Irreversible Event-Based Model for Thermal Emission of Electrons from Isolated Traps

Sima Dimitrijev*

Griffith School of Engineering and Queensland Microtechnology Facility,

Griffith University, Nathan, Qld. 4111, Australia

Abstract

In spite of the irreversible nature of macroscopic processes, our understanding of the fundamental physical phenomena remains limited to reversible models (the Loschmidt's paradox). We propose a direct irreversible model for the probability per unit time that an electron will be emitted from an isolated trap. This resolves a number of problems, including (1) the dubious link between emission measurements and the parameters of the independent capture process and (2) the elusive meaning of the degeneracy factor in the equilibrium Fermi-Dirac distribution.

---

*Electronic address: s.dimitrijev@griffith.edu.au

I. INTRODUCTION

Thermal emission of electrons from isolated traps in a depleted semiconductor is an illustrative example of microscopic irreversible events. The existing model for the emission probability per unit time is derived from the condition of thermal equilibrium with the opposite process, which is electron capture by the isolated traps [1–3]. Although generally accepted, the use of the equilibrium condition introduces serious problems that can be related to the Loschmidt's paradox. Following a brief introduction to the equilibrium approach, these problems are discussed to highlight the need for exploration beyond the limits of the equilibrium approach. To resolve the highlighted problems, we propose a direct irreversible model for the probability per unit time that an electron will be emitted from an isolated trap.

II. PROBLEMS WITH THE EQUILIBRIUM APPROACH

The free electrons in semiconductors (the electrons in the conduction band) can be modeled as gas particles that move without any obstruction through a perfect crystal but interact with crystal imperfections. An important type of crystal imperfections are isolated crystal defects (including impurity atoms) that have energy levels in the energy gap, referred to as the isolated traps. The probability that a moving electron will hit a trap per unit length, with a chance of being captured, is proportional to the trap concentration (N_t) and the capture cross section of the trap (σ_n). Given that the length traveled per unit time is the thermal velocity (v_{th}), the probability that the electron will hit a trap per unit time is $v_{th}\sigma_n N_t$. Multiplying this probability by the probability that the trap is empty $(1 - f_t)$ gives the probability that a single electron is captured per unit time and multiplying it further by the concentration of electrons (n) gives the capture rate (the concentration of electrons captured per unit time) [1]:

$$r_c = v_{th}\sigma_n N_t (1 - f_t)n \qquad (1)$$

There is no such physical model for the opposite process, which is the electron emission. To obtain the rate of electron emission, the concentration of trapped electrons $(N_t f_t)$ is multiplied by unknown emission probability per unit time $(1/\tau_e)$:

$$r_e = \frac{N_t f_t}{\tau_e} \tag{2}$$

Because $1/\tau_e$ is not known for the emission process itself, it is modeled by an equation that is derived from the condition of emission and capture equilibrium, $r_e = r_c$. This condition leads to the following equation for the emission probability per unit time: $1/\tau_e = v_{th}\sigma_n n(1 - f_t)/f_t$. The probability that a trap is occupied in equilibrium is given by the Fermi-Dirac distribution,

$$f_t = \frac{1}{1 + (1/g)e^{(E_t - E_F)/kT}} \tag{3}$$

where E_t is the energy level of the occupying electron, E_F is the Fermi level, and g is a degeneracy factor. With the Fermi-Dirac distribution, the equation for the emission probability per unit time becomes: $1/\tau_e = (v_{th}\sigma_n n/g)\exp[(E_t - E_F)/kT]$. Using the equation for concentration of electrons in non-degenerate semiconductors, $n = N_C \exp[-(E_C - E_F)/kT]$, the following equation for $1/\tau_e$ is obtained [1, 3]:

$$\frac{1}{\tau_e} = \frac{v_{th}\sigma_n N_C}{g} e^{-(E_C - E_t)/kT} \tag{4}$$

In Eq. (4), N_C is the effective density of states in the conduction band and $E_C - E_t$ is the energy needed to emit a trapped electron into the conduction band. The value of the degeneracy factor is not known, so it is ignored by setting $g = 1$ in Eq. (4) [3]. Sah *et al* [4] use the concept of "effective" trap level, which incorporates the impact of the degeneracy factor, but still have to use an incorrect assumption of a "unity statistical weight factor" because there is neither theoretical nor experimental estimate of its value.

