Transport Theory #9: The Drift-Diffusion Formalism

One of the key assumptions that the transport theory has rested upon to this point is the conservation of particles. This is what allowed us to relate the non-equilibrium distribution function to the condition,

$$\iint f(\vec{k},\vec{r})d\vec{k}d\vec{r} = N \tag{1}$$

But in solid-state devices, there is necessarily a deviation from the equilibrium particle concentration because devices, by definition, must exchange energy with the outside environment to be useful. This is true whether the device is electrical (electron and/or hole), mechanical, acoustical, thermal (i.e., phononic), radiative (i.e., photonic), or some combination of these (i.e., mixed domain). So it is incumbent on any practical transport theory to address deviations from (1) at every point in the device. First, we will seek and derive fundamental constraints and conservation laws that can be used instead of (1) to help in solving for the non-equilibrium state of the solid. Then we will inspect what common types of external forces or influences can create significant deviation from equilibrium. Finally, we will generalize the Boltzmann transport formalism to include the fundamental principles and forces. The result is drift-diffusion equations – arguably the most important formalism in the world today for simulating and designing semiconductor devices today.

Fundamental Constraints

Conservation of Charge

Along with energy and momentum, a third important conserved quantity in many natural processes is electric charge. This is generally true in solids[1] and can be analyzed simply using the most scalable of all physical laws – Maxwell's equations. We start with the sample cube of volume Vc inside the solid as shown in Fig. 1, assumed electrically homogeneous so that the unipolar (i.e., free) charge density $\rho_f(t)$ is continuous throughout. Then the total free charge inside is

$$Q_f = \int_{V_C} \rho_f dV \tag{2}$$

And the free charge can be related to Maxwell's displacement vector by

$$\vec{\nabla}\cdot\vec{D} = \rho_f \tag{3}$$

As shown in our coverage of electrostatics, a second form of charge must generally be considered in solids, particularly in solids having strong paraelectric response, such as semiconductors. This is the dipolar charge density, which we showed was generally neutral inside the solid, but could be described in finite samples by the unneutralized (or "un-paired") surface charge,. The magnitude is always describable by a surface charge σ_S which is related to an equivalent volumetric bound charge density $\rho_b = \vec{\nabla}\cdot\vec{P}$, where $\vec{P}$ is the polarization density vector. Hence, the total surface charge on the cube in Fig. 1 is just

[1] Assuming, of course, that there are no nuclear reactions occurring

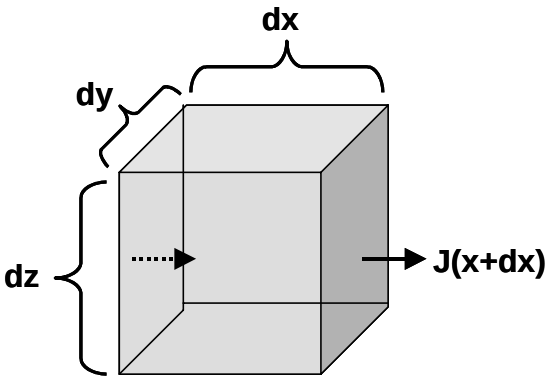

Fig. 1

$$Q_b = \oint_S \sigma_s \cdot dS = \oint_S \vec{P} \cdot d\vec{S} = \int_V (\vec{\nabla} \cdot \vec{P}) dV = \int_V \rho_b dV \tag{4}$$

where the step from the third term to the fourth term follows from Gauss' divergence theorem - arguably the most useful integral theorem of vector calculus.

The conservation of charge means that any change of the total free charge of (2) with time must be accountable to transport of an equivalent amount of charge into or out-of the cube. Mathematically, we can state this conditions using (2) in the following way:

$$\frac{\partial Q_f}{\partial t} = -I_f = \frac{\partial}{\partial t} \int_{V_C} \rho_f dV = \int_{V_C} \frac{\partial \rho_f}{\partial t} dV = -\oint_S \vec{J}_f \cdot d\vec{S} = -\int_{V_C} \vec{\nabla} \cdot \vec{J}_f dV \tag{5}$$

where I_f is the total (free) electrical current leaving the cube. From the third and sixth terms, we get

$$\frac{\partial \rho_f}{\partial t} = -\vec{\nabla} \cdot \vec{J}_f \tag{6}$$

an extremely important relation in device physics, called the free-charge current continuity equation.

