# Drift-Diffusion Model: Solution Details 

Prepared by:

## Dragica Vasileska

Associate Professor
Arizona State University

## Gummel's Iteration Method

Gummel's method solves the coupled set of semiconductor equations together with the Poisson equation via a decoupled procedure. If we choose the quasi-Fermi level formulation, we solve first a nonlinear Poisson's equation. The potential obtained from this solution is substituted into the continuity equations, which are now linear, and are solved directly to conclude the iteration step. The result in terms of quasi-Fermi levels is then substituted back into Poisson's equation and the process repeated until convergence is reached. In order to check for convergence, one can calculate the residuals obtained by positioning all the terms to the left hand side of the equations and substituting the variables with the iteration values. For the exact solution the residuals should be zero. Convergence is assumed when the residuals are smaller than a set tolerance. The rate of convergence of the Gummel method is faster when there is little coupling between the different equations. The computational cost of one Gummel iteration is one matrix solution for each carrier type plus one iterative solution for the linearization of Poisson's equation. Note that in conditions of equilibrium (zero bias) only the solution of Poisson's equation is necessary, since the equilibrium Fermi level is constant and coincides with both quasi-Fermi levels.

We give some examples of the quasi-linearization of Poisson equation, as necessary when Gummel's method is implemented. Let us consider the 1-D case in equilibrium first. As mentioned earlier, one has to solve only Poisson's equation, since the current is zero and the exact expressions for the carrier concentrations are known. In the non-degenerate case, the explicit expressions for the electron and hole densities are substituted into Poisson's equation to give
$\frac{d^{2} V}{d x^{2}}=\frac{q}{\varepsilon}\left[n_{i} \exp \left(-q \phi_{n}\right) \exp \left(\frac{q V}{k_{B} T}\right)-n_{i} \exp \left(q \phi_{p}\right) \exp \left(-\frac{q V}{k_{B} T}\right)+N_{A}-N_{D}\right]$
which is sometimes referred to as the non-linear Poisson equation due to the nonlinear terms involving $V$ on the RHS. In equilibrium, the quasi-Fermi energies are equal and spatially invariant, hence we may choose the Fermi energy as the reference energy, i.e. $\phi_{n}=\phi_{p}=0$.

Furthermore, the equation may be scaled by using the (minimum) extrinsic Debye length for the space coordinate $x$, and the thermal voltage $k_{B} T / q$ for the potential $V$. Writing $\bar{V}$ and $\bar{x}$ for the normalized potential and space coordinates, we obtain

$$
\begin{equation*}
\frac{d^{2} \bar{V}}{d \bar{x}^{2}}=\frac{n_{i}}{N}\left[\exp (\bar{V})-\exp (-\bar{V})+\frac{N_{A}-N_{D}}{n_{i}}\right] \tag{2}
\end{equation*}
$$

The equilibrium non-linear Poisson equation can be solved with the following quasi-linearization procedure

1. Choose an initial guess for the potential $\bar{V}$.
2. Write the potential at the next iteration step as $\bar{V}_{\text {new }}=\bar{V}+\delta V$, and substitute into Eq. (2) to solve for $\bar{V}_{\text {new }}$ to give

$$
\begin{equation*}
\frac{d^{2} \bar{V}}{d \bar{x}^{2}}+\frac{d^{2} \delta V}{d \bar{x}^{2}}=\frac{n_{i}}{N}\left[\exp (\bar{V}) \exp (\delta V)-\exp (-\bar{V}) \exp (-\delta V)+\frac{N_{A}-N_{D}}{n_{i}}\right] \tag{3}
\end{equation*}
$$

3. Use the linearization $\exp ( \pm \delta V) \approx 1+\delta V$ and discretize the resultant equation. This equation has a tridiagonal matrix form and is readily solved for $\delta V(i)$.

