Two-Dimensional Simulation of Thermal Runaway in a Nonplanar GTO-Thyristor

Hermann Brand and Siegfried Selberherr, *Fellow, IEEE*

Abstract—**The problem of electrothermal stability due to different cooling conditions has been investigated by computing the thermal transients in a nonplanar GTO-thyristor. In the first simulation, a steady state occurs with a heat sink removing all the dissipated power. In the second simulation severe thermal runaway is induced due to bad cooling conditions, allowing the analysis of destructive electrothermal interaction.**

The simulations are based on an advanced model for self-heating effects in silicon devices derived from first principles of irreversible thermodynamics, self-consistently incorporating a phenomenological model of band gap narrowing in order to take account of heavy doping effects. The system of governing equations is valid in both the steady state and the transient regimes. Four characteristic effects contributing to the heat generation can be identified: Joule heating, recombination heating, Thomson heating, and carrier source heating. Thermal runaway is significantly accelerated in the simulations based on the thermodynamic model of thermoelectric transport compared to a conventional heuristic theory of thermoelectricity.

The importance of the entropy balance equation is emphasized in order to derive the mathematical form of the heat flux and the current relations for electrons and holes. Limitations of underlying assumptions are discussed. It is shown that the heat generation implies the Thomson relations.

I. Introduction

THERMAL runaway is a phenomenon of electrothermal interaction where the dissipation of electrical energy causes a temperature rise over the entire area of the device, resulting in increased current flow and further dissipation, until unrecoverable device failure or "burn-out" of the device occur. This paper focuses on "global" thermal runaway due to bad cooling conditions rather than secondary breakdown which is often explained as "localized" thermal runaway [23].

Thermoelectric effects are associated with the simultaneous flow of charge carriers and heat in a system. Since the early works based on a heuristic approach [1], [5], [10] accurate simulation of transient self-heating effects in semiconductor devices has required the self-consistent solution of the semiconductor equations and the heat flow equation in space and time. However, recently thermoelectric effects have been treated rigorously from a thermodynamic [31] and a hydrodynamic [21] point of view.

Manuscript received October 31, 1994; revised June 29, 1995. The review of this paper was arranged by Associate Editor A. H. Marshak. This work was supported by Siemens Austria, PSE 3.

H. Brand was with the Institute for Microelectronics, Technical University of Vienna, A-1040 Vienna, Austria. He is now with Siemens AG-Austria, Vienna, A-1210 Austria

S. Selberherr is with the Institute for Microelectronics, Technical University of Vienna, A-1040 Vienna, Austria.

IEEE Log Number 9415413.

In Section II the entropy balance equation for the semiconductor is explicitly set up [4], [11], [22]. Identifying conjugate thermodynamic fluxes and forces in the actual mathematical form of the entropy source heat flux and current densities for electrons and holes can be derived [7], [14], [29], [30]. High doping effects are self-consistently accounted for by the incorporation of a phenomenological model of carrier degeneracy and band gap narrowing into the electrothermal transport due to principles of irreversible thermodynamics. Regarding underlying assumptions Section III provides comments on classical thermoelectricity and hydrodynamic theory of energy transport.

In Section IV, the numerical methods used to solve the governing equations for electrothermal transport are sketched. The results obtained in investigating thermal runaway of a GTO thyristor are presented in Section V. The scope is to analyze basic electrothermal interactions as well as to compare simulation results based on a heuristic [1] and the new thermodynamic approach to electrothermal transport.

II. The Mathematical Model

A. Governing Equations

The concept of local equilibrium involves the hypothesis that a continuous system may be regarded as the sum of cells in which thermostatic equilibrium conditions are fulfilled, despite various processes taking place between them. Thus, such quantities as temperature, entropy, internal energy, etc., can be used as field quantities depending on space and time, although according to classical heat theory they are only defined in equilibrium systems.

Basically the concept of internal energy in a solid contains only thermal energy due to random motion and short range forces. Thus the potential energy of built-in and applied electric fields causing long-range forces is excluded. On the other hand the boundary of a thermodynamic system is not certain apriori. The intrinsic semiconductor [30], the doped semiconductor [14] or the 'semiconductor together with the built-in and applied electric field' can be regarded as thermodynamic system to be investigated.

In the latter case the thermodynamic system can be characterized by the total internal energy per unit volume u, for which the first axiom of thermodynamics applies. The conservation equation for u can be written as follows, $\vec{J}_u$ being the total energy flux vector

$$\frac{\partial u}{\partial t} + \operatorname{div} \vec{J}_u = 0. \tag{1}$$

0018–9383/95$04.00 © 1995 IEEE

Note that a conservation equation neither holds for the total internal energy of the separated electron, hole or lattice system, nor for specific energy forms like kinetic or potential energy.

Due to the postulate of local equilibrium the Gibbs function of the thermodynamic system "semiconductor-electric field" $u(\vec{r},t) = u(s(\vec{r},t), n(\vec{r},t), p(\vec{r},t))$ can be applied in each cell of a continuum. $s(\vec{r},t)$ denotes the entropy density, $n(\vec{r},t), p(\vec{r},t)$ the electron and hole concentrations, respectively, in space and time. Actually, by definition of the complete set of independent extensive variables s, n, and p in the Gibbs function it is specified how energy can be exchanged in the semiconductor. Note that the entropy density s contains contributions of the electron, the hole, and the phonon system.

Defining intensive variables $T \equiv \frac{\partial u(s,n,p)}{\partial s}$, $-q \cdot \varphi_n \equiv \frac{\partial u(s,n,p)}{\partial n}$, $q \cdot \varphi_p \equiv \frac{\partial u(s,n,p)}{\partial p}$ the Gibbs fundamental equation of the semiconductor results when the time derivative of the total internal energy u is calculated

$$\frac{\partial u}{\partial t} = T \cdot \frac{\partial s}{\partial t} - q \cdot \varphi_n \cdot \frac{\partial n}{\partial t} + q \cdot \varphi_p \cdot \frac{\partial p}{\partial t}. \tag{2}$$

The quasi-Fermi potentials φ_n, φ_p for electrons and holes, respectively, associated with the thermodynamic system "semiconductor-electric field" reflect that the change of chemical energy of charged particles is always coupled with the change of electrical energy. T is the temperature. Note that equality of the lattice temperature and the carrier temperature has been assumed in the Gibbs function u. Thus, hot carrier effects are neglected.

The Gibbs fundamental equation (2) plays the role of the missing link between electrodynamics and thermodynamics. The actual form of the entropy balance equation is obtained by substitution of continuity equations for n, p [24] and u (1) into the Gibbs fundamental equation (2)

$$\begin{aligned}\frac{\partial s}{\partial t} + \operatorname{div}\left(\frac{1}{T} \cdot \vec{J}_q\right) = \vec{J}_q \cdot \left(-\frac{1}{T^2} \cdot \operatorname{grad} T\right) \\ + \vec{J}_n \cdot \left(-\frac{1}{T} \cdot \operatorname{grad} \varphi_n\right) \\ + \vec{J}_p \cdot \left(-\frac{1}{T} \cdot \operatorname{grad} \varphi_p\right) \\ + q \cdot R \cdot \left(\frac{1}{T} \cdot (\varphi_p - \varphi_n)\right).\end{aligned} \tag{3}$$

The right-hand side is the entropy production G_s per unit volume. $\vec{J}_n, \vec{J}_p$ represent the current densities for electrons and holes, respectively. It is remarkable that the mathematical form of G_s is not unique. Due to the first postulate of irreversible thermodynamics G_s can be expressed as formally equivalent sums of products of different pairs of conjugate thermodynamic fluxes and driving thermodynamic forces.