Similarly any time dependent change of the bound charge can be written as:

$$\frac{\partial Q_b}{\partial t} = -I_b = -\frac{\partial}{\partial t} \int_V \rho_b dV = -\int_V \frac{\partial \rho_b}{\partial t} dV = -\frac{\partial}{\partial t} \int_V (\vec{\nabla} \cdot \vec{P}) dV = -\int_V (\vec{\nabla} \cdot \frac{\partial \vec{P}}{\partial t}) dV \tag{7}$$

where I_b is the total bound electrical current leaving the cube. From the fourth and sixth terms, we get

$$\frac{\partial \rho_b}{\partial t} = \vec{\nabla} \cdot \frac{\partial \vec{P}}{\partial t} \equiv -\vec{\nabla} \cdot \vec{J}_P \ , \tag{8}$$

where $\vec{J}_P$ is the polarization current density. Eqn (8) is called the polarization current continuity relation and it has a central role in quantum electronics for connecting atomic and molecular quantum transitions back to Maxwell's equations.

Students who have had a thorough course in quantum electronics or electromagnetism may think that this is the only purpose of (8). But it has a central role in device physics too, defining a key component of the reactive, or capacitive current current.[2] once we add it to the vacuum term. The vacuum current is given by the time derivative of the vacuum energy storage term.

$$\frac{\partial \rho_0}{\partial t} = \frac{\partial}{\partial T} \vec{\nabla} \cdot (\varepsilon_0 \vec{E}) \equiv \vec{\nabla} \cdot (\varepsilon_0 \frac{\partial \vec{E}}{\partial T}) = \vec{\nabla} \cdot \vec{J}_T \ , \tag{9}$$

where $\vec{J}_T$ is the "total" current. A good way to think about (8) is the current associated with the change of charge outside the sample that is needed to create $\vec{E}_0$. Adding up all three terms, we have

$$\frac{\partial \rho_f}{\partial t} + \frac{\partial \rho_b}{\partial t} + \frac{\partial \rho_0}{\partial t} \equiv \frac{\partial \rho_t}{\partial t} = -\vec{\nabla} \cdot \vec{J}_f - \vec{\nabla} \cdot \frac{\partial \vec{P}}{\partial t} - \varepsilon_0 \vec{\nabla} \cdot \frac{\partial \vec{E}}{\partial t} \tag{10}$$

now by using the electrostatic relation $\vec{P} = \varepsilon_0 \chi_e \vec{E}$ this becomes

$$\frac{\partial \rho_t}{\partial t} = -\vec{\nabla} \cdot \vec{J}_f - \vec{\nabla} \cdot \left(\varepsilon_0 (\chi_e + 1) \frac{\partial \vec{E}}{\partial t} \right) = -\vec{\nabla} \cdot \vec{J}_f - \vec{\nabla} \cdot \left(\varepsilon_r \varepsilon_0 \frac{\partial \vec{E}}{\partial t} \right) = -\vec{\nabla} \cdot \vec{J}_f - \vec{\nabla} \cdot \frac{\partial \vec{D}}{\partial t}$$

The right side is just the divergence of one side of Maxwell's dynamic equations

$$\vec{\nabla} \times \vec{H} = \vec{J}_f + \frac{\partial \vec{D}}{\partial t}$$

By taking the divergence of both sides and utilizing a basic theorem of vector calculus, we get

$$\vec{\nabla} \cdot \vec{\nabla} \times \vec{H} = 0 = \vec{\nabla} \cdot (\vec{J}_f + \frac{\partial \vec{D}}{\partial t}) = -\frac{\partial \rho_T}{\partial t}$$

[2] Another current could be added to the mix corresponding to the magnetic dipole, or magnetization current. This is very important in magnetic devices, but is left out of the present analysis for brevity. Since there are no magnetic monopoles , there is no magnetic current analogous to (5) and (6)

This expression emphasizes two very important facts: (1) the conserved quantity in space-charge neutrality is really the total charge density ρ_T, and (2) the corresponding total current $\vec{J}_f + \partial\vec{D}/\partial t$ must be a constant independent of location in the solid.

These are powerful concepts that come up repeatedly in solid-state device analysis. They are particularly useful when the electric field in the solid is a pure sinusoid or a sum of sinusoids, as often occurs in electrical engineering. Then since Ohm's law $\vec{J}_f = \sigma\vec{E}$ is so often valid in solids, we can write $\vec{\nabla}\cdot(\sigma\vec{E} + j\omega\varepsilon\vec{E}) = 0$ which means that the quantity $\sigma\vec{E} + j\omega\varepsilon\vec{E} = C$, a constant independent of position in the solid.