$$
\begin{align*}
& \delta V(i-1)-\left[2+\frac{n_{i}}{N} \Delta^{2} x[\exp (\bar{V}(i))-\exp (-\bar{V}(i))]\right] \delta V(i)+\delta V(i+1)= \\
& -\bar{V}(i-1)+2 \bar{V}(i)-\bar{V}(i+1)+\frac{n_{i}}{N} \Delta^{2} x\left[\exp (\bar{V}(i))-\exp (-\bar{V}(i))+\frac{N_{A}-N_{D}}{n_{i}}\right] \tag{4}
\end{align*}
$$

4. Check for convergence. The residual of Eq. (4) is calculated and convergence is achieved if the norm of the residual is smaller than a preset tolerance. If convergence is not achieved, return to step 2. In practice one might simply check the norm of the error

$$
\begin{equation*}
\|\delta V\|_{2} \leq T o l \text { or }\|\delta V\|_{\infty} \leq T o l \tag{5}
\end{equation*}
$$

Note that for the solution of the nonlinear Poisson's equation, the boundary conditions are referenced to the equilibrium Fermi level. One may use the separation between the Fermi level and the intrinsic Fermi level at the contacts for the boundary conditions.

After the solution in equilibrium is obtained, the applied voltage is increased gradually in steps $\Delta V \leq k_{B} T / q$ to avoid numerical instability. The scaled nonlinear Poisson equation under nonequilibrium conditions now becomes

$$
\begin{equation*}
\frac{d^{2} V}{d x^{2}}=\frac{n_{i}}{N}\left[\exp \left(-\phi_{n}\right) \exp (V)-\exp \left(-\phi_{p}\right) \exp (-V)+\frac{N_{A}-N_{D}}{n_{i}}\right] \tag{6}
\end{equation*}
$$

where the quasi-Fermi levels are also normalized. Assuming Einstein's relations still hold, the current density equation may be re-written as

$$
\begin{align*}
& J_{n}=-q \mu_{n} n \frac{\delta V}{\delta x}+q \mu_{n} \frac{k_{B} T}{q} \frac{\partial}{\partial x}\left[n_{i} \exp \left(\frac{q\left(V-\phi_{n}\right)}{k_{B} T}\right)\right] \\
& =-q \mu_{n} n \frac{\delta V}{\delta x}+q \mu_{n} \frac{k_{B} T}{q} n \frac{q}{k_{B} T}\left[\frac{\partial V}{\partial x}-\frac{\partial \phi_{n}}{\partial x}\right] \\
& =-q \mu_{n} n \frac{\delta \phi_{n}}{\delta x} \\
& =-q \mu_{n} n_{i} \exp \left[\frac{q\left(V-\phi_{n}\right)}{k_{B} T}\right] \frac{\delta \phi_{n}}{\delta x} \\
& =-q \mu_{n} n_{i} \exp \left(\frac{q V}{k_{B} T}\right) \frac{-k_{B} T}{q} \frac{\partial}{\delta x} \exp \left(\frac{\left.-q \phi_{n}\right)}{k_{B} T}\right) \tag{7}
\end{align*}
$$

which may be written more compactly, including quasi-Fermi level normalization, as

$$
\begin{equation*}
J_{n}=a_{n}(x) \frac{\delta}{\delta x} \exp \left(-\phi_{n}\right) \tag{8}
\end{equation*}
$$

A similar formula is obtained for the holes

$$
\begin{equation*}
J_{p}=a_{p}(x) \frac{\delta}{\delta x} \exp \left(\phi_{p}\right) \tag{9}
\end{equation*}
$$

and the continuity equations are therefore given by

$$
\begin{align*}
& \frac{\delta}{\delta x}\left[a_{n}(x) \frac{\delta}{\delta x} \exp \left(-\phi_{n}\right)\right]=q U(x)  \tag{10}\\
& \frac{\delta}{\delta x}\left[a_{p}(x) \frac{\delta}{\delta x} \exp \left(\phi_{p}\right)\right]=q U(x) \tag{11}
\end{align*}
$$

