A Compact Diode Model for the Simulation of Fast Power Diodes including the Effects of Avalanche and Carrier Lifetime Zoning

A.T. Bryant
Engineering Department
University of Cambridge
Trumpington Street
Cambridge CB2 1PZ, UK
a.t.bryant.97@cantab.net

P.R. Palmer
Electrical & Computer Engineering
University of British Columbia
2356 Main Mall, Vancouver
BC Canada V6T 1Z4
ppr@ece.ubc.ca

E. Santi
Dept. of Electrical Engineering
University of South Carolina
Columbia, SC 29208, USA

J.L. Hudgins
Dept. of Electrical Engineering
University of Nebraska
Lincoln, NE 68588-0511, USA

Abstract—This paper presents the development and implementation of a compact diode model which can simulate aspects of high-voltage diodes such as snappy recovery during punch-through and the modified carrier density profile due to local lifetime control. It can be used in both circuit simulators and the formal optimisation of devices and circuits. A Fourier-based solution is used to solve the ambipolar diffusion equation (ADE) and describe the carrier dynamics. The model is shown to capture the required aspects of high-voltage diode recovery successfully, including the use of local lifetime control to eliminate snappy recovery.

I. INTRODUCTION

The current trend towards high-voltage power devices is placing increasing demands on device simulation. High-voltage power diodes, in particular, exhibit phenomena which are rarely captured in existing models suitable for circuit simulators. This leads to poor prediction of switching behaviour and hence power dissipation, thermal performance and EMC considerations.

High-voltage diodes are characterised by a lightly-doped wide N drift region sandwiched between highly-doped P and N emitter regions. The width of the drift region, in conjunction with its low doping, allows a high blocking voltage to be achieved. The carrier lifetime must be high enough to allow sufficient excess carriers to accumulate in the drift region to achieve a reasonable on-state (forward) voltage drop, but low enough to allow a quick removal of stored charge at turn-off. Practical devices have relatively long lifetimes, which give quicker recovery times; however under some conditions the reduced level of charge allows the space-charge region to occupy the whole drift region before the recovery current has ceased, causing classic ‘snappy’ recovery (also known as punch-through) [1]. Snappy recovery is detrimental to the device as it increases the time of its destruction due to the excessive electric field strength. Furthermore, it produces large-amplitude high-frequency oscillations which cause excessive amounts of electromagnetic interference. These can also be reflected in the IGBT gate waveforms, which again can cause device destruction [2]. The likelihood of snappy recovery occurring can be avoided by reducing the P emitter [3]–[5].

In higher-voltage devices snappy recovery is more likely to occur, and therefore local lifetime control is used as a means of eliminating this. In typical lumped-charge models [6], the lack of spatial information within the solution to the drift region dynamics precludes the modelling of local lifetime. In addition, most compact diode models either have no punch-through capability at all or only with one depletion layer.

More detailed models solve the ambipolar diffusion equation (ADE), which describes excess carrier dynamics under high-level injection. These are frequently based on finite difference approximations [7]–[9] or decomposition to specific basis functions such as exponentials or sinusoids [10]–[13]. None are as yet able to account for local lifetimes and few are able to account for punch-through. The lack of local lifetime modelling coupled with punch through means that existing high voltage diode models are of limited value in designing circuits and cannot be used at all in the design of the diodes themselves.

The Fourier-based solution of the ADE [14] has recently been shown to be suitable for providing an accurate, compact and rapid model for the simulation of power semiconductor devices [15]–[19]. This paper describes the modification of the standard Fourier-based solution to account for local lifetimes, and the revision of the boundary conditions necessary to accommodate the modelling of punch-through.

II. MODELLING HIGH VOLTAGE DIODES

The general arrangement of the carrier storage region (CSR) in the power diode is shown in fig. 1. In the on-state the depletion layers effectively vanish, allowing the CSR to occupy the whole width of the drift region. During turn-off the depletion layers expand from the ends of the drift region as the excess carriers are removed from the CSR by the reverse recovery current. This allows the diode to support a reverse voltage. If the depletion layers meet – i.e. punch through occurs – the device capacitance drops significantly and this interacts strongly with the circuit stray inductance, causing the large-amplitude high-frequency oscillations characteristic of snappy recovery.

