TRANSPORT PHYSICS FOR MODELING SEMICONDUCTOR DEVICES

S.SELBERHERR, W.GRIEBEL and H.PÖTZL

Institut für Allgemeine Elektrotechnik und Elektronik
Abteilung für Physikalische Elektronik
Technische Universität Wien
Gußhausstraße 29, 1040 WIEN, AUSTRIA

Abstract.

In recent years the numerical simulation of semiconductor devices has become a state-of-the art technique for the development and characterization of new devices, particularly for very large scale integrated circuits. The miniaturization of the single device, which is the major prerequisite of VLSI, brought about the necessity to deal with fairly expensive numerical models in order to obtain high accuracy.

The basis of numerical models for semiconductor devices are the well known Poisson equation, current continuity equations and current relations. This physical model has historically proved to be adequate for the simulation of many categories of devices. The reason why these models are adequate is that integral quantities (like current or voltage) are rather insensitive to quite strong deviations of local quantities. Thus, even when local current density or electric field are poorly predicted, e.g., in hot electron areas, device characteristics are predicted quite accurately even by classical models. In future submicron devices, however, areas in which classical models fail, cover a large part of the device. This leads to a poor prediction of integral quantities also. Therefore it is necessary to review the derivation and associated assumptions of the traditional basic semiconductor device equations.

The goal of this presentation is to highlight the assumptions which may cause the traditional model to fail in predicting various effects which occur prominently in miniaturized (submicron) devices. The theoretical background of current transport in semiconductors will be reviewed to help device engineers in properly applying simulation programs.

1. Some Fundamental Properties.

To analyze a semiconductor, a mathematical model has to be given. The equations which form this model are commonly called the basic semiconductor equations, which can be derived from Maxwell's equations (1.1), (1.2), (1.3) and (1.4), several relations obtained from solid-state physics and various - sometimes over simplified - assumptions.

$$\text{rot } \bar{H} = \bar{J} + \frac{\partial \bar{D}}{\partial t} \tag{1.1}$$

$$\text{rot } \bar{E} = -\frac{\partial \bar{B}}{\partial t} \tag{1.2}$$

$$\text{div } \bar{D} = \rho \tag{1.3}$$

$$\text{div } \bar{B} = 0 \tag{1.4}$$

$\bar{E}$ and $\bar{D}$ are the electric field and displacement vector; $\bar{H}$ and $\bar{B}$ are the magnetic field and induction vector, respectively. $\bar{J}$ denotes the conduction current density, and ρ is the electric charge density.

The next sections will be devoted entirely to an outline of the procedures which have to be carried out in order to derive the basic semiconductor equations.

2. Poisson's Equation.

Poisson's equation is essentially the third Maxwell equation (1.3). However, to make this equation directly applicable to semiconductor problems, some manipulations have to be undertaken. We first introduce a relation for the electric displacement vector $\bar{D}$ and the electric field vector $\bar{E}$ (2.1).

$$\bar{D} = \varepsilon \cdot \bar{E} \tag{2.1}$$

ε denotes the permittivity tensor. This relation is valid for all materials which have a time independent permittivity. Furthermore, polarization by mechanical forces is neglected [10]. Both assumptions hold relatively well considering the usual applications of semiconductor devices. However, an investigation of piezoelectric phenomena, ferroelectric phenomena and nonlinear optics is impossible when using only (2.1).

As the next step it is desirable to relate the electric field vector $\bar{E}$ to the electrostatic potential ψ. For that purpose we solve (1.4) by introducing a vector field $\bar{A}$ and remembering that "div rot" applied to any vector quantity is always zero.

$$\bar{B} = \text{rot } \bar{A} \qquad \text{div } \bar{A} = \frac{-1}{c^2} \cdot \frac{\partial \psi}{\partial t} \tag{2.2}$$

The second term of (2.2) is the well-known Lorentz convention. We substitute (2.2) into (1.2) and we obtain readily (2.3).

$$\text{rot} \ (\ \overline{E} + \frac{\partial \overline{A}}{\partial t}) = 0 \tag{2.3}$$

If "rot $\overline{z}$ = 0" holds for a vector field $\overline{z}$ we know from basic differential calculus that $\overline{z}$ can be expressed as a gradient field. Therefore, the electric field vector $\overline{E}$ can be expressed as:

$$\overline{E} = -\frac{\partial \overline{A}}{\partial t} - \text{grad} \ \psi \tag{2.4}$$

Now we substitute (2.4) into (2.1) and then the result into (1.3).

$$\overline{D} = -\epsilon \cdot \frac{\partial \overline{A}}{\partial t} - \epsilon \cdot \text{grad} \ \psi \tag{2.5}$$

$$\text{div} \ (\epsilon \cdot \frac{\partial \overline{A}}{\partial t}) + \text{div} \ (\epsilon \cdot \text{grad} \ \psi) = -\rho \tag{2.6}$$

$$\text{div} \ (\epsilon \cdot \frac{\partial \overline{A}}{\partial t}) = -\frac{\epsilon}{c^2} \cdot \frac{\partial^2 \psi}{\partial t^2} \tag{2.7}$$

The divergence term in (2.7) accounts for the effects of wave propagation. The permittivity ϵ will be treated here in all further investigations as a scalar quantity. In principle it has to be represented as a tensor of rank two. The structure of ϵ depends on crystal symmetry. If the material has one of the cubic symmetry groups (as Si and Ge) , ϵ is a scalar matrix. In other cases ϵ has different components. If ,however, the anisotropy is not very pronounced, the approximation of a scalar ϵ is not too bad. Inhomogeneity effects of the permittivity have been neglected in (2.7). There does not exist any pronounced experimental evidence for inhomogeneity effects. Considering the small size of the active area of modern devices wave propagation phenomena may be neglected even at high frequencies. Thus, we finally end up with (2.8) which is the well known form of Poisson's equation.

$$\text{div} \ (\epsilon \cdot \text{grad} \ \psi) = -\rho \tag{2.8}$$

The space charge density ρ can be further broken apart (2.9) into the product of the elementary charge q times the sum of the positively charged hole density p, the negatively charged electron density n and an additional concentration C which will be subject of later investigations.

$$\rho = q \cdot (\ p - n + C \ ) \tag{2.9}$$

From a purely mathematical point of view (2.9) represents a substitution only, without introducing any assumptions. However, additional assumptions are brought about by modeling the quantities n, p etc. as will become clearly apparent in chapter 3.

