Basic Semiconductor Devices

BJT, MosFET, LED, Laser and Solar Cells and photo-detectors.

Fig 6.1

Basic BJT structure is a three-terminal device used for switching and amplification.

Two flavors npn and pnp – both back to back diodes with depletion regions.
Band Diagrams (In equilibrium)

- In equilibrium $E_F$ must be flat – this creates depletion regions and barriers.
- For normal operation
  - EBJ forward biased
    - Barrier reduced and so electrons diffuse into the base
    - Electrons get swept across the base into the collector
  - CBJ reverse biased
    - Electrons roll down the hill (high E-field)
Minority Carrier Concentration Profiles

• Current dominated by injection of electrons from emitter to base (by design)
  – Back injection from base to emitter is small as we use a N^+PN diode

• The forward bias EBJ and diffusion of the minority carrier concentration causes a gradient (diffusion) through the base:
  – some recombination causes bowing of electron concentration (in the base)
  – base is designed to be fairly short (minimize recombination)

• Drift currents are usually small and neglected
• Diffusion of electrons through the base is set by concentration profile at the EBJ

• “Law of the junction” sets the concentrations of the minority carriers at edges of the base region

\[ n_p(0) = n_{p0} e^{v_{BE}/V_T} \]

• Diffusion current of electrons through the base is (assuming an ideal straight line case):

\[ I_n = A_E q D_n \frac{dn_p(x)}{dx} = A_E q D_n \left( -\frac{n_p(0)}{W} \right) \]

• Due to back injection and recombination in the base, the current at the EBJ and current at the CBJ are not equal and differ by a base current
Collector Current

- Electrons that diffuse across the base to the CBJ junction are swept across the CBJ depletion region to the collector b/c of the higher potential applied to the collector.

\[ i_C = I_s e^{v_{BE}/V_T} \] where the saturation current is \( I_S = qA_E D_n n_p / W \)

and we can rewrite the saturation current as:

\[ I_S = \frac{qA_E D_n n_i^2}{N_A W} \]

- Note that \( i_C \) is independent of \( v_{CB} \) (potential bias across CBJ) ideally
- Saturation current is
  - inversely proportional to \( W \) and directly proportional to \( A_E \)
    - Want short base and large emitter area for high currents
  - dependent on temperature due to \( n_i^2 \) term
Base Current

• Base current $i_B$ composed of two components:
  – holes back injected from the base region into the emitter region
    $$i_{B1} = \frac{q A_E D_P n_i^2}{N_D L_P} e^{v_{BE}/V_T}$$
  – holes supplied due to recombination in the base with diffusing electrons and depends on minority carrier lifetime $\tau_b$ in the base
    $$i_{B2} = \frac{Q_n}{\tau_b}$$
    And the Q in the base is
    $$Q_n = \frac{q A_E W n_i^2}{N_A} e^{v_{BE}/V_T}$$
    So, current is
    $$i_{B2} = \frac{q A_E W n_i^2}{N_A \tau_b} e^{v_{BE}/V_T}$$

• Total base current is
    $$i_B = \left( \frac{q A_E D_P n_i^2}{N_D L_P} + \frac{q A_E W n_i^2}{N_A \tau_b} \right) e^{v_{BE}/V_T}$$
Beta and Emitter Current

- Can relate $i_B$ and $i_C$ by the following equation
  
  
  \[
  i_B = \frac{i_C}{\beta} = \frac{I_S}{\beta} e^{v_{BE}/V_T}
  \]

  and \(\beta\) is
  
  - Beta is constant for a particular transistor
  - On the order of 100-200 in modern devices (but can be higher)
  - Called the **common-emitter current gain**

- For high current gain, want small \(W\), low \(N_A\), high \(N_D\)

- Emitter current is the sum of \(i_C\) and \(i_B\alpha\)
  - is called the **common-base current gain**

\[
\beta = \frac{1}{\frac{D_p}{D_n} \frac{N_A}{N_D} \frac{W}{L_p} + \frac{1}{2} \frac{W^2}{D_n \tau_b}}
\]

\[
i_E = i_C + i_B
\]

\[
i_E = \frac{\beta + 1}{\beta} i_C
\]

\[
i_C = \alpha i_E \text{ where } \alpha = \frac{\beta}{\beta + 1}
\]
Neutral region resistances

- $r_C$ decrease the slope of the curve in the saturation region.
- In active mode, the current is dependent on $I_B$.
- $r_C$ can limit the current-handling capability and $f_{\text{max}}$.

The main effect of $r_E$ is a reduction of the $V_{BE}$ by a factor of $I_E r_E$.

$r_E$ is equivalent to a base resistance of $(1 + \beta_F) r_E$. Therefore it affects the collector current as well.

$r_E$ also complicates the determination of $r_B$, and $I_C$ if $r_C$ is low.