The difficulty of an adequate definition of the heat flux is the consequence of the fact that in contrast to u and s heat is not a Gibbs function. Heat only appears when energy is exchanged. One possibility to define the heat flux is to generalize the second axiom of classical thermodynamics. Thus the heat flux $\vec{J}_q$ is proportional to the entropy flux $\vec{J}_s$ which is well defined because the entropy density is a Gibbs function

$$\vec{J}_q \equiv T \cdot \vec{J}_s = \vec{J}_u - \varphi_n \cdot \vec{J}_n - \varphi_p \cdot \vec{J}_p. \tag{4}$$

Note that it is equivalent to express $\vec{J}_q$ in terms of fluxes of the total energy and electrochemical energy [4], [14] or in terms of kinetic and chemical energy current densities [30]. An alternative definition of the heat flux is based on the separation of the total internal energy into different energy forms [11], [13], [22]. Apart from the flux of chemical energy the heat flux resulting from this Ansatz is equivalent with (4).

The first postulate of irreversible thermodynamics allows for identifying conjugate thermodynamic fluxes and forces in (3). Due to the second postulate of irreversible thermodynamics linear relations may be assumed. Thus phenomenological equations (5) can be set up for the semiconductor

$$\begin{pmatrix} \vec{J}_n \\ \vec{J}_p \\ \vec{J}_q \\ q \cdot R \end{pmatrix} = \begin{pmatrix} L_{nn} & 0 & L_{nT} & 0 \\ 0 & L_{pp} & L_{pT} & 0 \\ L_{Tn} & L_{Tp} & L_{TT} & 0 \\ 0 & 0 & 0 & L_{RR} \end{pmatrix} \cdot \begin{pmatrix} -\frac{1}{T}\operatorname{grad}\varphi_n \\ -\frac{1}{T}\operatorname{grad}\varphi_p \\ -\frac{1}{T^2}\operatorname{grad} T \\ \frac{1}{T} \cdot (\varphi_p - \varphi_n) \end{pmatrix}. \tag{5}$$

Isotropy has been assumed. q denotes the elementary charge, R the net recombination rate. Recombination processes are represented as a quasichemical reaction in (5), $q \cdot R$ being the reaction velocity, $(\varphi_p - \varphi_n)$ the corresponding driving affinity. Incomplete ionization of impurities can be taken into account by introducing additional reaction velocities and affinities, as can be necessary for low temperature simulations. Equation (5) already accounts for Curie's principle which states that in isotropic systems phenomena which are described by thermodynamic forces and currents of different tensorial order do not interfere with one another.

Due to Onsager's symmetric reciprocal relations $L_{ik} = L_{ki}$, known as the third postulate of irreversible thermodynamics, the number of independent kinetic coefficients in (5) can be reduced. However, the L_{ik} are parametrized in terms of conductivities $\sigma_c \equiv \frac{L_{cc}}{T}$, thermoelectric powers $P_c \equiv \frac{L_{cT}}{T^2 \cdot \sigma_c}$ for the carrier concentrations $c = n, p$, respectively, and the thermal conductivity $\kappa \equiv \frac{L'_{TT}}{T^2} - T \cdot \sum_{c=n,p} \sigma_c \cdot P_c^2$ since the latter are accessable by measurement and can hence be used to characterize material properties. The definition of the heat conductivity κ is usually given under the condition that no net transfer of charge carriers to the surrounding takes place. The semiconductor equations [24], however, require to describe the electron and hole current flow separately. Therefore it is adequate to define the heat conductivity κ under the assumption that either $\vec{J}_n$ or $\vec{J}_p$ is zero instead of the total current $\vec{J} = \vec{J}_n + \vec{J}_p$. Thus κ does not explicitely contain the ambipolar contribution to the heat conductivity, which is, however, implicitly included in (6) through (8) (details in Section III). L'_{TT} subsumes a lattice, a hole, and an electron contribution. Both limiting cases, flow of electric charge due to the imposition of the quasi-Fermi potential (Ohm's law), and flow of heat caused by a temperature gradient (Fourier's law) are included in the resulting current relations for the thermoelectric transport in a semiconductor

$$\vec{J}_n = \sigma_n \cdot (-\operatorname{grad}\varphi_n - P_n \cdot \operatorname{grad} T) \tag{6}$$

$$\vec{J}_p = \sigma_p \cdot (-\operatorname{grad}\varphi_p - P_p \cdot \operatorname{grad} T) \tag{7}$$

$$\vec{J}_q = T \cdot P_n \cdot \vec{J}_n + T \cdot P_p \cdot \vec{J}_p - \kappa \cdot \operatorname{grad} T. \tag{8}$$

Note, that the form of the current relations is derived with this Ansatz, whereas in [31] it is assumed.

With the heat flux (8) the entropy balance equation (3) can be transformed into the heat flow equation. Heat capacities of the semiconductor must be determined considering also the contributions of the carriers to the entropy. Due to the concept of local equilibrium not only the Gibbs fundamental equation is valid but also all thermostatic relations [13], allowing to utilize Maxwell's relations to eliminate the entropy density in the left hand side of (3) without further assumptions [4], [13], [17]

$$T \cdot \frac{\partial s}{\partial t} = C_{n,p} \cdot \frac{\partial T}{\partial t} + q \cdot T \cdot \frac{\partial \varphi_n}{\partial T} \cdot \frac{\partial n}{\partial t} - q \cdot T \cdot \frac{\partial \varphi_p}{\partial T} \cdot \frac{\partial p}{\partial t}. \tag{9}$$

$C_{n,p} \equiv T \cdot \frac{\partial s}{\partial T}$ is a definition of the heat capacity per unit volume.

In order to gain a system of partial differential equations which is reduced to the semiconductor equations in the isothermal case, the dependencies of the quasi-Fermi potentials on ψ, n, p, T, and the effective intrinsic carrier concentration n_{ie}, in order to fit heavy doping effects, have to be explicated. Using Boltzmann statistics the expansion of the gradient of the quasi-Fermi potentials with respect to ψ, n_{ie}, n, p, T in (6) and (7) yields the final form of the current relations

$$\vec{J}_n = q \cdot \mu_n \cdot n \cdot \left(-\operatorname{grad} \psi - \frac{k \cdot T}{q} \cdot \operatorname{grad} (\ln n_{ie}) + \frac{k \cdot T}{q} \cdot \frac{1}{n} \cdot \operatorname{grad} n + P_n^{\text{eff}} \cdot \operatorname{grad} T \right) \tag{10}$$

$$\vec{J}_p = q \cdot \mu_p \cdot p \cdot \left(-\operatorname{grad} \psi + \frac{k \cdot T}{q} \cdot \operatorname{grad} (\ln n_{ie}) - \frac{k \cdot T}{q} \cdot \frac{1}{p} \cdot \operatorname{grad} p - P_p^{\text{eff}} \cdot \operatorname{grad} T \right). \tag{11}$$

In (10) and (11) the conductivity $\sigma_{n,p}$ has been expressed in terms of the mobility $\mu_{n,p}$ and carrier concentrations. k is the Boltzmann constant. $P_{n,p}^{\text{eff}}$ denotes the effective thermoelectric powers which are defined as follows:

$$\left(\frac{k}{q} \cdot \ln \frac{c}{n_{ie}} - \frac{k \cdot T}{q} \cdot \frac{1}{n_{ie}} \cdot \frac{\partial n_{ie}}{\partial T} \right) \mp P_c \equiv P_c^{\text{eff}} \qquad c = n, p; P_n < 0. \tag{12}$$

Equations (10) and (11) are said to be the extended-drift extended-diffusion approximations of the current relations. That is, an effective electric field is introduced because of the possible dependence of the intrinsic carrier concentration on position, and the diffusion not only accounts for concentration gradients but also for temperature gradients. Note that including a single quasi-field corrective term in the current relations (10) and (11) allows for accurately modeling the combined phenomenon of carrier degeneracy and band gap narrowing within the context of classical statistics, provided that this correction is based on measurements of the higher $n \cdot p$-product being the phenomenological effect of heavy doping, inherently including both phenomenons [2]. Equation (12) reflects the dependence of the quasi-Fermi potentials on lattice temperature.