If we introduce (2.9) and the assumption of a homogeneous scalar permittivity into (2.6), we obtain the final form of Poisson's equation to be used for semiconductor device modeling.

$$\text{div grad } \Psi = \frac{q}{\epsilon} \cdot (n - p - C) \tag{2.10}$$

3. Continuity Equations.

The continuity equations can be derived in a straightforward manner from the first Maxwell equation (1.1). If we apply the operator "div" on this equation we obtain:

$$\text{div rot } \overline{H} = \text{div } \overline{J} + \frac{\partial \varrho}{\partial t} = 0 \tag{3.1}$$

Now we split the conduction current density $\overline{J}$ into a component $\overline{J}_p$ caused by holes and a component $\overline{J}_n$ caused by electrons:

$$\overline{J} = \overline{J}_p + \overline{J}_n \tag{3.2}$$

Furthermore, we assume that all charges in the semiconductor, except the mobile carriers electrons and holes, are time invariant. This is an acceptable assumption if carrier lifetimes are long. Large errors occur in the vicinity of boundaries or in areas of strong generation or recombination, specially, under rapidly varying transient conditions.

$$\frac{\partial C}{\partial t} = 0 \tag{3.3}$$

If we substitute (2.8) and (3.2) into (3.1) and if we make use of (3.3) we obtain:

$$\text{div } (\overline{J}_p + \overline{J}_n) + q \cdot \frac{\partial}{\partial t} (p - n) = 0 \tag{3.4}$$

This result is interpreted fairly trivially. It just means that sources and sinks of the total conduction current are fully compensated by the time variation of the mobile charge. In order to obtain two continuity equations a few formal steps have to be carried out. We first define a quantity R in (3.5) and, secondly, we rewrite (3.4) by making use of the definition R.

$$\text{div } \overline{J}_n - q \cdot \frac{\partial n}{\partial t} = q \cdot R \tag{3.5}$$

$$\text{div } \overline{J}_p + q \cdot \frac{\partial p}{\partial t} = -q \cdot R \tag{3.6}$$

It is obvious that we can not gain information by writing one equation (3.4) in two different ways (3.5), (3.6). However, these formal steps enable us to interprete the equation more easily. The quantity R can be understood as a function describing the net recombination of electrons and holes. Positive R means predominance of recombination and negative R means predominance of generation. So far we have no information about the structure of R except equations (3.5) and (3.6). R has to be modeled carefully using knowledge from the solid-state physics of semiconductors. If we have a model for R, equations (3.5) and (3.6) can really be considered as two equations. It seems worthwhile to note explicitly here that there is no necessity or even evidence that R can be expressed as a function depending only upon local quantities and not upon integral quantities; non-local generation or recombination phenomena may certainly occur in semiconductor devices considering only the derivation of the continuity equations.

4. Carrier Transport Equations.

The derivation of current relations for the semiconductor equations is a very cumbersome task. It is not the intention of this paper to cover the extraordinarily wide field of physics behind all the considerations necessary to derive the current relations in detail. Therefore, some of the required relations will be given without proof, but with reference to a text more specialized in that field.

Without loss of generality the current density of charged particles is the product of the charge constant per particle, the particle density and the average velocity (drift velocity) of the particles. So the hole current density and the electron current density can be written as (4.1) and (4.2), respectively.

$$\overline{J}_p = q \cdot p \cdot \overline{v}_p \tag{4.1}$$

$$\overline{J}_n = -q \cdot n \cdot \overline{v}_n \tag{4.2}$$

The major problem is to find expressions which relate the average carrier velocities to the electric field vector $\bar{E}$ and to the carrier density. In order to obtain information about the drift velocity we have to describe the carrier density by means of a distribution function $f_{\boldsymbol{y}}$ in phase space which is the space of spatial coordinates $\bar{x}=(x,y,z)^T$, momentum coordinates $\bar{k}=(k_x,k_y,k_z)^T$ and time t, thus a seven dimensional space. The distribution function determines the carrier density per unit volume of phase space. By integrating the distribution function over the entire momentum volume V_k we obtain the carrier density $\boldsymbol{y}(\bar{x},\ t)$. $\boldsymbol{y}$ stands for n or p, denoting electrons or holes.

$$\frac{1}{4\cdot\pi^3}\cdot\int_{V_k} f_{\boldsymbol{y}}(\bar{x},\bar{k},t)\cdot d\bar{k} = \boldsymbol{y}(\bar{x},t) \tag{4.3}$$

This normalization (4.3) defines $f_{\boldsymbol{y}}$ as a probability. In the literature various different normalizations can be found, e.g. [8], [11].

The distribution function has the property that its derivative along a subsystem trajectory $\bar{x}_{\boldsymbol{y}}(t)$, $\bar{k}_{\boldsymbol{y}}(t)$ with respect to time vanishes in the entire phase space in compliance with the Liouville theorem about the invariance of the phase volume for a system moving along the phase paths or an account of the conservation of the number of states [11]. The Liouville theorem, unfortunately, holds for a constant number of particles only. The following considerations are therefore excluding effects of generation or recombination.

$$\frac{d}{dt}f_{\boldsymbol{y}}(\bar{x}_{\boldsymbol{y}}(t),\bar{k}_{\boldsymbol{y}}(t),t) = 0 \tag{4.4}$$

By expanding the total derivative we obtain:

$$\frac{\partial f_{\boldsymbol{y}}}{\partial t} + grad_k\ f_{\boldsymbol{y}}\cdot\frac{d\bar{k}_{\boldsymbol{y}}}{dt} + grad_x\ f_{\boldsymbol{y}}\cdot\frac{d\bar{x}_{\boldsymbol{y}}}{dt} = 0 \tag{4.5}$$

Here $grad_k$ denotes the gradient operator with respect to the momentum coordinates $\bar{k}$; $grad_x$ is the gradient operator with respect to the spatial coordinates $\bar{x}$. Equation (4.5) shows that the variation of the distribution function at each point of phase space $(\bar{x},\bar{k})$ with time is caused by the motion of particles in normal space $(\bar{x})$ and in momentum space $\bar{k}$.