The solution of the ADE in the carrier storage region is critical to modelling the device successfully. The excess carrier density is described by the ADE as:

\[ D \frac{\partial^2 p(x, t)}{\partial x^2} = \frac{p(x, t)}{\tau} + \frac{\partial p(x, t)}{\partial t} \]  

(1)

where \( p(x, t) \) is the excess carrier density, \( D \) is the ambipolar diffusivity and \( \tau \) is the high-level lifetime. \( p(x, t) \) can be expressed as a Fourier series in space \( x \), with the coefficients \( p_k \) varying in time \( t \):

\[ p(x, t) = \sum_{k=0}^{\infty} p_k(t) \cos \left( \frac{\pi k(x-x_1)}{x_2-x_1} \right) \]  

(2)

This work was supported by the U.S. Office of Naval Research under grant N00014-00-1-0131 and the Schiff Foundation, Cambridge University.
The use of this representation allows a compact solution of the ADE to be obtained. This is achieved by multiplying each term in the ADE by:

$$\cos\left(\frac{\pi k(x - x_1)}{(x_2 - x_1)}\right)$$

and integrating from $x_1$ to $x_2$. The standard solution may be found in [14], [15]. The boundary conditions are the carrier density gradients $\partial p/\partial x|_{x_1}$, $\partial p/\partial x|_{x_2}$ and the positions of the boundaries $x_1$, $x_2$. The carrier density gradients are obtained from the boundary currents, which take account of the minority carrier recombination in the emitter regions and the displacement current due to depletion layer expansion and contraction. The positions are derived from the depletion layer behaviour.

A. Solution of the ADE with Variable Lifetime

While the standard solution of the ADE is suitable for fixed lifetime $\tau$, the solution for variable lifetime must clearly take account of variation of $\tau$ with position $x$. Hence the $p/\tau$ term in the lifetime must be re-integrated with this consideration. In addition, the lifetime profile observed within the CSR will depend on where its boundaries $x_1$, $x_2$ are.

Firstly, the lifetime $\tau(x)$ will be considered, assuming that the profile of the lifetime between the boundaries $x_1$, $x_2$ is known. This is most convenient if the reciprocal of the lifetime is expressed as a Fourier series between the boundaries, given by:

$$\frac{1}{\tau(x)} = \tau'(x) = \sum_{k=0}^{\infty} \tau'_k(t) \cos\left(\frac{k\pi(x - x_1)}{(x_2 - x_1)}\right)$$

This gives Fourier coefficients which will vary in time since the boundaries are moving. Including this in the integration of the $p/\tau$ term of the ADE gives the following expressions for calculating the Fourier coefficients:

$$k > 0: \frac{dp_k}{dt} = \frac{2D}{(x_2 - x_1)} \left[ \frac{\partial p}{\partial x|_{x_2}} (1)^k - \frac{\partial p}{\partial x|_{x_1}} + p_k \ \tau'_0 + \frac{D\pi^2k^2}{(x_2 - x_1)^2} \right] p_{2k} + p_n = \sum_{n=k}^{\infty} \tau'_n (p_{n+k} + p_{n-k})$$

When the lifetime is constant across the drift region, i.e. $\tau'_0 = 1/\tau$ and $\tau'_k = 0$ for all $k > 0$, these expressions reduce to the same as the standard case.

Calculation of the Fourier coefficients of $\tau'$ proceeds as follows. Firstly, the lifetime must be expressed analytically, i.e. the lifetime $\tau$ as a function of $x$. The positions of the boundaries then give the positions within the CSR at which the lifetime is calculated. Reciprocals of the lifetimes are then taken, and these values used to generate the Fourier series coefficients of $\tau'$ which are then used in the solution of the ADE.

