. The continuous expressions

\[ H_J = \sum_{\Delta(n;p), \Delta(l;\alpha)} \frac{1}{2} |\mathbf{j}_l|_2^2 = - \sum_{\Delta(n;p), \Delta(l;\alpha)} \mathbf{j}_l \cdot (\nabla \varphi_x + P_J \nabla T), \quad H_{T,P} = -T \sum_{\Delta(n;p)} \mathbf{j}_l \cdot \nabla P_J. \]
are all of the type $\tilde{H} = \kappa (j, \nabla \psi)$, where $j$ is an electrical current density vector and $\nabla \psi$ is the gradient of a particular scalar field with $\psi \in \{\varphi, T, P\}$ and $\kappa \in \{1, P, T\}$ for $\lambda \in \{n, p\}$. The essential problem in the discretization of $\tilde{H}$ is that the full vector field $j$, and the full gradient $\nabla \psi$ are required on the cell $\Omega_K$, whereas naturally only approximations of the normal components are available. There are several methods to tackle this problem, see Ref. [86] for a discussion. Here, we follow the approach suggested by Eymard & Gallouët [88], who introduced a weakly converging gradient. The discrete expressions (27a) and (27b) are obtained along the following lines:

We seek for a finite volume discretization of the term $\tilde{H} = \kappa (j, \nabla \psi)$ by integration over the $K$-th cell

$$\int_{\Omega_K} dV \tilde{H} = \frac{1}{2} \sum_{L \in N(K)} \int_{D_{KL}} dV \tilde{H} = \frac{1}{2} \sum_{L \in N(K)} |D_{KL}| \kappa_{KL} (j, \nabla \psi)|_{KL},$$

where the integral was recast in an integral over the adjacent bi-hyperpyramids $D_{KL}$ (see Fig. 3) with volumes

$$|D_{KL}| = \frac{1}{d} ||\mathbf{r}_L - \mathbf{r}_K|| \text{area} K \cap \partial \Omega_L| = \frac{1}{d} ||\mathbf{r}_L - \mathbf{r}_K||^2 s_{KL}. \quad (C.1)$$

Here, $d$ is the dimensionality of the computational domain and $s_{KL}$ is the edge factor (26). Moreover, $\kappa_{KL}$ is a suitably chosen average of $\kappa$ on the domain $D_{KL}$ (see below). Following Eymard & Gallouët [88, Definition 2], we approximate $\nabla \psi$ by the weakly converging discrete gradient

$$\nabla \psi|_{KL} \approx \frac{d}{\mathbf{n}_{KL} \cdot (\mathbf{r}_L - \mathbf{r}_K)} (\psi_L - \psi_K), \quad (C.2)$$

where $\mathbf{n}_{KL} = (\mathbf{r}_L - \mathbf{r}_K)/||\mathbf{r}_L - \mathbf{r}_K||$ is the (outward-oriented) normal vector and $d \in \{1, 2, 3\}$ is the dimensionality of the domain. The technique has been adopted for the discretization of Joule heating in similar models [86, 100, 101] and for the computation of contact currents [83]. Finally, we combine Eqs. (C.1), (C.2) and the normal projection of the current $J_{i,K,L} = (\mathbf{r}_L - \mathbf{r}_K) \cdot j_i$ (see Eq. (25), where $J_{i,K,L}$ is given by a Scharfetter–Gummel type discretization scheme) and obtain

$$\int_{\Omega_K} dV \tilde{H} \approx \frac{1}{2} \sum_{L \in N(K)} s_{KL} \kappa_{KL} J_{i,K,L} (\psi_L - \psi_K),$$

cf. Eqs. (27a)–(27b). In order to ensure the non-negativity of the discrete Joule heating and dissipation rate (see Sec. 3.3.5), $\kappa_{KL}$ is taken as the logarithmic mean temperature (29) or the edge-averaged Seebeck coefficient (38), respectively.

Appendix D. Power balance

The balance equation for the total power is derived by integrating the heat transport Eq. (4) over the domain $\Omega$:

$$\int_{\Omega} dV c_V \partial_t T - \int_{\Omega} dV \nabla \cdot k \nabla T = \int_{\Omega} dV H.$$

Using Eqs. (2)–(3) and partial integration, the expression for the heat generation rate (21) is rearranged as

$$H = -\nabla \cdot \left( (\varphi_n + \Pi_n) j_n + (\varphi_p + \Pi_p) j_p \right) + q (\varphi_n + \Pi_n) \partial_n \varphi_n - q (\varphi_p + \Pi_p) \partial_p \varphi_p.$$