The derivative of $\overline{k}_\nu$ with respect to time multiplied with Planck's constant h equals the sum of all forces $\overline{F}$. These forces have to be devided into two classes (4.7).

$$\frac{d\overline{k}_\nu}{dt} = \frac{\overline{F}_\nu}{\hbar} , \qquad \hbar = \frac{h}{2 \cdot \pi} \tag{4.6}$$

$$\overline{F}_\nu = \overline{F}_{\nu e} + \overline{F}_{\nu i} \tag{4.7}$$

$\overline{F}_{\nu e}$ comprises forces due to macroscopic external fields and $\overline{F}_{\nu i}$ denotes forces due to internal localized crystal attributes like impurity atoms or ions, vacancies, and thermal lattice vibrations. It is (quite) impossible to calculate the effect of internal forces $\overline{F}_{\nu i}$ upon the distribution function from the laws of dynamics [11]. Statistical laws have to be invoked instead. By introducing the quantity $S_\nu(\overline{k}, \overline{k}') \cdot d\overline{k}'$ which is the probability per unit time that a carrier in the state $\overline{k}$ will be scattered into the momentum volume $d\overline{k}'$, we can write the internal collision term as follows:

$$\text{grad}_k \, f_\nu \cdot \frac{\overline{F}_{\nu i}}{\hbar} =$$

$$= \int_{Vk'} \{ \, f_\nu(\overline{x}, \overline{k}, t) \cdot [1 - f_\nu(\overline{x}, \overline{k}', t)] \cdot S_\nu(\overline{k}, \overline{k}') -$$

$$- f_\nu(\overline{x}, \overline{k}', t) \cdot [1 - f_\nu(\overline{x}, \overline{k}, t)] \cdot S_\nu(\overline{k}', \overline{k}) \, \} \cdot d\overline{k}' \tag{4.8}$$

(4.8) is termed the collision integral. The first term in the integrand describes the number of carriers scattered from the state $\overline{k}$ into the volume element $d\overline{k}'$ per unit time. Correspondingly, the second term in the integrand of (4.8) equals the number of carriers scattered from volume element $d\overline{k}'$ into state $\overline{k}$ per unit time. $f_\nu(\overline{x}, \overline{k}, t)$ gives the probability that a carrier initially occupies the state $\overline{k}$. $[1 - f_\nu(\overline{x}, \overline{k}', t)]$ gives the probability that the volume element $d\overline{k}'$ is initially unoccupied and can, therefore, accept a carrier (no volume of the phase space can be occupied by more than two carriers (these having antiparallel spin) because of the Pauli exclusion principle). $S_\nu(\overline{k}, \overline{k}')$ gives an a priori probability of the scattering event. There are different scattering mechanisms in a semiconductor; for each a $S_\nu(\overline{k}, \overline{k}')$ is present. In nonpolar semiconductors (like Ge and Si) the collision of carriers with phonons are randomizing and lead to an isotropic behavior. The collisions between carriers and ionized impurities, however, are dominated by

small scattering angles, thus leading to an anisotropic behavior of the scattering. In polar semiconductors, however, even carrier-phonon scattering is partially anisotropic. Thorough investigations about the scattering probability $S_\nu(\overline{k},\overline{k}')$ can be found in, e.g., [6].

The derivative of $\overline{x}_\nu$ with respect to time represents the group velocity of the carriers.

$$\frac{d\overline{x}_\nu}{dt} = \overline{u}_\nu \tag{4.9}$$

We have now to substitute the relations (4.6) to (4.9) into (4.5), and we obtain the Boltzmann transport equation in explicit form.

$$\frac{\partial f_\nu}{\partial t} + \frac{\overline{F}_{\nu e}}{\hbar}\cdot grad_k\; f_\nu + \overline{u}_\nu\cdot grad_x\; f_\nu =$$

$$-\int_{V\overline{k}'} \{\; f_\nu(\overline{x},\overline{k},t)\cdot[1-f_\nu(\overline{x},\overline{k}',t)]\cdot S_\nu(\overline{k},\overline{k}')-$$

$$-\; f_\nu(\overline{x},\overline{k}',t)\cdot[1-f_\nu(\overline{x},\overline{k},t)]\cdot S_\nu(\overline{k}',\overline{k})\;\}\cdot d\overline{k}' \tag{4.10}$$

A fairly accurate approach would be to directly solve (4.10) in order to calculate carrier densities and drift velocities. However, this is an extraordinarily difficult task to accomplish. (4.10) represents an integro-differential equation with seven independent variables. This equation does not admit a closed solution. It rather requires the use of iterative procedures which, moreover, are scarcely suitable for numerical approaches [3], or additionally, invoke very stringent assumptions [8].

An alternative approach to solving the Boltzmann equation consists in simulating the motion of one or more carriers at microscopic level with Monte Carlo methods, e.g. [9]. However, this category of simulations is very computationally intensive [15], [16] and therefore, with a few exceptions only, not suitable for engineering application.

One should be aware of the fact that the validity of the Boltzmann equation (4.10) implies already several assumptions (cf. [3], [4]).

- The Boltzmann equation is a quasi-classic equation, thus not covering any quantum mechanical effects. Quantum effects to be found in actual devices are tunneling phenomena, quantization of energy levels (e.g. in sub-μ field-effect devices or due to strong magnetic fields (Landau levels)) and various contact phenomena.
- The Boltzmann equation in this form is valid for a constant number of particles only, thus not covering generation or recombination effects.
- The scattering probability is independent of external forces.
- The duration of a collision is much shorter than the average time of motion of a particle; collisions are instantaneous.
- Carrier-carrier interaction is negligible. This effect would change the integrand of the right hand side integral in (4.10) highly nonlinear in f_ν [1].
- External forces are almost constant over a length comparable to the physical dimensions of the wave packet describing the motion of a carrier.
- The band theory and the effective mass theorem apply to the semiconductor under consideration [20]. The band structure is valid strictly in an endless, undisturbed crystal only. It can be shown, however, that the band structure is a good approximation for a real crystal. In areas where the periodic lattice is strongly disturbed, like in the vicinity of boundaries, contacts or in areas of high doping concentration, the band structure cannot be expected to hold.

However, it is our intention here to outline the derivation of the classical current relations and only to pinpoint the problems associated with much more basic and error-prone models.