If $M$ coefficients of $\tau'$ are required, these can be found as a 'best fit' from $M$ points of the values of $\tau'$. This is achieved by finding the inverse of the original Fourier series expression which, when truncated to $M$ terms, is simply a matrix equation in $M$ dimensions, shown in equation (6). The matrix on the right-hand side of the equation is symmetric and invertible and is used to find $\tau'_k$ in terms of $\tau(x)$.

$$\begin{bmatrix} \tau'(x_1, t) \\ \tau'(x_1 + \Delta x, t) \\ \tau'(x_1 + 2\Delta x, t) \\ \vdots \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & \cdots \\ 1 & \cos \frac{\pi}{M-1} & \cos \frac{2\pi}{M-1} & \cdots \\ 1 & \cos \frac{2\pi}{M-1} & \cos \frac{4\pi}{M-1} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} \begin{bmatrix} \tau'_0(t) \\ \tau'_1(t) \\ \tau'_2(t) \\ \vdots \end{bmatrix}$$

(6)

B. Depletion Layer Behaviour at Punch-Through

When considering depletion layer behaviour at punch-through, the correct modelling of displacement current is essential. Once the depletion layers have met and the CSR proceeds as

$$I_{\text{disp}} = \varepsilon A \frac{\partial E}{\partial t}$$

(7)

When the depletion layers meet they merge to form one depletion layer extending across the whole drift region. This can be split into two adjacent regions, characterised by positive and negative electric field gradients $E' = \varepsilon A \frac{\partial E}{\partial x}$. Using straight-line approximations to the electric field – i.e. constant $E'$ – is valid, and so the existence of two distinct but interdependent depletion layers can be assumed. The simplest way to define where the depletion layers meet, and hence define the value of the electric field at the meeting point $E_{\text{m}}$, is with the consideration of the junction electric fields $E_{01}, E_{02}$.
respectively can be described using the following equations:

\[ E_m = \frac{E_{01} |E'_1| + E_{02} |E'_2| + W_B |E'_1| |E'_2|}{E'_1 + |E'_2|} \]  

(8)

The assumption of zero electric field in the CSR is not strictly true, however, due to its resistive voltage drop. The boundary electric fields at \( x_1 \) and \( x_2 \) based on \( p_{s1} \) and \( p_{s2} \) respectively can be described using the following equations:

\[ E_{x1} = \frac{I_A}{qA (\mu_n N_B + (\mu_n + \mu_p) p_{s1})} \]  

(9)

\[ E_{x2} = \frac{I_A}{qA (\mu_n N_B + (\mu_n + \mu_p) p_{s2})} \]  

(10)

Defining \( E_{n1} \) and \( E_{n2} \) as the electric fields seen by the depletion layers at \( x_1 \) and \( x_2 \) respectively,

\[ E_{n1} = \begin{cases} E_{x1} & E_m > E_{x1} \\ E_m & E_m < E_{x1} \end{cases} \]  

(11)

\[ E_{n2} = \begin{cases} E_{x2} & E_m > E_{x2} \\ E_m & E_m < E_{x2} \end{cases} \]  

(12)

\[ W_{d1} = \frac{E_{n1} - E_{01}}{|E'_1|} \]  

(13)

\[ W_{d2} = \frac{E_{n2} - E_{02}}{|E'_2|} \]  

(14)

**Fig. 2.** Punch-through determination using the centre boundary field \( E_m \).

and the electric field gradients \( E'_1, E'_2 \). Fig. 2 shows the action of the two depletion layers before, at and after punch-through.

The following expression gives the centre boundary electric field \( E_m \) in terms of the drift region width, the field gradients and the boundary fields (noting that the electric fields are negative):

\[ E_m = \frac{E_{01} |E'_1| + E_{02} |E'_2| + W_B E'_1 |E'_2|}{E'_1 + |E'_2|} \]  

Dynamic avalanche plays an important role in limiting the peak reverse voltage during snappy recovery. Effectively it increases the depletion layer capacitance because generation of electron-hole pairs decreases the electric field gradient.

Fig. 3 shows the effect of impact ionisation on the depletion layer. Generated holes are swept at saturated velocity towards the P emitter, and generated electrons towards the N+ emitter. This gives a reduced electric field slope \( \partial E/\partial x \) as \( x \) increases, but a steeper slope towards the PN- junction. The electric field at the PN- junction is much larger than that at the N-N+ junction, and therefore it may be assumed for the purposes of this model that dynamic avalanche only takes place at the PN- junction.

