7 Appendix: Derivation of some important model equations

The derivation of the *basic* formulation of the generalized ICCR (GICCR) [36] requires the following assumptions:

- A one-dimensional transistor structure is considered, with the emitter contact at the mono-silicon surface and the collector contact at the transition from the “lightly”-doped collector region to the buried layer. The vertical ordinate is designated with x and is omitted in the following derivation from the indices for reasons of better legibility and understanding.
- The volume recombination in above region is negligible.
- The time derivative is zero; in general, it is sufficient though to assume quasi-static operation.
- Effects such as thermionic emission and tunneling across the junctions are neglected; they need to be accounted for by separate terms and can - in principle - be combined with the GICCR solution.

It can be shown that a GICCR can also be derived for the a three-dimensional case, so that the assumption of a one-dimensional region can be dropped for a actual applications.

First, consider the electron continuity equation which, together with the second and third assumption, reads

$$\frac{dJ_n}{dx} = q\left(R + \frac{\partial n}{\partial t}\right) = 0 \ . \tag{7.0.0-1}$$

This corresponds to a spatially independent electron current density,

$$J_n = const(x) = -J_T = \frac{-I_T}{A_E} \ , \tag{7.0.0-2}$$

which can be expressed by the transfer current and area of the (1D) transistor. Note, that under DC operation, $I_T = I_C$.

The derivation starts with the electron transport equation:

$$J_n = -q\mu_n n \frac{d\varphi_n}{dx} \tag{7.0.0-3}$$

Inserting the electron density,

$$n = n_i \exp\left(\frac{\psi - \varphi_n}{V_T}\right) \qquad (7.0.0\text{-}4)$$

gives

$$J_n = -q\mu_n n_i \exp\left(\frac{\psi - \varphi_n}{V_T}\right)\frac{d\varphi_n}{dx} = -q\mu_n n_i \exp\left(\frac{\psi}{V_T}\right)\exp\left(-\frac{\varphi_n}{V_T}\right)\frac{d\varphi_n}{dx} . \qquad (7.0.0\text{-}5)$$

Note, that n_i is the effective intrinsic carrier density; it contains for the various transistor regions possible bandgap differences, that are caused by high-doping effects and bandgap-engineering. This topic will be discussed later.

Using

$$\frac{d\left(\exp\left(-\frac{\varphi_n}{V_T}\right)\right)}{dx} = -\frac{1}{V_T}\exp\left(-\frac{\varphi_n}{V_T}\right)\frac{d\varphi_n}{dx} \qquad (7.0.0\text{-}6)$$

leads to

$$J_n = qV_T\mu_n n_i \exp\left(\frac{\psi}{V_T}\right)\frac{d\left(\exp\left(-\frac{\varphi_n}{V_T}\right)\right)}{dx} . \qquad (7.0.0\text{-}7)$$

Separation of the differential terms gives

$$\frac{J_n}{qV_T\mu_n n_i}\exp\left(-\frac{\psi}{V_T}\right)dx = d\left(\exp\left(-\frac{\varphi_n}{V_T}\right)\right) . \qquad (7.0.0\text{-}8)$$

The inconvenient term $\exp(-\psi/V_T)$ can be replaced by more useful variables such as the hole density and an $\exp(\varphi_p/V_T)$ term via the transformation,

$$\exp\left(-\frac{\psi}{V_T}\right) = \frac{n_i \exp\left(\frac{\varphi_p - \psi}{V_T}\right)}{n_i \exp\left(\frac{\varphi_p}{V_T}\right)} = \frac{p}{n_i \exp\left(\frac{\varphi_p}{V_T}\right)} , \qquad (7.0.0\text{-}9)$$

yielding

$$\frac{J_n}{qV_T}\frac{p}{\mu_n n_i^2 \exp\left(\frac{\varphi_p}{V_T}\right)}dx = d\left(\exp\left(-\frac{\varphi_n}{V_T}\right)\right) \tag{7.0.0-10}$$

Integration of the above equation over the interval $[x_l, x_u]$ gives:

$$\int_{x_l}^{x_u}\frac{J_n}{qV_T}\frac{p}{\mu_n n_i^2 \exp\left(\frac{\varphi_p}{V_T}\right)}dx = \int_{\exp(\varphi_n(x_l)/V_T)}^{\exp(\varphi_n(x_u)/V_T)} d\left(\exp\left(-\frac{\varphi_n}{V_T}\right)\right) \tag{7.0.0-11}$$