By assuming that all scattering processes are elastic and by neglecting all effects caused by degeneracy the scattering integral can be approximated and the Boltzmann equation is reduced to a pure differential equation [6], [8], [20]. The assumption of elastic scattering can not be justified for hot electrons or holes. Even at moderate carrier temperatures some inelastic scattering mechanisms are present(depending on the material). Nondegeneracy is a good approximation only if doping concentration and carrier densities are not too high.

$$\frac{\partial f_\nu}{\partial t} + \frac{\bar{F}_\nu e}{\hbar}\cdot\text{grad}_k\ f_\nu + \bar{u}_\nu\cdot\text{grad}_x\ f_\nu = -\frac{f_\nu - f_{\nu 0}}{\tau_\nu} \qquad (4.11)$$

The physical motivation for the right hand side of (4.11) is as follows: Suppose that at some moment of time t=0 all external forces are switched off and f_ν is homogenous.

$$\frac{\bar{F}_\nu e}{\hbar}\cdot\text{grad}_k\ f_\nu + \bar{u}_\nu\cdot\text{grad}_x\ f_\nu = 0 \qquad (4.12)$$

It follows from (4.11) that the distribution function will change as a result of collisions only. (4.13) will reduce to:

$$\frac{\partial f_{\nu}}{\partial t} = - \frac{f_{\nu} - f_{\nu 0}}{\tau_{\nu}} \tag{4.13}$$

The solution of this differential equation is quite simple.

$$f_{\nu}(\overline{x},\overline{k},t) = f_{\nu 0}(\overline{x},\overline{k}) + [f_{\nu}(\overline{x},\overline{k},0) - f_{\nu 0}(\overline{x},\overline{k})] \cdot e^{-t/\tau_{\nu}} \tag{4.14}$$

$f_{\nu 0}$ is the equilibrium distribution function, and the quantity τ_{ν} shows the rate of return to the state of equilibrium from the disturbed state, therefore, it is termed the relaxation time. It should be noted that the relaxation time depends on the momentum. This is a very simplified model; in the general case τ_{ν} is a tensor of rank two. In case of both inelastic and anistropic scattering no meaningful relaxation time can be declared at all. In other cases a meaningful relaxation time may be defined. Under the very restrictive assumptions stated above the problem of solving the Boltzmann equation can be eased drastically by modeling the relaxation time as only a function of energy [8].

In order to obtain the current relations from (4.11) we multiply this equation with the group velocity $\overline{u}_{\nu}$, and then we integrate the equation over momentum space. In the following derivation the most general case is assumed, implying that M_{ν} (the generalized effective mass tensor) and τ_{ν} (the relaxation time tensor) are tensorial quantities which are $\overline{k}_{\nu}$ dependent. Furthermore, no special form of the distribution function is assumed (meaning that the distribution function may be any function making physical sense, and not only a Maxwellian or Fermi-Dirac distribution, which cover special cases only).

$$\int_{V_k} \overline{u}_{\nu} \cdot \frac{\partial f_{\nu}}{\partial t} \cdot d\overline{k} + \int_{V_k} \overline{u}_{\nu} \cdot (\frac{\overline{F}_{\nu e}}{\hbar} \cdot \text{grad}_k \ f_{\nu}) \cdot d\overline{k} + \int_{V_k} \overline{u}_{\nu} \cdot (\overline{u}_{\nu} \cdot \text{grad}_x \ f_{\nu}) \cdot d\overline{k} =$$

$$- \int_{V_k} \overline{u}_{\nu} \cdot \frac{f_{\nu} - f_{\nu 0}}{\tau_{\nu}} \cdot d\overline{k} \tag{4.15}$$

The first integral of (4.15) is quite easy to evaluate, because the group velocity depends on $\overline{k}_{\nu}$ only. Therefore integration over the momentum space and differentiation with respect to time may be swapped and the integral evaluated.

$$\int\limits_{Vk} \overline{u}_{y} \frac{\partial f_{y}}{\partial t} \cdot d\overline{k} \;=\; \frac{\partial}{\partial t} \cdot \int\limits_{Vk} \overline{u}_{y} \, f_{y} \, d\overline{k} \;=\; 4 \cdot \pi^3 \cdot \frac{\partial}{\partial t} (\mathcal{V} \cdot \langle \overline{u}_{y} \rangle) \qquad (4.16)$$

The $< >$ brackets denote the mean value of a physical quantity (scalar or vectorial) weighted by the distribution function. The definition of the mean of a physical quantity x is:

$$\langle x \rangle \;=\; \frac{\int\limits_{Vk} x \cdot f_{y} \, d\overline{k}}{\int\limits_{Vk} f_{y} \, d\overline{k}} \;=\; \frac{\int\limits_{Vk} x \cdot f_{y} \, d\overline{k}}{4 \cdot \pi^3 \cdot \mathcal{V}} \qquad (4.17)$$

The second integral of the Boltzmann equation is called 'force term' for obvious reasons. The external forces are independent of $\overline{k}_{y}$ (if no magnetic field is present). The gradient operator and the multiplication with $\overline{F}_{ye}$ may be summed up in one divergence operation giving (4.18).

$$\int\limits_{Vk} \overline{u}_{y} \, (\frac{\overline{F}_{ye}}{\hbar} \cdot \mathrm{grad}_k f_{y}) \cdot d\overline{k} \;=\; \frac{1}{\hbar} \cdot \int\limits_{Vk} \overline{u}_{y} \, \mathrm{div}_k (\overline{F}_{ye} \cdot f_{y}) \cdot d\overline{k} \qquad (4.18)$$

The multiplication with the group velocity in (4.18) may be concentrated within the divergence operator by the relation (4.19):

$$\mathrm{div}_k (\overline{u}_{y} \, (\overline{F}_{ye} \cdot f_{y})^*) \;=\; \overline{u}_{y} \, \mathrm{div}_k (\overline{F}_{ye} \cdot f_{y}) + \mathrm{grad}_k \overline{u}_{y} \, (\overline{F}_{ye} \cdot f_{y}) \qquad (4.19)$$

So we obtain two new integrals for the second part of (4.18).

