Exact closure relations for the maximum entropy moment system in semiconductor using Kane's dispersion relation

Michael Junk[1] and Vittorio Romano[2]

[1] FB Mathematik, Universität des Saarlandes, Postfach 151150, 66041 Saarbrücken, Germany junk@num.uni-sb.de
[2] Dipartimento di Matematica e Informatica, Università di Catania, viale A. Doria 6 - 95125 Catania, Italy romano@dmi.unict.it

Summary. The maximum entropy moment systems of the Boltzmann equation is only solvable with unphysical restrictions on the choice of the macroscopic variables. We show that no such difficulties appear in the semiconductor case if Kane's dispersion relation is used for the energy band of electrons. As an application the 5-moment model is discussed.

Key words: maximum entropy moment closure, semiconductor Boltzmann equation, Kane's dispersion relation

1 The maximum entropy moment systems for electrons in semiconductors

In a semi classical approximation, a kinetic description of electrons in a semiconductor is given by a transport equation for the one particle distribution function $f(t, \boldsymbol{x}, \boldsymbol{k})$, which represents the probability of finding an electron at time t in an elementary volume $d\boldsymbol{x}d\boldsymbol{k}$, around position $\boldsymbol{x}$ and with crystal momentum $\boldsymbol{k}$,

$$\frac{\partial f}{\partial t} + v_i(\boldsymbol{k})\frac{\partial f}{\partial x_i} - \frac{e}{\hbar}E_i\frac{\partial f}{\partial k_i} = \mathcal{C}[f]. \tag{1}$$

Here e is the absolute value of the electron charge, $\boldsymbol{k}$ represents the crystal momentum of the electron and $\boldsymbol{E}$ is the electric field which is related to the electron distribution by Poisson's equation:

$$\boldsymbol{E} = -\nabla\phi, \qquad \epsilon\Delta\phi = -e(N_D - N_A - n),$$

where ϕ is the electric potential, ϵ is the permittivity of the semiconductor, N_D and N_A are respectively the donor and acceptor density, and n is the electron density. The latter is related to f by $n = \int_B f d\boldsymbol{k}$, B being the first Brillouin zone. The right hand side $\mathcal{C}[f]$ in (1) is the collision operator, which takes into account scattering of the electrons with acoustical and optical phonons and with impurities. The electron velocity $\boldsymbol{v}(\boldsymbol{k})$ depends on the electron energy $\mathcal{E}$ by the relation $\boldsymbol{v}(\boldsymbol{k}) = \frac{1}{\hbar}\nabla_{\boldsymbol{k}}\mathcal{E}$. In general, the expression of $\mathcal{E}$ (the so called band structure) depends on the material and is very complicated. A rough approximation is given by the *parabolic band* while a more refined model is given by *Kane's dispersion relation* which takes into account the non-parabolicity at high energies

$$\mathcal{E}(\boldsymbol{k}) = \frac{1}{1 + \sqrt{1 + 2\frac{\alpha}{m^*}\hbar^2|\boldsymbol{k}|^2}}\frac{\hbar^2|\boldsymbol{k}|^2}{m^*} = \sqrt{\frac{1}{4\alpha^2} + \frac{\hbar^2|\boldsymbol{k}|^2}{2\alpha m^*}} - \frac{1}{2\alpha}, \qquad \boldsymbol{k} \in \mathbb{R}^3 \tag{2}$$

where α is the non-parabolicity parameter. The corresponding electron velocity is

$$\boldsymbol{v}(\boldsymbol{k}) = \frac{1}{\sqrt{1+\frac{2\alpha}{m^*}\hbar^2|\boldsymbol{k}|^2}}\frac{\hbar}{m^*}\boldsymbol{k}.$$

In the mathematical modelling of electron transport in semiconductors the Kane dispersion relation is considered as one of the best analytical approximation to the real energy band.