The continuity equations may be discretized with a straightforward finite difference approach (here for simplicity with uniform mesh)

$$
\begin{equation*}
\frac{\frac{a_{\alpha}\left(i+\frac{1}{2}\right)\left[\Phi_{\alpha}(i+1)-\Phi_{\alpha}(i)\right]}{\Delta x}-\frac{a_{\alpha}\left(i-\frac{1}{2}\right)\left[\Phi_{\alpha}(i)-\Phi_{\alpha}(i-1)\right]}{\Delta x}}{\Delta x}=U \tag{12}
\end{equation*}
$$

where the Slotboom variables have been used for simplicity of notation. Note that the inner derivative has been discretized with centered differences around the points $\left(i \pm \frac{1}{2}\right)$ of the interleaved mesh. Variables on the interleaved mesh must be determined very carefully, using consistent interpolation schemes for potential and carrier density, as discussed later. The discretized continuity equations lead to the tridiagonal system

$$
\begin{align*}
& a_{n}\left(i-\frac{1}{2}\right) \Phi_{n}(i-1)-\left[a_{n}\left(i+\frac{1}{2}\right)+a_{n}\left(i-\frac{1}{2}\right)\right] \Phi_{n}(i)+a_{n}\left(i+\frac{1}{2}\right) \Phi_{n}(i+1)=\Delta^{2} x U(i)  \tag{13}\\
& a_{p}\left(i-\frac{1}{2}\right) \Phi_{p}(i-1)-\left[a_{p}\left(i+\frac{1}{2}\right)+a_{p}\left(i-\frac{1}{2}\right)\right] \Phi_{p}(i)+a_{p}\left(i+\frac{1}{2}\right) \Phi_{p}(i+1)=-\Delta^{2} x U(i) \tag{14}
\end{align*}
$$

The decoupled iteration now solves Poisson's equation, initially with a guess for the quasi-Fermi levels. The voltage distribution obtained for the previous voltage considered is normally a good initial guess for the potential. Since the quasi-Fermi levels are inputs for Poisson's equation, the quasi-linearization procedure for equilibrium can be used again. The potential is then used to update the $a_{n}(i)$ and ${ }^{a_{p}(i)}$, and Eqs. (13) and (14) are solved to provide new quasi-Fermi level values for Poisson's equation, and the process is repeated until convergence is reached. The
generation-recombination term depends on the electron and hole concentrations, therefore it has to be updated at each iteration. It is possible to update the generation-recombination term also intermediately, using the result for the electron concentration.

The examples given bellow illustrates the Gummel's approach that is limited to the nondegenerate case. If field dependent mobility and diffusion coefficients are introduced, minimal changes should be necessary, as long as it is still justified the use of Einstein's relations. Extension to nonuniform mesh is left as an exercise for the reader. In the 2-D case, the quasilinearized Poisson's equation becomes

$$
\begin{align*}
& -\left(4+h^{2} \frac{n_{i}}{N}\left[\Phi_{n}(i, j) \exp (V(i, j))+\Phi_{p}(i, j) \exp (-V(i, j))\right]\right) \partial V(i, j)+[\partial V(i-1, j)+\partial V(i+1, j)+\partial V(i, j-1)+\partial V(i, j+1)]= \\
& 4 V(i, j)-V(i-1, j)-V(i+1, j)-V(i, j-1)-V(i, j+1)+h^{2} \frac{n_{i}}{N}\left[\Phi_{n}(i, j) \exp (V(i, j))+\Phi_{p}(i, j) \exp (-V(i, j))+\frac{N_{A}+N_{B}}{n_{i}}\right] \tag{15}
\end{align*}
$$

The normalized mesh size is $h=\Delta x=\Delta y$. As before, the thermal voltage $k_{B} T / q$ has been used to normalize the potential V and the quasi-Fermi levels $\phi_{n}$ and $\phi_{p}$ included in the Slotboom variables $\Phi_{n, p}=\exp \left( \pm \phi_{n, p}\right)$.