$$\frac{1}{\hbar} \cdot \int\limits_{Vk} \overline{u}_{y} \, \mathrm{div}_k (\overline{F}_{ye} \cdot f_{y}) \cdot d\overline{k} = \frac{1}{\hbar} \cdot \int\limits_{Vk} \mathrm{div}_k (\overline{u}_{y} \, (\overline{F}_{ye} \cdot f_{y})^*) \, d\overline{k} \; -$$

$$\frac{1}{\hbar} \cdot \int\limits_{Vk} \mathrm{grad}_k \overline{u}_{y} \, (\overline{F}_{ye} \cdot f_{y}) \, d\overline{k} \qquad (4.20)$$

Applying the Gauss integral theorem to the second integral of (4.20) and evaluating the gradient operator in the third yields (4.21).

$$\frac{1}{\hbar} \cdot \int\limits_{Vk} \overline{u}_{y} \, \mathrm{div}_k (\overline{F}_{ye} \cdot f_{y}) \, d\overline{k} \;=\; \frac{1}{\hbar} \cdot \int\limits_{\overline{A}} \overline{u}_{y} \, (\overline{F}_{ye} \cdot f_{y})^* \cdot d\overline{A} \; -$$

$$-\int\limits_{Vk} M_{y}^{-1} \cdot (\overline{F}_{ye} \cdot f_{y}) \cdot d\overline{k} \qquad (4.21)$$

The surface integral vanishes by virtue of the periodicity of the first Brillouin zone in the momentum space (shown by elementary arithmetic). The second term of (4.21) may be evaluated by the mean value theorem (4.17).

$$-\int_{Vk} \underline{M}\underline{y}^{-1}\cdot\overline{F}_{\!\!\chi e}\cdot f_{\!\chi r}\,d\overline{k} \;=\; -4\cdot\pi^3\cdot\underline{y}\cdot<\underline{M}^{-1}\cdot\overline{F}_{\!\!\chi e}> \tag{4.22}$$

If we take a magnetic field into account, the external forces become dependent on the group velocity. Provided that the thermal energy of the carriers is much greater than the energy in the magnetic field (meaning that quantization effects are neglible, which is a reasonable assumption at room temperature and not too high fields) we can express the force acting on a carrier by

$$\overline{F}_{\!\!\chi e} \;=\; q\cdot(\overline{E}+\overline{u}_{\!y}\times\overline{B}) \tag{4.23}$$

The electrical component is still independent of the group velocity and may be treated as in the above derivation with

$$\overline{F}_{\!\!\chi e} \;=\; q\cdot\overline{E} \tag{4.24}$$

as the external force and will therefore not be covered here. The magnetic term, however, is treated seperately now. For the force term with a magnetic field only we obtain

$$\frac{1}{\hbar}\cdot\int_{Vk} \overline{u}_{\!y}\,(q\cdot(\overline{u}_{\!y}\times\overline{B})\cdot\mathrm{grad}_k f_{\!y})\cdot d\overline{k} \tag{4.25}$$

Now we make use of the identity

$$\mathrm{div}_k((\overline{u}_{\!y}\times\overline{B})\cdot f_{\!y}) \;=\; (\overline{u}_{\!y}\times\overline{B})\cdot\mathrm{grad}_k f_{\!y} + f_{\!y}\cdot\mathrm{div}_k(\overline{u}_{\!y}\times\overline{B}) \tag{4.26}$$

The third term of (4.26) is zero. This can be proved by the relation (4.27).

$$\mathrm{div}_k(\overline{u}_{\!y}\times\overline{B}) \;=\; \overline{B}\cdot\mathrm{rot}_k\overline{u}_{\!y} - \overline{u}_{\!y}\,\mathrm{rot}_k\overline{B} \tag{4.27}$$

The group velocity is related to the momentum vector by a symmetric matrix (the generalized effective mass tensor). Therefore the rot operator applied to the group velocity gives a zero vector. The magnetic field is independent of the momentum, giving zero as the result of the third term of (4.27) also. Now we use (4.26) in (4.25) and obtain (4.28).

$$\frac{q}{\hbar}\cdot\int_{Vk} \overline{u}_{\!y}\,((\overline{u}_{\!y}\times\overline{B})\cdot\mathrm{grad}_k f_{\!y})\cdot d\overline{k} \;=\; \frac{q}{\hbar}\cdot\int_{Vk} \overline{u}_{\!y}\,\mathrm{div}_k((\overline{u}_{\!y}\times\overline{B})\cdot f_{\!y})\cdot d\overline{k} \tag{4.28}$$

By exploiting the fact that

$$\mathrm{div}_k(\overline{u}_{\!y}\,(\overline{u}_{\!y}\times\overline{B})^{*}f_{\!y}) \;=\; \overline{u}_{\!y}\,\mathrm{div}_k((\overline{u}_{\!y}\times\overline{B})f_{\!y}) + \mathrm{grad}_k\overline{u}_{\!y}\,(\overline{u}_{\!y}\times\overline{B})f_{\!y} \tag{4.29}$$

we rewrite (4.28) as

$$\frac{q}{\hbar} \cdot \int_{Vk} \overline{u}_{y} \cdot div_k ((\overline{u}_{y} \times \overline{B}) \cdot f_{y}) \cdot d\overline{k} = \frac{q}{\hbar} \cdot \int_{Vk} div_k ((\overline{u}_{y} \cdot (\overline{u}_{y} \times \overline{B})^{*} \cdot f_{y}) \cdot d\overline{k}$$

$$\frac{q}{\hbar} \cdot \int_{Vk} grad_k \overline{u}_{y} \cdot (\overline{u}_{y} \times \overline{B}) \cdot f_{y} \cdot d\overline{k} \qquad (4.30)$$

Using the Gauss integral theorem and evaluating the gradient operator yields:

$$\frac{q}{\hbar} \cdot \int_{Vk} \overline{u}_{y} \cdot div_k ((\overline{u}_{y} \times \overline{B}) \cdot f_{y}) \cdot d\overline{k} = \frac{q}{\hbar} \cdot \int_{\overline{A}} \overline{u}_{y} \cdot ((\overline{u}_{y} \times \overline{B})^{*}) \cdot f_{y} \cdot d\overline{A} -$$

$$-q \cdot \int_{Vk} M_{y}^{-1} \cdot (\overline{u}_{y} \times \overline{B}) \cdot f_{y} \cdot d\overline{k} \qquad (4.31)$$