Other than the small region adjacent to the junction, the electric field slope \( \partial E/\partial x \) is given by:

\[ E'_1 = \frac{qN_B}{\varepsilon} + \frac{|I_{p1}|}{\varepsilon \alpha_v \varepsilon} - \frac{|I'_{gen}|}{\varepsilon m_0 \alpha_v \varepsilon} \]  

(17)

\( I_{p1} \) and \( I_{gen} \) are the hole and electron currents at the depletion layer edge during avalanche away from the junction.

The generated current \( I_{gen} \) consists entirely of holes at the PN-junction, and can therefore be added onto the hole current \( I_{p1} \) to give the total hole current at this junction. Since the electric field slope only varies in the small region adjacent to the junction, equation (17) may be used to calculate the depletion layer widths and voltages.

Solution of \( I_{gen} \) proceeds as in [20]. Assuming that \( \partial p/\partial t, \partial n/\partial t \approx 0 \), the current is integrated across the depletion layer width to give:

\[ I'_{gen} = \frac{m_0 I_{p1}}{m_1 E'_1 E'_2 \exp \left( m_1 E'_1 W_{d1} \right) \exp \left( m_1 E'_2 W_{d2} \right) - 1} \]  

(18)

\[ I_{gen} = G(s) I'_{gen} \]  

(19)
where $G(s)$ is defined in equation (26) and the constants are defined in equations (20-25):

$$\alpha_0 = A_t \exp \left( -b_t / E_0 \right) \quad (20)$$

$$m_1 = b_t / E_0^3 \quad (21)$$

$$m_2 = b_t / E_0 \quad (22)$$

$$A_t = 1.07 \times 10^6 \text{cm}^{-1} \quad (23)$$

$$b_t = 1.65 \times 10^6 \text{Vcm}^{-1} \quad (24)$$

$$E_0 = 1.9 \times 10^5 \text{Vcm}^{-1} \quad (25)$$

$m_0=1.3$ is chosen so that this expression matches the exact solution (which is not used due to its poor convergence [20]). $E_{n1}$ and $W_{d1}$ are calculated in equations (11,13). To ensure generality applicability of the model, $I_{gen}$ is added to the emitter recombination current $I_{n1}$ to give the total electron current at the edge of the CSR (at $x=x_1$).

The transit time across the depletion layer is given by approximately $2W_{d1}/v_{sat}$ [20]. This is modelled as a simple first-order lag applied to the calculation of the electric field slope $E'_x$. To simplify the model the time constant is approximated to $W_{B}/v_{sat}$; the lag transfer function $G(s)$ is then given by:

$$G(s) = \frac{1}{\left( \frac{W_B}{v_{sat}} \right) s + 1} \quad (26)$$

**D. Behaviour of the Depletion Layer at the N-N+ Junction**

It should be noted that the depletion layer at the N-N+ junction only forms under specific circumstances. The voltage drop across the N-N+ junction is much lower than that of the PN junction. The former becomes more significant with higher reverse currents and decreasing doping [21]. Under low-field and low-level injection conditions, electrons can move freely between the $N$- and N+ regions and the behaviour is ohmic [22]. Hence $J_n = q\mu_n E$ with $n \approx N_B$. At high reverse currents during recovery, the field increases and the electrons reach velocity saturation. This causes the electron carrier density $n$ to change to a level set by the electron current density $J_n$, with $n = J_n / (q\mu_n E)$, giving a negative $E'_x$ and the formation of a depletion layer.

The standard expression for field-dependent electron mobility [23] can be inserted into the drift current equation $J_n = q\mu_n E$ to obtain:

$$n = \frac{J_n}{q\mu_n E} \left( 1 + \left( \frac{\mu_n E}{v_{sat}} \right)^2 \right) \quad (27)$$

The condition $n \approx N_B$ for low electric fields gives $J_n / (q\mu_n E) = N_B$, and inserting this and equation (27) into Poisson’s equation gives:

$$E'_x = -\frac{qN_B}{\varepsilon} \left( 1 + \left( \frac{I_{n2}}{qAN_B \varepsilon v_{sat}} \right)^2 - 0.999 \right) \quad (28)$$

At low fields and currents this reduces to $\partial E / \partial x \approx 0$, and at high currents it reduces to $E'_x \approx J_n / (e v_{sat})$. The value of 0.999 is used instead of 1 to avoid a divide-by-zero error when using $E'_x$ in equation (14).