The result for the right-hand-side is

$$\int_{\exp(\varphi_n(x_l)/V_T)}^{\exp(\varphi_n(x_u)/V_T)} d\left(\exp\left(-\frac{\varphi_n}{V_T}\right)\right) = -\left[\exp\left(-\frac{\varphi_n(x_u)}{V_T}\right) - \exp\left(-\frac{\varphi_n(x_l)}{V_T}\right)\right] . \tag{7.0.0-12}$$

The exact value of the electron quasi-fermi potentials depends on the choice of the integration limits and will be discussed later. First, the left-hand-side of (7.0.0-12) is integrated, giving after extension with the exp of the controlling (internal) base-emitter voltage, $\exp(-V_{BEi}/V_T)$, and with the transfer current density as defined in (7.0.0-2) the general expression

$$\int_{x_l}^{x_u}\frac{J_n}{qV_T}\frac{p}{\mu_n n_i^2 \exp\left(\frac{\varphi_p}{V_T}\right)}dx = \frac{J_T \exp\left(-\frac{V_{BEi}}{V_T}\right)}{qV_T}\int_{x_l}^{x_u}\frac{p}{\mu_n n_i^2}\frac{J_n}{J_T}\exp\left(\frac{V_{BEi}-\varphi_p}{V_T}\right)dx , \tag{7.0.0-13}$$

that contains the desired terminal variables of the 1D transistor. The reason for not using (7.0.0-2) and $\varphi_p = V_{BEi}$ directly is the yet undefined integration interval. Equating again (7.0.0-13) and (7.0.0-12), and keeping only the desired variable J_T on the l.h.s., results in the general basic formulation

$$J_T = qV_T \frac{\exp\left(\frac{V_{BEi}}{V_T}\right)\left[\exp\left(-\frac{\varphi_n(x_l)}{V_T}\right) - \exp\left(-\frac{\varphi_n(x_u)}{V_T}\right)\right]}{\int_{x_l}^{x_u} \frac{p}{\mu_n n_i^2} \frac{J_n}{J_T} \exp\left(\frac{V_{BEi} - \varphi_p}{V_T}\right) dx} , \tag{7.0.0-14}$$

from which different forms of the ICCR can be derived. At this point, the various terms and variables as well as the choice of the integration limits need to be discussed in order to enable the derivation of a practically more applicable relation than the above one.

The electron fermi-potentials in the numerator assume their known 1D terminal values only at the contacts, i.e. if the entire 1D transistor region $[x_E, x_C]$ is chosen as integration interval,

$$\varphi_n(x_l) = \varphi_n(x_E) = 0 \quad \text{and} \quad \varphi_n(x_u) = \varphi_n(x_C) = V_{CEi} , \tag{7.0.0-15}$$

making this an attractive choice, the effect of which on the denominator is discussed next.

Defining the spatially dependent weighting function in normalized form,

$$h(x) = \underbrace{\frac{\mu_{n0r} n_{ir}^2}{\mu_n(x) n_i^2(x)}}_{h_g} \underbrace{\frac{J_n(x)}{J_T}}_{h_i} \underbrace{\exp\left(\frac{V_{BEi} - \varphi_p(x)}{V_T}\right)}_{h_v} = h_g\ h_i\ h_v g \tag{7.0.0-16}$$

with μ_{n0r} and n_{ir} as mobility and intrinsic carrier density, respectively, of a reference material (e.g. the base or a point in the base region), permits to write the denominator as

$$\int_{x_E}^{x_C} \frac{p}{\mu_n n_i^2} \frac{J_n}{J_T} \exp\left(\frac{V_{BEi} - \varphi_p}{V_T}\right) dx = \frac{1}{\mu_{n0r} n_{ir}^2} \int_{x_E}^{x_C} h(x) p(x) dx . \tag{7.0.0-17}$$

For any bias point, the integral on the r.h.s., multiplied by q, can *always* be written as

$$q \int_{x_E}^{x_C} h(x) p(x) dx = \bar{h} \bar{Q}_p , \tag{7.0.0-18}$$

with $\bar{h}$ as an average value of the weighting function and

$$\bar{Q}_p = q\int_{x_l}^{x_u} p(x)dx \qquad (7.0.0\text{-}19)$$

is the 1D (i.e. per area) hole charge stored in the selected integration interval. The latter can be measured via the terminals if the integration interval is chosen properly.