Two additional parameters of the pre-exponential term in Eq. (4) — v_{th} and σ_n — are problematic. Both v_{th} and σ_n are parameters of the capturing process that are involved into Eq. (4) through the equilibrium condition. A problem with this is that it is not logical the emission of a trapped electron to depend on the parameters of an independent capturing process. For example, the capture cross section for electrons is larger for positively-charged traps than for neutral traps because of the Coulomb attraction. The fact that a positively-charged trap can capture an electron easier does not mean that the electron can be released easier from that trap, as implied by Eq. (4). Likewise, the fact that it is harder to emit an electron from a deeper trap does not mean that it is harder for an electron to fall into a

deeper trap. In general, it is easier to fall into a wider trap and it is harder to get out of a deeper trap; however, it is not easier to get out of a wider trap and it is not harder to fall into a deeper trap.

This logical problem is not properly recognized and the existence of a direct link between the emission probability per unit time and the cross section is generally accepted. This leads to serious problems in explaining why the experimental data do not confirm the direct link between the emission probability per unit time and the capture cross section. Lang *et al* [5] presented a very illustrative set of experimental data. To their surprise, they found that measured emission probabilities, shown by the symbols in Fig. 1, did not differ for different types of defects with mid-gap levels (Co, Rh, Ag, and Au acceptors and process-induced donors). Importantly, the measured capture cross sections of these defects were significantly different, even by a factor as large as 30. This type of experimental data, obtained by different authors using different techniques, clearly shows that the link between the emission probability per unit time and the capture cross section, set by Eq. (4) with $g = 1$, is incorrect. An effort to correct this link employed thermodynamic analysis and introduced a degeneracy factor [3, 4, 10–12]. The involvement of thermodynamic concepts did not go beyond thermal equilibrium, so it did not remove the need to model the emission rate directly, which means without the link between the emission probability per unit time and the capture parameters.

Perhaps because of the obscure meaning and the unknown value of the degeneracy factor, the thermodynamic correction is largely ignored. Moreover, a significant number of researchers continue to use the relationship between the emission probability per unit time and the capture cross section as given by Eq. (4). In particular, it is very popular to measure the emission probability per unit time by different techniques of deep-level transient spectroscopy (DLTS) and to calculate the capture cross section using Eq. (4) with $g = 1$. The calculated values for the capture cross section are then used to make conclusions about the nature of the defects and to derive theories about the apparent temperature dependencies of the capture cross section. Frequently, unreasonably small (much smaller than the atom size) or unreasonably large capture cross sections are obtained. This is clearly causing a significant degree of confusion.
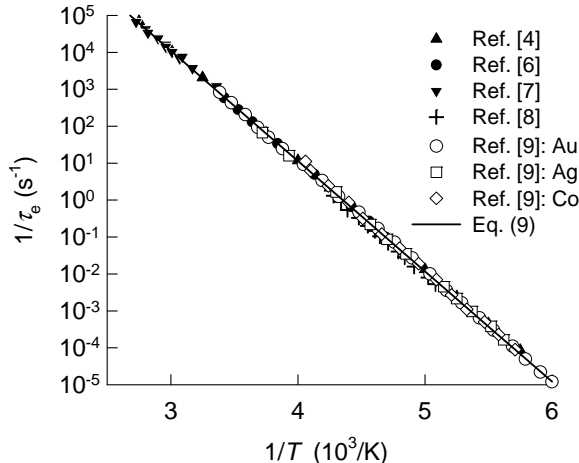
FIG. 1: Experimental data (the symbols), collected by Lang *et al* [5], for electron emission from approximately midgap traps with different optical and capture properties. The best fit by Eq. (9) (the line) is achieved with $E_C - E_t = 0.593$ eV and $\eta = 0.07$.