**E. CSR Formulation during Punch-through**

When punch-through occurs, the formulation for the width of the CSR is limited to approximately one diffusion length, since this allows the hole and electron concentrations to vary in a physical manner across the CSR. The diffusion length $L_a$ is calculated from the reduced hole and electron mobilities, which depend on the electric field in this region ($\approx (E_{n1} + E_{n2})/2$):

$$\mu_n = \frac{\mu_{n0}}{\sqrt{1 + \left( \frac{\mu_{n0}(E_{n1} + E_{n2})}{2v_{sat,n}} \right)^2}} \quad (29)$$

$$\mu_p = \frac{\mu_{p0}}{\sqrt{1 + \left( \frac{\mu_{p0}(E_{n1} + E_{n2})}{2v_{sat,p}} \right)^2}} \quad (30)$$

$$D = \frac{2V_f \mu_n \mu_p}{\mu_n + \mu_p} \quad (31)$$

$$L_a = \sqrt{D} \quad (32)$$

The CSR width is then limited to the diffusion length as follows:

$$W = \begin{cases} W_B - W_{d1} - W_{d2} & \text{if } W_B - W_{d1} - W_{d2} > 0.9675L_a \\ 0.9675L_a & \text{otherwise} \end{cases} \quad (33)$$

Retaining this width in the model removes the need for a model change. However, the accuracy of the model – in particular the frequency and amplitude of the snappy oscillations following punch-through – is sensitive to the scale factor applied to $L_a$. A value of 0.9675 was chosen to obtain the best matching. Also, the limiting of the width $W$ during punch-through applies only to the Fourier series ADE solution: the electric fields, depletion layer widths and voltages are still calculated assuming that the CSR has zero width.

When modelling punch-through, the input and output requirements of the CSR are reversed with respect to that provided by the Fourier series solution. A simple PI control loop is used to adapt this accordingly:

$$\frac{\partial p}{\partial x} \bigg|_{x_1} = -\left( K_{FP} + \frac{K_{FI}}{s} \right) (p_{x1,in} - p_{x1,calc}) \quad (34)$$

$$\frac{\partial p}{\partial x} \bigg|_{x_2} = \left( K_{FP} + \frac{K_{FI}}{s} \right) (p_{x2,in} - p_{x2,calc}) \quad (35)$$

In a similar fashion to the simple diode model, the boundary carrier densities are related to the depletion layers to provide the feedback around the CSR. However, in this case the electric fields are used to calculate the carrier densities. Since the electric fields will be negative when the depletion layers exist, the feedback must only produce carrier densities when the electric fields are positive.

$$p_{x1} = \begin{cases} K_F E_{01} & \text{if } K_F E_{01} > -N_B \\ -N_B & \text{otherwise} \end{cases} \quad (36)$$

$$p_{x2} = \begin{cases} K_F E_{02} & \text{if } K_F E_{02} > -N_B \\ -N_B & \text{otherwise} \end{cases} \quad (37)$$

The electric fields are generated by integrating the displacement currents. The equations used are given as follows, with the emitter recombination equations (38,39) given by [24]:
The Fourier series must be truncated in order to realise a practical implementation. The number of terms used \((M)\) will depend on the relative sizes of the ambipolar diffusion length \((L_a = \sqrt{D\tau})\) and the drift region width \(W_B\). For high-voltage diodes, where the diffusion length is significantly smaller than the drift region width, \(M\) may be as high as 11 to 15. Due to Gibb’s phenomenon (arising from the Fourier series truncation) it was found necessary to include a filter to reduce spatial oscillations in the carrier density. This was achieved using a triangular window function applied to the Fourier coefficients [26], [27]. Note that as the coefficients of a finite cosine Fourier series are even, any values of \(p_{n+k}\) and \(p_{n-k}\) that are required in equation (4) when \(n + k\) or \(n - k\) are less than zero or greater than \(M - 1\) are obtained by reflecting the coefficients about the end indices \(p_0\) and \(p_{M-1}\). The constants required for the boundary carrier density feedback loops (equations (34-37)) are \(K_F=10^{13}\), \(K_{FP}=10^6\) and \(K_{FP}=10^{14}\).