A brief discussion of the weighting function follows in order to provide a better idea of the assumptions made for the ICCR and GICCR.

The weighting function h_v

The hole quasi-fermi potential is equal to V_{BEi} at least over the base and its adjacent space charge regions as well as far into the emitter and collector region, resulting in $h_v = 1$ within the respective integration interval. This is also true in the very thin emitter region of advanced transistors. In conventional processes with larger emitter junction depth though, a significant portion of the holes injected back into the neutral emitter recombine there, leading to a slight decrease of φ_p. However, the resulting deviation of h_v from 1 within this region coincides with a rapid decrease of holes, so that the contribution of the product $h*p$ to the integral is negligible. Therefore, $h_v = 1$ is a very good assumption under all relevant bias conditions.

The weighting function h_i

The transfer current density (at the 1D collector contact) equals the electron current density in the base and collector region. Only in the BE space-charge region and the neutral emitter region, J_n can increase slightly as a result of the back injection of holes and the corresponding recombination. The maximum increase occurs at the emitter contact x_E:

$$J_n(x_E) = J_T + J_p(x_E)\,. \qquad (7.0.0\text{-}20)$$

The maximum possible increase occurs for opaque emitters, where $J_p(x_E) = J_B$, and is given by

$$max\{h_i\} = h_i(x_E) = 1 + \frac{J_B}{J_T} = 1 + \frac{1}{B} \qquad (7.0.0\text{-}21)$$

with B as the DC current gain. As a consequence, for the vast majority and, particularly advanced bipolar transistors, $h_i = 1$ is a good assumption under all relevant bias conditions.

The weighting function h_g

High-doping effects as well as intentional bandgap grading in heterojunction transistors can cause n_i^2 values to differ by orders of magnitude in the various transistor regions. In addition, also the mobility varies significantly within the transistor as a function of both doping and bias (via the electric field). The variation caused by the latter is most pronounced in the BC junction and collector region. In general though, μ_n and n_i possess an opposite dependence on doping, leading to a partial compensation within h_g. However, the influence of n_i still remains much stronger than that of μ_n. As a consequence, the weighting function h_g deviates strongly from 1 and has to be considered for all processes. Note, that the index "g" has been chosen to indicate the major factor of influence, namely the bandgap in the effective intrinsic carrier concentration. Nevertheless, a spatially dependent function h_g can still be accommodated in a compact model formulation as shown below.

Overall, for properly designed transistors the weighting function h consists only of the contribution of h_g, so that the average value in (7.0.0-18) is given by

$$\bar{h} = \bar{h}_g = \frac{\mu_{n0r} n_{ir}^2}{\overline{\mu_n n_i^2}} \ . \tag{7.0.0-22}$$

Note that the bias dependence of $\bar{h}$ can also be caused by a bias dependent hole distribution within the transistor, even if $h(x)$ does not depend on bias; e.g., at high current densities, $p(x)$ can spread into the collector in which n_i^2 might be few orders of magnitude smaller than in the base, so that the collector portion of the hole charge has to be weighted much stronger than the base portion and can, therefore, still cause a significant contribution to the overall integral, although the charge in the collector might still be very small [36]!

Since the major contribution to the integral comes from the base region, where the hole density is largest, the only other relevant contributions will come from those regions, where the hole den-

sity has not dropped to insignificant values. In other words, the regions towards the contacts do not contribute to the integral and are irrelevant for the discussion on how to chose the integration interval for the denominator of (7.0.0-14). Of the infinite number of possible integration intervals, the two most important choices are:

- If the base region is chosen as integration interval, $\bar{h}$ can be shown to be only very weakly bias dependent for homojunction transistors. The main disadvantages, however, are (i) the unknown and strongly bias dependent values for the electron quasi-fermi potentials at the base region boundaries, which are not easily accessible by measurements, and (ii) the difficulty to measure the base hole charge separately (and to extract the corresponding model parameters). In addition, the bias dependence of $\bar{h}$ will not be as weakly dependent anymore for drift HBTs.
- If the entire 1D transistor region is chosen as integration interval, $\bar{h}$ is expected to exhibit a larger bias dependence than for homojunction transistors. It has been shown though for bipolar technologies over (at least) the past 25 years that the bias dependence of $\bar{h}$ is negligible in the bias region of practical interest. As an advantage of choosing the 1D contacts for boundary conditions, the electron quasi-fermi potentials are well-known and given by (7.0.0-15).