III.   EVENT-BASED MODEL

Irreversible processes can be modeled directly and independently of equilibrium conditions if we assume that the irreversible processes consist of irreversible events. This defines discrete intervals of time, τ_i $(i = 1, \ldots, M)$, corresponding to each of the M events that form a process. A process that results in an irreversible outcome becomes an event, which will be referred to as a composite event to distinguish it from the constituting events. If a process consists of consecutive events, then the time needed for the composite event is simply

$$\tau = \sum_{i=1}^{M} \tau_i \tag{5}$$

If a process consists of simultaneous and independent events, then the probabilities per unit time for each of these events are added to obtain the probability per unit time for the composite event:

$$\frac{1}{\tau} = \sum_{i=1}^{M} \frac{1}{\tau_i} \tag{6}$$

This approach is applicable to both macroscopic and microscopic events and processes. Furthermore, if the time intervals for the constituting events are not known, this approach can be applied hierarchically to simpler levels until events with known times are identified.

Consider the example of light-induced emission of electrons from isolated traps. In this case, photons provide the energy to a trapped electron so that it can be emitted into the conduction band. The energy of a photon (E_p) cannot be transferred to an electron within a shorter time interval than the wave period (h/E_p). This means that h/E_p is the time interval of an elementary event that involves energy change of E_p. This quantum effect is the reason for considering photons as elementary particles, at least in terms of their interactions with other particles (including electrons). The use of the term "elementary event" instead of "elementary particle" does not change the underlying quantum effect, it just links it to time as the variable of interest.

In the case of thermal emission of electrons from isolated traps, phonons provide the thermal energy to a trapped electron to emit it into the conduction band. Analogously to the case of photons, the energy of a phonon cannot be transferred to an electron within a shorter time interval than the wave period (h/E_p). Differently from the case of photons, however, the thermal emission can occur even when the energy of a single phonon is smaller than the energy needed for the trapped electron to be emitted into the conduction band. The relationship between the energy taken by the electron during the emission process (E) and the energy supplied by a single particle (E_p) can be expressed as $E_p = \eta_1 E$, where η_1 is a parameter satisfying the following condition: $\eta_1 \leq 1$. The condition $\eta_1 = 1$ corresponds to the case when the energy E is supplied by a single particle ($E = E_p$), such as the case of photon-induced emission. As distinct from this, the condition $\eta_1 < 1$ corresponds to the case when the energy E is supplied by $1/\eta_1$ particles, the energy transferred by each particle being E_p: $E = (1/\eta_1)E_p$. Therefore, the time interval of an elementary event involving energy change of E can be expressed by $h/E_p = h/(\eta_1 E)$.

The reciprocal value of the time limit for an emission event, $\eta_1 E/h$, corresponds to the probability per unit time that a trapped electron will be emitted under the condition that the energy E is transferred to the electron. The condition that the energy E is transferred to the electron can be modeled by two parts: (1) the probability that the trap is in the energy state E, labeled by $p(E)dE$, and (2) the probability that the trap will transfer the energy to the electron, labeled by η_2. Accordingly, the probability per unit time for an emission event involving energy E is $E\eta p(E)dE/h$, where $\eta = \eta_1 \eta_2$.

The probability $p(E)dE$ that a system will be in a state with energy E is established by the methods of statistical physics:

$$p(E)dE = \frac{e^{-E/kT}}{Z}dE \tag{7}$$

where the constant Z is determined from the normalization condition. In this case, the energy quantization can be neglected so that the integral form of the normalization condition can be used: $\int_0^\infty p(E)dE = 1$. This condition is satisfied by $Z = kT$ in Eq. (7). Consequently, the probability that a considered trap will be in a state with energy E is $p(E)dE = (1/kT)\exp(-E/kT)dE$.

With these considerations, the probability per unit time that an electron will be excited from E_t to an energy level $E + E_t$ in the conduction band is $(\eta/h)(E/kT)\exp(-E/kT)dE$. An emission event occurs when the electron is excited to any of the energy levels in the conduction band. This means that the minimum-energy emission corresponds to $E = E_C - E_t$ and that higher-energy emissions also occur. With the assumption of continuous energy in the conduction band, the sum in Eq. (6) is replaced by the following integral:

$$\frac{1}{\tau_e} = \int_{E_C - E_t}^\infty \frac{\eta}{h}\frac{E}{kT}e^{-E/kT}dE \tag{8}$$