Besides the electron density n, other physically relevant quantities are the average electron velocity, energy and energy-flux

$$\boldsymbol{u} = \frac{1}{n}\int_{\mathbb{R}^3}\boldsymbol{v}(\boldsymbol{k})f\,d\boldsymbol{k}, \quad W = \frac{1}{n}\int_{\mathbb{R}^3}\mathcal{E}(\boldsymbol{k})f\,d\boldsymbol{k}, \quad \boldsymbol{S} = \frac{1}{n}\int_{\mathbb{R}^3}\boldsymbol{v}(\boldsymbol{k})\mathcal{E}(\boldsymbol{k})f\,d\boldsymbol{k}.$$

To generalize this observation, we introduce general weight functions $a_i : \mathbb{R}^3 \mapsto \mathbb{R}$ and the corresponding moments $\rho_i = \langle f, a_i\rangle$, $i = 1,\ldots,m$ where $\langle\cdot,\cdot\rangle$ denotes $\boldsymbol{k}$ integration. We split the vector of weight functions $\boldsymbol{a}$ into two subgroups. The first m_1 components of $\boldsymbol{a}$ are chosen as $(P_1(\boldsymbol{v}(\boldsymbol{k})),\ldots,P_{m_1}(\boldsymbol{v}(\boldsymbol{k})))$ where $P_1,\ldots,P_{m_1}$ are linearly independent polynomials with $P_1(\boldsymbol{v}) = 1$, and the remaining m_2 components give rise to energy moments $(\mathcal{E}(\boldsymbol{k})Q_1(\boldsymbol{v}(\boldsymbol{k})),\ldots,\mathcal{E}(\boldsymbol{k})Q_{m_2}(\boldsymbol{v}(\boldsymbol{k})))$ where, again, $Q_1,\ldots,Q_{m_2}$ are linearly independent polynomials and $Q_1(\boldsymbol{v}) = 1$.

Since the direct numerical approximation of the kinetic equation (1) is very expensive due to the high dimensionality of the problem, and in view of the fact that one is rather interested in moments of f than in f itself, it is a natural idea to derive equations directly for the averaged quantities. Multiplying (1) with weight functions $\boldsymbol{a} = (a_1,\ldots,a_m)^T$ and integrating over $\boldsymbol{k}$, we obtain equations for the moments

$$\frac{\partial\boldsymbol{\rho}}{\partial t} + \frac{\partial}{\partial x_j}\langle f, v_j\boldsymbol{a}\rangle = \langle\mathcal{C}[f] + \gamma\boldsymbol{E}\cdot\nabla_{\boldsymbol{k}}f, \boldsymbol{a}\rangle, \qquad \gamma = e/\hbar. \tag{3}$$

The system would be closed if the particle distribution could be expressed in terms of the moment vector $\boldsymbol{\rho}$ as $f(t,\boldsymbol{x},\boldsymbol{k}) = F(\boldsymbol{\rho}(t,\boldsymbol{x}),\boldsymbol{k})$ A method to obtain such a relationship is the maximum entropy approach where $F(\boldsymbol{\rho},\boldsymbol{k})$ is taken as solution of the problem

$$\text{maximize } H(f) = -\langle f, \log f - 1\rangle \text{ with } f \geq 0 \text{ and } \langle f, \boldsymbol{a}\rangle = \boldsymbol{\rho} \tag{4}$$

It is important to remark that the maximum entropy distribution represents, in a statistical sense [Jaynes (1957)], the least biased estimator of the exact distribution f on the base of the knowledge of a finite number of moments of f.

For general a_i, the formal solution of (4) is obtained with the method of Lagrange multipliers. We introduce the Lagrange functional $L(f,\boldsymbol{\lambda}) := H(f) - \boldsymbol{\lambda}\cdot(\boldsymbol{\rho} - \langle f, \boldsymbol{a}\rangle)$ where $\boldsymbol{\lambda}$ is the vector of Lagrange multipliers. The necessary condition that all directional derivatives vanish in the maximum $f_{\boldsymbol{\lambda}}$ leads to

$$f_{\boldsymbol{\lambda}} = \exp(\boldsymbol{\lambda}\cdot\boldsymbol{a}). \tag{5}$$

Finally, the Lagrange multipliers $\boldsymbol{\lambda}$ are chosen in such a way (if possible) that the moment constraints $\boldsymbol{\rho} = \langle f_{\boldsymbol{\lambda}}, \boldsymbol{a}\rangle$ are satisfied which gives rise to a function $\boldsymbol{\lambda} = \boldsymbol{\lambda}(\boldsymbol{\rho})$. We then introduce $F(\boldsymbol{\rho},\boldsymbol{k}) = f_{\boldsymbol{\lambda}(\boldsymbol{\rho})}(\boldsymbol{k})$.