Applying Boltzmann statistics again (9) together with (8) and (3) and some algebraic operations yield the following final heat flow equation:

$$\begin{aligned} C_{n,p} \cdot \frac{\partial T}{\partial t} &+ \operatorname{div} (-\kappa \cdot \operatorname{grad} T) \\ &= \frac{\vec{J}_n \cdot \vec{J}_n}{q \cdot \mu_n \cdot n} + \frac{\vec{J}_p \cdot \vec{J}_p}{q \cdot \mu_p \cdot p} \\ &\quad - \vec{J}_n \cdot T \operatorname{grad} \left(-P_n^{\text{eff}} + \left(\frac{k_B}{q} \cdot \ln \frac{n}{n_{ie}} - \frac{k_B \cdot T}{q} \cdot \frac{1}{n_{ie}} \cdot \frac{\partial n_{ie}}{\partial T} \right) \right) \\ &\quad - \vec{J}_p \cdot T \cdot \operatorname{grad} \left(P_p^{\text{eff}} - \left(\frac{k_B}{q} \cdot \ln \frac{p}{n_{ie}} - \frac{k_B \cdot T}{q} \cdot \frac{1}{n_{ie}} \cdot \frac{\partial n_{ie}}{\partial T} \right) \right) \\ &\quad + q \cdot R \cdot \left(2 \cdot T \cdot \frac{k \cdot T}{q} \cdot \frac{1}{n_{ie}} \cdot \frac{\partial n_{ie}}{\partial T} \right) \\ &\quad + T \cdot P_n^{\text{eff}} \cdot \operatorname{div} \vec{J}_n - T \cdot P_p^{\text{eff}} \cdot \operatorname{div} \vec{J}_p. \end{aligned} \tag{13}$$

Equation (13) is valid in both the steady state and the transient regimes. Its right-hand side represents the heat generation which describes all local energy conversion processes contributing to self-heating of the device due to dissipation of electrochemical energy. Four characteristic heat sources can be distinguished. The first two terms represent Joule heating due to electron and hole current, respectively. As Joule heating is proportional to $J_{n,p}^2$ it must always be positive. The next two expressions take account of Thomson heating. Due to its proportionality to $J_{n,p}$ the sign depends on the current's direction and the carrier type. The subsequent term is recognized as generalized recombination heat. The last two contributions are interpreted as carrier source heat due to electron and hole conduction, respectively. It represents the change of the energy of the system due to the net change of the concentration of charged particles in a volume element with a certain energy expressed in terms of the thermoelectric power times the temperature. The variation of the total internal energy due to the local charging rate is especially significant as a transient phenomenon.

The governing equations comprising Poisson's equation, continuity equations for n and p [24] supplemented with the current relations (10) and (11) and the heat flow equation (13) depend nonlinearly on the lattice temperature. If lattice heating is significant, the thermal system becomes tightly coupled to the electrical system. For the electrical subsystem, either ohmic contacts or homogeneous Neumann boundaries are assumed. Mixed boundary conditions (14) for the heat flow equation are mandatory in order to be able to model realistic imperfect cooling conditions

$$-\kappa \cdot \operatorname{grad} T \cdot \vec{N} = h \cdot (T - T_{\text{sink}}). \tag{14}$$

h denotes the heat sink thermal conductance and T_{sink} the ambient temperature of the heat sink. $\vec{N}$ represents the unit normal vector. Note that the left hand side in (14) results

from the total energy flux density in the limiting case when surface recombination processes of carriers are neglected. The proper modeling of imperfect cooling conditions is of special importance for transient electrothermal simulations, as the characteristic time for self-heating increases with increasing external thermal resistance.

B. Physical Parameters

The dependence of the recombination rate and the mobility on temperature has been taken from [24]. Auger recombination and carrier-carrier scattering—known as limiting physical effects for high injection conditions in power semiconductor devices [1]—have been taken into account. The phenomenological mobility model includes velocity saturation due to high electric fields. Thus the regime of applicability of the drift current expression in (10) and (11) can be extended phenomenologically despite the fact that doing so means withdrawing the principle of linear transport theory requiring the independence of kinetic coefficients from driving forces. It is noteworthy that it is not an infringement of linear transport theory to comprehend the mobility as a function of the carrier temperature.

In the high temperature range the intrinsic carrier concentration n_i as a function of temperature [25] has been fitted to the data from [28]. As heavy doping effects limit thyristor operation, the effective intrinsic carrier concentration is computed using [26]. Calculating the change of n_{ie} with temperature gives the result

$$\frac{\partial n_{ie}}{\partial T} = n_{ie}\cdot\left(\frac{E_g}{2\cdot k\cdot T^2}+\frac{3}{2\cdot T} - \frac{q\cdot V_1\cdot\left(\ln\left(\frac{N_D^+ + N_A^-}{N_0}\right)+\sqrt{\ln^2\left(\frac{N_D^+ + N_A^-}{N_0}\right)+0.5}\right)}{2\cdot k\cdot T^2}\right) + n_{ie}\cdot\left(\frac{9.025\cdot 10^{-5}\frac{\mathrm{eV}}{\mathrm{K}}+6.10\cdot 10^{-7}\frac{\mathrm{eV}}{\mathrm{K}}\cdot\left(\frac{T}{\mathrm{K}}\right)}{2\cdot k\cdot T}\right) + n_{ie}\cdot\left(\frac{3}{4}\cdot m_0\cdot\left(\frac{1.4\cdot 10^{-4}\frac{1}{\mathrm{K}}\cdot m_n^* + 4.5\cdot 10^{-4}\frac{1}{\mathrm{K}}\cdot m_p^*}{m_n^*\cdot m_p^*}\right)\right). \tag{15}$$

In order to determine $P_{n,p}$ and $P_{n,p}^{\mathrm{eff}}$ (12) either $P_{n,p}$ or $P_{n,p}^{\mathrm{eff}}$ has to be interpreted physically and modeled adequately. It can be shown that with the help of the effective thermoelectric power thermal diffusion due to kinetic excess energy in hot regions is described. Utilizing results of kinetic transport theory one obtains [3], [7], [18], [29]

$$P_{n,p}^{\mathrm{eff}} = \frac{1}{T}\cdot\frac{1}{q}\cdot\left[\left(r_{n,p}+\frac{5}{2}\right)\cdot k_B\cdot T - \frac{3}{2}\cdot k_B\cdot T\right] = \frac{k_B}{q}\cdot(r_{n,p}+1) = \frac{D_{n,p}^T}{\mu_n}. \tag{16}$$

The numerical parameter $r_{n,p}$ depends on the dominating scattering mechanism and can be computed as a function of mobilities $r_{n,p} = r_{n,p}(\mu^L, \mu^I)$ [6]. μ^L is the mobility accounting for lattice scattering, μ^I for impurity scattering only. As the effective thermoelectric power $P_{n,p}^{\mathrm{eff}}$ relates the thermal diffusivity $D_{n,p}^T$ to the mobility it may be called "Soret factor."