For the reasons given above, the entire (1D) transistor region has been chosen as integration interval. The resulting transfer current expression then reads

$$J_T = qV_T \frac{\exp\left(\frac{V_{BEi}}{V_T}\right) - \exp\left(\frac{V_{BCi}}{V_T}\right)}{\int_{x_E}^{x_C} \frac{p}{\mu_n n_i^2} dx} , \tag{7.0.0-23}$$

or after introducing the weighting function h_g according to (7.0.0-16)

$$J_T = q^2 V_T \mu_{n0r} n_{ir}^2 \frac{\exp\left(\frac{V_{BEi}}{V_T}\right) - \exp\left(\frac{V_{BCi}}{V_T}\right)}{q\int_{x_E}^{x_C} \frac{\mu_{n0r} n_{ir}^2}{\mu_n n_i^2} p dx} . \tag{7.0.0-24}$$

7.1 The ICCR (homojunction transistors)

Inserting (7.0.0-15), (7.0.0-18) with $\bar{h} = \bar{h}_g$ as well as $V_{BCi} = V_{BEi}$-V_{CEi} into (7.0.0-14) yields

$$J_T = q^2 V_T \overline{\mu_n n_i^2} \frac{\exp\left(\frac{V_{BEi}}{V_T}\right) - \exp\left(\frac{V_{BCi}}{V_T}\right)}{\bar{Q}_p} . \tag{7.1.0-25}$$

This equation, with

$$c_{10} = q^2 V_T \overline{\mu_n n_i^2} \tag{7.1.0-26}$$

is the well-known ICCR, which is very accurate for most homojunction transistors, provided that the hole charge (density),

$$\bar{Q}_p = \bar{Q}_{p0} + \bar{Q}_{jEi} + \bar{Q}_{jCi} + \bar{Q}_f + \bar{Q}_r \tag{7.1.0-27}$$

is accurately modeled as a function of bias. The latter is a prerequisite for the description of high-speed applications in any way. Also, since the hole charge has to be continuously differentiable with respect to bias, the transfer current is also automatically continuously differentiable over all bias regions and is modeled via a single-piece formulation.

Fig. illustrates the impact of the various hole charges on the transfer current during a sweep of the BE voltage. At low current densities, the depletion charges Q_{jEi} and Q_{jCi} are the only bias dependet contributors to Q_p. For constant V_{BC}, the bias dependence of the BE depletion charge determines the ideality of the transfer current charectristic; for constant V_{BC}, the bias dependence of the BC depletion charge determines the slope of the output characteristics and, thus, automatically takes into account the bias dependent Early voltage. At high current densities, the minority charge Q_f starts to increase rapidly exceeding the depletion charges around peak f_T and then dominating the bias dependence of Q_p beyond peak f_T. As a consequence, the ideality factor of the transfer current characteristic also increases rapidly leading to a string decrease in the transcondutance. A detailed discussion of the ICCR for BJT processes can be found in, e.g., [32].

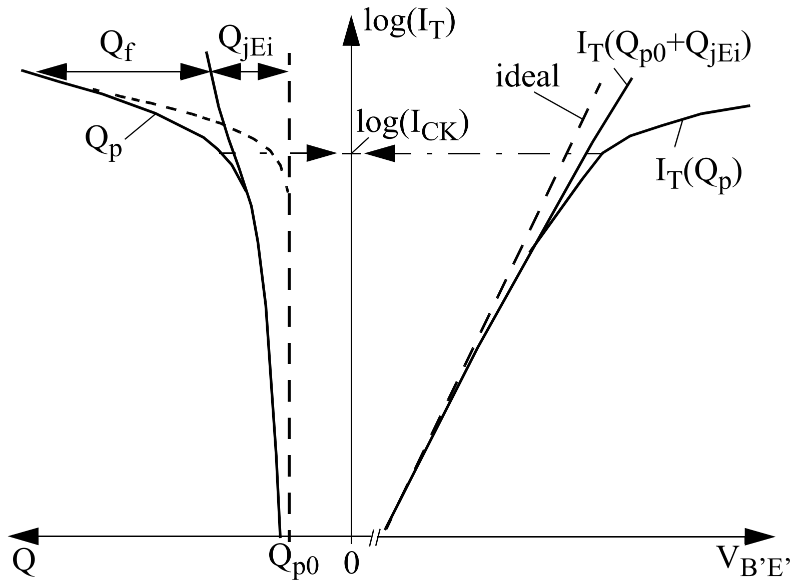

Fig. 7.1.0/1: Illustration of the impact of the various charge components on the transfer current, calculated by the ICCR.