Depending on the choice of weight functions a_i, it can happen that problem (4) is not always solvable, i.e. that there exist moment vectors $\boldsymbol{\rho}$ which cannot be written as $\boldsymbol{a}$-moments of any exponential density $f_{\boldsymbol{\lambda}} = \exp(\boldsymbol{\lambda}\cdot\boldsymbol{a})$ [Junk (1997), Junk (1998), Dreyer et al (2001) , Junk et al (2002)]. Our main goal is to show that, when Kane's model is considered for the energy band, problem (4) is always solvable which implies that the corresponding moment system is a symmetric hyperbolic system with an open and convex domain of definition.

2 Solvability of the maximum entropy problem

In order to state our main result, we first reformulate (4). For notational convenience, we measure $\mathcal{E}$, $\boldsymbol{k}, \boldsymbol{v}$ in units $1/(2\alpha)$, $\sqrt{m^*/(2\alpha\hbar^2)}$, and $1/\sqrt{2\alpha m^*}$ which leads to

$$\mathcal{E}(\boldsymbol{k}) = \sqrt{1+|\boldsymbol{k}|^2} - 1, \qquad \boldsymbol{v}(\boldsymbol{k}) = \frac{\boldsymbol{k}}{\sqrt{1+|\boldsymbol{k}|^2}}. \tag{6}$$

Note that for large $\boldsymbol{k}$, $\boldsymbol{v}(\boldsymbol{k})$ is bounded and $\mathcal{E}(\boldsymbol{k})$ grows only linearly due to the estimates

$$|\boldsymbol{v}(\boldsymbol{k})| < 1, \qquad |\boldsymbol{k}| - 1 \le \mathcal{E}(\boldsymbol{k}) \le 2|\boldsymbol{k}| + 1. \tag{7}$$

Based on $\mathcal{E}$ and $\boldsymbol{v}$ and two sets $\{P_1, \dots, P_{m_1}\}$, $\{Q_1, \dots, Q_{m_2}\}$ of linearly independent polynomials with $P_1 = Q_1 = 1$, we define the weight functions as

$$\boldsymbol{a} = (P_1(\boldsymbol{v}), \dots, P_{m_1}(\boldsymbol{v}), \mathcal{E}Q_1(\boldsymbol{v}), \dots, \mathcal{E}Q_{m_2}(\boldsymbol{v}))^T. \tag{8}$$

Since the assumption of a three dimensional $\boldsymbol{k}$-space is not relevant for our argument, we assume $\boldsymbol{k} \in \mathbb{R}^d$. The moment set related to the weights a_i is generated by the functions in $\mathcal{F} = \{f \ge 0 : f \not\equiv 0, |\boldsymbol{a}|f \in \mathbb{L}^1(\mathbb{R}^d)\}$. The corresponding moments are collected in $\mathcal{M} = \{\langle f, \boldsymbol{a}\rangle : f \in \mathcal{F}\}$. Using this notation and the definition of the entropy functional $H(f) = -\langle f, \log f - 1\rangle$, we can restate (4) as

$$\text{maximize } H(f) \text{ subject to } f \in \mathcal{F} \text{ and } \langle f, \boldsymbol{a}\rangle = \boldsymbol{\rho} \tag{9}$$

Our main result is

Theorem 1. *The maximum entropy moment problem* (9) *is uniquely solvable for any* $\boldsymbol{\rho}$ *inside the open, convex cone* $\mathcal{M}$*. The solution is an exponential density* $\exp(\boldsymbol{\lambda}\cdot\boldsymbol{a})$ *for some* $\boldsymbol{\lambda} \in \mathbb{R}^m$ *depending on* $\boldsymbol{\rho}$*.*

First we observe that, up to normalization, every $f \in \mathcal{F}$ can be viewed as a probability density. The normalization $f^* = f/\langle f, 1\rangle$ is abbreviated by a $*$-superscript and its image of $\mathcal{F}$ is denoted $\mathcal{F}^*$. Since we assume $a_1 = 1$, the moment vector of f^* has the structure $\langle f^*, \boldsymbol{a}\rangle = (1, \rho_2/\rho_1, \dots, \rho_m/\rho_1)^T$, $\boldsymbol{\rho} = \langle f, \boldsymbol{a}\rangle$, which gives rise to a normalization operation acting on vectors in $\mathbb{R}^m$ by $\boldsymbol{\alpha}^* = (\alpha_2/\alpha_1, \dots, \alpha_m/\alpha_1)^T$, $\boldsymbol{\alpha} \in \mathbb{R}^m$, $\alpha_1 > 0$.