Inserting (15) and (16) into (12) gives the thermoelectric powers P_n, P_p as a result. Due to the third term in the right hand side of (15) they decrease in highly doped regions, which is a very remarkable property of the derived expressions for the thermoelectric powers compared to conventional models of P_n, P_p [28], [29], [31]. Furthermore the dependence of the band gap $E_g$and the effective masses m_n^*, m_p^* on temperature is accounted for in P_n, P_p. It can be shown that the thermoelectric powers physically have to be interpreted as entropy per elementary charge of one electron or hole, respectively [3], [14], [29].

III. Comments on Thermoelectric Transport

The contribution of the carriers to the overall entropy can be regarded as negligible in the low temperature range, compared to the Debye temperature, which is 645 K for silicon [16]. The entropy of the lattice even dominates in the high temperature range. Omitting the last two terms in (9) the heat flow equation (13) can be reduced. Under steady state conditions the divergence of the electron and hole current densities equals the recombination rate of carriers times the elementary charge. So recombination heat and carrier source heat can be summarized resulting in the quasisteady state approximation of the heat generation, which is often used for transient simulations too [15], [31], [32].

In order to demonstrate that the heat flow equation implies Thomson's laws, it has to be written in terms of the total current density $\vec{J} = \vec{J}_n + \vec{J}_p$ and the difference of quasi-Fermi potentials $\delta\varphi = \varphi_p - \varphi_n$ [20]

$$\begin{aligned} C_{n,p}\cdot\frac{\partial T}{\partial t} + \operatorname{div}(-\kappa'\cdot\operatorname{grad} T) &= \frac{\vec{J}\cdot\vec{J}}{\sigma} + P\cdot\vec{J}\cdot\operatorname{grad} T - \vec{J}\cdot\operatorname{grad}(T\cdot P) \\ &\quad - (T\cdot P)\cdot\operatorname{div}\vec{J} + q\cdot R\cdot\delta\varphi \\ &\quad + \frac{\Omega}{T}\cdot\operatorname{grad}\delta\varphi\cdot\operatorname{grad} T + \frac{\sigma_n\cdot\sigma_p}{\sigma_n+\sigma_p}\cdot(\operatorname{grad}\delta\varphi)^2 \\ &\quad + \operatorname{grad}(\Omega\cdot\operatorname{grad}\delta\varphi). \end{aligned} \tag{17}$$

In (17) $P = \frac{P_n\cdot\sigma_n + P_p\cdot\sigma_p}{\sigma_n+\sigma_p}$ denotes the total thermoelectric power. $\sigma = \sigma_n + \sigma_p$ represents the total electrical conductivity. Note that $\kappa' = \left(\kappa + T\cdot\frac{\sigma_n\cdot\sigma_p}{\sigma_n+\sigma_p}\cdot(P_p - P_n)^2\right)$ comprises the ambipolar contribution of the thermal conductivity. $\Omega = \frac{\sigma_n\cdot\sigma_p}{\sigma_n+\sigma_p}\cdot T\cdot(P_p - P_n)$ is an auxiliary quantity.

Regarding (17) in the limiting case $\delta\varphi = 0$, all contributions to the heat generation due to two-band conduction vanish

$$\begin{aligned} C_{n,p}\cdot\frac{\partial T}{\partial t} + \operatorname{div}(-\kappa\cdot\operatorname{grad} T) &= \frac{\vec{J}\cdot\vec{J}}{\sigma} - \vec{J}\cdot(\operatorname{grad}\Pi - P\cdot\operatorname{grad} T) \\ &\quad - \Pi\cdot\operatorname{div}\vec{J}. \end{aligned} \tag{18}$$

The Peltier coefficient $\Pi \equiv T \cdot P$ is defined in conformance with Thomson's second law. In the classical theory of thermoelectricity steady state conditions are usually presumed. Hence the divergence of the current density $\vec{J}$ vanishes. Furthermore it is assumed that the Peltier coefficient is only a function of the temperature, which is true in a homogeneous material. So finally

$$\operatorname{div}(-\kappa \cdot \operatorname{grad} T) = \frac{\vec{J} \cdot \vec{J}}{\sigma} + \Theta \cdot \vec{J} \cdot \operatorname{grad} T \qquad (19)$$

is obtained in perfect agreement with Thomson's first law being a definition of the Thomson coefficient $\Theta \equiv \left(\frac{\partial \Pi}{\partial T} - P\right)$ [27].

Within the theory of thermoelectricity neither individual contributions of the electron, the hole, and the phonon system to the overall energy, nor various forms of energy, such as kinetic or potential energy, e.g., in the conduction band are treated separately. On the other hand the hydrodynamic energy balance equation represents a balance equation of kinetic energy of carriers only, neglecting their position dependent potential energy due to built-in potentials in an inhomogeneously doped semiconductor. Furthermore the change of potential energy of carriers due to recombination processses of carriers is not described. Thus the thermoelectric model based on hydrodynamics [21] is restricted to the special case of a homogeneous one band conductor. The constant thermoelectric power only accounts for the contributions due to the transport energy of carriers with the assumption of an energy independent relaxation time. The dependence of the thermoelectric power on the Fermi energy is neglected at all.

The theory of thermoelectricity of either the electron, the hole, or the phonon system delivers an energy balance equation accounting for the total energy of carriers as a sum of the kinetic and potential energy in analogy to the microscopic picture of an electron in the band diagram. The continuity equation of potential energy can be derived from the continuity equation of carriers. It accounts for inhomogenities of potential energy due to doping and interband transitions in a two-band conducting solid. Neglecting the contributions of the drift energy to the mean kinetic energy of carriers the hydrodynamic energy balance equation, supplemented by a continuity equation of potential energy, delivers the heat flow equation which also results from principles of irreversible thermodynamics [3].

The limitations of the assumption of local equilibrium and the approximation of linear transport theory can hardly be discussed separately. By solving the Boltzmann equation with Enskog's method it can be shown theoretically that at least in the case when transport phenomena can be described by linear phenomenological laws the hypothesis of local equilibrium is also fulfilled [12].

Beyond linear irreversible thermodynamics three regimes can be distinguished. First, system properties (e.g., transport parameters) turn out to be functions of driving forces. Second, if the system is perturbed further qualitatively new phenomena may occur e.g. turbulences, shock waves or fast plasma processes. Finally, it is possible that the local equilibrium does not hold any more [13]. The concept of local equilibrium breaks down when continuum theory of matter cannot be applied any more. That is, characteristic spacial and temporal variations of macroscopic quantities exhibit the same order of magnitude as characteristic microscopic lengths and times. If the variation of temperature and velocity within the mean free path is small compared to the absolute temperature and the velocity of sound [13], [27], relaxation processes are strong enough to guarantee adjustment of local equilibrium.