Base Resistance, $r_B$

In the Ebers-Moll model (SPICE’s default model for BJT), $r_B$ is assumed to be constant.

$$r_{be} = \frac{v_{be}}{i_b} = \frac{\delta V_{BE}}{\delta I_B} \approx \beta \frac{\delta V_{BE}}{\delta I_E} \approx \frac{\beta 25}{I_C} \text{(mA)}$$
Early Effect

\[ I_C = I_S e^{V_{BE}/V_T} \left( 1 + \frac{V_{CE}}{V_A} \right) \]

- Early Effect
  - Current in active region depends (slightly) on \( V_{CE} \)
  - \( V_A \) is a parameter for the BJT (50 to 100) and called the Early voltage
  - Due to a decrease in effective base width \( W \) as reverse bias increases
  - Account for Early effect with additional term in collector current equation
  - Nonzero slope means the output resistance is NOT infinite, but…
    - \( I_C \) is collector current at the boundary of active region
Early Effect

- What causes the Early Effect?
  - Increasing $V_{CB}$ causes depletion region of CBJ to grow and so the effective base width decreases (base-width modulation)
  - Shorter effective base width $\rightarrow$ higher $\frac{dn}{dx}$

$$I_n = A_E q D_n \frac{dn_p(x)}{dx} = A_E q D_n \left( -\frac{n_p(0)}{W} \right)$$
(a) An npn transistor operated in the active region in the common emitter (CE) configuration. The dc voltage across the BE junction, $V_{BE}$, controls the current $I_E$ and hence $I_B$ and $I_C$. The input current is the current that flows between $V_{BE}$ and the base which is $I_B$. The output current is the current flowing between $V_{CE}$ and the collector which is $I_C$.

(b) DC $I$-$V$ characteristics of the npn bipolar transistor in the CE configuration (exaggerated to highlight various effects).
Ebors Moll Model

\[ i_E = \frac{l_S}{\alpha_F} \left[ \exp\left(\frac{v_{BE}}{V_T}\right) - 1 \right] - l_S \left[ \exp\left(\frac{v_{BC}}{V_T}\right) - 1 \right] \]

\[ i_C = l_S \left[ \exp\left(\frac{v_{BE}}{V_T}\right) - 1 \right] - \frac{l_S}{\alpha_R} \left[ \exp\left(\frac{v_{BC}}{V_T}\right) - 1 \right] \]

\[ i_B = \frac{l_S}{\beta_F} \left[ \exp\left(\frac{v_{BE}}{V_T}\right) - 1 \right] + \frac{l_S}{\beta_R} \left[ \exp\left(\frac{v_{BC}}{V_T}\right) - 1 \right] \]

\[ \beta_R = \frac{\alpha_R}{1 - \alpha_R} \quad \beta_F = \frac{\alpha_F}{1 - \alpha_F} \]

Need to add depletion and diffusion capacitances and worry about breakdown.

Fig 6.26

MOSFET

As with BJT three terminal device for switching and amplification and two flavours (n-channel and p-channel).

Drift based device!

Fig 6.36

The field effect. (a) In a metal-air-metal capacitor, all the charges reside on the surface. (b) Illustration of field penetration into a $p$-type semiconductor. (c) As the field increases eventually when $V > V_{th}$ an inversion layer is created near the surface in which there are conduction electrons.

**Fig 6.35**

Basic Operation of MOSFET

“Off” – no channel

(a) Below threshold $V_{GS} < V_{th}$ and $V_{DS} > 0$

“Triode Region” – simple channel voltage controlled resistor

(b) Above threshold $V_{GS} > V_{th}$ and $V_{DS} < V_{DS(sat)}$

The MOSFET $I_D$ vs. $V_{DS}$ characteristics

Fig 6.37

"Pinch-off" – channel is reduced to zero at drain

(c) Above threshold $V_{GS} > V_{th}$ and saturation, $V_{DS} = V_{DS(sat)}$

"Saturation" – pinch-off limits $I_D$ to constant value

(d) Above threshold $V_{GS} > V_{th}$ and saturation region, $V_{DS} > V_{DS(sat)}$

The MOSFET $I_D$ vs. $V_{DS}$ characteristics

• Basic Mosfet model is voltage controlled current source plus:
  - Two diodes S->B and D->B
  - Basic gate capacitances.
  - Parasitic capacitances
• More on this later.
LEDs and Semiconductor lasers

To create light source using a semiconductor we can use the simple PN diode in a direct semiconductor

Forward biasing the diode creates a inversion of the “normal” carrier densities in the C and V bands. These carriers then recombine producing photons.

(a) The energy band diagram of a $p$-$n^+$ (heavily $n$-type doped) junction without any bias. Built-in potential $V_0$ prevents electrons from diffusing from $n^+$ to $p$ side.

