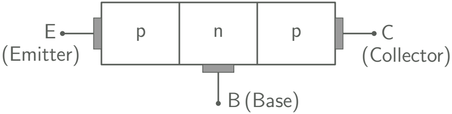
SEMICONDUCTOR DEVICES

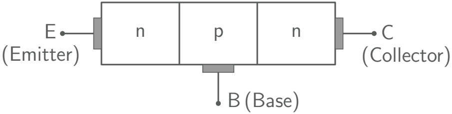
Bipolar Junction Transistors: Part 1

M. B. Patil
mbpatil@ee.iitb.ac.in
www.ee.iitb.ac.in/~sequel

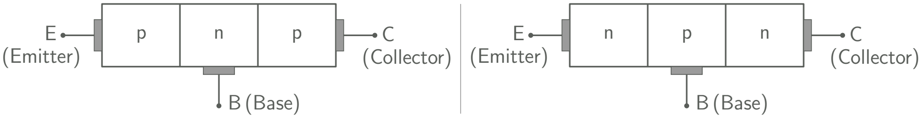
Department of Electrical Engineering
Indian Institute of Technology Bombay

Bipolar junction transistors



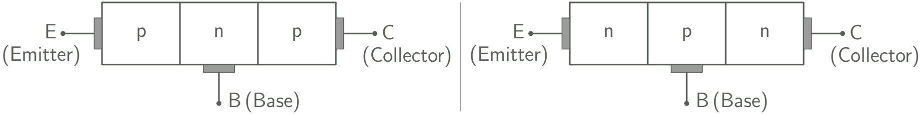


Bipolar junction transistors



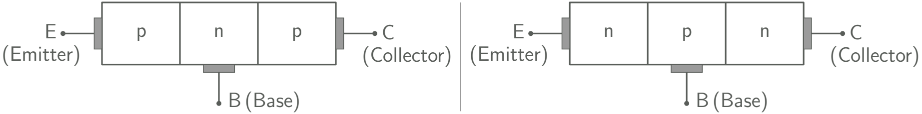
- A bipolar junction transistor (BJT) consists of two *pn* junction diodes connected back to back.

Bipolar junction transistors



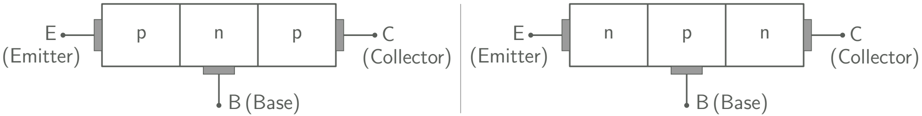
- A bipolar junction transistor (BJT) consists of two *pn* junction diodes connected back to back.
- A BJT is a three-terminal device; the terminals are called emitter, base, and collector.

Bipolar junction transistors



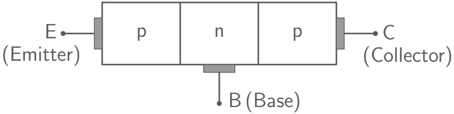
- ∗ A bipolar junction transistor (BJT) consists of two *pn* junction diodes connected back to back.
- ∗ A BJT is a three-terminal device; the terminals are called emitter, base, and collector.
- ∗ There are two types of BJTs: *pnp* and *npn*.

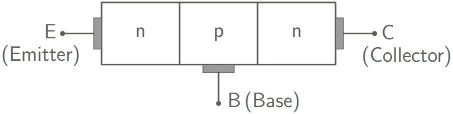
Bipolar junction transistors



- A bipolar junction transistor (BJT) consists of two *pn* junction diodes connected back to back.
- A BJT is a three-terminal device; the terminals are called emitter, base, and collector.
- There are two types of BJTs: *pnp* and *npn*.
- The actual device construction is different than the above schematic diagram (to be discussed).

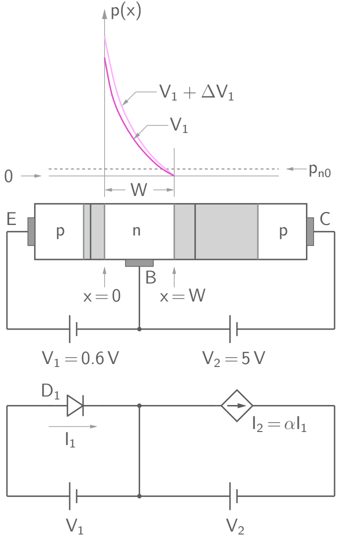
Bipolar junction transistors



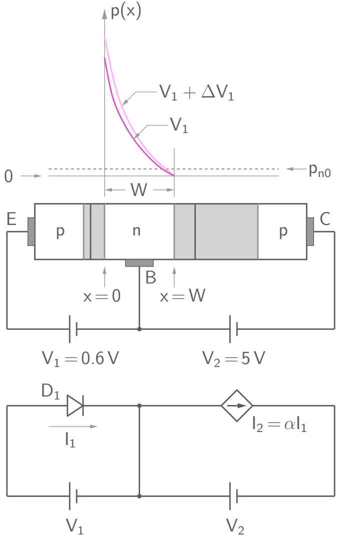


- A bipolar junction transistor (BJT) consists of two *pn* junction diodes connected back to back.
- A BJT is a three-terminal device; the terminals are called emitter, base, and collector.
- There are two types of BJTs: *pnp* and *npn*.
- The actual device construction is different than the above schematic diagram (to be discussed).
- For the device to work as a transistor (rather than two independent diodes), the two junctions must be "close."

Basic operation

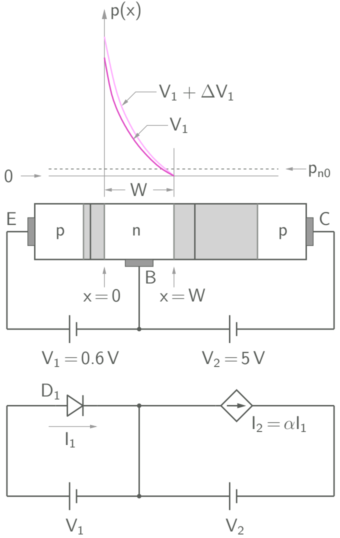


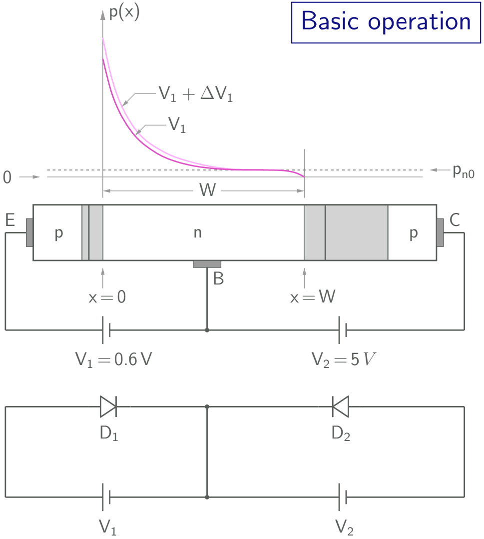
Basic operation



- If V_1 is varied, $p(x)$ varies $\rightarrow I_p(W) \propto \dfrac{dp}{dx}(W)$ varies, i.e., by changing V_1, I_2 can be controlled. This is the basic transistor action.

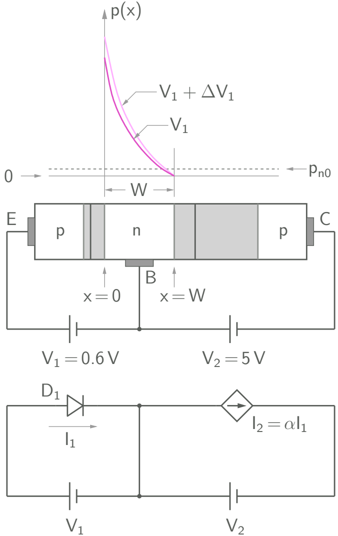
Basic operation

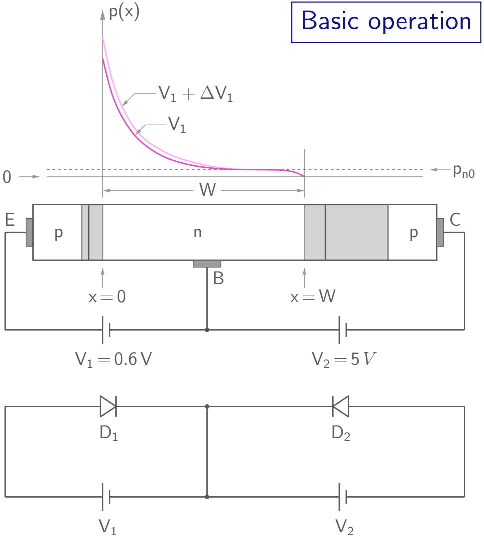




- If V_1 is varied, $p(x)$ varies → $I_p(W) \propto \dfrac{dp}{dx}(W)$ varies, i.e., by changing V_1, I_2 can be controlled. This is the basic transistor action.