7.2 The GICCR (heterojunction transistors)

In homojunction transistors bandgap differences are caused by high-doping effects only, so that the variation in n_i^2 over the integration interval $[x_E, x_C]$ is moderate. In contrast, large bandgap differences are intentional in HBTs, leading to n_i^2 variations of orders of magnitude. As demonstrated in [36], the transfer current of (SiGe) HBTs cannot be described anymore with the ICCR. The latter can be extended though, resulting in the generalized ICCR (GICCR) which maintains the advantages of a single-piece continuously differentiable description of the main transistor current. While the integration interval and, thus, the numerator of (7.0.0-14) remain the same, the denominator of (7.0.0-14) has to be considered more carefully for HBTs.

The starting point for the extension is eq. (7.0.0-17), with $h(x) \approx h_g(x)$ for the same reasons as discussed before, resulting in

$$\int_{x_E}^{x_C} \frac{p}{\mu_n n_i^2} dx = \frac{1}{\mu_{n0r} n_{ir}^2} \int_{x_E}^{x_C} \frac{\mu_{n0r} n_{ir}^2}{\mu_n n_i^2} p(x) dx \ . \tag{7.2.0-28}$$

The choice of the reference material (or reference transistor region) and its associated parameter values for $\mu_{n0r}n_{ir}^2$ is arbitrary and will be discussed later. Due to the large variation of n_i^2 between emitter, base and collector region, the definition of a sufficiently bias independent average value $\overline{h_g}$ is not always possible and might lead to unacceptable errors in modeling the transfer current high current densities. The goal is, therefore, to formulate the integral in terms of single charges in particular transistor regions, multiplied with proper average values of the respective weighting functions, that are both suitable for compact modeling.

In HBTs, the bandgap is significantly different in the emitter, base and collector region, while the transition from one bandgap to another usually takes place around the junction. As a consequence, partitioning of the integral into the neutral emitter, base and collector region as well as into depletion regions looks like a reasonable choice in order to obtain separate integrals in which the weighting functions are sufficiently independent on location and bias. Defining the respective average values,

$$\bar{h}_k = \bar{h}_{gk} = \frac{\int h_{gk} p(x) dx}{\int p(x) dx} = \frac{\mu_{n0r} n_{ir}^2}{\overline{\mu_{nk} n_{ik}^2}} \quad \text{with } k = (e, jei, b, jci, c) , \tag{7.2.0-29}$$

the denominator integral of the ICCR reads after multiplying with q

$$\bar{Q}_{p,T} = q \int_{x_E}^{x_C} \frac{\mu_{n0r} n_{ir}^2}{\mu_n n_i^2} p(x) dx = \bar{Q}_{p0} + h_{jEi}\bar{Q}_{jEi} + h_{jCi}\bar{Q}_{jCi} + \bar{Q}_{f,T} + \bar{Q}_{r,T} . \tag{7.2.0-30}$$

where the base region has been chosen as reference, i.e. $\mu_{n0r}n_{ir}^2 = \overline{\mu_{nb}n_{ib}^2}$. According to [36], assuming bias independent average values for the separate weighting functions leads to a significant improvement over the conventional ICCR with a sufficiently accurate description of the transfer current characteristics and the respective derivatives over the entire bias range of interest (up to very high current densities). The average weighting factors are new model parameters.

Application of the GICCR also requires the modeling of the separate portions, $\bar{Q}_{pE}$ and $\bar{Q}_{pC}$, of the total hole charge. The GICCR has been evaluated for many advanced SiGe bipolar technolo-

gies. It has been found that a constant factor $\bar{h}$ can in fact be used for more than one bias point up to quite high current densities, thus maintaining all the advantages of the ICCR with the added benefit of being able to model HBTs in physics-based way.

For a 2D/3D GICCR, see [64].