"Local" Thermodynamic Equilibrium

Devices force us to abandon the notion of equilibrium over the entire volume of the device. However, they are often operated under conditions that are close enough to equilibrium that the non-equilibrium distribution function throughout the device is a linear deviation from the equilibrium function or, at least, traceable back to the equilibrium function is some way. This justifies a second pervasive constraint in device physics and engineering in which the device is "globally" in non-equilibrium but "locally" in thermodynamic equilibrium. In other words, at any point in the solid, we can define local macroscopic thermodynamic variables that can be considered constant over some dimension much greater than the atomic scale. The most important such variables are the *intensive* ones:[3] the temperature, the chemical potential, the electric field, and the elastic strain. Being able to define these uniquely at each and every point in the solid has profound consequences. We have already seen an example of this in semiconductors, where knowledge of the temperature and chemical potential (synonymous with chemical potential in electronic devices) allowed us to write expressions for the electron and hole densities in the conduction and valence bands, respectively:

$$n_c = N_C(T)e^{(U_F - U_C)/kT} \quad \text{and} \quad p_v = N_V e^{(U_V - U_F)/kT}$$

These result in a very useful constraint

$$n_c p_v = N_C N_V e^{-(U_C - U_V)/k_B T} = N_C N_V e^{-U_G/k_B T} \equiv n_i^2$$

were n_i is the intrinsic carrier concentration. This is called the law of mass action.

External Forces and Influences for the Special Case of a Semiconductor

There are a variety of external "forces" that can significantly alter the charge carrier densities from their equilibrium values – a process called charge "injection." With metals or semimetals, these are all frustrated by the fact the are many ways to do this, but the following two are the most common.

[3] See Chapter 1 for review of thermodynamics in solids.

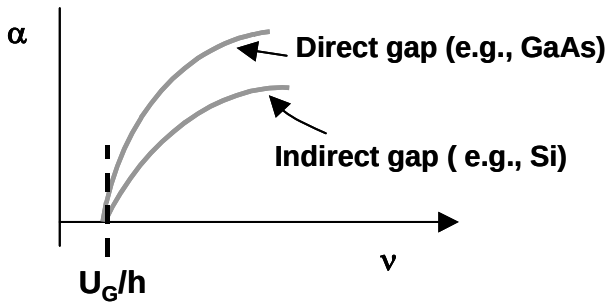

Fig. 2. Qualitative behavior of absorption coefficient in semiconductors or insulators as a function of radiation frequency, ν. Once above threshold, the a of direct-band-gap semiconductors (e.g., GaAs) is generally much larger than it is for indirect band-gap semiconductors Much larger for direct-gap materials (e.g., GaAs) than indirect-gap semiconductors (e.g., Si).

Radiation

Perhaps the easiest to understand is cross-gap generation in semiconductors. Semiconductors have the have the wonderful property of internal photoemission, which is very similar to surface photoemission or photoelectric effect that led to the original concept of representing electromagnetic radiation as photons.[4] The absorption coefficient of radiation is represented by a threshold absorption coefficient

$$\alpha \simeq \alpha(\nu)\theta[h\nu - U_G]$$

where θ is the Heaviside step function and α(ν) is generally sub-linear as shown in Fig. 2. Once in the solid, the intensity associated with the radiation tends to vary according to Beer's universal exponential decay law.

$$I \simeq I_0 e^{-\alpha z}$$

Every photon making up I is absorbed in a "corpuscular" fashion, meaning that the power absorbed per unit volume at each point in the solid is just

$$P'(z) \simeq -\frac{dI}{dz} = \alpha I_0 e^{-\alpha z} = \alpha I(z)$$

and the corresponding photon absorption rate is P'/ην.

Each photon generates an exciton – an electron-hole pair – that is initially highly correlated or even bound, but then rapidly de-correlates and annihilates to an independent electron and hole. Assuming that this happens to every exciton generated, the electron and hole generation rates, G_n and G_p, are given by

[4] A representation that won Einstein his only Nobel prize in 1921.

$$G_n = G_p = \frac{\alpha I(z)}{h\nu}$$

By the principle of quantum reversibility, photon absorption has an inverse process called spontaneous emission whereby a free electron and free hole annihilate, or "recombine", emitting a photon in the process. This process has associated recombination rates for electrons and holes, R_n and R_p, respectively, each of which is proportional to the "natural" radiative lifetime $1/\tau_{eh}$. In very pure semiconductors, these two rates are approximately equally,

$$R_n \approx R_p$$

In impure semiconductors having many energy levels in the band gap, these two rates can be substantially different.