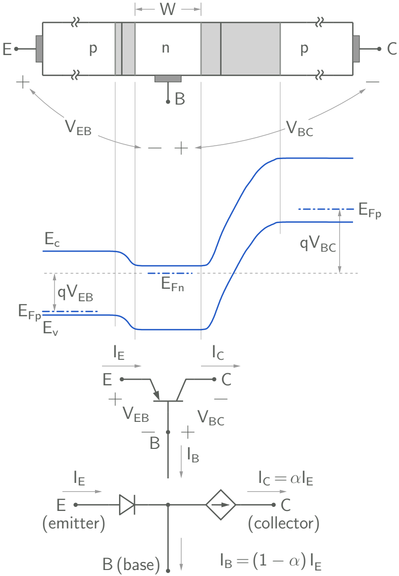
Basic operation

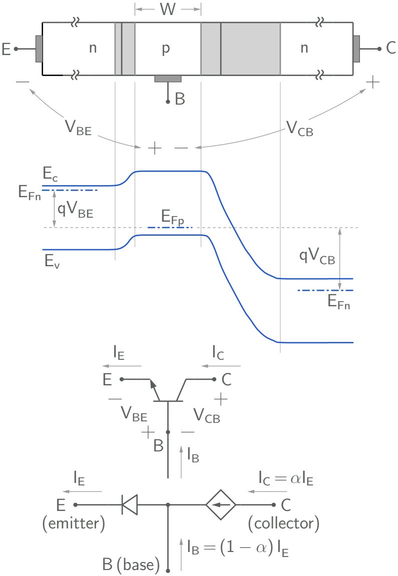




- If V_1 is varied, $p(x)$ varies → $I_p(W) \propto \dfrac{dp}{dx}(W)$ varies, i.e., by changing V_1, I_2 can be controlled. This is the basic transistor action.
- If the two junctions are not sufficiently close, the device behaves like two *independent* diodes connected back-to-back, and there is no transistor action.

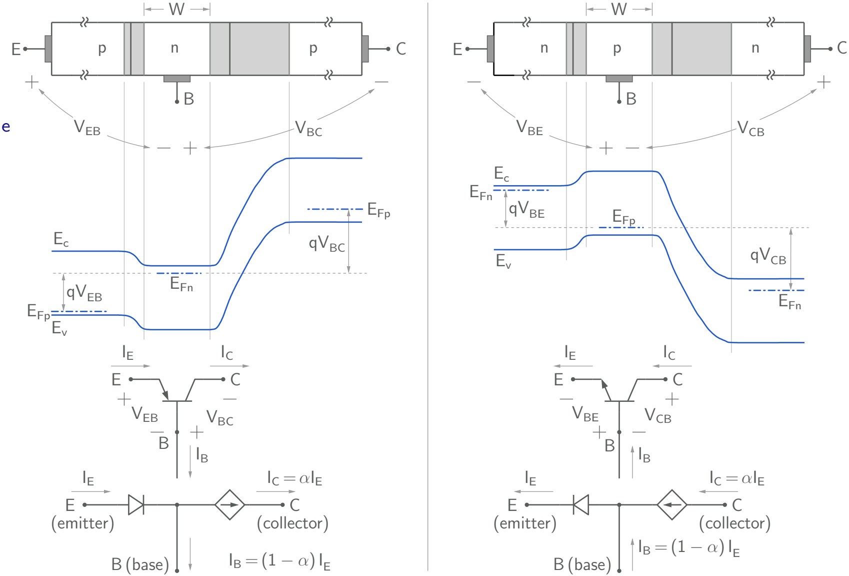
pnp and *npn* transistors





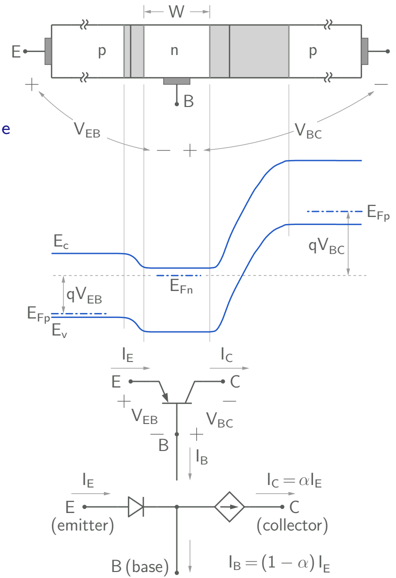
pnp and *npn* transistors

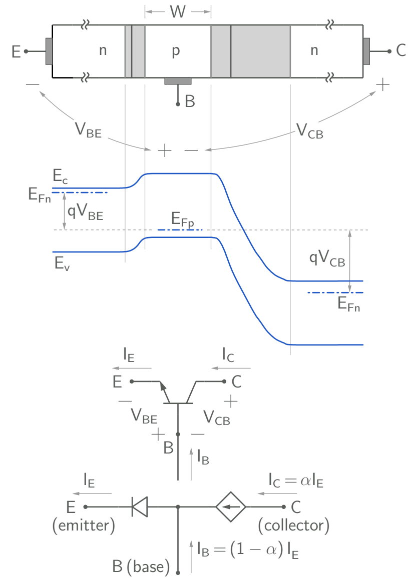
- ∗ In the "active" or "linear" mode, the B-E junction is under forward bias, the B-C junction is under reverse bias.



pnp and *npn* transistors

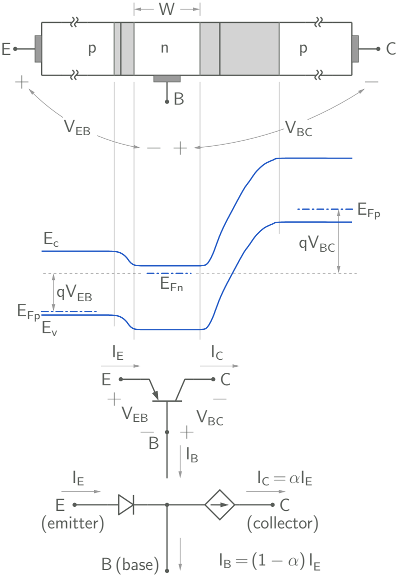
- In the "active" or "linear" mode, the B-E junction is under forward bias, the B-C junction is under reverse bias.
- The B-E voltage (magnitude) is restricted to about 0.8 V in a low-power silicon BJT, as in a forward-biased diode.

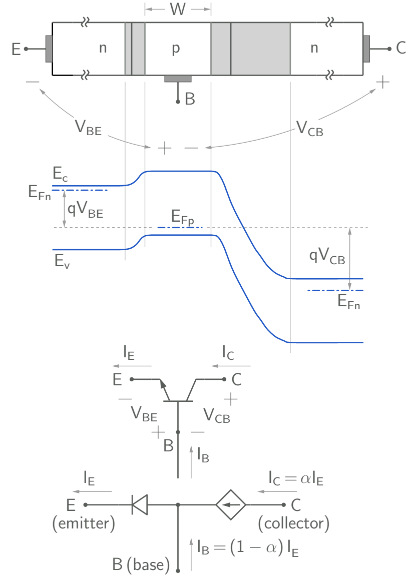




pnp and *npn* transistors

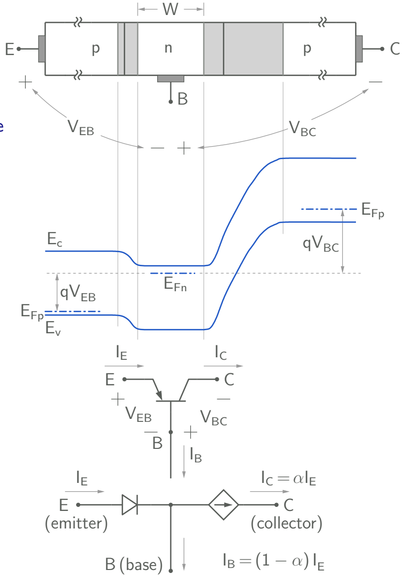
- In the “active” or “linear” mode, the B-E junction is under forward bias, the B-C junction is under reverse bias.
- The B-E voltage (magnitude) is restricted to about 0.8 V in a low-power silicon BJT, as in a forward-biased diode.
- The B-C voltage (magnitude) can be much larger than the B-E voltage and is limited by breakdown phenomena.