Note that $\boldsymbol{a}^* = (a_2, \dots, a_m)^T$ because $a_1 = 1$ and thus $\langle f, \boldsymbol{a}\rangle = \boldsymbol{\rho}$ implies $\langle f^*, \boldsymbol{a}^*\rangle = \boldsymbol{\rho}^*$. Apart from the passage to probability measures, we need to introduce the functional of relative entropy. If P and R are probability measures on the Borel sets $\mathcal{B}$ on $\mathbb{R}^d$, such that P has a density with respect to R, i.e. $P(A) = \int_A p_R\, dR$ with $A \in \mathcal{B}$, the relative entropy (or I-divergence) is defined as $I(P||R) = \int p_R \log p_R\, dR$. As measure R we are going to use $R(A) = \int_A g^*\, d\boldsymbol{k}$, $g(\boldsymbol{k}) = \exp(-\mathcal{E}(\boldsymbol{k}))$ where g is integrable since $\mathcal{E}(\boldsymbol{k})$ grows linearly (see (7)). Then, if P_{f^*} has density $f^* \in \mathcal{F}^*$ with respect to the Lebesgue measure, it has density f^*/g^* with respect to R and

$$I(P_{f^*}||R) = \int \frac{f^*}{g^*} \log \frac{f^*}{g^*}\, dR = \int f^* \log \frac{f^*}{g^*}\, d\boldsymbol{k}$$

Using the definition of H and $\log f^* = \log f - \log\langle f, 1\rangle$, $\quad \log g^* = -\mathcal{E} - \log\langle g, 1\rangle$, we obtain the relation

$$I(P_{f^*}||R) = -\frac{1}{\langle f, 1\rangle} H(f) + 1 + \log\frac{\langle g, 1\rangle}{\langle f, 1\rangle} + \frac{\langle f, \mathcal{E}\rangle}{\langle f, 1\rangle}. \tag{10}$$

Since $\langle f, 1\rangle$ and $\langle f, \mathcal{E}\rangle$ are constant on the set of densities $f \in \mathcal{F}$ with $\langle f, \boldsymbol{a}\rangle = \boldsymbol{\rho}$, we see that maximizing H subject to $\langle f, \boldsymbol{a}\rangle = \boldsymbol{\rho}$ is equivalent to

$$\text{minimize } I(P_{f^*}||R) \quad \text{subject to } f^* \in \mathcal{F}^* \text{ and } \langle f^*, \boldsymbol{a}^*\rangle = \boldsymbol{\rho}^* \tag{11}$$

Therefore the main theorem can be reformulated as follows.

Proposition 1. *Let $\boldsymbol{\rho} \in \mathcal{M}$. Then problem* (9) *has a unique solution $f \in \mathcal{F}$ if and only if* (11) *has a unique solution $f^* \in \mathcal{F}^*$. The relation between f and f^* is given by $f = \rho_1 f^*$. In particular, if $f^* = c\exp(\boldsymbol{\xi}\cdot\boldsymbol{a}^*)$ for some $\boldsymbol{\xi} \in \mathbb{R}^{m-1}$ and some $c > 0$, then $f = \exp(\boldsymbol{\lambda}\cdot\boldsymbol{a})$ with $\boldsymbol{\lambda} = (\log(c\rho_1), \xi_1, \ldots, \xi_{m-1})^T$.*

Using a results by Csiszar's [Csiszár (1975)] for measurable spaces $(X, \mathcal{H})$ with weight functions $\boldsymbol{a}^* = (a_2, \ldots, a_m)$ being $\mathcal{H}$-measurable, we are able to shows that (11) is uniquely solvable with an exponential density if P_{f^*} is replaced by general probability measures P on $\mathbb{R}^d$ that have the correct moments. In connection with Proposition 1 this immediately yields Theorem 1. Here we skip all the technical details (the interested reader is referred to [Junk et al (2004)]) but want only remark that the important point in the proof of Theorem 1 is the boundedness of $|\boldsymbol{v}(\boldsymbol{k})|$. Therefore a similar result can be expected for more general dispersion relations which exhibit an effect of saturation for the modulus of $\boldsymbol{v}(\boldsymbol{k})$.