Electrical Contacts

Much more common in devices but also more complicated from a solid-state physics standpoint is electrical injection. When the solid is an insulator or semiconductor, the injection is usually done by contacts. They are usually made with metal and fall into one of the following categories: (1) *ohmic* contacts for injecting the same type charge carriers as are predominant in the solid in the equilibrium state,[5] and (2) *rectifying* contacts for injecting the opposite type carrier. The are several types of rectifiying contacts with two being very common n semiconductor devices: (1) p-n junctions, and (2) Schottky barriers. The p-n junctions are easy to fabricate internally to the solid, creating the ability to do the minority carrier injection that is crucial to the function of bipolar transistors and semiconductor injection lasers. The Schottky barriers are easy to fabricate at a surface of the solid, creating the ability to do majority carrier injection in a manner controlled by the difference in the metal and semiconductor (or insulator) Fermi energies[6] and the difference in their electron affinities.

Like radiation, electrical injection can also be characterized by generation rates for majority and minority carriers, G_n for electrons and G_p for holes. But these rates are not necessarily equal at each point in the solid as they are with photonic injection. However, they generally have well-defined boundary conditions that get used to solve real devices. There are corresponding recombination rates, R_n and R_p which, again, are not necessarily equal.

Generalization of Current Continuity Equations

Given a semiconductor solid and the existence of generation and recombination mechanisms that perturb both free carrier types away from their equilibrium densities, we need to generalize the continuity equations. The first issue we face is how to generalize relation (6), $\partial\rho_f / \partial t = -\vec{\nabla} \cdot \vec{J}_f$, for two types of free carriers. The total electrical current can be thought of as independent electron and hole currents taken as statistical averages over the conduction and valence bands respectively, i.e.,

[5] In a semiconductor or semimetal, the predominant charge carrier type in equilibrium is called colloquially the "majority" carrier. The other carrier type is called the "minority" carrier.
[6] A quantity called the "work" function.

$\vec{J}_f = \vec{J}_n + \vec{J}_p$. Hence, a necessary but not sufficient condition to satisfy (6) is to decouple the electron and hole densities too, leading to

$$-e\frac{\partial n}{\partial t} = -\vec{\Delta}\cdot\vec{J}_n - e\left(G_n - R_n\right) \tag{11}$$

$$+e\frac{\partial p}{\partial t} = -\vec{\Delta}\cdot\vec{J}_p + e\left(G_p - R_p\right) \tag{12}$$

where we have paid careful attention to the signs of the charge-polarity-dependent terms.
To apply (11) and (12) we need expressions for $\vec{J}_n, \vec{J}_p, G_n, R_n, G_p$ *and* R_p in terms of n and p.
For $\vec{J}_n$ *and* $\vec{J}_p$ we go back to the semiclassical *Boltzmann* equation in a uniform electric field of force $F_e = qE$

$$\frac{df}{dt} = -\frac{\partial f}{\partial \vec{r}}\frac{d\vec{r}}{dt} - \frac{\partial f}{\partial \vec{k}}\frac{d\vec{k}}{dt} - \left.\frac{\partial f}{\partial t}\right|_{scatt}$$

$$\frac{d\vec{k}}{dt} = \frac{1}{\hbar}\vec{F} = \frac{q}{\hbar}\vec{E}\ ;\ \frac{d\vec{r}}{dt} = \vec{\mathrm{v}} = \frac{1}{\hbar}\vec{\nabla}_k U$$

so,
$$\frac{df}{dt} = -\vec{\mathrm{v}}\cdot\frac{\partial f}{\partial \vec{r}} - \frac{\vec{F}_e}{\hbar}\frac{\partial f}{\partial \vec{k}} - \left.\frac{\partial f}{\partial t}\right|_{scatt} \tag{13}$$

This can be reduced to special forms by various methods, all dependent on ascribing a temperature to the carriers consistent with local thermodynamic equilibrium. The following discussion only reviews these methods. For details, the student is referred to an excellent book by C. M Snowden. *Semiconductor Device Modelling*, Chapter 3, "Carrier Transport Equations," Peter Peregrinus Ltd. on behalf of IEE, 1998.