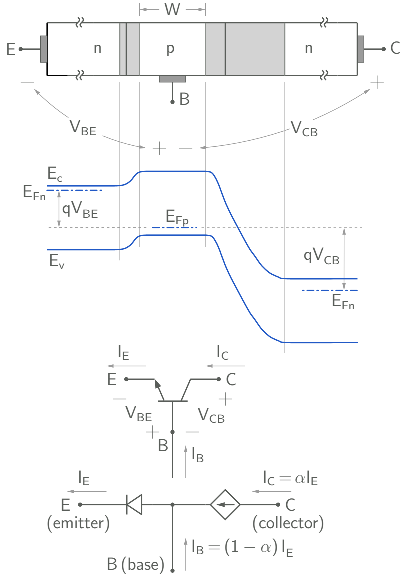




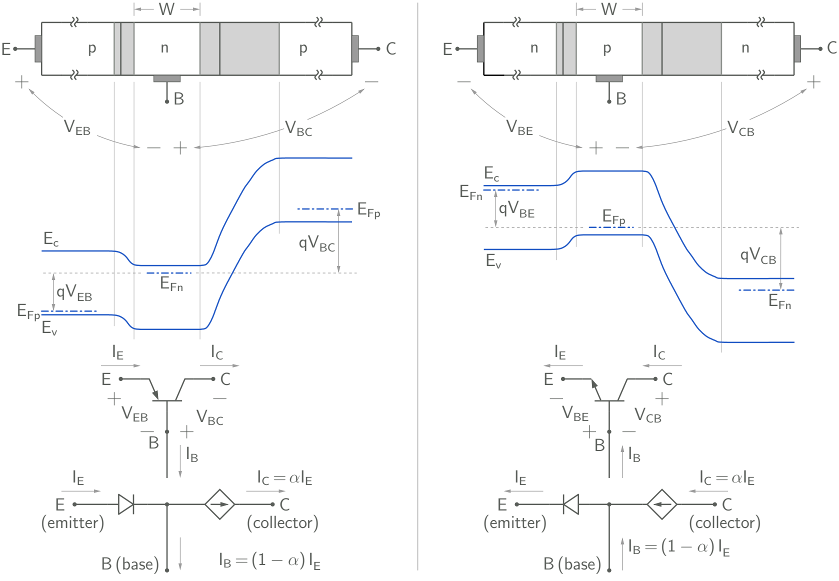
pnp and *npn* transistors

- In the "active" or "linear" mode, the B-E junction is under forward bias, the B-C junction is under reverse bias.
- The B-E voltage (magnitude) is restricted to about 0.8 V in a low-power silicon BJT, as in a forward-biased diode.
- The B-C voltage (magnitude) can be much larger than the B-E voltage and is limited by breakdown phenomena.
- *pnp* transistor: Holes are injected from the emitter. Most of them reach the B-C depletion layer, get swept away by the field there, and get collected by the collector.



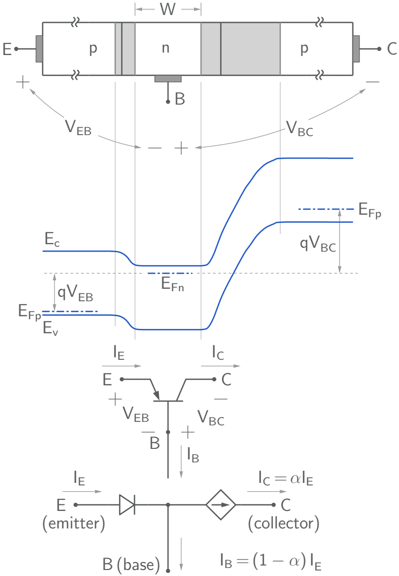


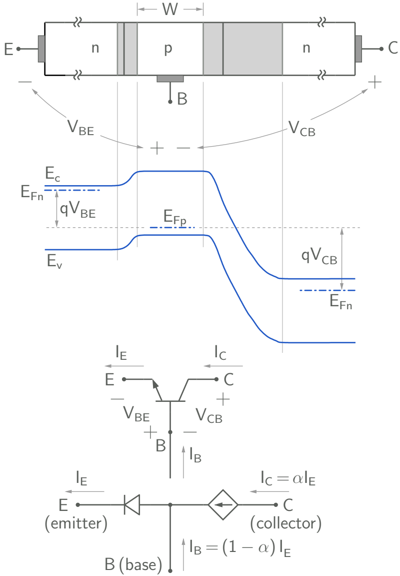
pnp and *npn* transistors



- In the "active" or "linear" mode, the B-E junction is under forward bias, the B-C junction is under reverse bias.
- The B-E voltage (magnitude) is restricted to about 0.8 V in a low-power silicon BJT, as in a forward-biased diode.
- The B-C voltage (magnitude) can be much larger than the B-E voltage and is limited by breakdown phenomena.
- *pnp* transistor: Holes are injected from the emitter. Most of them reach the B-C depletion layer, get swept away by the field there, and get collected by the collector.
- *npn* transistor: Electrons are injected from the emitter. Most of them reach the B-C depletion layer, get swept away by the field there, and get collected by the collector.

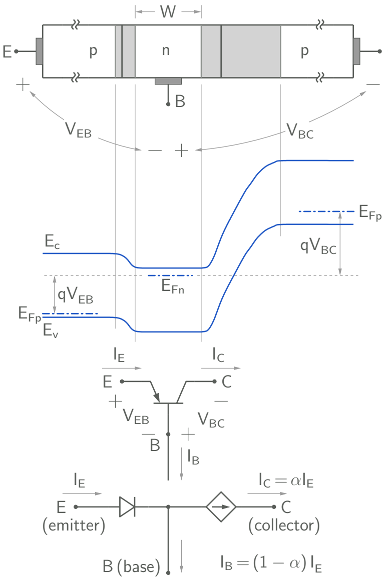
pnp and *npn* transistors

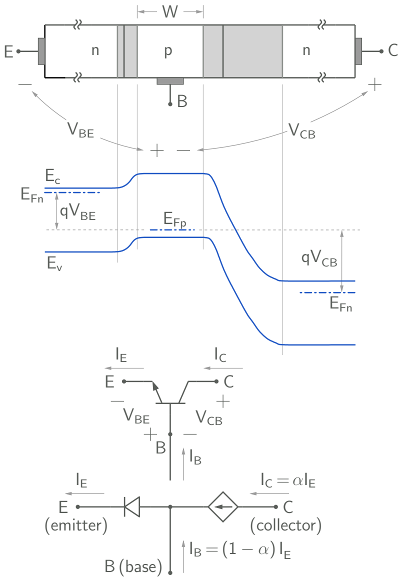




pnp and *npn* transistors

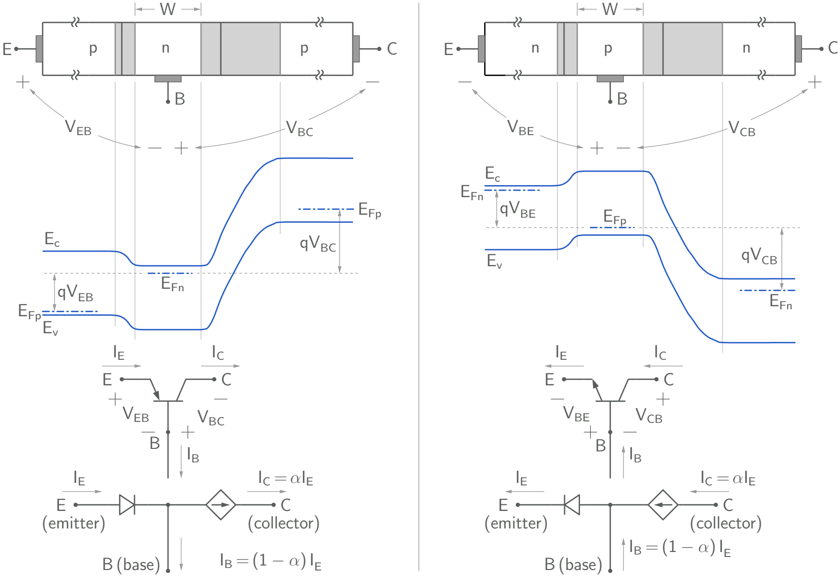
- ∗ The emitter arrow in the BJT symbol indicates the direction of the emitter current when the BJT is operating in the active mode (for both *pnp* and *npn* BJTs).





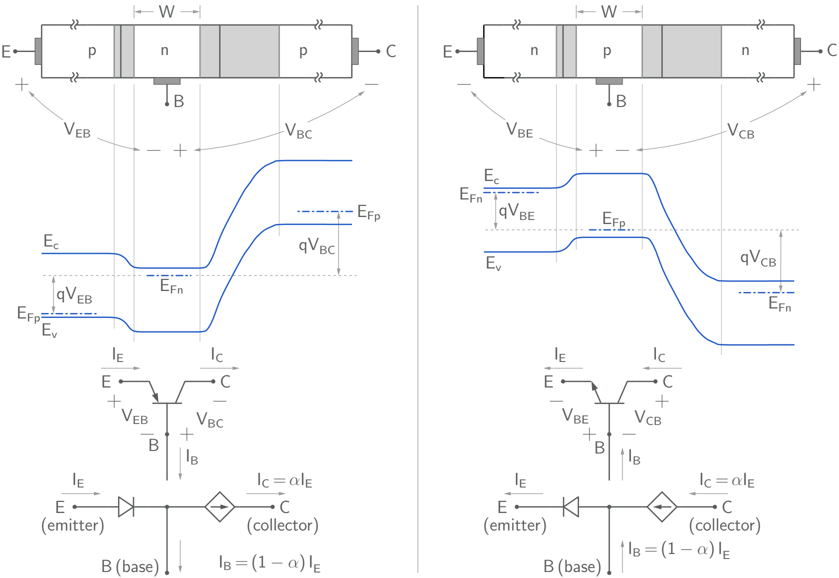
pnp and *npn* transistors

- The emitter arrow in the BJT symbol indicates the direction of the emitter current when the BJT is operating in the active mode (for both *pnp* and *npn* BJTs).
- The collector current I_C is a fraction of the emitter current: $I_C = \alpha I_E$. For a good transistor, $\alpha \approx 1$.