(b) The applied bias reduces $V_0$ and thereby allows electrons to diffuse, be injected, into the $p$-side. Recombination around the junction and within the diffusion length of the electrons in the $p$-side leads to photon emission.

Fig 6.43

A schematic illustration of one possible LED device structure. First $n^+$ is epitaxially grown on a substrate. A thin $p$ layer is then epitaxially grown on the first layer.

An LED is a PN diode optimized to produce and emit light.
(a) Photon emission in a direct bandgap semiconductor.
(b) GaP is an indirect bandgap semiconductor. When doped with nitrogen there is an electron recombination center at $E_N$. Direct recombination between a captures electron at $E_N$ and a hole emits a photon.

Band structure (material) must be chosen to produce the light of the right wavelength (color)

Need direct band gap.

Traps can be used to produce direct transitions.

Fig 6.45

We can use band gap engineering to create hetero-junctions with optimized performance.

They can confine and “hold” the carriers in the active (depletion) region.

This produces better efficiencies due to lower ohmic losses.

(a) A double heterostructure diode has two junctions which are between two different bandgap semiconductors (GaAs and AlGaAs). (b) A simplified energy band diagram with exaggerated features. EF must be uniform. (c) Forward biased simplified energy band diagram. (d) Forward biased LED. Schematic illustration of photons escaping reabsorption in the AlGaAs layer and being emitted from the device.

Fig 6.46

Due to $3/2kT$ spread in electron/hole energies there are variety of transitions that can occur.

The light is therefore not absolutely monochromatic but has a band-width.

IV characteristic is just the ideal diode equation.

Light production is linearly related to $I$.

Fig 6.47

The LED structure creates an inversion of the carriers which will provide optical gain from stimulated emission.

All we need to do is add optical feedback in the form of mirrors to create a semiconductor laser.

Semiconductor lasers have an optical cavity to build up the required electromagnetic oscillations. In this example, one end of the cavity has a Bragg distributed reflector, a reflection grating, that reflects only certain wavelengths back into the cavity.

Fig 6.64

Semiconductor lasers come in many geometries:

Early lasers were slab lasers made from a single chip

Ridge lasers are formed using a ridge optical waveguide integrated with PN diode

Vertical Cavity Lasers (VCSELS) are perpendicular to the surface and typically very short and formed using deposited layers.
(a) A double heterostructure diode has two junctions which are between two different bandgap semiconductors (GaAs and AlGaAs).

(b) Simplified energy band diagram under a large forward bias. Lasing recombination takes place in the p-GaAs layer, the active layer.

(c) The density of states and energy distribution of electrons and holes in the conduction and valence bands in the active layer.

As with LEDs to make efficient lasers we use band gap engineering to create well defined waveguides and active regions where the carriers are concentrated and we have a strong inversion.

---

Laser Characteristics

Lasers exhibit a threshold current below which they do not lase.

They also have a much narrower band width due to the presence of a optical mode structure.

(a) Typical optical power output vs. forward current for a laser diode and an LED.
(b) Comparison of spectral output characteristics.

Fig 6.65

Many types of lasers: slab, ridge, VCSEL
- Complicated geometry
- Complicated band-structures
- Complicated operation
Laser Equations

- EM solution to provide modes
- Drift-diffusion equation for carriers
- Diffusion equation for heat
- QM for gain expressions and rate equations.

\[ \nabla \cdot \mathbf{D} = \rho \]
\[ \nabla \cdot \mathbf{B} = 0 \]
\[ \nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \]
\[ \nabla \times \mathbf{H} = \mathbf{J} + \frac{\partial \mathbf{D}}{\partial t} \]

\[ \Delta \varphi = -(q/\varepsilon_0)(p - n + N_d - N_a) \]
\[ \frac{\nabla \cdot \mathbf{J}_n}{q} = U, \quad \mathbf{J}_n = q(-\mu_n n \nabla \varphi + D_n \nabla n) \]
\[ \frac{\nabla \cdot \mathbf{J}_p}{q} = -U, \quad \mathbf{J}_p = q(-\mu_p p \nabla \varphi - D_p \nabla p) \]

\[ \kappa \nabla^2 T = \rho(\hat{r}) \]

Coupling of the Equations

- Optical propagation equations are primarily determined by $n_{\text{eff}}(x,y,z)$ (index of refraction) and $G(x,y,z)$ (optical gain fct). Both of these are a function of $N(x,y,z)$ (electron density).
- $N(x,y,z)$ is fct of $T$, $J$ (current density) and photon density.
- Temperature equation is driven by heat generation which is fct of $J$.
- $J$ is fct of applied potential and temperature.
- Rather complicated!
- Approaches usually make approximations and simplifications.