The solution of this integral leads to the following equation for the emission probability per unit time:

$$\frac{1}{\tau_e} = \frac{\eta(E_C - E_t + kT)}{h}e^{-(E_C - E_t)/kT} \tag{9}$$

Equation (9) fits the experimental data shown in Fig. 1 (the solid line) with the following set of parameters: $E_C - E_t = 0.593$ eV and $\eta = 0.07$. In silicon, the maximum energy of phonons is $E_{p-max} = 0.066$ eV (this energy corresponds to the minimum phonon wavelength set by the crystal-lattice constant). This means that $E_{p-max}/(E_C - E_t) = 0.11$ sets the maximum value of η_1 to 0.11 and the minimum value of $\eta_2 = \eta/\eta_1$ to 0.64. However, if we assume that $\eta_2 = 1$, the average energy of phonons contributing to the thermal emission is estimated as $E_p = \eta(E_C - E_t) = 0.041$ eV. These numbers indicate that it is quite possible that the value of the parameter η is set by the average energy of the phonons responsible for the thermal emission process.

IV. IMPLICATIONS FOR FERMI-DIRAC DISTRIBUTION

The establishment of the equilibrium-independent equation for the emission probability per unit time — Eq. (9) — opens the question of its implication on the equilibrium condition $r_c = r_e$. Based on Eqs. (1), (2), and (9), the equilibrium condition $r_c = r_e$ becomes

$$v_{th}\sigma_n N_t (1 - f_t)n = N_t f_t \frac{\eta(E_C - E_t + kT)}{h} e^{-(E_C - E_t)/kT} \tag{10}$$

With $n = N_C \exp[-(E_C - E_F)/kT]$, Eq. (10) is transformed into

$$v_{th}\sigma_n N_C (1 - f_t) = f_t \frac{\eta(E_C - E_t + kT)}{h} e^{(E_t - E_F)/kT} \tag{11}$$

The dependent variable in this equation is the trap-occupancy probability f_t. From Eq. (11), the following equation is obtained for the trap-occupancy probability:

$$f_t = \frac{1}{1 + [\eta(E_C - E_t + kT)/(hv_{th}\sigma_n N_C)]e^{(E_t - E_F)/kT}} \tag{12}$$

A comparison with Eq. (3) shows that this is the Fermi-Dirac distribution with the degeneracy factor $g = hv_{th}\sigma_n N_C/\eta(E_C - E_t + kT)$. According to Eq. (12), the trap-occupancy probability depends on both capture and emission parameters and, for example, is higher for larger capture cross sections. This result is logical because a larger capture cross section and an increased capture rate should shift the equilibrium balance toward a higher population of the traps. Therefore, the mystery associated with the degeneracy factor is resolved by the new equation for the emission probability per unit time.

V. MODELING GENERATION AS A NONEQUILIBRIUM PROCESS

The direct modeling of nonequilibrium emission events by Eq. (9) enables direct modeling of nonequilibrium processes. Modeling the generation rate in a fully depleted semiconductor, such as a reverse-biased P–N junction, is a good example. A generation event can be considered as a composite event that consists of two consecutive emission events: an electron emission and a hole emission in either order [14]. In full analogy with macroscopic events, the times taken by each consecutive event are added to obtain the time for the composite event [Eq. (5)]. Labeling the times for the electron and hole emissions by τ_e and τ_h, respectively, the time for the composite generation event is

$$\tau_g = \tau_e + \tau_h \tag{13}$$

Given that $1/\tau_e$ and $1/\tau_h$ are probabilities per unit time that electron and hole emission events will occur, the times τ_e and τ_h are the expected times for electron and hole emissions, as distinct from precisely determined time intervals. Because consecutive events of electron and hole emissions by a single trap can be continuously repeated, the time τ_g can be considered as the average time that it takes a single trap to generate an electron-hole pair [14]. Note that this concept has a clear physical meaning as distinct from the commonly used generation lifetime [3], which can be orders of magnitude shorter from the time needed to generate a single electron-hole pair [14].