The surface integral vanishes for the same reason than in (4.21) The third term may be evaluated by our well-known mean value theorem, giving (4.32) as the final result for the 'force term' with a magnetic field only.

$$\frac{q}{\hbar} \cdot \int_{Vk} \overline{u}_{y} \cdot ((\overline{u}_{y} \times \overline{B}) \cdot grad_k f_{y}) \cdot d\overline{k} = -4\pi^3 \cdot \boldsymbol{\mathcal{V}} \cdot q \cdot < M_{y}^{-1} \cdot (\overline{u}_{y} \times \overline{B}) > \qquad (4.32)$$

The result obtained in (4.32) may be simply added to the electrical force term, giving (4.33) as a general result covering an electric and a moderate magnetic field simultanously.

$$\frac{1}{\hbar} \int_{Vk} \overline{u}_{y} \cdot (\overline{F}_{y e} grad_k f_{y}) \cdot d\overline{k} = -4\pi^3 \boldsymbol{\mathcal{V}} q \cdot (< M_{y}^{-1} \cdot \overline{E}> + < M_{y}^{-1} \cdot (\overline{u}_{y} \times \overline{B}) >) \qquad (4.33)$$

The third term of the integral of the Boltzmann equation may be, at first, treated in a similar way as in (4.18). $\overline{u}_{y}$ is solely dependent on $\overline{k}$; therefore the second multiplication with the group velocity and the gradient operator with respect to $\overline{x}$ may be put in one divergence operation, giving (4.34).

$$\int_{Vk} \overline{u}_{y} \cdot (\overline{u}_{y} \cdot grad_x f_{y}) \cdot d\overline{k} = \int_{Vk} \overline{u}_{y} \cdot (div_x (\overline{u}_{y} \cdot f_{y})) \cdot d\overline{k} \qquad (4.34)$$

Even the first multiplication with $\overline{u}_{y}$ in (4.34) may be concentrated within the divergence operator by rewriting the product of the group velocities as a tensorial quantity.

$$\int_{Vk} div_x (\overline{u}_{y} \cdot \overrightarrow{u}_{y}^{*}) \cdot f_{y} \cdot d\overline{k} = div_x \int_{Vk} (\overline{u}_{y} \cdot \overrightarrow{u}_{y}^{*}) \cdot f_{y} \cdot d\overline{k} \qquad (4.35)$$

The divergence operator with respect to $\overline{x}$ and the integration over the momentum space are independent operations which therefore may

be exchanged. The evaluation of the remaining integral is carried out by the mean value theorem.

$$\text{div}_x \int\limits_{Vk} (\vec{u}_y \vec{u}_y^*) \cdot f_y \, d\vec{k} = \text{div}_x (4 \cdot \pi^3 \cdot \mathcal{V} \cdot \langle \vec{u}_y \vec{u}_y^* \rangle) \tag{4.36}$$

Now let us consider the physical significance of the mean value of the tensorial product of the group velocities. The elements of this product have the dimension of energy divided by a mass. This fact motivates us to assign temperatures to the tensor of the group velocities by (4.37).

$$\langle \vec{u}_y \vec{u}_y^* \rangle = \underline{M}_0^{-1} \cdot k_b \cdot \underline{T} \tag{4.37}$$

The quantities $\underline{M}_0$ and $\underline{T}$ denote the effective mass tensor at energy minimum and the temperature tensor, respectively. We admit that thinking of temperature as a tensorial quantity is a bit unfamiliar; but in crystal physics few physical quantities retain their simple classical meaning because of the inherently quantum mechanical nature of a solid state problem. If the effective mass tensor at minimum energy is independent of $\overline{x}$ (meaning that the band structure is the same in the entire crystal), it may be written before the divergence operator. If, furthermore, the temperature tensor is a scalar matrix (a case strictly true for thermodynamic equilibrium only), (4.36) turns out to be the 'diffusion term' of classical models.

$$4 \cdot \pi^3 \cdot \text{div}_x (\mathcal{V} \cdot k_b \cdot \underline{M}_0^{-1} \cdot \underline{T}) = 4 \cdot \pi^3 \cdot \underline{M}_0^{-1} \cdot \text{grad}_x (\mathcal{V} \cdot k_b \cdot T) \tag{4.38}$$

It should be noted that the approximation of a scalar carrier temperature is a very poor one for high fields; this means that the diffusion of so-called hot carriers is uncorrectly described. The last integral of the Boltzmann equation may be split up in two integrals. The integral containing the equilibrium distribution function turns out to be zero for reasons of symmetry.

$$\int\limits_{Vk} \underline{\tau}_y^{-1} \cdot \vec{u}_y (f_y - f_{y0}) \cdot d\vec{k} = \int\limits_{Vk} \underline{\tau}_y^{-1} \cdot \vec{u}_y \, f_y \, d\vec{k} - \int\limits_{Vk} \underline{\tau}_y^{-1} \cdot \vec{u}_y \, f_{y0} \cdot d\vec{k} \tag{4.39}$$

The remaining integral is evaluated by the mean value theorem, giving (4.40).

$$\int\limits_{Vk} \underline{\tau}_y^{-1} \cdot \vec{u}_y \, f_y d\vec{k} = 4 \cdot \pi^3 \cdot \mathcal{V} \cdot \langle \underline{\tau}_y^{-1} \cdot \vec{u}_y \rangle \tag{4.40}$$

Now let us rewrite the the integral of the Boltzmann equation in terms of our evaluated integrals.

$$\frac{\partial}{\partial t} \cdot (\mathcal{V} \langle \overline{u}_y \rangle) - \mathcal{V} \langle M^{-1} \rangle \overline{F}_e + M_0^{-1} \operatorname{grad}_x (\mathcal{V} k_b T) = \mathcal{V} \langle \underline{\tau}_y^{-1} \overline{u}_y \rangle \tag{4.41}$$