The continuity equations with the form $\nabla \cdot(a(x, y) \nabla \Phi)= \pm U(x, y)$ are discretized as

$$
\begin{align*}
& -\left[a\left(i+\frac{1}{2}, j\right)+a\left(i-\frac{1}{2}, j\right)+a\left(i, j+\frac{1}{2}\right)+a\left(i, j-\frac{1}{2}\right)\right] \Phi(i, j)+a\left(i+\frac{1}{2}, j\right) \Phi(i+1, j) \\
& +a\left(i-\frac{1}{2}, j\right) \Phi(i-1, j)+a\left(i, j+\frac{1}{2}\right) \Phi(i, j+1)+a\left(i, j-\frac{1}{2}\right) \Phi(i, j-1)= \pm h^{2} U(i, j) \tag{16}
\end{align*}
$$

## Newton's Method

Newton's method is a coupled procedure which solves the equations simultaneously, through a generalization of the Newton-Raphson method for determining the roots of an equation. We rewrite Eqs. $(6,10,11)$ in the residual form

$$
\begin{equation*}
W_{V}(V, n, p)=0 \quad W_{n}(V, n, p)=0 \quad W_{p}(V, n, p)=0 \tag{17}
\end{equation*}
$$

Starting from an initial guess $V_{o}, n_{o}$, and $p_{o}$, the corrections $V, \Delta n$, and $\Delta p$ are calculated from the Jacobian system

$$
\left(\begin{array}{lll}
\frac{\delta W_{V}}{\delta V} & \frac{\delta W_{V}}{\delta n} & \frac{\delta W_{V}}{\delta p}  \tag{18}\\
\frac{\delta W_{n}}{\delta V} & \frac{\delta W_{n}}{\delta n} & \frac{\delta W_{n}}{\delta p} \\
\frac{\delta W_{p}}{\delta V} & \frac{\delta W_{p}}{\delta n} & \frac{\delta W_{p}}{\delta p}
\end{array}\right)\left(\begin{array}{l}
\Delta V \\
\Delta n \\
\Delta p
\end{array}\right)=-\left(\begin{array}{l}
W_{V} \\
W_{n} \\
W_{p}
\end{array}\right)
$$

which is obtained by Taylor expansion. The solutions are then updated according to the scheme

$$
\begin{align*}
& V(k+1)=V(k)+\Delta V(k) \\
& n(k+1)=n(k)+\Delta n(k) \\
& p(k+1)=p(k)+\Delta p(k) \tag{19}
\end{align*}
$$

where $k$ indicates the iteration number. In practice, a relaxation approach is also applied to avoid excessive variations of the solutions at each iteration step.

The system (18) has 3 equations for each mesh point on the grid. This indicates the main disadvantage of a full Newton iteration, related to the computational cost of matrix inversion (one may estimate that a $3 N \times 3 N$ matrix takes typically 20 times longer to invert than an analogous $N \times N$ matrix). On the other hand convergence is usually fast for the Newton method, provided that the initial condition is reasonably close to the solution, and is in the neighborhood where the solution is unique. There are several viable approaches to alleviate the computational requirements of the Newton's method. In the Newton-Richardson approach, the Jacobian matrix is updated only when the norm of the error does not decrease according to a preset criterion. In general, the Jacobian matrix is not symmetric positive definite, and fairly expensive solvers are necessary. Iterative schemes have been proposed to solve each step of Newton's method by reformulating Eq. (18) as

$$
\left(\begin{array}{ccc}
\frac{\delta W_{V}}{\delta V} & 0 & 0  \tag{20}\\
\frac{\delta W_{n}}{\delta V} & \frac{\delta W_{n}}{\delta n} & 0 \\
\frac{\delta W_{p}}{\delta V} & \frac{\delta W_{p}}{\delta n} & \frac{\delta W_{p}}{\delta p}
\end{array}\right)\left(\begin{array}{l}
\Delta V \\
\Delta n \\
\Delta p
\end{array}\right)_{k+1}=-\left(\begin{array}{ccc}
W_{V} \\
W_{n} \\
W_{p}
\end{array}\right)-\left(\begin{array}{ccc}
0 & \frac{\delta W_{V}}{\delta n} & \frac{\delta W_{V}}{\delta p} \\
0 & 0 & \frac{\delta W_{n}}{\delta p} \\
0 & 0 & 0
\end{array}\right)\left(\begin{array}{c}
\Delta V \\
\Delta n \\
\Delta p
\end{array}\right)_{k}
$$