The model can be implemented in a circuit simulator such as PSpice, where the differential equations in \(dp/dt\) are represented as R-C cells [14], [15]. The implementation adopted here uses MATLAB/Simulink, which takes advantage of MATLAB’s matrix capabilities to produce a compact model [16]. This is particularly useful in calculating the extra terms used for modelling variable lifetime. The number of terms \(M\) can easily be modified since the matrices necessary for calculating the differential equations in \(dp/dt\) are generated using a script file.

### III. Simulation Results

The diode model was applied to the freewheel diode in an IGBT-based chopper cell with an inductive load [15], [16], shown in fig. 5. The diode model was also simulated in Silvaco ATLAS [28] to validate the dynamic operation of the compact diode model. The comparison between ATLAS and the compact model may be seen in fig. 7.

Fig. 6 shows the lifetime profile of the diode with local lifetime control compared with the original diode. This is a hypothetical profile which similar in shape to that estimated for practical diodes with local lifetime control. Fig. 8 shows the simulation results.

The agreement between the compact model and ATLAS is excellent for both diodes. In the case of the diode with uniform lifetime, the compact model clearly captures the electric field during punch-through and, since the peak reverse recovery voltage closely matches that from the ATLAS simulation, correctly simulates dynamic avalanche. In the diode with local lifetime control, the carrier density is clearly higher towards the cathode end of the drift region as a result of the longer lifetime. During turn-off the charge is removed more quickly.

\[
V_{j1} = V_T \ln \left( \frac{p_{x1} N_B}{n_i^2} \right) \quad (50)
\]

\[
V_{j2} = V_T \ln \left( \frac{p_{x2}}{N_B} \right) \quad (51)
\]

\[
V_{AK} = V_{j1} + V_{j2} + V_D - V_{d1} - V_{d2} \quad (52)
\]
from the anode end of the drift region, allowing the depletion layer to expand and the diode reverse voltage to build up. The charge at the cathode takes longer to be removed, buffering the depletion layer from the N emitter and preventing punch-through from occurring. The oscillation seen in the ATLAS result in fig. 8(a) does not seem realistic. The shape of the reverse recovery current is typical of that expected of modern fast soft recovery diodes. This is achieved only using geometrical information and is due to the correct solution of the ADE.

**IV. CONCLUSIONS**

The model has clearly captured the effects of local lifetime control, punch-through and dynamic avalanche in the performance of fast power diodes. Their inclusion allows the use of a more accurate diode model in both circuit simulators and in formal optimisation. The latter is particularly promising since it can be used to investigate a detailed optimisation of the lifetime profile, and hence find the optimum trade-off between stored charge, recovery softness and forward voltage drop. The compact model allows many operating conditions to be analysed rapidly, to ensure reliable operation.

**REFERENCES**


Fig. 7. Waveforms, excess carrier density and electric field plots from the ATLAS (Silvaco) and compact diode model (Simulink) simulations for the diode without lifetime zoning. The diode used had area $A=0.75\text{cm}^2$, drift region width $W_D=200\mu\text{m}$, drift region doping $N_D=2\times10^{13}\text{cm}^{-3}$, lifetime $\tau=0.6\mu\text{s}$. The on-state current was $75\text{A}$ and the off-state voltage was $750\text{V}$.

Fig. 8. Waveforms, excess carrier density and electric field plots from the ATLAS (Silvaco) and compact diode model (Simulink) simulations for the diode with local lifetime control. The diode used had area $A=0.75\text{cm}^2$, drift region width $W_D=200\mu\text{m}$, drift region doping $N_D=2\times10^{13}\text{cm}^{-3}$, lifetime: mean $\tau=0.6\mu\text{s}$, max. $\tau=1.0\mu\text{s}$, min. $\tau=0.2\mu\text{s}$. The on-state current was $75\text{A}$ and the off-state voltage was $750\text{V}$.