If we assume that the mean of the effective mass is a constant (meaning that the energy bands are strictly parabolic and that all valleys (if more than one) are equally populated) and it does, therefore, not depend on the distribution function $\langle \underline{M} \rangle$ and $\underline{M}_0$ are the same. If furthermore, the effective mass is a scalar matrix, we denote it by m_y^*. The assumptions made above do not hold for hot electrons. If we approximate the relaxation time tensor by a scalar independent of the distribution function, it may be written outside the $\langle \rangle$ bracket. The relaxation mechanisms are, in a first order approximation, dependent on the energy of the carriers, and therefore dependent on the distribution function. For low-energy carriers the assumption of a constant relaxation time is justified. If we now identify the mean value of the group velocity with a macroscopic drift velocity $\overline{v}$ and assume that the external force applied to a particle is of merely electrical origin, we end up with two ordinary differential equations.

$$\frac{\partial}{\partial t}(n \cdot \overline{v}_n) + \frac{q}{m_n^*} \cdot n \cdot \overline{E} + \frac{1}{m_n^*} \cdot \operatorname{grad}(n \cdot k \cdot T) = -\frac{n \cdot \overline{v}_n}{\tau_n} \tag{4.42}$$

$$\frac{\partial}{\partial t}(p \cdot \overline{v}_p) - \frac{q}{m_p^*} \cdot p \cdot \overline{E} + \frac{1}{m_p^*} \cdot \operatorname{grad}(p \cdot k \cdot T) = -\frac{p \cdot \overline{v}_p}{\tau_p} \tag{4.43}$$

q denotes the elementary charge. These equations can also be regarded as macroscopic force balance equations. A "closed solution" of these equations is, unfortunately, not possible. In order to obtain an approximate solution we introduce effective carrier mobilities μ_n and μ_p.

$$\mu_n = \frac{q \cdot \tau_n}{m_n^*} \tag{4.44}$$

$$\mu_p = \frac{q \cdot \tau_p}{m_p^*} \tag{4.45}$$

We rewrite (4.42) and (4.43) after multiplication with the corresponding average collision times τ_y and charge constant $\pm q$, and - remembering (4.1) and (4.2) - we end up with:

$$\tau_n \cdot \frac{\partial \bar{J}_n}{\partial t} + \bar{J}_n = q \cdot \mu_n \cdot n \cdot (\bar{E} + \frac{1}{n} \cdot grad(n \cdot \frac{k \cdot T}{q})) \tag{4.46}$$

$$\tau_p \cdot \frac{\partial \bar{J}_p}{\partial t} + \bar{J}_p = q \cdot \mu_p \cdot p \cdot (\bar{E} - \frac{1}{p} \cdot grad(p \cdot \frac{k \cdot T}{q})) \tag{4.47}$$

The average collision times τ_ν are very small, usually in the order of tenth of picoseconds. Therefore, equations (4.46) and (4.47) can be understood as being singularly perturbed. This suggests to expand the solution into powers of the perturbation parameter which is the collision time.

$$\bar{J}_n(\tau_n) = \sum_{i=0}^{\infty} \bar{J}_{ni} \cdot (\tau_n)^i \tag{4.48}$$

$$\bar{J}_p(\tau_p) = \sum_{i=0}^{\infty} \bar{J}_{pi} \cdot (\tau_p)^i \tag{4.49}$$

We have an algebraic equation for the zero order term of the current density.

$$\bar{J}_{no} = q \cdot \mu_n \cdot n \cdot (\bar{E} + \frac{1}{n} \cdot grad(n \cdot \frac{k \cdot T}{q})) \tag{4.50}$$

$$\bar{J}_{po} = q \cdot \mu_p \cdot p \cdot (\bar{E} - \frac{1}{p} \cdot grad(p \cdot \frac{k \cdot T}{q})) \tag{4.51}$$

These equations are formal approximations of order τ_ν

$$\bar{J}_n = \bar{J}_{no} + O(\tau_n) \tag{4.52}$$

$$\bar{J}_p = \bar{J}_{po} + O(\tau_p) \tag{4.53}$$

We further assume that the lattice temperature is constant.

$$T = const. \tag{4.54}$$

Then we can use the substitutions (4.55) and (4.56) which by means of physical interpretation are termed the Einstein relations

$$D_n = \mu_n \cdot \frac{k \cdot T}{q} \tag{4.55}$$

$$D_p = \mu_p \cdot \frac{k \cdot T}{q} \tag{4.56}$$

to define the diffusion constants D_ν and, finally, we are able to write down the current relations in the well known established form as sums of a drift and a diffusion component.

$$\bar{J}_n \cong q \cdot n \cdot \boldsymbol{\mu}_n \cdot \bar{E} + q \cdot D_n \cdot \text{grad } n \tag{4.57}$$

$$\bar{J}_p \cong q \cdot p \cdot \boldsymbol{\mu}_p \cdot \bar{E} - q \cdot D_p \cdot \text{grad } p \tag{4.58}$$

In the following we should like to summarize the most important assumptions which had to be performed over and above to the ones necessary for the validity of the Boltzmann equation to obtain the current relations (4.57) and (4.58).

- All scattering processes have been assumed to be either elastic or isotropic in order to justify the use of a relaxation time. Therefore, for instance, polar optical phonon scattering which is a major scattering mechanism in GaAs has been excluded.
- The spatial variations of the band structure are neglected. This implies a slowly varying impurity concentration over a carrier mean free path.
- Effects of degeneracy have been neglected in the approximation for the scattering integral.
- The spatial variation of the external forces is neglected which implies a slowly varying electric field vector.
- The influence of the Lorentz force is ignored by assuming zero magnetic induction.
- The time and spatial variation of carrier temperature is neglected and, furthermore, lattice and carrier temperature are assumed to be equal. Therefore, the diffusion of hot carriers is improperly described. Several authors have tried to overcome this problem by using modified Einstein relations [2], [12], [13], [17], [18], [19], [21].
- Parabolic energy bands are assumed which is an additional reason why degenerate semiconductor materials cannot be treated properly. Calculations of the realistic band structure of various semiconductors can be found in, e.g., [5]. However, for a realistic band structure it can become necessary to use a system of Boltzmann equations to describe the carrier distribution instead of just one (cf. [23]).
- The zero order term of the series expansions of $\bar{J}_n$ and $\bar{J}_p$ into powers of the collision time only has been taken into account. Thus, all time dependent mobility phenomena are not included.
- The semiconductor has been assumed to be infinitely large. In a real device the distribution function is changed in a complex, highly irregular manner in the vicinity of boundaries, for instance contacts [14] and interfaces [7]. It can be expected that the drift-diffusion approximation fails within a few carrier mean free paths of boundaries.