Detailed Laser Modeling

• Detailed device modeling usually attempts to solve a particular aspect:
  – Optical modes
  – Electron transport and photon generation
  – Thermal effects
• Example is optical modes for different geometries.

In the present paper we investigate optical near fields in semiconductor lasers. We perform finite element simulations for two different laser types, namely a super large optical waveguide (SLOW) laser, which is an edge emitter, and a vertical cavity surface emitting laser (VCSEL). We give the mathematical formulation of the different eigenvalue problems that arise for our examples and explain their numerical solution with the finite element method (FEM). Thereby, we also comment on the usage of transparent boundary conditions, which have to be applied to respect the exterior environment, e.g., the very large substrate and surrounding air. For the SLOW laser we compare the computed near fields to experimental data for different design parameters of the device. For the VCSEL example a comparison to simplified 1D mode calculations is carried out.
Finite element simulation of the optical modes of semiconductor lasers

Figure 6  Finite element triangulation of cross-section of SLOW laser. The meaning of the parameters w and d is also indicated.

Figure 10 Intensity of fundamental lasing mode for varying w1/42.2, 9.0, 15.0 mm. Gold trenches and active layer are shown in green and red, respectively.
Coupled models: with approximations

- Model makes approximations but does couple 3 equations:
  - 1D/2D/3D thermal
  - 1D optical (simple gain no mode)
  - 1D distributed electrical
Equations

- Optical intensity ($I$) equation is 1D. With 2 equation for forward and backwards

- Electron density equation ($n$) includes diffusion (1D).

- Gain expression

- Thermal equation is 3D
- Thermal heating includes ohmic generation and optical heating at the facet.
2D and 3D temperature distributions

1D and 2D heat flow models are bad. Generally heat flow is 3D processes and it is difficult to approximate the geometry.
Plots along the laser axis

![Graphs showing total photon density, temperature, and carrier density along the laser axis.](image-url)
Rate Equations

- The simplest effective model is a three rate equation model.
- Three variables are:
  - N – Electron density
  - P – Photon density
  - \( \Phi \) – Optical phase
- Driving by diode current.
- Three coupled non-linear 1\textsuperscript{st} order differential equations.
- Very complicated behaviour
- Characterized by photon and electron lifetimes, optical gain and offset current

\[
\frac{dN}{dt} = \frac{\eta (I_d - I_{off})}{qV_i} - \frac{N}{\tau_n} - \frac{G_0(N - N_0)P_i}{1 + \epsilon P}
\]

\[
\frac{dP}{dt} = \frac{\Gamma G_0(N - N_0)P}{1 + \epsilon P} - \frac{P}{\tau_p} + \frac{\Gamma \beta N}{\tau_n}
\]

\[
\frac{d\Phi}{dt} = \frac{1}{2} \alpha (\Gamma G_0(N - N_0) - \frac{1}{\tau_p})
\]
Laser Compact Modeling

Add electrical diode, temperature sub-circuit. More later when we look at optical spice models.
The principle of operation of the solar cell (exaggerated features to highlight principles)

Solar Cells and Photodiodes

- Optical absorption of photons can be used either to:
  - Generate electrical power (solar cell)
  - Sense optical signal (photo-detector)
- Same device different optimization of circuit and structure.
- Generate hole/electron pairs and separate them using E field in depletion region.

IV curve

- $P_0$ is the dark current (light off)
- $P_1$ and $P_2$ are with two different light levels.
- The Photo-voltaic region is used for solar cell operation
- The Photo-conductive region for photo detection of optical signals.

Fig 6.53

Series and shunt resistances and various fates of photo-generated EHPs.

- Basic model is current $I_{ph}$ linearly related to photon flux.
- Solar Cell model has additional resistors to capture resistive losses.
- Photodiode also would need to capture high frequency effects:
  - Parasitic Ls and Cs
  - Intrinsic transient response of $I_{ph}$ – not easy to model
A heterojunction solar cell between two different bandgap semiconductors (GaAs and AlGaAs)

A tandem cell. Cell 1 has a wider bandgap and absorbs energetic photons with \( h\omega > E_{g1} \). Cell 2 absorbs photons that pass cell 1 and have \( h\omega > E_{g2} \).
Photodetector structure can be optimized by use of intrinsic layer to maximize gain.

\[ C_{\text{dep}} = \frac{\varepsilon_o \varepsilon_r A}{W} \]

Can also take advantage of avalanche behaviour.

\[ E = E_o + \frac{V_r}{W} \approx \frac{V_r}{W} \]

Trade of signal/noise and speed.

(a) The schematic structure of an idealized pin Photodiode
(b) The net space charge density across the photodiode.
(c) The built-in field across the diode.
(d) The pin photodiode in photodetection is reverse biased.

Fig 6.62

Summary

Fig 6.62