3 The Euler-Poisson model

As an example of application we analyze the Euler-Poisson model in the case of Kane's dispersion relation. It is based on the same moments employed in ideal gas dynamics, that is density n, average velocity $\boldsymbol{u}$ and average energy W

$$\frac{\partial n}{\partial t} + \frac{\partial (nu^i)}{\partial x^i} = 0,\ \frac{\partial (nu^i)}{\partial t} + \frac{\partial (nU^{ij})}{\partial x^j} = -enE_j H^{ij} + nC_u^i, \tag{12}$$

$$\frac{\partial (nW)}{\partial t} + \frac{\partial (nS^j)}{\partial x^j} = -enu_k E^k + nC_W, \tag{13}$$

$$\text{where } U^{ij} = \frac{1}{n}\int_{\mathbb{R}^3} f v^i v^j d\boldsymbol{k}, \quad H^{ij} = \frac{1}{n}\int_{\mathbb{R}^3} \frac{1}{\hbar} f \frac{\partial v_i}{\partial k_j} d\boldsymbol{k}, \quad C_u^i = \frac{1}{n}\int_{\mathbb{R}^3} \mathcal{C}[f] v^i d\boldsymbol{k},$$

$$C_W = \frac{1}{n}\int_{\mathbb{R}^3} \mathcal{C}[f]\mathcal{E}(k) d\mathbf{k}.$$

For the 5-moment case the weight function vector is $\boldsymbol{a} = (1, \boldsymbol{v}, \mathcal{E})$ and the corresponding Lagrange multipliers are given by the vector $\boldsymbol{\lambda} = -\left(\lambda, \boldsymbol{\lambda}^v, \lambda^W\right)$. The MEP distribution function reads $f_{\boldsymbol{\lambda}} = \exp\left(-\lambda - \lambda_i^v v^i - \lambda^W \mathcal{E}\right)$ and one has the straightforward characterization of the cone Λ (which is obviously convex and open) $\Lambda = \left\{\boldsymbol{\lambda} = -\left(\lambda, \boldsymbol{\lambda}^v, \lambda^W\right) :\ \boldsymbol{\lambda} \in \mathbb{R}^5, \lambda^W > 0\right\}$. By writing $d\boldsymbol{k} = \frac{m^*}{\hbar^3}\sqrt{2m^*\mathcal{E}(1+\alpha\mathcal{E})}(1+2\alpha\mathcal{E}) d\mathcal{E}\, d\Omega$ with $d\Omega$ elementary solid angle, the explicit relation between the Lagrange multipliers and the macroscopic variables are given by

$$u_3 = \frac{1}{d_0}\int_0^\infty v(\mathcal{E}) e^{-\lambda^W \mathcal{E}} \sqrt{\mathcal{E}(1+\alpha\mathcal{E})}(1+2\alpha\mathcal{E}) \left[\frac{\sinh z}{z^2} - \frac{\cosh z}{z}\right] d\mathcal{E} \tag{14}$$

$$n = 2\sqrt{2}\pi \frac{(m^*)^{3/2}}{\hbar^3}\ e^{-\lambda} d_0,\ W = \frac{1}{d_0}\int_0^\infty \mathcal{E} e^{-\lambda^W \mathcal{E}} \sqrt{\mathcal{E}(1+\alpha\mathcal{E})}(1+2\alpha\mathcal{E})\ \frac{\sinh z}{z}\ d\mathcal{E} \tag{15}$$

$$z = \lambda_3^v\ v(\mathcal{E}),\ d_0 = \int_0^\infty e^{-\lambda^W \mathcal{E}} \sqrt{\mathcal{E}(1+\alpha\mathcal{E})}(1+2\alpha\mathcal{E})\ \frac{\sinh z}{z}\ d\mathcal{E} \tag{16}$$