Method of Moments

The Moment Method multiplies both sides the Boltzmann equation (13) by the group velocity (electron or hole), and then integrates over each band in k space. The resulting equations are:

Electrons:
$$\frac{\partial \vec{\mathrm{v}}_e}{\partial t} + \vec{\mathrm{v}}_e\cdot\vec{\nabla}\mathrm{v}_e + \frac{q\vec{E}}{m_e^*} + \frac{\vec{\nabla}(nk_BT_e)}{m_e^* n} = \left.\frac{\partial \vec{\mathrm{v}}_e}{\partial t}\right|_{scatt} \tag{14}$$

Holes:
$$\frac{\partial \vec{\mathrm{v}}_h}{\partial t} + \vec{\mathrm{v}}_h\cdot\vec{\nabla}\mathrm{v}_h + \frac{q\vec{E}}{m_h^*} + \frac{\vec{\nabla}\left(pk_BT_h\right)}{m_h^* p} = \left.\frac{\partial \vec{\mathrm{v}}_h}{\partial t}\right|_{scatt} \tag{15}$$

$T_e \rightarrow$ local temperature, electrons

$T_h \rightarrow$ local temperature, holes

$n \rightarrow$ non-equilibrium concentration, electrons

$p \rightarrow$ non-equilibrium concentration, holes

$m_e^* \rightarrow$ effective mass, electrons (conductivity mass)

$m_h^* \rightarrow$ effective mass, holes (conductivity mass)

Drift Diffusion Approximation

The application of (14) and (15) are fairly straightforward on computers, but not simple by analytic means. So approximations are usually made pertaining to the spatial and time dependence of the temperatures and velocities

1) Uniform temperature: $T_e = T_0,\ T_h = T_0,\ \vec{\nabla} T_e \approx 0,\ \vec{\nabla} T_h \approx 0$

2) Nearly-uniform velocities: $\vec{v}_e \cdot \vec{\nabla} v_e$ and $\vec{v}_h \cdot \vec{\nabla} v_h$ both small

3) Quasi-steady state:. $\frac{\partial \vec{v}_e}{\partial t} \approx \frac{\partial \vec{v}_h}{\partial t} \approx 0$

(4) Relaxation time approximation: $\left.\frac{\partial \vec{v}_e}{\partial t}\right|_{scatt} \approx -\frac{\vec{v}_e}{\tau_e}, \left.\frac{\partial \vec{v}_h}{\partial t}\right|_{scatt} \approx -\frac{\vec{v}_h}{\tau_h}$

These approximations yield

$$\vec{v}_e = \underbrace{-\frac{e\tau_e}{m_e^*}\vec{E}}_{\text{Drift term}} \underbrace{-\frac{k_B T_0 \tau_e}{m_e^* n}\vec{\nabla} n}_{\text{Diffusion term}} \equiv -\mu_e \vec{E} - \frac{k_B T_0 \mu_e}{ne}\vec{\nabla} n \qquad (16)$$

And with the understanding that all the terms in (16) are averaged over the distribution function by the Moment Method, we can write the electrical current directly as

$$\vec{J}_e = -en\vec{v}_e = en\mu_e \vec{E} + k_B T_0 \mu_e \vec{\nabla} n \qquad (17)$$

Similarly we get for the holes:

$$\vec{\mathrm{v}}_h = \frac{e\tau_h}{m_h^*}\vec{E} - \frac{k_B T_o \tau_h}{m_h^* p}\vec{\nabla} p$$

and

$$\vec{J}_h = ep\vec{\mathrm{v}}_h = ep\mu_h \vec{E} - k_B T_0 \mu_h \vec{\nabla} p \tag{18}$$

The key results (17) and (18) can be written

$$\vec{J}_n = eD_n \vec{\nabla} n + e\mu_n n\vec{E} \tag{19}$$

$$\vec{J}_p = -eD_p \vec{\nabla} p + e\mu_p p\vec{E} \tag{20}$$

where

$$D_n = \frac{\mu_n k_B T}{e}; D_p = \frac{\mu_p k_B T}{e}$$

and where we have changed the labeling from "e" and "h" to "n" and "p" for electrons and hole, consistent convention the semiconductor device literature.