In principle phenomena of all three categories can be found in submicrometer semiconductor devices (e.g., electron shock waves [9]). The fact that, within linear transport theory transport parameters may be functions of local intensive variables but must not depend on driving forces, represents the most serious restriction with respect to device simulation. Due to the spacial extensions of power semiconductor device structures, continuum theory of matter and consequently the concept of local equilibrium can be applied. As high dissipation capability is a design goal for power semiconductor devices, dissipative processes due to scattering are strong enough in order to compensate perturbations by external forces. Actually, exchange processes between subsystems of the semiconductor are magnitudes of order slower than those within a single subsystem. Thus it can be assumed that each subsystem locally is in a variable thermostatic equilibrium state, although—regarding a specific cell—the subsystems are not in equilibrium with one another.

IV. Solution Methods

Usually, the characteristic time for the electrical and the thermal subsystem differs by several orders of magnitude. Thus a device is often in steady state from the electrical point of view, whereas thermally it is still in transition. Provided the applied voltages are time independent, the electrical transient can be neglected with respect to the thermal transient, an assumption which is well satisfied in practice [5], [10].

Spatial discretization is obtained using finite boxes, a generalization of finite differences [8], [24], while time is discretized with the backward Euler method. The electrothermal problem is computed self-consistently following a decoupled approach. At each time step the electrical subsystem of equations is solved first, the lattice temperature being regarded as an independent variable. Then, the temperature distribution is updated by solving the heat flow equation. Newton's method and LU-decomposition are used to solve the electrical and thermal subsystem alternately until convergence is attained.

V. Results

One half of a symmetric GTO-thyristor cell has been investigated. Geometry and doping profiles are specified following [19]. Fig. 1 is a sketch of the simulation area, which is 254 μm long and 160 μm wide. The anode contact covers the whole right side in Fig. 1. The $2 \cdot 10^{14}$ cm^{-3} doped n-base region is 150 μm long. The surface doping concentration of the emitter is $1 \cdot 10^{20}$ cm^{-3}, the doping concentration in the anode and p-base amounts to $1.5 \cdot 10^{18}$ cm^{-3}. Note that the quadrangle at the left bottom corner is not part of the thyristor.

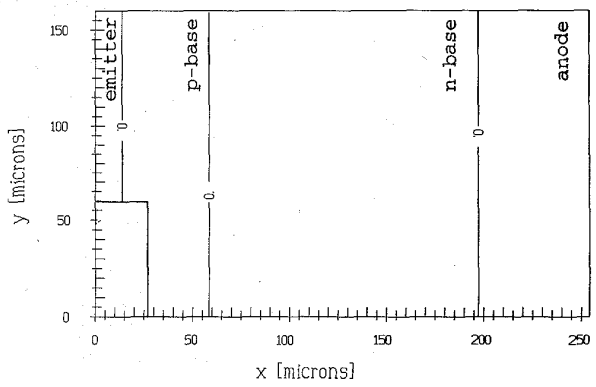


Fig. 1. Geometry of simulation area in μm (one half of symmetric GTO-thyristor cell).

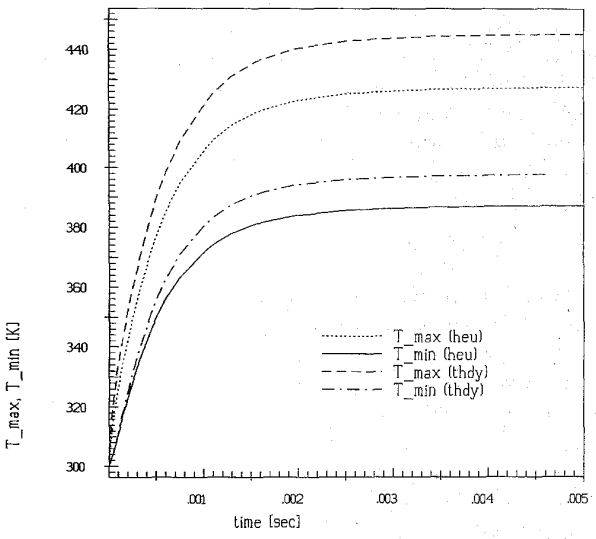


Fig. 2. Time evolution of maximum and minimum temperature [K] when the heat sink thermal conductance h is 50 $\mathrm{Wcm^{-2}K^{-1}}$, (heat sink temperature T_{sink}: 300 K).

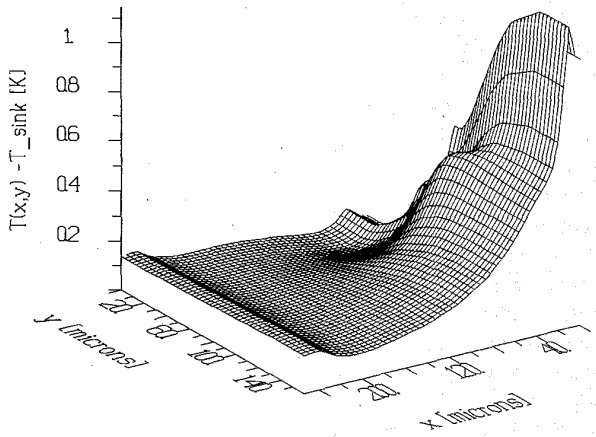


Fig. 3. Increase of lattice temperature [K] after 10^{-5} seconds.

First, the electrical steady state is computed for the thyristor in the on-state, where significant current filamentation occurs. The anode voltage is 3.3 V, the anode current 28 A. In [19] switching off the thyristor under an ohmic load has been simulated. In this paper it is investigated if and how a dynamic electrothermal equilibrium evolves due to given cooling conditions. The primary interest focuses on the thermal behavior and thermally induced electrical behavior of the thyristor.

With the isothermal electrical steady state solution as a starting point, the subsequent thermal transients are calculated. Although the thyristor is in a steady state from an electrical point of view (time independent applied voltages), it is in transition thermally due to self-heating because of energy dissipation. In order to show the physical significance of various contributions to the heat generation in (13), simulations have also been performed using a heuristic model of heat generation, allowing to compare results based on a rigorous and a conventional model on thermoelectric transport. The heuristic Ansatz only contains a simplified expression for Joule heat and recombination heat [1], [24].

A. The Electrothermal Steady State

In the first series of computations for the heat sink thermal conductance $h = 50$ W/cm^{2}K has been chosen [1]. The heat sink temperature T_{sink} is 300 K. Double sided cooling is employed.

Fig. 2 shows the evolution of the internal maximum and minimum temperatures of the thyristor in time. After about three milliseconds self-heating ceases. An electrothermal steady state adjusts representing a dynamic equilibrium of the electrical and thermal subsystem of the device. The thyristor consumes electrical energy and delivers heat energy to the heat sinks. Within the device energy conversion takes place.

The numerical results agree very well with the analytical solution of an equivalent simplified thermal network, consisting of a thermal capacitor C_{th} and a thermal resistor R_{th}. If at the time t_0 the dissipated power P_d is applied to the thermal network, following temporal increase of the temperature results, provided the initial condition is $T = T_{\mathrm{sink}}$

$$\Delta T = T - T_{\mathrm{sink}} = R_{\mathrm{th}} \cdot P_d \cdot \left(1 - \exp\left(-\frac{t}{R_{\mathrm{th}} C_{\mathrm{th}}}\right)\right). \quad (20)$$

P_d can be computed by integration of the heat generation in the whole volume of the device.

Figs. 3 and 4 are snapshots of the internal lattice temperature distribution after 10^{-5} and 0.005 seconds. There are significant temperature gradients only at the beginning of self-heating, the absolute temperature differences $T - T_{\mathrm{sink}}$, however, are low. As the thyristor warms up the temperature profile flattens.