It is relevant only to study the dependence of λ_3^v and λ^W on u_3 and W because λ plays only the role of a normalization factor. We want to investigate whether the moment cone, that is the set of moment for which the MEP distribution there exists, is sufficiently large for concrete applications. To this aim we have numerically checked the invertibility of the rectangle $\left\{(W, u_3) \in [0.04, 0.35] \times [-1.2 \times 10^5, 1.2 \times 10^5]\right\}$ under the mapping $(u_3, W) \mapsto (\lambda_3^v, \lambda^W)$ implicitly defined by the relations (14)-$(15)_2$. W is expressed in eV, u_3 in m/sec, $\lambda_3^v/\sqrt{m^*}$ in $1/\sqrt{eV}$ and λ^W in $1/eV$.

We remark that these values are those usually encountered in the simulations of electron devices. The numerical analysis (see figure) shows that the moment cone contains the above rectangle and therefore it is sufficiently wide to enclose the relevant physical region of velocity and energy.

Acknowledgements

The author V.R. acknowledges that this work has been partially supported by M.I.U.R. (P.R.I.N. 2004 *Problemi matematici delle teorie cinetiche*) and by P.R.A. (ex fondi 60%).

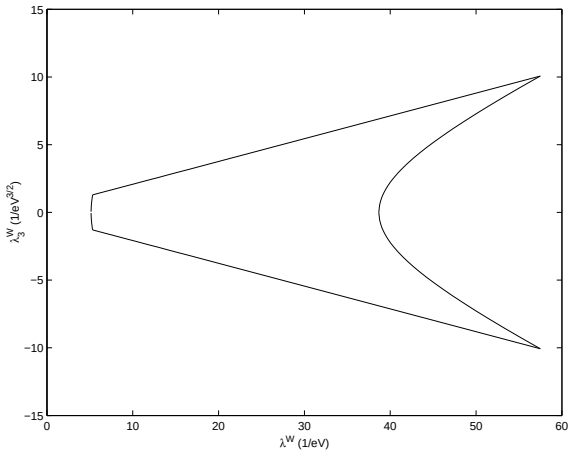

Fig. 1. image of the rectangle $\left\{(\lambda_3^v/\sqrt{m^*}, \lambda^W) \in [-10, 10] \times [1, 65]\right\}$ under the mapping $(\lambda_3^v, \lambda^W) \mapsto (u_3, W)$ defined by the relations (14)-(15_2).

References

[Jaynes (1957)] Jaynes, E. T.: Information theory and Statistical Mechanics. Phys. Rev., **106**, 620–630 (1957)

[Dreyer (1987)] Dreyer, W.: Maximization of the entropy in non–equilibrium. J. Phys. A: Math. Gen., **20**, 6505–6517 (1987)

[Levermore (1996)] Levermore, C. D.: Moment closure hierarchies for kinetic theories. J. Stat. Phys., **83**, 1021–1065 (1996)

[Anile et al. ((2003)] Anile, A.M., Mascali, G., Romano, V.: Recent developments in hydrodynamical modeling of semiconductors. 1:54. In: Anile, A.M. (ed) Mathematical Problems in Semiconductor Physics. Lecture Notes in Mathematics. Springer (2003)

[Junk (1997)] Junk,M.: Domain of definition of Levermore's five-moment system. J. Stat. Phys., **93**, 1143–1167 (1998)

[Junk (1998)] Junk, M.: Maximum entropy for reduced moment problems. Models Methods Appl. Sci., **10**, 1001–1025 (2000)

[Dreyer et al (2001)] Dreyer, W., Junk, M. and Kunik, M.: On the approximation of the Fokker-Planck equation by moment systems. Nonlinearity, , 881–906 (2001)

[Junk et al (2002)] Junk, M. and Unterreiter, A.: Maximum entropy moment systems and Galilean invariance. Continuum Mech. Thermodyn., **14**, 563–576 (2002)

[Csiszár (1975)] Csiszár, I.: I-divergence geometry of probabaility distributions and minimization problems. Ann. of Prob., **3**, 146–158 (1975)

[Junk et al (2004)] Junk, M. and Romano, V.: Maximum entropy moment system of the semiconductor Boltzmann equation using Kane's dispersion relation. Preprint (2004), accepted for publication in Continuum Mech. Thermodyn.