The internal energy density of the system reads $u(n, p, T) = f(n, p, T) + T s(n, p, T)$, where the free energy density is given in Eq. (13). Using the defining relations for the quasi-Fermi potentials and the entropy density (11) as well as the Kelvin formula (10), the differential internal energy per carrier is obtained as

$$\frac{\partial u}{\partial n} = \frac{\partial f}{\partial n} + T \frac{\partial s}{\partial n} = -q (\varphi_n + \Pi_n), \quad \frac{\partial u}{\partial p} = \frac{\partial f}{\partial p} + T \frac{\partial s}{\partial p} = +q (\varphi_p + \Pi_p),$$

where the Peltier coefficients $\Pi_{n/p}$ are defined via the Kelvin relation $\Pi_{n/p} = T \Pi_{n/p}$. Using the definition of the heat capacity $c_V = \partial_T u$, the (integrated) heat transport equation is rearranged as a balance equation for the internal energy:

$$\int_{\Omega} dV \left( \frac{\partial u}{\partial t} + \frac{\partial u}{\partial n} \partial_n + \frac{\partial u}{\partial p} \partial_p \right) = \frac{d}{dt} \int_{\Omega} dV u = -\int_{\partial \Omega} d\Sigma n \cdot (\kappa \nabla T + (\varphi_n + \Pi_n) j_n + (\varphi_p + \Pi_p) j_p). \quad (D.1)$$
In the following, we will recast the surface integral on the right hand side into an expression for the injected electrical power and the dissipated heat.

The local continuity equation for the internal energy (first law of thermodynamics, cf. Eq. (D.1)) reads

\[ \partial_t u + \nabla \cdot j_u = 0, \quad j_u = -\kappa \nabla T + (\varphi_n + \Pi_n) j_n + (\varphi_p + \Pi_p) j_p, \quad (D.2) \]

where \( j_u \) is the internal energy flux density. The corresponding continuity equation for the entropy density is obtained from Gibb’s fundamental thermodynamic relation \( \text{d} s = T \text{d}s + \mu_e \text{d} n - \mu_i \text{d} p \). By substituting Eqs. (2)–(3) and (D.2), one arrives at

\[ \partial_t s = \frac{1}{T} \partial_t u - \frac{\mu_e}{T} \partial_t n + \frac{\mu_i}{T} \partial_t p = -\nabla \cdot j_s + \dot{s}_{\text{int}}, \quad j_s = \frac{1}{T} \left( j_u - \varphi_n j_n - \varphi_p j_p \right), \quad (D.3) \]

where \( j_s \) is the entropy flux density. The entropy production rate reads (cf. Sec. 3.3.5)

\[ \dot{s}_{\text{int}} = \frac{1}{T} (\mu_e - \mu_i) R + \frac{\kappa}{T^2} ||\nabla T||^2 + \frac{1}{T} H_J, \quad (D.4) \]

Here, \( H_J = j_u \cdot (\nabla \varphi_n + \Pi_n \nabla T) + j_p \cdot (\nabla \varphi_p + \Pi_p \nabla T) \) is the Joule heat term. The entropy flux density is closely connected with the heat flux density

\[ j_Q = T j_s = j_u - \varphi_n j_n - \varphi_p j_p = -\kappa \nabla T + \Pi_n j_n + \Pi_p j_p. \quad (D.5) \]

We return to the power balance equation (D.1) and rewrite the surface integral on the right hand side as

\[ \oint_{\partial \Omega} dA \cdot j_Q = \oint_{\partial \Omega} dA \cdot j_s + \oint_{\partial \Omega} dA \cdot (\varphi_n j_n + \varphi_p j_p). \]

The first term is the heat flux density leaving the device. The second term is evaluated for a device with mixed boundary conditions \( \partial \Omega = \Gamma_D \cup \Gamma_N \), where the segments \( \Gamma_D = \bigcup_i \Gamma_{D,i} \) are the electrical contacts with (ideal) Ohmic boundary conditions \( \varphi_n = \varphi_p = \varphi_{D,i} = \text{const.} \) on \( \Gamma_{D,i} \) (Dirichlet conditions). The remaining facets \( \Gamma_N \) are semiconductor-insulator interfaces or artificial boundaries with “no-flux” boundary conditions \( n \cdot j_{n,p} = 0 \). One obtains

\[ \oint_{\partial \Omega} dA \cdot (\varphi_n j_n + \varphi_p j_p) = \sum_i (\varphi_{D,i} \oint_{\Gamma_{D,i}} dA \cdot j_n + \varphi_{p,i} \oint_{\Gamma_{D,i}} dA \cdot j_p) = \sum_i \varphi_{D,i} I_i, \]

where \( I_i = \oint_{\Gamma_{D,i}} dA \cdot (j_n + j_p) \) is the electrical current flux across the \( i \)-th electrical contact. For a two-terminal device with total current \( I = I_2 = -I_1 \) (Kirchhoff’s current law) and applied voltage \( U = \varphi_{D,1} - \varphi_{D,2} \), this is (cf. Ref. [12])

\[ \oint_{\partial \Omega} dA \cdot (\varphi_n j_n + \varphi_p j_p) = I_1 \varphi_{D,1} + I_2 \varphi_{D,2} = -UI. \]