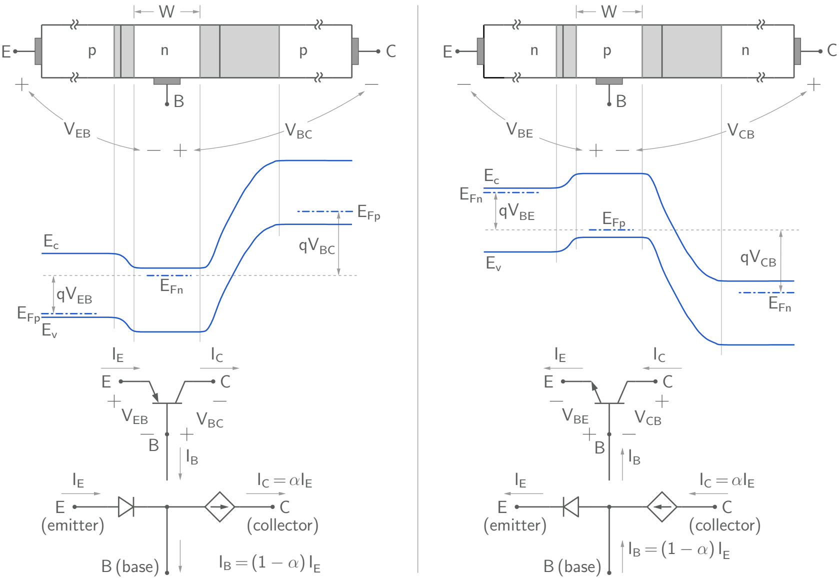


pnp and *npn* transistors

- The emitter arrow in the BJT symbol indicates the direction of the emitter current when the BJT is operating in the active mode (for both *pnp* and *npn* BJTs).
- The collector current I_C is a fraction of the emitter current: $I_C = \alpha I_E$. For a good transistor, $\alpha \approx 1$.
- The three currents satisfy KCL, i.e., $I_E = I_C + I_B$. Substituting for I_C, we get $I_B = (1 - \alpha)\, I_E$.

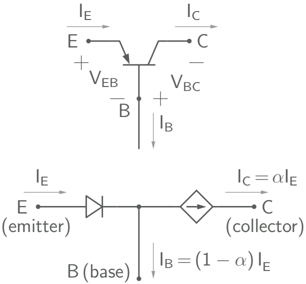


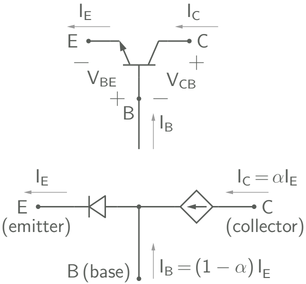
pnp and *npn* transistors



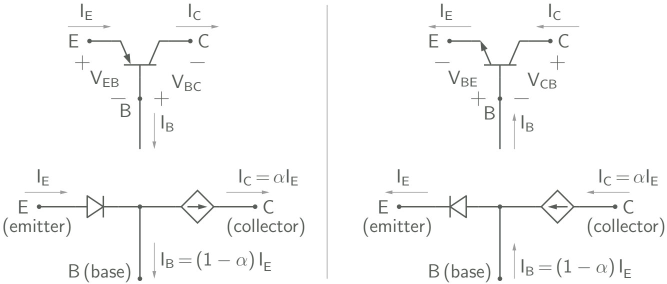
- ∗ The emitter arrow in the BJT symbol indicates the direction of the emitter current when the BJT is operating in the active mode (for both *pnp* and *npn* BJTs).
- ∗ The collector current I_C is a fraction of the emitter current: $I_C = \alpha I_E$. For a good transistor, $\alpha \approx 1$.
- ∗ The three currents satisfy KCL, i.e., $I_E = I_C + I_B$. Substituting for I_C, we get $I_B = (1 - \alpha)\, I_E$.
- ∗ Note that the diagrams are not drawn to scale. In reality, for a silicon BJT, the depletion regions would be narrower.

BJT in active mode



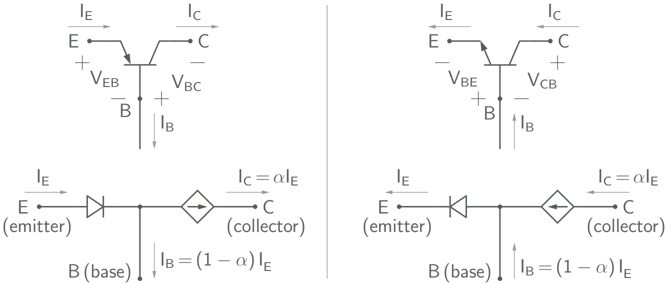


BJT in active mode



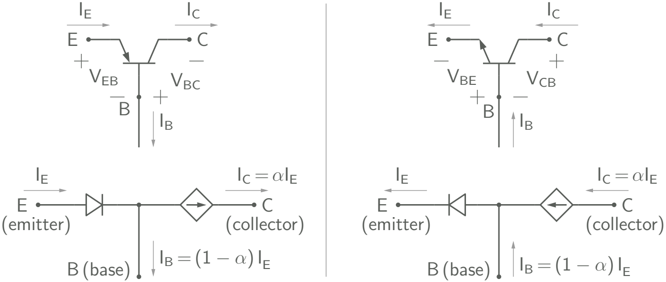
* The “common-emitter current gain” β – a figure of merit of a BJT – is defined as $\beta = I_C/I_B$.

BJT in active mode



- The "common-emitter current gain" β – a figure of merit of a BJT – is defined as $\beta = I_C/I_B$.
- $\beta = \dfrac{I_C}{I_B} = \dfrac{\alpha I_E}{(1-\alpha) I_E} = \dfrac{\alpha}{1-\alpha}$.

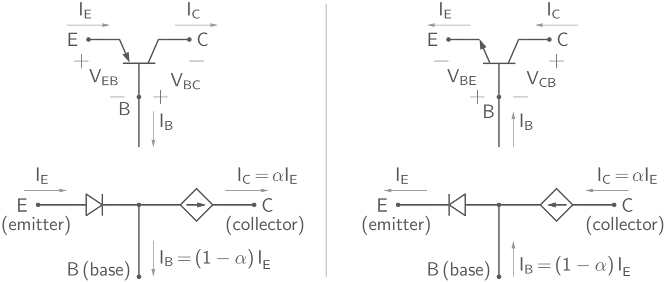
BJT in active mode



α	$1-\alpha$	$\beta=\alpha/(1-\alpha)$
0.9	0.1	9
0.95	0.05	19
0.99	0.01	99
0.995	0.005	199

- The “common-emitter current gain” β – a figure of merit of a BJT – is defined as $\beta = I_C/I_B$.
- $\beta = \dfrac{I_C}{I_B} = \dfrac{\alpha I_E}{(1-\alpha)I_E} = \dfrac{\alpha}{1-\alpha}$.

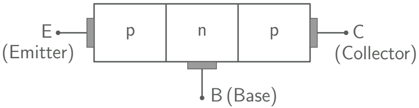
BJT in active mode

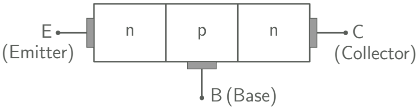


α	$1-\alpha$	$\beta = \alpha/(1-\alpha)$
0.9	0.1	9
0.95	0.05	19
0.99	0.01	99
0.995	0.005	199

- The "common-emitter current gain" β – a figure of merit of a BJT – is defined as $\beta = I_C/I_B$.
- $\beta = \dfrac{I_C}{I_B} = \dfrac{\alpha I_E}{(1-\alpha)I_E} = \dfrac{\alpha}{1-\alpha}$.
- For a typical discrete low-power transistor such as BC107A, β is in the range of 100 to 200.