B. Thermal Runaway

In order to induce severe thermal runaway, bad cooling conditions are defined by choosing 5 W/cm^{2}K for the heat sink thermal conductance. It can be seen from Fig. 5 that the increase of the computed temperature flattens after a few mil-

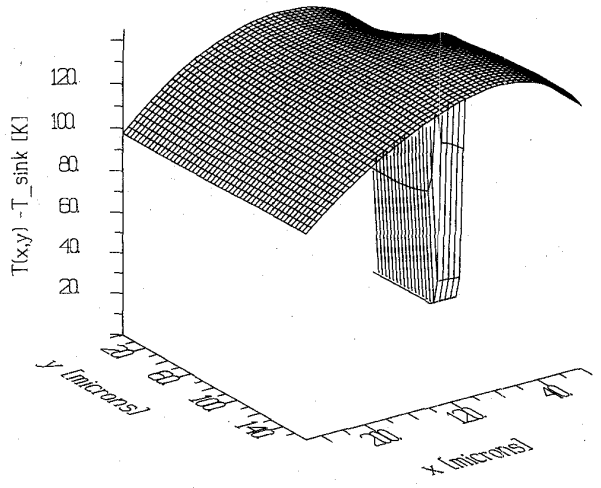


Fig. 4. Increase of lattice temperature [K] after $5 \cdot 10^{-3}$ seconds.

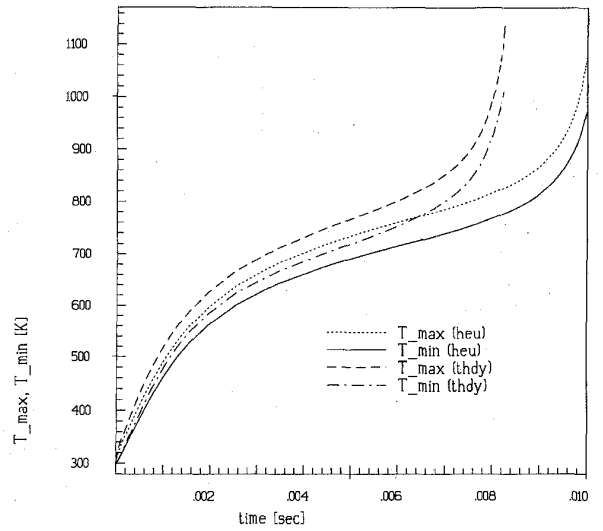


Fig. 5. Time evolution of maximum and minimum temperature [K], ($h = 5$ Wcm^{-2}K^{-1}, $T_{\text{sink}} = 300$ K).

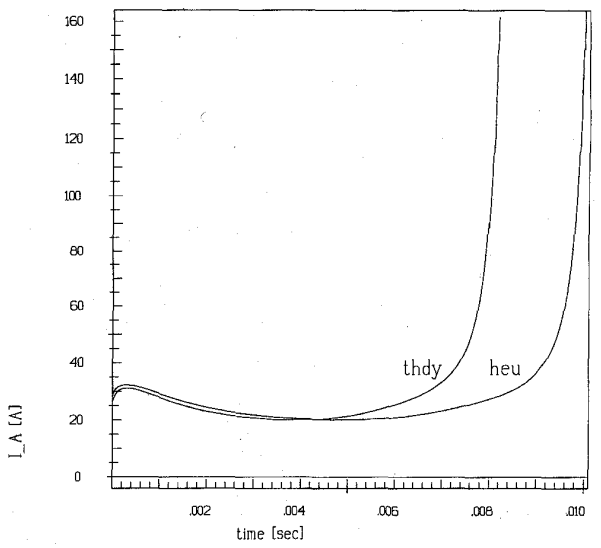


Fig. 6. Time evolution of anode current [A] in case of severe thermal runaway.

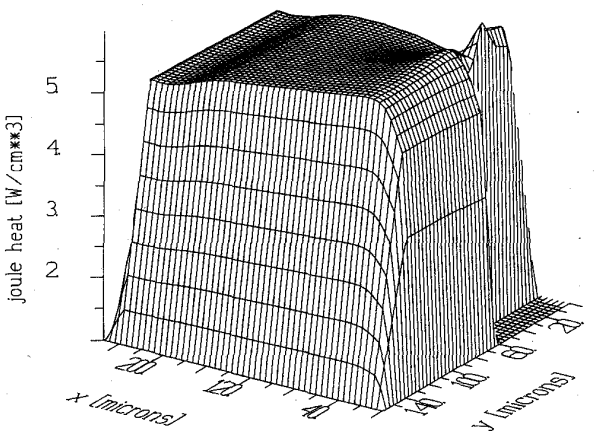


Fig. 7. Joule heat [Wcm^{-3}] (in log scale) within the GTO-thyristor in case of severe thermal runaway after $5.0 \cdot 10^{-3}$ seconds, ($h = 5$ Wcm^{-2} K^{-1}, $T_{\text{sink}} = 300$ K).

liseconds. The thyristor seems to approach a steady state. After 6 ms, however, the temperature curve precipitously bends upwards. The maximum thyristor temperature exponentially ascends about 100 K within 0.1 ms. On the time scale of self-heating the temperature increase until the melting point is reached takes place quasi instantanously. The evolution of the temperature in time obtained with a heuristic model of heat transport is qualitatively equal. The onset of the exponential increase of temperature, however, occurs about 2 ms later (see Fig. 5).

The progress of the thyristor's thermoelectric destabilization in the on-state can also be traced by the evolution of the anode current in time. Three periods can be distinguished in Fig. 6. At the beginning the anode current rises for a short time because in the boundary region of the nonplanar p-base more carriers are generated, flowing off to the anode and emitter contact. Because the carrier mobility decreases with increasing temperature the second period until about 4 ms is characterized by a decline of the anode current. The dependence of the mobility on temperature limits the anode current thus stabilizing thyristor operation. Finally the mobility effect is overcome by thermal generation of carriers. The electrical conductivity increases with temperature again because more mobile carriers are available for current flow. At last the exponential dependence of the effective intrinsic carrier concentration on temperature causes an exponential rise of the anode current. As the removal of heat to the heat sinks is only proportional to the difference of the device temperature and the heat sink temperature there is a misrelation of produced and removed heat causing an exponential increase of the temperature too.

Figs. 7–10 show Joule, Thomson, recombination, and carrier source heat (in log scale) as four separated contributions

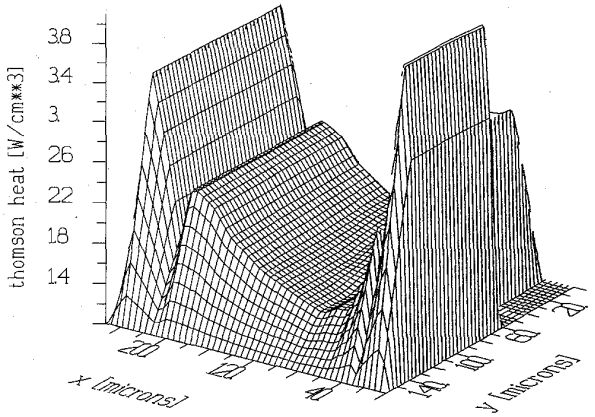


Fig. 8. Thomson heat [Wcm^{-3}] (in log scale) within the GTO-thyristor in case of severe thermal runaway after $5.0 \cdot 10^{-3}$ seconds, ($h = 5$ $\mathrm{Wcm}^{-2}\mathrm{K}^{-1}, T_{\mathrm{sink}} = 300$ K).