Since the matrix on the left hand side is lower triangular, one may solve Eq. (20) by decoupling into three systems of equations solved in sequence. First, one solves the block of equations (again, one for each grid point)

$$
\begin{equation*}
\frac{\delta W_{V}}{\delta V}(\Delta V)_{k+1}=-W_{V}-\frac{\delta W_{V}}{\delta n}(\Delta n)_{k}-\frac{\delta W_{V}}{\delta p}(\Delta p)_{k} \tag{21}
\end{equation*}
$$

and the result is used in the next block of equations

$$
\begin{equation*}
\frac{\delta W_{n}}{\delta n}(\Delta n)_{k+1}=-W_{n}-\frac{\delta W_{n}}{\delta V}(\Delta V)_{k+1}-\frac{\delta W_{n}}{\delta p}(\Delta p)_{k} \tag{22}
\end{equation*}
$$

Similarly, for the third block

$$
\begin{equation*}
\frac{\delta W_{p}}{\delta p}(\Delta p)_{k+1}=-W_{p}-\frac{\delta W_{p}}{\delta V}(\Delta V)_{k+1}-\frac{\delta W_{p}}{\delta n}(\Delta n)_{k+1} \tag{23}
\end{equation*}
$$

The procedure achieves a decoupling of the equations as in a block Gauss-Seidel iteration, and can be intended as a generalization of the Gummel method. A block-SOR method is obtained if the left hand sides are premultiplied by a relaxation parameter. This iteration procedure has better performance if the actual variables are $\left(V, \phi_{n}, \phi_{p}\right)$.

In general, Gummel's method is preferred at low bias because of its faster convergence and low cost per iteration. At medium and high bias the Newton's method becomes more convenient, since the convergence rate of Gummel's method becomes worse as the coupling between equations becomes stronger at higher bias. But since Gummel's method has a fast initial error reduction, it is often convenient to couple the two procedures, using Newton's method after
several Gummel's iterations. Remember that it is very important for the Newton's iteration to start as close as possible to the true solution. Close to convergence, the residual in Newton's iteration should decrease quadratically from one iteration to the other.

## Generation and Recombination

The Shockley-Reed-Hall model is very often used for the generation-recombination term due to trap levels

$$
\begin{equation*}
U_{\text {SRH }}=\frac{n p-n_{i}^{2}}{\tau_{p}\left[n+n_{i} \exp \left(\frac{q\left(E_{t}-E_{i}\right)}{k_{B} T}\right)\right]+\tau_{n}\left[p+n_{i} \exp \left(\frac{q\left(E_{i}-E_{t}\right)}{k_{B} T}\right)\right]} \tag{24}
\end{equation*}
$$

where $E_{t}$ is the trap energy level involved and $\tau_{n}$ and $\tau_{p}$ are the electron and hole lifetimes. Surface rates may be included with a similar formula, in which the lifetimes are substituted by $\frac{1}{S_{n, p}}$ where $S_{n, p}$ is the surface recombination velocity.

The Auger recombination may be accounted for by using the formula

$$
\begin{equation*}
U_{A u g}=C_{n}\left[p n^{2}-n n_{i}^{2}\right]+C_{p}\left[n p^{2}-p n_{i}^{2}\right] \tag{25}
\end{equation*}
$$

where $C_{n}$ and $C_{p}$ are appropriate constants. The Auger effect is for instance very relevant in the modeling of highly doped emitter regions in bipolar transistors.

The generation process due to impact ionization can be included using the fielddependent rate

$$
\begin{equation*}
U_{I}=\frac{a_{n}^{\infty} \exp \left(\frac{-E_{n}^{c r i t}}{E}\right)^{\beta n}\left|J_{n}\right|+a_{p}^{\infty} \exp \left(\frac{-E_{p}^{c r i t}}{E}\right)^{\beta_{p}}\left|J_{p}\right|}{q} \tag{26}
\end{equation*}
$$