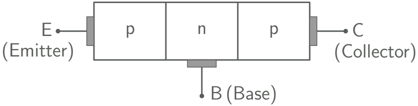
BJT modes of operation

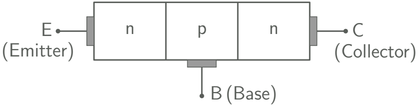




Mode	*B*-*E* junction	*B*-*C* junction
Active (linear)	forward	reverse
Cutoff	reverse	reverse
Saturation	forward	forward
Inverse active	reverse	forward

BJT modes of operation

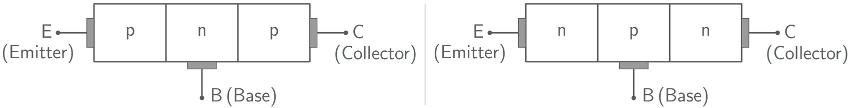




Mode	*B*-*E* junction	*B*-*C* junction
Active (linear)	forward	reverse
Cutoff	reverse	reverse
Saturation	forward	forward
Inverse active	reverse	forward

- In analog circuits, BJTs are generally biased to operate in the active mode.

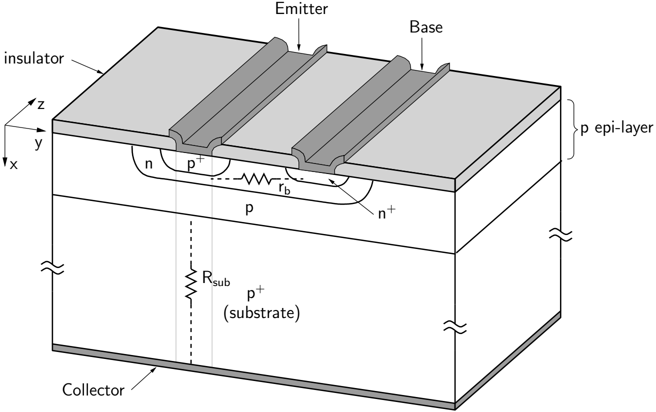
BJT modes of operation



Mode	*B*-*E* junction	*B*-*C* junction
Active (linear)	forward	reverse
Cutoff	reverse	reverse
Saturation	forward	forward
Inverse active	reverse	forward

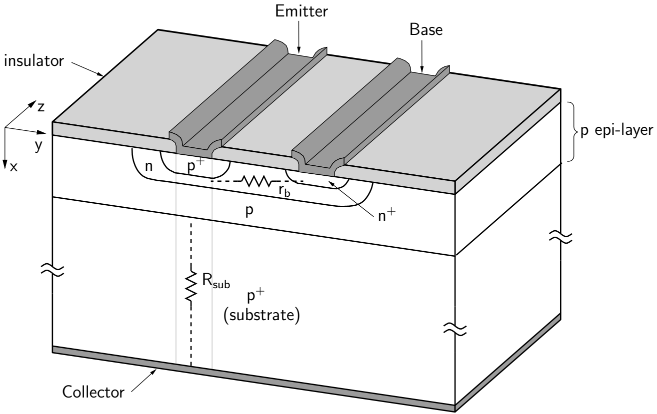
- In analog circuits, BJTs are generally biased to operate in the active mode.
- BJT as a switch:
  - Closed: saturation mode
  - Open: cutoff mode

Device structure (*pnp* transistor)



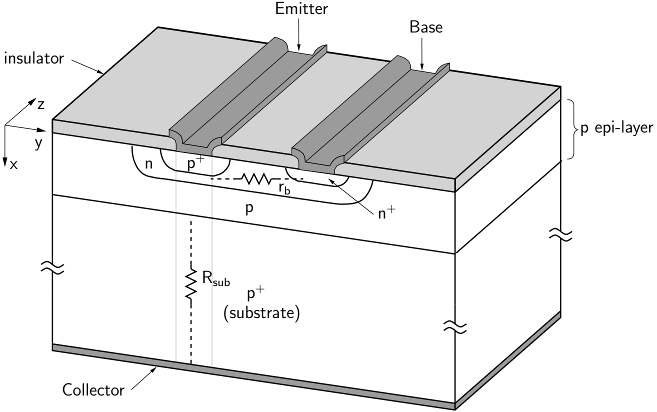
Device structure (*pnp* transistor)

- The substrate thickness is hundreds of microns whereas the *p* epi-layer and the rest of the device structure is confined to a few microns.



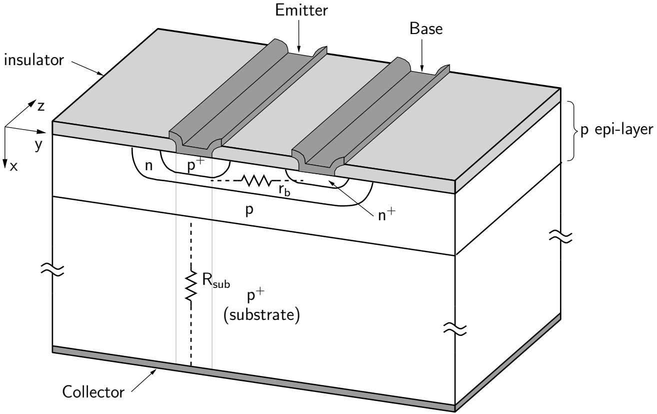
Device structure (*pnp* transistor)

- The substrate thickness is hundreds of microns whereas the *p* epi-layer and the rest of the device structure is confined to a few microns.
- The p^+ substrate can be treated as a small resistance R_{sub}.



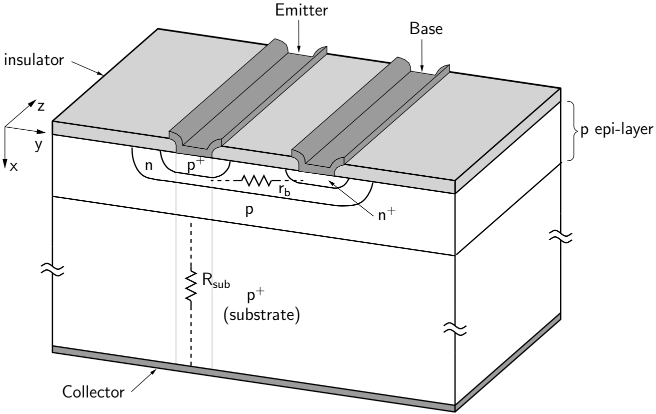
Device structure (*pnp* transistor)

- The substrate thickness is hundreds of microns whereas the *p* epi-layer and the rest of the device structure is confined to a few microns.
- The p^+ substrate can be treated as a small resistance R_{sub}.
- The “active” region of the device is only the top few microns (the *p* epi-layer).



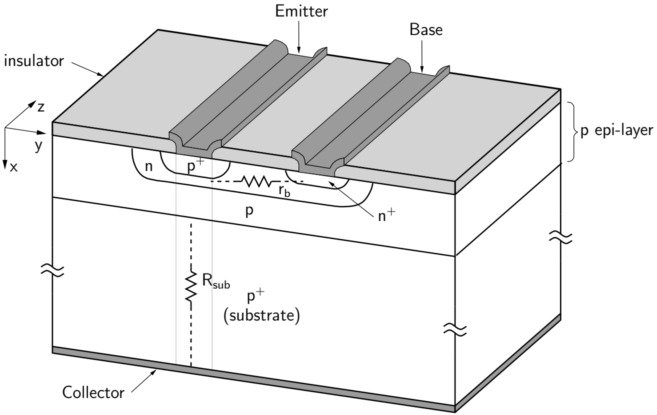
Device structure (*pnp* transistor)

- The substrate thickness is hundreds of microns whereas the *p* epi-layer and the rest of the device structure is confined to a few microns.
- The p^+ substrate can be treated as a small resistance R_{sub}.
- The "active" region of the device is only the top few microns (the *p* epi-layer).
- The figure is not drawn to scale. Typically, the y and z dimensions are much larger than the *p* epi thickness.



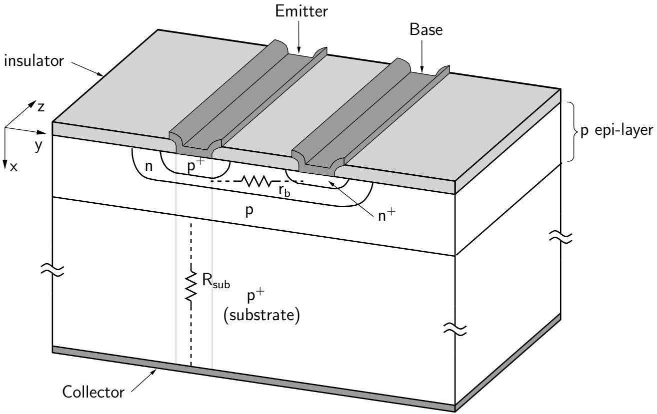
Device structure (*pnp* transistor)

- The substrate thickness is hundreds of microns whereas the *p* epi-layer and the rest of the device structure is confined to a few microns.
- The p^+ substrate can be treated as a small resistance R_{sub}.
- The "active" region of the device is only the top few microns (the *p* epi-layer).
- The figure is not drawn to scale. Typically, the *y* and *z* dimensions are much larger than the *p* epi thickness.
- Because the contact region dimensions are much larger than the base width (∼ the width of the *n* base region), the device can be considered to be one-dimensional in nature.