For $\mathrm{R_n}$ and $\mathrm{R_p}$, we make the assumption that recombination always occurs in pairs and is exponential (similar in spirit to the relaxation-time approximation), so that

$$\mathrm{R_n} \approx \mathrm{R_p}$$

Substitution of $\vec{J}_n$, $\vec{J}_{p1}$, R_n & R_p into the current continuity relations (11) and (12) yields:

$$\frac{\partial n}{dt} = D_n \vec{\nabla}^2 n + \mu_n \vec{\nabla} \cdot \left(n\vec{E}\right) + G_n - \frac{n}{\tau_n} \tag{21}$$

$$\frac{\partial p}{\partial t} = D_p \vec{\nabla}^2 p - \mu_p \vec{\nabla} \cdot \left(p\vec{E}\right) + G_p - \frac{p}{\tau_p} \tag{22}$$

Eqns (21) and (22) are the famous "drift-diffusion" equations. The are so pervasive in semiconductor device analysis and modeling that are also sometimes called the "master" equations. The can be further simplified with a theorem from vector calculus for the divergence of a product of a scalar and a vector: $\vec{\nabla} \cdot (\phi \vec{A}) = \vec{A} \cdot \vec{\nabla}\phi + \phi\left(\vec{\nabla} \cdot \vec{A}\right)$. The leads to the forms:

$$\frac{\partial n}{\partial t} = D_n \vec{\nabla}^2 n + \mu_n \vec{E} \cdot \vec{\nabla} n + n\mu_n \vec{\nabla} \cdot \vec{\mathrm{E}} + G_n - \frac{n}{\tau_n} \tag{23}$$

$$\frac{\partial p}{\partial t} = D_p \vec{\nabla}^2 p - \mu_p \vec{E} \cdot \vec{\nabla} p - p \mu_p \vec{\nabla} \cdot \vec{E} + G_p - \frac{p}{\tau_p} \tag{24}$$

Even with all the assumptions, (23) and (24) are difficult to solve analytically, but relatively easy to solve numerically by modern finite element techniques. However, analytic results become simple for small deviations from equilibrium – the analog of the "small-signal" approximation in circuit theory: $\delta p = p - p_0$, δp < majority carrier concentration

$\delta n = n - n_0$, δn < majority carrier concentration

We further assume space-charge neutrality and complete ionization of any donors or acceptors (concentrations N_A and N_D, respectively, so that so that

$$n + N_A = p + N_D$$

or

$$n_0 + \delta n + N_A = p_o + \delta p + N_D$$

Now we know that in thermodynamic equilibrium $n_0 + N_A = p_0 + N_D$, which yields the crucial relation $\delta n = \delta p$ (25)

Substitution into the drift-diffusion equations (23) and (24) result in:

$$\frac{\partial (n_0 + \delta n)}{\partial t} = D_n \vec{\nabla}^2 (n_0 + \delta n) + \mu_n \vec{E} \cdot \vec{\nabla} (n_0 + \delta n) + \mu_n (n_0 + \delta n) \vec{\nabla} \cdot \vec{E} + G_n - \frac{n_0 + \delta n}{\tau_n} \tag{26}$$

$$\frac{\partial (p_0 + \delta p)}{\partial t} = D_p \vec{\nabla}^2 (p_0 + \delta p) - \mu_p \vec{E} \cdot \vec{\nabla} (p_0 + \delta p) - (p_0 + \delta p) \mu_p \vec{\nabla} \cdot \vec{E} + G_p - \frac{p_0 + \delta p}{\tau_p} \tag{27}$$

Since n_0 and p_0 are spatial and time independent (by definition), we thus get

$$\frac{\partial \delta n}{\partial t} = D_n \vec{\nabla}^2 \delta n + \mu_n \vec{E} \cdot \vec{\nabla} \delta n + \mu_n (n_0 + \delta n) \vec{\nabla} \cdot \vec{E} + G_n - \frac{n_0 + \delta n}{\tau_n} \tag{28}$$

$$\frac{\partial \delta p}{\partial t} = D_p \vec{\nabla}^2 \delta p - \mu_p \vec{E} \cdot \vec{\nabla} \delta p - \mu_p (p_0 + \delta p) \vec{\nabla} \cdot \vec{E} + G_p - \frac{p_0 + \delta p}{\tau_p} \tag{29}$$