Thus, at stationary conditions, we finally obtain the power balance equation as

\[ UI = \oint_{\partial \Omega} dA \cdot j_Q. \quad (D.6a) \]

In conclusion, the injected electrical power \( UI \) is equal to the heat flux density leaving the device. With the divergence of the heat flux density \( \nabla \cdot j_Q = H + \nabla \cdot (\Pi_n j_n + \Pi_p j_p) \), this can also be written as

\[ UI = \int_\Omega dV H + \int_{\Gamma_{D}} dA n \cdot (\Pi_n j_n + \Pi_p j_p), \quad (D.6b) \]

where the first term is the total heat generated on the full domain and the second term describes the “Peltier power”, that can be either positive or negative, depending on the direction of the current flow. In the case of optoelectronic devices (with open optical cavities), the power balance must be supplemented by additional terms describing the emitted radiation power [29]. The power balance equation (D.6) is consistent with the results previously reported in the literature, i.e., the result persists when assuming the Kelvin formula for the Seebeck coefficient.
References

[1] J. Lutz, H. Schlangenotto, U. Scheuermann, R. D. Donker, Semiconductor Power Devices — Physics, Characteristics, Reliability, Springer, Berlin, Heidelberg, 2011. doi:10.1007/978-3-642-11125-9.

[2] H.-J. Schulze, F.-J. Niedernostheide, F. Pfitzsch, R. Baburske, Limiting factors of the safe operating area for power devices, IEEE Trans. Electron Devices 60 (2012) 551–562. doi:10.1109/ted.2012.2225148.

[3] A. Fischer, T. Kopruckki, K. Gätter, M. L. Tietz, J. Brückner, B. Lüssem, K. Leo, A. Glitzky, R. Scholz, Feel the heat: Nonlinear electrothermal feedback in organic LEDs, Adv. Funct. Mater. 24 (2014) 3367–3374. doi:10.1002/adfm.201303066.

[4] A. Fischer, M. Pfalz, K. Vandelau, S. Lenk, M. Liéro, A. Glitzky, S. Reineke, Full electrothermal OLEn model including nonlinear self-heating effects, Phys. Rev. Appl 10 (2018) 014023. doi:10.1103/physrevapplied.10.014023.

[5] M. Osifski, W. Nakwaski, Thermal effects in vertical-cavity surface-emitting lasers, Int. J. High Speed Electron. Syst. 5 (1994) 667–730. doi:10.1142/s0129156494000267.

[6] J. Piprek, J. K. White, A. J. SpringThorpe, What limits the maximum output power of long-wavelength AlGaN/InP laser diodes?, IEEE J. Quantum Electr. 38 (2002) 1253–1259. doi:10.1109/jqe.2002.802441.

[7] M. Streifl, W. Fichtner, A. Witzig, Vertical-Cavity Surface-Emitting Lasers: Single-Mode Control and Self-Heating Effects, Springer, New York, 2005, pp. 217–247. doi:10.1007/0-387-27256-9_8.

[8] H. Wenzel, P. Cump, A. Pietrzak, X. Wang, G. Erbert, G. Tränkle, Theoretical and experimental investigations of the limits to the maximum output power of laser diodes, New J. Phys. 12 (2010) 085007. doi:10.1088/1367-2630/12/8/085007.

[9] G. K. Wachutka, Rigorous thermodynamic treatment of heat generation and conduction in semiconductor device modeling, Trans. Comput.-Aided Design Integr. Circuits Syst. 9 (1990) 1141–1149. doi:10.1109/43.62751.

[10] U. Lindefelt, Heat generation in semiconductor devices, J. Appl. Phys. 75 (1994) 942–957. doi:10.1063/1.356450.

[11] H. Brand, S. Selberherr, Two-dimensional simulation of thermal runaway in a nonplanar GTO-thyristor, IEEE Trans. Electron Devices 42 (1995) 2137–2146. doi:10.1109/16.477772.