To model M simultaneous and independent events, the probabilities per unit time for each of these events are added to obtain the probability per unit time for the composite event [Eq. (6)]. Therefore, if M identical traps are generating electron-hole pairs, the probability that an electron-hole pair is generated per unit time is

$$\frac{1}{\tau_{gM}} = \sum_{i=1}^{M} \frac{1}{\tau_g} = \frac{M}{\tau_g} \tag{14}$$

To convert the number of traps (M) into trap concentration (N_t), the number of traps is divided by the volume they occupy: $N_t = M/V$. Dividing Eq. (14) by the volume V, the number of generation events per unit time and unit volume, which is the generation rate, is obtained

$$G = \frac{N_t}{\tau_g} = \frac{N_t}{\tau_e + \tau_h} = \frac{N_t}{A_e e^{(E_C - E_t)/kT} + A_h e^{(E_t - E_V)/kT}} \tag{15}$$

In Eq. (15), $A_e = h/\eta(E_C - E_t)$, $A_h = h/\eta_h(E_t - E_V)$, E_V is top of the valence band, and η_h is analogous parameter to the parameter η . The only difference between Eq. (15) and the widely used Shockley-Read-Hall equation [1, 2], derived from the equilibrium model, is in the parameters A_e and A_h. This difference does not have experimental implications because both equations contain unknown parameters: η and η_h in this case and the capture cross sections in the case of Shockley-Read-Hall equation. In other words, the widely established experimental evidence for the Shockley-Read-Hall equation is just as valid for Eq. (15).

VI.   CONCLUSIONS AND WIDER IMPLICATIONS

We have demonstrated that the thermal emission of electrons from isolated traps in a depleted semiconductor can be modeled by a direct irreversible equation. The proposed direct-modeling approach utilizes a condition for the minimum time of an emission event, which is set by the period of the phonons supplying the energy needed for the emission. This enables to remove the parameters of the independent capturing events from the pre-exponential term of the Arrhenius-type equation for the emission probability per unit time.

The proposed direct modeling of irreversible events and processes is fundamentally different from the existing practice of modeling fundamental processes by continuous-time equations. Events are inherently irreversible but consecutive irreversible events can establish the state of macroscopic equilibrium. This is demonstrated by the example of balanced emission and capture rates, which shows the need for the degeneracy factor in the Fermi-Dirac distribution. As opposed to this, macroscopic irreversibility could not be explained by the concept of reversible continuous-time processes (the Loschmidt's paradox). To model composite events and processes, the irreversible models of the constituting events are combined in a way that is common sense for macroscopic events: Eq. (5) is the generic model for a composite event consisting of two consecutive events, whereas Eq. (6) is the generic model for a composite event consisting of M simultaneous events. This demonstrates that the concept of events and the associated models, including the hierarchical relationship between the constituting and composite events, have the same meaning and analogous applicability from the fundamental microscopic level to complex macroscopic levels.

---

[1] W. Shockley and W.T. Read, Phys. Rev. **87**, 835 (1952).

[2] R.N Hall, Phys. Rev. **87**, 387 (1952).

[3] D.K. Schroder, *Semiconductor Material and Device Characterization*, 3rd ed. (Wiley, New York, 2006).

[4] C.T. Sah *et al.*, Appl. Phys. Lett. **15**, 145 (1969).

[5] D.V. Lang, H.G. Grimmeiss, E. Meijer, and M. Jaros, Phys. Rev. B **22**, 3917 (1980).

[6] O. Engström and H.G. Grimmeiss, J. Appl. Phys. **46**, 831 (1975).

[7] L.C. Parrillo and W.C. Johnson, Appl. Phys. Lett. **20**, 104 (1972).

[8] S.D. Brotherton and J. Bicknell, J. Appl. Phys. **49**, 667 (1978).

[9] L.D. Yau and C.T. Sah, Appl. Phys. Lett. **21**, 157 (1972).

[10] C.D. Thurmond, J. Electrochem. Soc. **122**, 1133 (1975).

[11] J.A. Van Vechten and C.D. Thurmond, Phys. Rev. B **14**, 3539 (1976)

[12] O. Engström and A. Alm, Solid-St. Electron. **21**, 1571 (1978).

[13] P.A.M. Dirac, *The Principles of Quantum Mechanics*, 4th ed. (Oxford University Press, Oxford, 1958).

[14] S. Dimitrijev, *Principles of Semiconductor Devices* (Oxford University Press, New York, 2006).