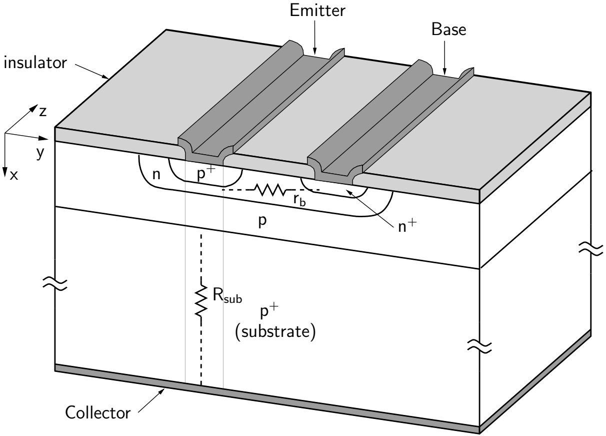


Device structure (*pnp* transistor)

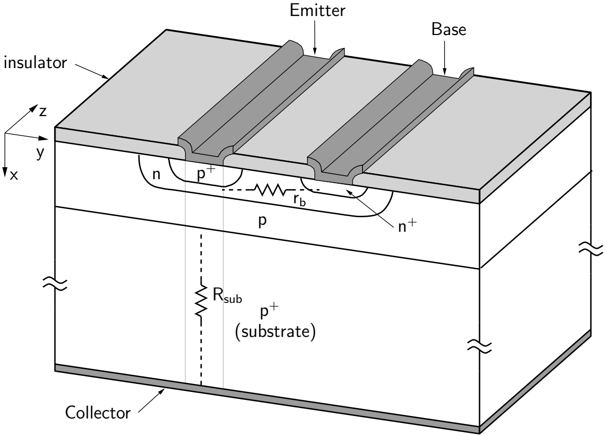
- The substrate thickness is hundreds of microns whereas the p epi-layer and the rest of the device structure is confined to a few microns.
- The p^+ substrate can be treated as a small resistance R_{sub}.
- The "active" region of the device is only the top few microns (the p epi-layer).
- The figure is not drawn to scale. Typically, the y and z dimensions are much larger than the p epi thickness.
- Because the contact region dimensions are much larger than the base width ($\sim$ the width of the n base region), the device can be considered to be one-dimensional in nature.
- A "base resistance" r_b exists between the base region and the base contact. To keep r_b small, the base contact is made close to the emitter.

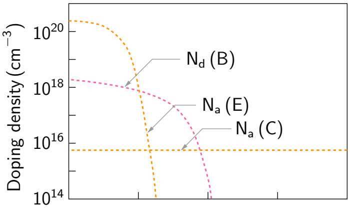


Device structure (*pnp* transistor)

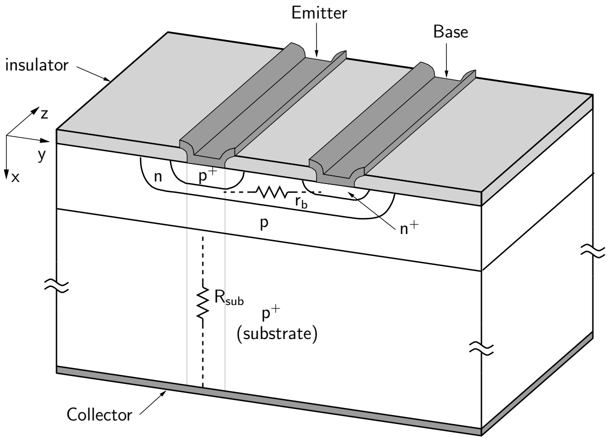


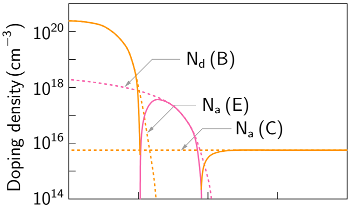
Device structure (*pnp* transistor)



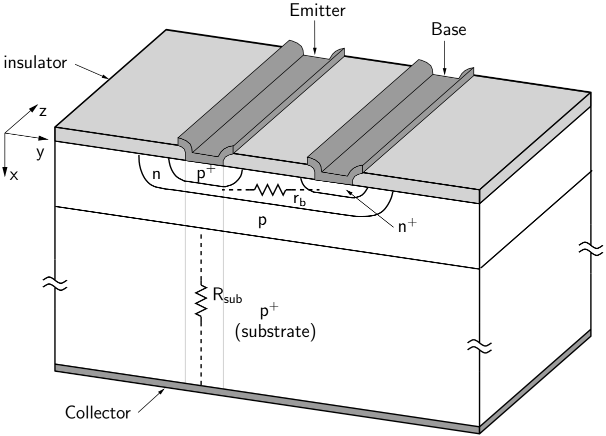


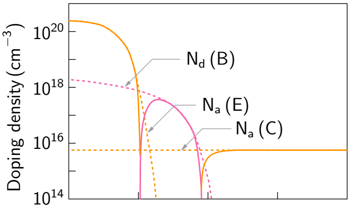
Device structure (*pnp* transistor)





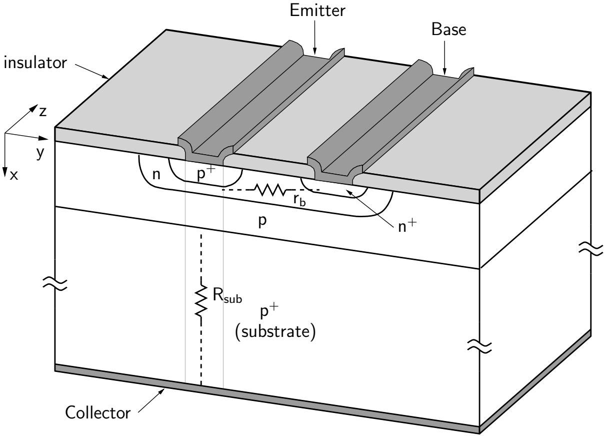
Device structure (*pnp* transistor)

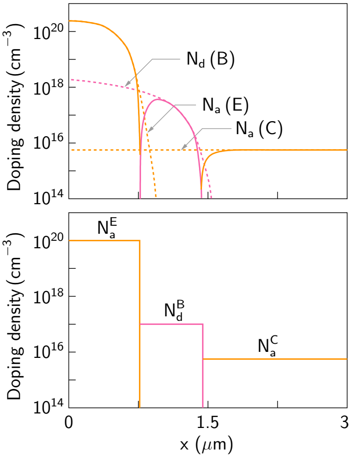




∗ For simplicity, we will assume the doping densities to be constant in the emitter, base, and collector regions.

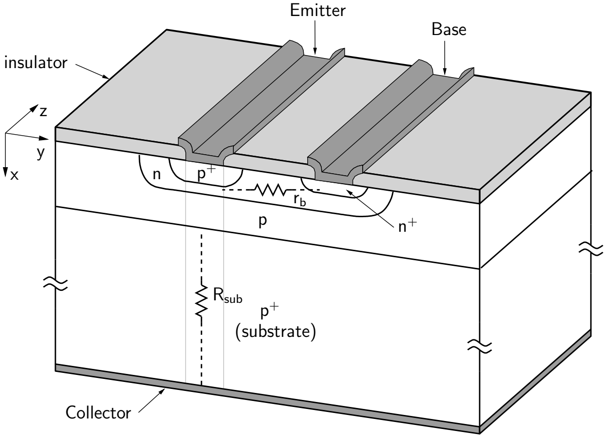
Device structure (*pnp* transistor)

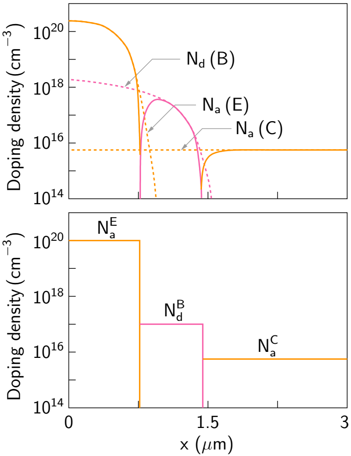




- ∗ For simplicity, we will assume the doping densities to be constant in the emitter, base, and collector regions.