[12] J. E. Parrott, Thermodynamic theory of transport processes in semiconductors, IEEE Trans. Electron Devices 43 (1996) 809–826. doi:10.1080/11191564.9783527338405.CH1.

[13] G. Albinus, H. Gajewski, R. Hünlich, Thermodynamic design of energy models of semiconductor devices, Nonlinearity 15 (2002) 367–383. doi:10.1088/0951-7715/15/2/307.

[14] U. Bandelow, H. Gajewski, R. Hünlich, Fabry–Perot lasers: Thermodynamics-based modeling, Springer, New York, 2005, pp. 63–85. doi:10.1007/0-387-27256-9_3.

[15] S. Selberherr, Analysis and Simulation of Semiconductor Devices, Springer, Vienna, 1984. doi:10.1007/978-3-7091-8752-4.

[16] P. A. Markowich, The stationary Semiconductor device equations, Springer, Vienna, 1986. doi:10.1007/978-3-7091-3678-2.

[17] H. J. Goldsmid, Introduction to Thermoelectricity, volume 121 of Springer Series in Materials Science.

[18] H. Brand, S. Selberherr, Two-dimensional simulation of thermal runaway in a nonplanar GTO-thyristor, IEEE Trans. Electron Devices 42 (1995) 2137–2146. doi:10.1109/16.477772.

[19] J. C. Freeman, Self-heating in semiconductors: A comparative study, Solid State Electron. 95 (2014) 8–14. doi:10.1007/978-3-642-00716-3.

[20] M. Osiński, W. Nakwaski, Thermal e.

[21] H. Wenzel, P. Cump, A. Pietrzak, X. Wang, G. Erbert, G. Tränkle, Theoretical and experimental investigations of the limits to the maximum output power of laser diodes, New J. Phys. 12 (2010) 085007. doi:10.1088/1367-2630/12/8/085007.

[22] G. K. Wachutka, Rigorous thermodynamic treatment of heat generation and conduction in semiconductor device modeling, IEEE Trans. Comput.-Aided Design Integr. Circuits Syst. 9 (1990) 1141–1149. doi:10.1109/43.62751.

[23] U. Lindefelt, Heat generation in semiconductor devices, J. Appl. Phys. 75 (1994) 942–957. doi:10.1063/1.356450.

[24] H. Brand, S. Selberherr, Two-dimensional simulation of thermal runaway in a nonplanar GTO-thyristor, IEEE Trans. Electron Devices 42 (1995) 2137–2146. doi:10.1109/16.477772.

[25] J. E. Parrott, Thermodynamic theory of transport processes in semiconductors, IEEE Trans. Electron Devices 43 (1996) 809–826. doi:10.1080/11191564.9783527338405.CH1.

[26] G. Albinus, H. Gajewski, R. Hünlich, Thermodynamic design of energy models of semiconductor devices, Nonlinearity 15 (2002) 367–383. doi:10.1088/0951-7715/15/2/307.

[27] U. Bandelow, H. Gajewski, R. Hünlich, Fabry–Perot lasers: Thermodynamics-based modeling, Springer, New York, 2005, pp. 63–85. doi:10.1007/0-387-27256-9_3.

[28] S. Selberherr, Analysis and Simulation of Semiconductor Devices, Springer, Vienna, 1984. doi:10.1007/978-3-7091-8752-4.

[29] P. A. Markowich, The stationary Semiconductor device equations, Springer, Vienna, 1986. doi:10.1007/978-3-7091-3678-2.

[30] H. J. Goldsmid, Introduction to Thermoelectricity, volume 121 of Springer Series in Materials Science.

[31] M. Osiński, W. Nakwaski, Thermal e.

[32] M. R. Peterson, B. S. Shastry, Kelvin formula for thermopower, Phys. Rev. B 82 (2010) 195105. doi:10.1103/physrevb.82.195105.

[33] B. S. Shastry, Electrothermal transport coefficients at finite frequencies, Rep. Prog. Phys. 72 (2009) 085007. doi:10.1088/0951-7715/15/2/307.

[34] J. C. Freeman, Self-heating in semiconductors: A comparative study, Solid State Electron. 95 (2014) 8–14. doi:10.1007/978-3-642-00716-3.

[35] M. Osiński, W. Nakwaski, Thermal e.

[36] M. R. Peterson, B. S. Shastry, Kelvin formula for thermopower, Phys. Rev. B 82 (2010) 195105. doi:10.1103/physrevb.82.195105.

[37] B. S. Shastry, Electrothermal transport coefficients at finite frequencies, Rep. Prog. Phys. 72 (2009) 085007. doi:10.1088/0951-7715/15/2/307.