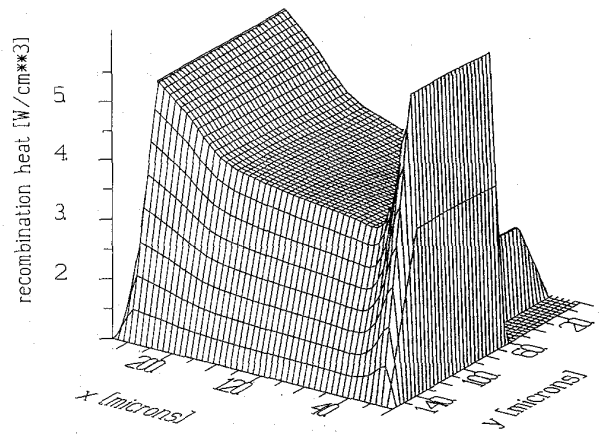


Fig. 9. Recombination heat [Wcm^{-3}] (in log scale) within the GTO-thyristor in case of severe thermal runaway after $5.0 \cdot 10^{-3}$ seconds, ($h = 5$ $\mathrm{Wcm}^{-2}\mathrm{K}^{-1}, T_{\mathrm{sink}} = 300$ K).

to the overall heat generation. The simulation time is 5 ms. In Fig. 7 a pronounced Joule heat production due to current crowding has to be confirmed in the emitter. The nonplanar device geometry between emitter and gate contacts represents a perturbation of electric field and current flow. Due to a field peak and local current crowding a striking maximum of Joule heat results. All together the Joule heat produces the main contribution to the overall losses. Thomson heat in Fig. 8 yields significant contributions only in the highly doped boundary regions. The reason is that the thermoelectric power exhibits pronounced gradients only in transition regions to the highly doped emitter and anode.

While Joule and Thomson heat are functions of the electrical current densities, the recombination and carrier source heat depend on the local increment of carrier concentrations. Due to high injection conditions the recombination heat is very pronounced in the whole conducting thyristor area. Fig. 9 shows that recombination heat exhibits maxima in the emitter and the anode region as well. These contributions are brought

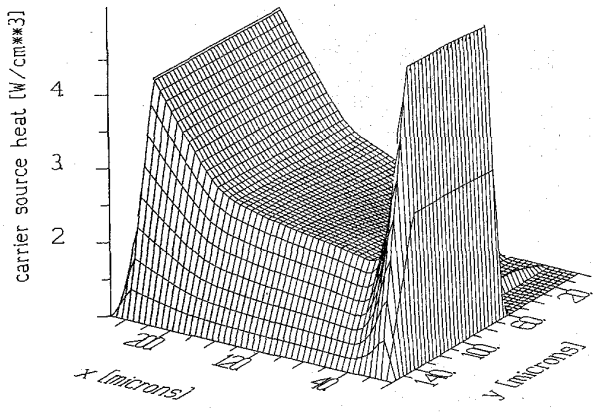


Fig. 10. Carrier source heat [Wcm^{-3}] (in log scale) within the GTO-thyristor in case of severe thermal runaway after $5.0 \cdot 10^{-3}$ seconds, ($h = 5$ $\mathrm{Wcm}^{-2}\mathrm{K}^{-1}, T_{\mathrm{sink}} = 300$ K).

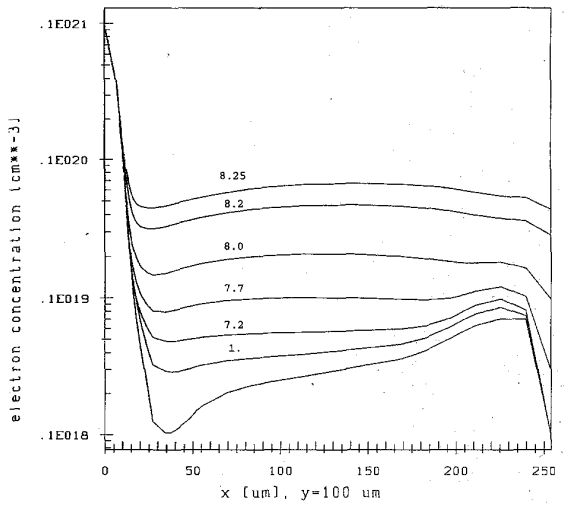


Fig. 11. Time evolution of the electron concentration [cm^{-3}] at the center of the emitter after 1 μs and 1, 7.2, 7.7, 8.0, 8.20, and 8.25 ms in case of thermal runaway.

about by excessive injection of minority carriers recombining with majority carriers there. Fig. 10 displays the logarithm of carrier source heat. Maxima can be found in the emitter and anode region too, on the one hand, because the effective thermoelectric power takes higher values there than in the n-base region, on the other hand for the reason, that the divergence of electrical current densities approximately equals the recombination rate under quasisteady state conditions.

In order to visualize the evolution of dependent variables n and p in time in Figs. 11 and 12, the thyristor has been cut at the center of the emitter. As lattice temperature increases the doping profile loses its implication in the majority carrier concentrations and hence in device operation. Only the electron concentration in the emitter is determined by the high doping level until irreparable device failure occurs. The rest of the device is flooded with electrons and holes which have been generated thermally. Note the high injection level

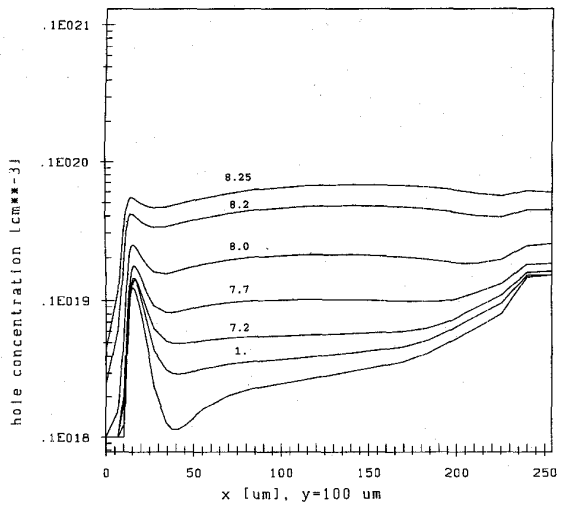


Fig. 12. Time evolution of the hole concentration [cm^{-3}] at the center of the emitter after 1 μs and 1, 7.2, 7.7, 8.0, 8.20, and 8.25 ms in case of thermal runaway.

of minority carriers in the highly doped boundary regions. In the high temperature regime both base regions are filled with a quasineutral electron hole plasma. The space charge disappears as the temperature profile rises. Thus, Poisson's equation degenerates to Laplace's equation and the potential drops linearly between the anode and emitter contacts. The GTO-thyristor behaves as an ohmic resistor which shorts under the given bias because of the excessive thermal conductivity modulation. Note that the boundary conditions for ψ, n and p change in time due to different values of the effective intrinsic carrier concentration determing the built-in potential and the equilibrium concentration of electrons and holes at the contacts.

VI. Conclusion

An advanced model for self-heating effects has been derived from principles of irreversible thermodynamics, selfconsistently including an empiric model for heavy doping effects. Being valid in both the stationary and the transient regimes it exhibits four characteristic contributions to the overall heat generation, Joule heat, Thomson heat, recombination heat, and carrier source heat.