5. The Basic Semiconductor Equations.

We shall now summarize the results which we have obtained in the previous sections in order to be able to write down a set of equations, the "basic" semiconductor equations. It is obvious that for the sake of transparency and efficiency, one has to perform a trade-off between accuracy and complexity of the model. The equations we concentrate on are valid for the major number of engineering applications, particularly for silicon devices. Certainly, conditions do exist for which their validity is not guaranteed, or at least in doubt. However, as we tried to express in the previous sections, the more sophisticated results in semiconductor physics are too complex to give a rigorous, generally applicable and still sufficiently simple model for the purpose of device simulation.

The basic semiconductor equations consist of Poisson's equation (5.1), the continuity equations for electrons (5.2) and holes (5.3) and the current relations for electrons (5.4) and holes (5.5).

$$\text{div grad } \psi = \frac{q}{\varepsilon} \cdot (n - p - C) \tag{5.1}$$

$$\text{div } \bar{J}_n - q \cdot \frac{\partial n}{\partial t} = q \cdot R \tag{5.2}$$

$$\text{div } \bar{J}_p + q \cdot \frac{\partial p}{\partial t} = -q \cdot R \tag{5.3}$$

$$\bar{J}_n = q \cdot n \cdot \mu_n \cdot \bar{E}_n + q \cdot D_n \cdot \text{grad } n \tag{5.4}$$

$$\bar{J}_p = q \cdot p \cdot \mu_p \cdot \bar{E}_p - q \cdot D_p \cdot \text{grad } p \tag{5.5}$$

To almost this level of completeness, these equations were first presented by VanRoosbroeck [22].

Models for C, the net doping concentration, for R, the net generation/recombination, for μ_n, μ_p, the carrier mobilities, are discussed in the literature. $\bar{E}_n$ and $\bar{E}_p$, the effective fields in the current relations are to first order the electric field, however, one may use supplementary correction terms to account for, e.g., heavy doping. For such mathematical investigations, relatively slight perturbations are of only secondary importance. Hence, for most applications, accounting for some specific effect is possible by properly modeling the parameters in the basic equations.

Conclusion .

We have shown the complete procedure how the basic semiconductor equations are derived from the Boltzmann transport equation plus the assumptions made. In view of some too restrictive assumptions an improvement of the current equations will be a crucial topic. The justification of all assumptions made in this derivation must, of course, be given by a more basic theory. Monte Carlo methods can be expected to serve as an excellent tool for testing classical models and new theories as well.

References.

1 Anselm A.I., "Einführung in die Halbleitertheorie", Akademie Verlag, Berlin, (1964).

2 Baccarani G., Mazzone A.M., "On the Diffusion Current in Heavily Doped Silicon", Solid-State Electron., Vol.18, pp.469-470, (1975).

3 Baccarani G., "Physics of Submicron Devices", Proc.VLSI Process and Device Modeling, pp.1-23, Katholieke Universiteit Leuven, (1983).

4 Capasso F., Pearsall T.P., Thornber K.K., "The Effect of Collisional Broadening on Monte Carlo Simulations of High-Field Transport in Semiconductor Devices", IEEE Electron Dev.Lett., Vol.EDL-2, pp.295, (1981).

5 Cohen M.L., Bergstresser T.K., "Band Structures and Pseudopotential Form Factors for Fourteen Semiconductors of the Diamond and Zinc-blende Structures", Physical Review, Vol.141, pp.789-796, (1966).

6 Conwell E.M., "High Field Transport in Semiconductors", Academic Press, New York, (1967).

7 Gnädinger A.P., Talley H.E., "Quantum Mechanical Calculations of the Carrier Distribution and the Thickness of the Inversion Layer of a MOS Field-Effect Transistor", Solid-State Electron., Vol.13, pp.1301-1309, (1970).

8 Heywang W., Pötzl H.W., "Bandstruktur und Stromtransport", Springer, Berlin, (1976).

9 Hillbrand H., "Untersuchungen des Transportverhaltens von III-V Halbleitern bei hohen elektrischen Feldern mit Monte-Carlo-Methoden", Dissertation, Technische Hochschule Wien, (1974).

10 Hofmann H., "Das elektromagnetische Feld", Springer, (1974).

11 Kireev P.S., "Semiconductor Physics", MIR Publishers, Moscow, (1978).

12 Landsberg P.T., Hope S.A., "Diffusion Currents in Semiconductors", Solid-State Electron., Vol.19, pp.173-174, (1976).

13 Landsberg P.T., Hope S.A., "Two Formulations of Semiconductor Transport Equations", Solid-State Electron., Vol.20, pp.421-429, (1977).

14 Lue J.T., "Theory of Schottky Barrier Heights of Amorphous MIS Solar Cells", Solid-State Electron., Vol.25, No.9, pp.869-874, (1982).

15 Moglestue C., Beard S.J., "A Particle Model Simulation of Field Effect Transistors", Proc.NASECODE I Conf., pp.232-236, (1979).

16 Moglestue C., "A Monte-Carlo Particle Model Study of the Influence of the Doping Profiles on the Characteristics of Field-Effect Transistors", Proc.NASESCODE II Conf., pp.244-249, (1981).

17 Nag B.R., "Diffusion Equation for Hot Electrons", Phys.Rev.B, Vol.11, No.8, pp.3031-3036, (1974).

18 Nag B.R., Chakravarti A.N., "Comments on the Generalized Einstein Relation for Semiconductors", Solid-State Electron., Vol.18, pp.109-110, (1975).

19 Nag B.R., "Parallel Diffusion Constant of Hot Electrons in Silicon", Appl.Phys.Lett., Vol.28, No.9, pp.550-551, (1976).

20 Seeger K., "Semiconductor Physics", Springer, Wien, (1973).

21 Stratton R., "Semiconductor Current-Flow Equations (Diffusion and Degeneracy)", IEEE Trans.Electron Devices, Vol.ED-19, No.12, pp.1288-1292, (1972).

22 Tihanyi J., "Integrated Power Devices", Proc.International Electron Devices Meeting, pp.6-10, (1982).

23 Zimmerl O., "Iterative Lösung der Boltzmanngleichung für heiße Elektronen in InSb", Dissertation, Technische Hochschule Wien, (1972).