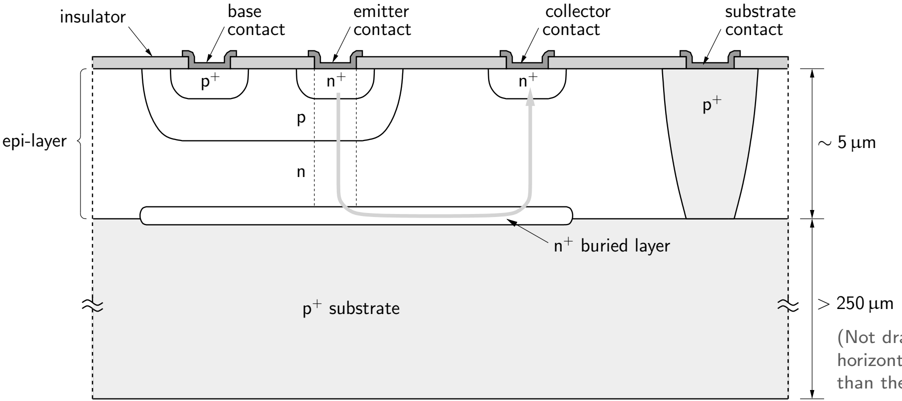
Device structure (*pnp* transistor)





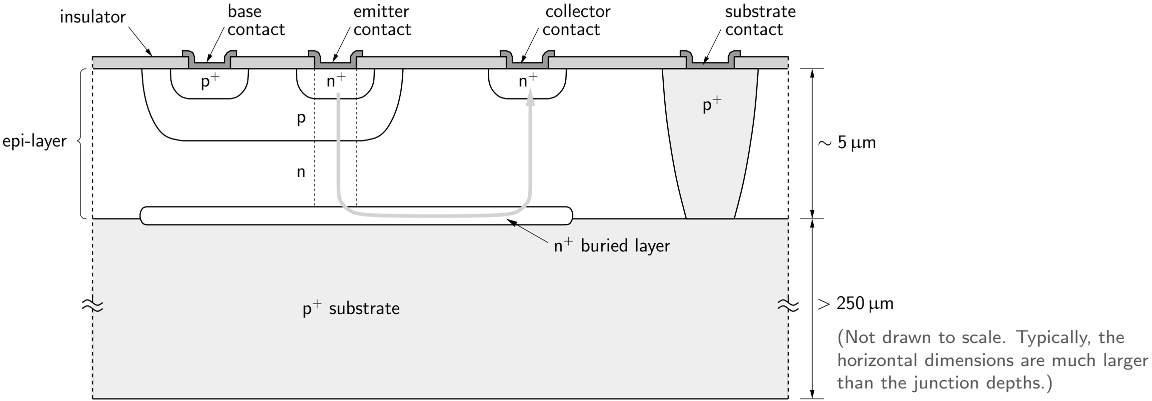
- For simplicity, we will assume the doping densities to be constant in the emitter, base, and collector regions.
- The relationship $N_a^E > N_d^B > N_a^C$, which is a consequence of the fabrication process, is also desirable from the device performance angle.

BJT structure in integrated circuits (*npn* transistor)



(Not drawn to scale. Typically, the horizontal dimensions are much larger than the junction depths.)

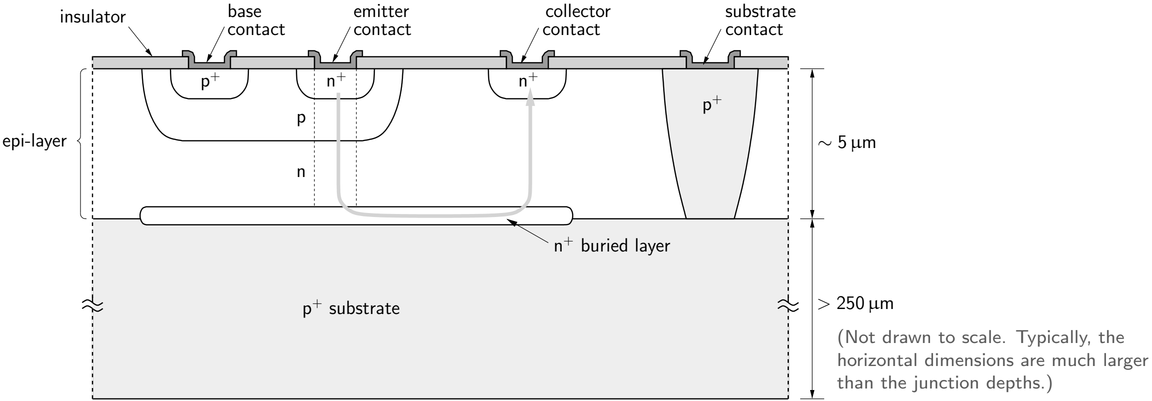
BJT structure in integrated circuits (*npn* transistor)



(Not drawn to scale. Typically, the horizontal dimensions are much larger than the junction depths.)

- To make an integrated circuit (IC), a large number of transistors are fabricated on a single silicon piece and interconnected as required.

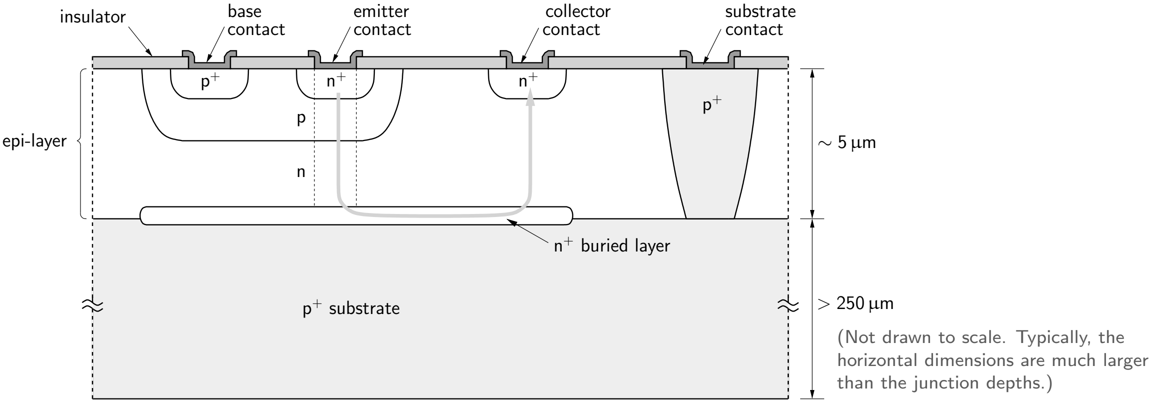
BJT structure in integrated circuits (*npn* transistor)



(Not drawn to scale. Typically, the horizontal dimensions are much larger than the junction depths.)

- To make an integrated circuit (IC), a large number of transistors are fabricated on a single silicon piece and interconnected as required.
- The BJTs are isolated from each other using reverse-biased *pn* junctions.

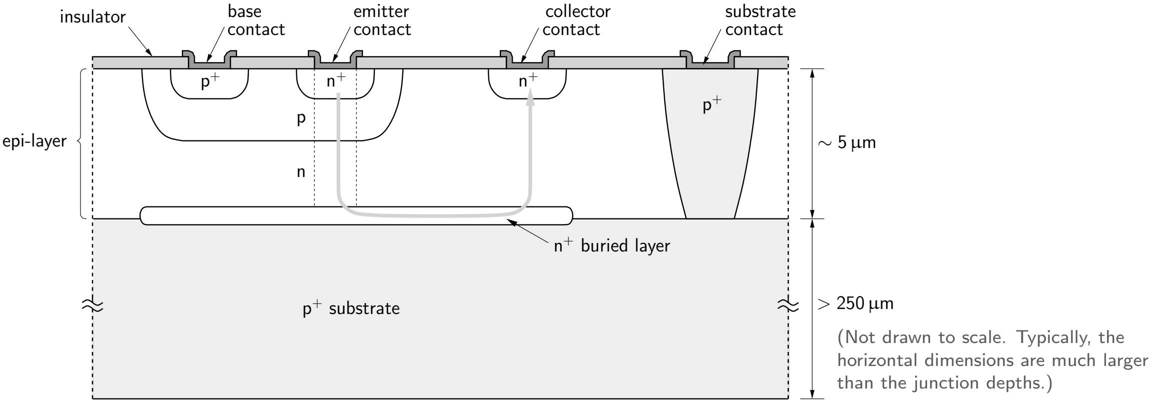
BJT structure in integrated circuits (*npn* transistor)



(Not drawn to scale. Typically, the horizontal dimensions are much larger than the junction depths.)

- To make an integrated circuit (IC), a large number of transistors are fabricated on a single silicon piece and interconnected as required.
- The BJTs are isolated from each other using reverse-biased *pn* junctions.
- Contacts (E, B, C) are made on the top surface for connecting to other transistors.

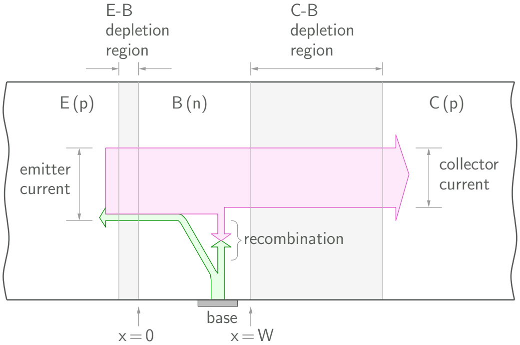
BJT structure in integrated circuits (*npn* transistor)



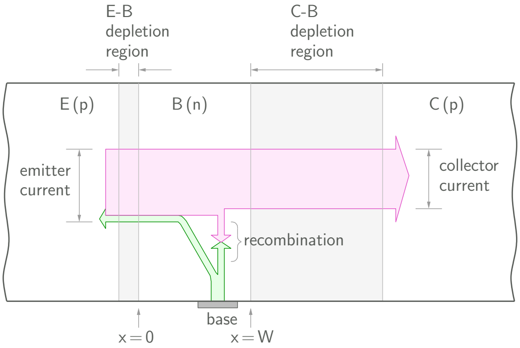
(Not drawn to scale. Typically, the horizontal dimensions are much larger than the junction depths.)