The importance of the entropy balance equation is emphasized in order to derive the mathematical form of the heat flux and the current relations for electrons and holes. Comments on thermoelectric transport are given relating the resulting electrothermal transport model to the theory of classical thermoelectricity and hydrodynamics. Limitations of underlying assumptions are discussed.

The problem of electrothermal stability due to different cooling conditions has been investigated by computing the thermal transients in a nonplanar GTO-thyristor. Simulations based on the thermodynamic model of electrothermal transport show that thermal runaway is significantly accelerated by carrier source heat, Thomson heat and additional contributions in the generalized expressions describing Joule heating and recombination heating, compared to results obtained by usage of a heuristic theory of thermoelectricity.

The simulation results allow extraction of the thermal relaxation time, and the value of the total thermal resistor and capacitor for the equivalent thermal circuit model of the device under investigation.

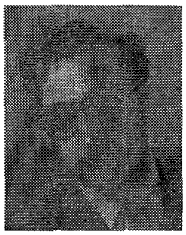
Acknowledgment

The authors give special thanks to PSE 362 for the provision of computer resources.

References

[1] M. S. Adler, "Accurate calculations of the forward drop and power dissipation in thyristors," *IEEE Trans. Electron Devices*, vol. ED-25, pp. 16–22, Jan. 1978.
[2] ——, "An operational method to model carrier degeneracy and band gap narrowing," *Solid-State Electron.*, vol. 26, pp. 279–283, May 1983.
[3] H. Brand, "Thermoelektrizität und hydrodynamik," Ph.D. thesis, in German, Technical University of Vienna, Austria, 1994.
[4] H. B. Callen, *Thermodynamics*. New York: Wiley, 1966.
[5] A. Chryssafis and W. Love, "A computer-aided analysis of one-dimensional thermal transients in n-p-n power transistors," *Solid-State Electron.*, vol. 22, pp. 249–256, 1978.
[6] J. M. Dorkel, "On electrical transport in nonisothermal semiconductors," *Solid-State Electron.*, vol. 26, pp. 819–821, Aug. 1983.
[7] J. R. Drabble and H. J. Goldsmid, *Thermal Conduction in Semiconductors*. New York: Pergamon Press, 1961.
[8] A. F. Franz, G. A. Franz, S. Selberherr, and P. Markowich, "Finite boxes—A generalization of the finite-difference method suitable for semiconductor device simulation," *IEEE Trans. Electron Devices*, vol. ED-30, pp. 1070–1083, Sept. 1983.
[9] C. L. Gardner, "Numerical simulation of a steady-state electron shock wave in a submicron semiconductor device," *IEEE Trans. Electron Devices*, vol. 38, pp. 392–398, Feb. 1991.
[10] S. P. Gaur and D. H. Navon, "Two-dimensional carrier flow in a transistor structure under nonisothermal conditions," *IEEE Trans. Electron Devices*, vol. ED-23, pp. 50–57, Jan. 1976.
[11] P. Glansdorf and I. Prigogine, *Thermodynamic Theory of Structure, Stability and Fluctuation*. New York: Wiley, 1971.
[12] S. R. De Groot and P. Mazur, *Grundlagen der Thermodynamik Irreversibler Prozesse*. Mannheim-Zürich: Bibliographisches Institut, 1969.
[13] I. Gyarmati, *Non-Equilibrium Thermodynamics*. Berlin: Springer, 1970.
[14] T. C. Harman and J. M. Honig, *Thermoelectric and Thermomagnetic Effects and Applications*. New York: McGraw-Hill, 1967.
[15] K. Kells, S. Müller, G. Wachutka, and W. Fichtner, "Simulation of self-heating effects in a power p-i-n diode," in *Proc. 5th SISDEP Conf.*, Vienna, Austria, Sept. 1993, pp. 41–44.
[16] C. Kittel, *Einführung in die Festkörperphysik*. München-Wien: R. Oldenbourg Verlag, 1976.
[17] L. D. Landau and E. M. Lifschitz, *Lehrbuch der Theoretischen Physik, Band 8, Elektrodynamik der Kontinua*. Berlin: Akademie Verlag, 1980.
[18] O. Madelung, *Grundlagen der Halbleiterphysik*. Berlin-Heidelberg-New York: Springer-Verlag, 1970.
[19] A. Nakagawa and D. H. Navon, "A time-and temperature-dependent 2-D simulation of the GTO thyristor turn-off process," *IEEE Trans. Electron Devices*, vol. ED-31, pp. 1156–1163, Sept. 1984.
[20] J. E. Parrot, "Transport theory of semiconductor energy conversion," *J. Appl. Phys.*, vol. 53, pp. 9105–9111, Dec. 1982.
[21] A. Pierantoni, P. Ciampolini, and G. Baccarani, "Accurate modeling of electro-thermal effects in silicon devices," in *Proc. 22nd ESSDERC Conf.*, Leuven, Belgium, Sept. 1992, *Microelectron. Eng.* 19. Amsterdam: Elsevier, pp. 769–772.
[22] I. Prigogine, *Thermodynamics of Irreversible Processes*. New York: Interscience Publishers, 1961.
[23] H. A. Schafft, "Second breakdown—A comprehensive review," *Proc. IEEE*, Aug. 1967, vol. 55, pp. 1272–1288.
[24] S. Selberherr, *Analysis and Simulation of Semiconductor Devices*. New York: Springer, 1984.
[25] ——, "MOS device modeling at 77 K," *IEEE Trans. Electron Devices*, vol. 36, pp. 1464–1474, Aug. 1989.

[26] J. W. Slotboom and H. C. DeGraaff, "Bandgap narrowing in silicon bipolar transistors," *IEEE Trans. Electron Devices*, vol. ED-24, pp. 1123–1125, Aug. 1977.
[27] A. Sommerfeld, *Thermodynamik und Statistik*. Frankfurt: Verlag Harri Deutsch, 1977.
[28] S. M. Sze, *Physics of Semiconductor Devices*. New York: Wiley, 1981.
[29] J. Tauc, *Photo and Thermoelectric Effects in Semiconductors*. New York: Pergamon, 1962.
[30] K. M. Van Vliet, "Entropy production and choice of heat flux for degenerate semiconductors and metals," *Phys. Stat. Sol. (b)*, vol. 78, pp. 667–676, Feb. 1976.
[31] G. Wachutka, "Rigorous thermodynamic treatment of heat generation and conduction in semiconductor device modeling," *IEEE Trans. Computer-Aided Design*, vol. 9, pp. 1141–1149, Nov. 1990.
[32] P. B. M. Wolbert, G. Wachutka, B. H. Krabbenborg, and T. J. Mouthaan, "Nonisothermal device simulation using the 2-D numerical process/device simulator TRENDY and application to SOI-devices," *IEEE Trans. Computer-Aided Design*, vol. 13, pp. 293–302, Mar. 1994.



Hermann Brand was born in Mauthausen, Austria, in 1955. He received the degree of "Diplomingenieur" from the Technical University of Graz, Austria, in 1982, and the Ph.D. degree from the Technical University of Vienna, Austria, in 1994, both in electrical engineering.

He has held various positions in many industrial research and software development projects including computer aided design, operating systems, switching software, compilers for formal description languages, and protocol engineering. His research interests comprise computational engineering especially process-and device simulation, software engineering and telecommunications.

Siegfried Selberherr (M'70–SM'84–F'93), for a photograph and biography, see p. 1755 of the October 1995 issue of this TRANSACTIONS.