Inspection of (28) and (29) indicate that we can utilize the constraints δp = δn, and p – n = p_0 – n_0 to eliminate δp in favor of δn, or vice versa. We choose the first option here, getting

$$\frac{\partial \delta n}{\partial t} = D_n \vec{\nabla}^2 \delta n + \mu_n \vec{E} \cdot \vec{\nabla} \delta n + \mu_n n \vec{\nabla} \cdot \vec{E} + G_n - \frac{n}{\tau_n} \tag{30}$$

$$\frac{\partial \delta n}{\partial t} = D_p \vec{\nabla}^2 \delta n - \mu_p \vec{E} \cdot \vec{\nabla} \delta n - \mu_p p \vec{\nabla} \cdot \vec{E} + G_p - \frac{p}{\tau_p} \tag{31}$$

We now carry out a step aimed at eliminating the subtle $\vec{\nabla} \cdot \vec{E}$ term – a term that is not necessarily zero in the present analysis even though we started with a uniform field in the Boltzmann equation. We will come back to this term later. To eliminate it for now, we multiply (30) by $\mu_p p$ and (31) by $\mu_n n$, and then add them to get

$$\left(\mu_p p + \mu_n n\right)\frac{\partial \delta n}{\partial t} = \left(\mu_p p D_n + \mu_n n D_p\right)\vec{\nabla}^2 \delta n + \left(\mu_n \mu_p (p_0 - n_0)\right)\vec{E} \cdot \vec{\nabla} \delta n + \mu_p p\left(G_n - \frac{n}{\tau_n}\right) + \mu_n n\left(G_p - \frac{p}{\tau_p}\right) \tag{32}$$

By dividing both sides by ($\mu_n p + \mu_e n$), we finally approach a beautiful result,

$$\frac{\partial \delta n}{\partial t} = \left(\frac{\mu_p p D_n + \mu_n n D_p}{\mu_p p + \mu_n n}\right)\vec{\nabla}^2 \delta n + \frac{\mu_n \mu_p (p_0 - n_0)\vec{E} \cdot \vec{\nabla} \delta n}{(\mu_p p + \mu_n n)} + \frac{\mu_p p(G_n - n/\tau_n)}{(\mu_p p + \mu_n n)} + \frac{\mu_n n(G_p - p/\tau_p)}{(\mu_p p + \mu_n n)} \tag{33}$$

We can eliminate simplify the first term on the right side with Einstein's relation $D = k_B T \mu / e$

So
$$\frac{\partial \delta n}{\delta t} = \frac{D_n D_p (p+n)}{D_p p + D_n n}\vec{\nabla}^2 \delta n + \frac{(p_0 - n_0)\mu_n \mu_p}{(\mu_p p + \mu_n n)}\vec{E} \cdot \vec{\nabla} \delta n + \frac{\mu_p p(G_n - n/\tau_n) + \mu_n n(G_p - p/\tau_p)}{(\mu_p p + \mu_n n)} \tag{34}$$

The term,
$$D^* \equiv \frac{D_n D_p (p+n)}{D_p p + D_n n} \tag{35}$$

is called the ambipolar diffusion coefficient and labeled D*. The term

$$\mu^* \equiv \frac{(p_0 - n_0)\mu_n \mu_p}{(\mu_p p + \mu_n n)} \tag{36}$$

is called the ambipolar mobility. Both of these have very interesting properties related to minority-carrier "bottlenecking." To see this clearly we consider an n-type semiconductor having $n_0 >> p_o$, $D_p \approx D_n$ and $\delta n = \delta p << n_0$. With these assumptions, (35) and (36) become

$$D^* = \frac{D_n D_p (p+n)}{D_p p + D_n n} \approx \frac{D_n D_p n}{D_n n} \approx D_p \quad \text{and} \quad \mu^* = \frac{(p_0 - n_0)\mu_n \mu_p}{(\mu_p p + \mu_n n)} \approx \frac{-n_0 \mu_n \mu_p}{\mu_n n} \approx \frac{-n_0 \mu_p}{n} \approx -\mu_p$$

This elucidates the following key point about the drift-diffusion behavior in two-band (i.e., bipolar) semiconductors:

The overall drift and diffusion processes are limited by the minority-carrier diffusion and drift coefficients. This can be linked back to the requirement that both carrier types maintain space-charge neutrality. Since the minority carriers are so much less dense, their drift and diffusion currents are much smaller and thus act as the bottleneck in the neutralization maintenance.