- To make an integrated circuit (IC), a large number of transistors are fabricated on a single silicon piece and interconnected as required.
- The BJTs are isolated from each other using reverse-biased *pn* junctions.
- Contacts (E, B, C) are made on the top surface for connecting to other transistors.
- An n^+ buried layer is used to provide a low-resistance path for the electron current.

Dependence of α on device parameters



Dependence of α on device parameters

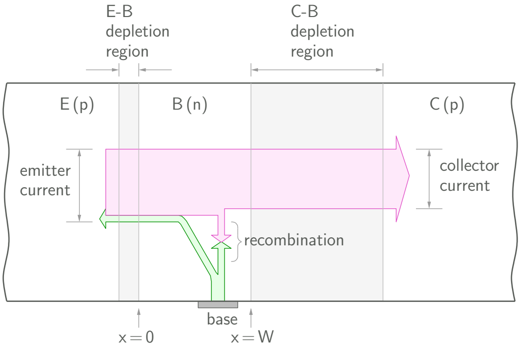


Consider a *pnp* transistor.

- I_E has a hole component and an electron component. Of these, only the hole component contributes to I_C.

  We define "emitter injection efficiency" (or simply "injection efficiency") as $\gamma = \dfrac{I_{pE}}{I_E} = \dfrac{I_{pE}}{I_{pE} + I_{nE}}$.

Dependence of α on device parameters



Consider a *pnp* transistor.

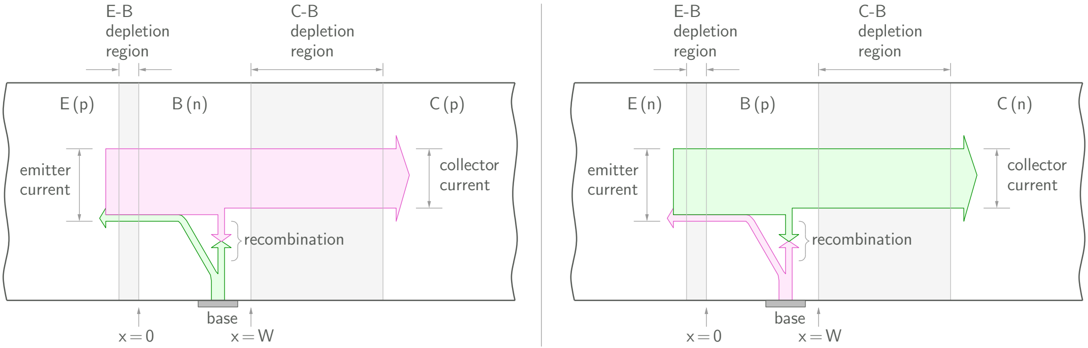
- I_E has a hole component and an electron component. Of these, only the hole component contributes to I_C.

  We define "emitter injection efficiency" (or simply "injection efficiency") as $\gamma = \dfrac{I_{pE}}{I_E} = \dfrac{I_{pE}}{I_{pE} + I_{nE}}$.

- Some of the injected holes are lost due to recombination in the neutral base region (we will ignore G-R in the B-C depletion regions).

  We define "base transport factor" as $\alpha_T = \dfrac{I_{pC}}{I_{pE}}$.

Dependence of α on device parameters



Consider a *pnp* transistor.

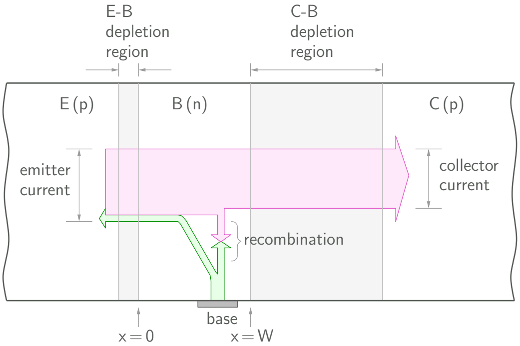
- I_E has a hole component and an electron component. Of these, only the hole component contributes to I_C.

  We define "emitter injection efficiency" (or simply "injection efficiency") as $\gamma = \dfrac{I_{pE}}{I_E} = \dfrac{I_{pE}}{I_{pE} + I_{nE}}$.

- Some of the injected holes are lost due to recombination in the neutral base region (we will ignore G-R in the B-C depletion regions).

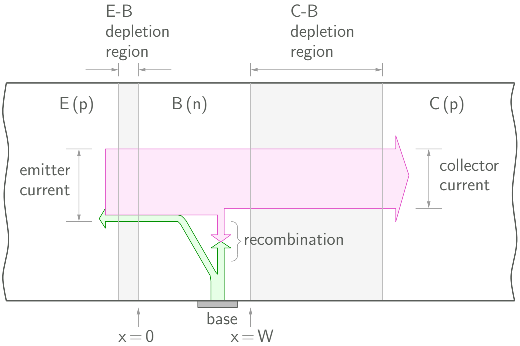
  We define "base transport factor" as $\alpha_T = \dfrac{I_{pC}}{I_{pE}}$.

Dependence of α on device parameters



Since the C-B junction is reverse biased, the *pn* junction current arising because of V_{CB} is negligibly small.

Dependence of α on device parameters

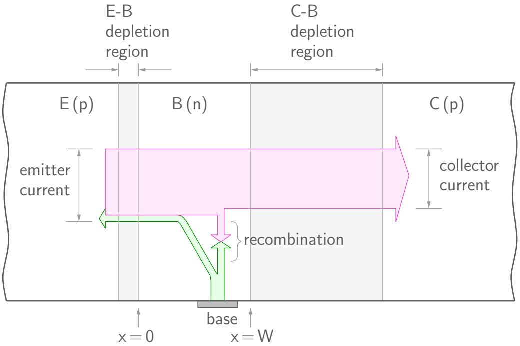


Since the C-B junction is reverse biased, the *pn* junction current arising because of V_{CB} is negligibly small.

→ I_C is entirely due to the holes injected by the emitter which make it to the C-B depletion boundary ($x = W$), i.e.,

$$I_C \approx I_{pC} = \alpha_T I_{pE} = \alpha_T(\gamma\, I_E) \rightarrow \alpha = \frac{I_C}{I_E} = \gamma\, \alpha_T\,.$$

Dependence of α on device parameters

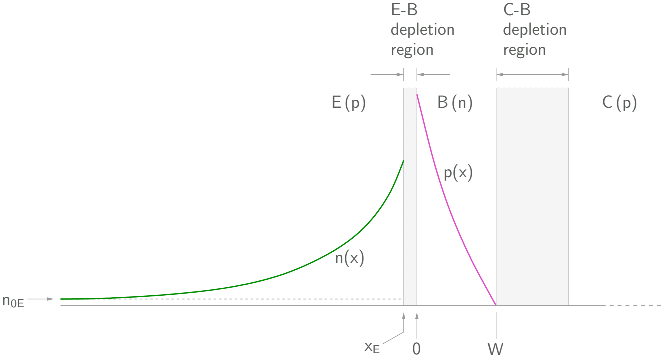


Since the C-B junction is reverse biased, the *pn* junction current arising because of V_{CB} is negligibly small.

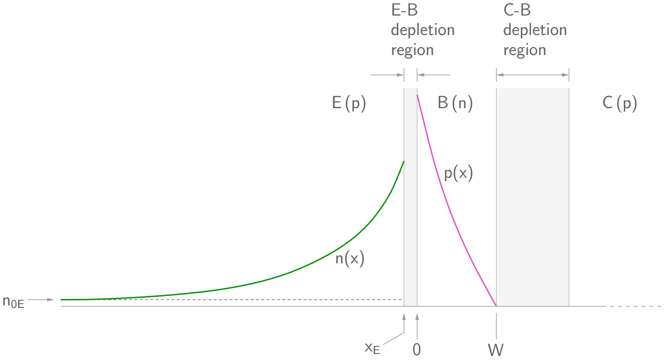
→ I_C is entirely due to the holes injected by the emitter which make it to the C-B depletion boundary ($x = W$), i.e.,

$$I_C \approx I_{pC} = \alpha_T I_{pE} = \alpha_T (\gamma \, I_E) \rightarrow \alpha = \frac{I_C}{I_E} = \gamma \, \alpha_T .$$

→ For $\alpha \approx 1$, both γ and α_T must be close to 1.



We assume that the emitter width is greater than 5 L_n.



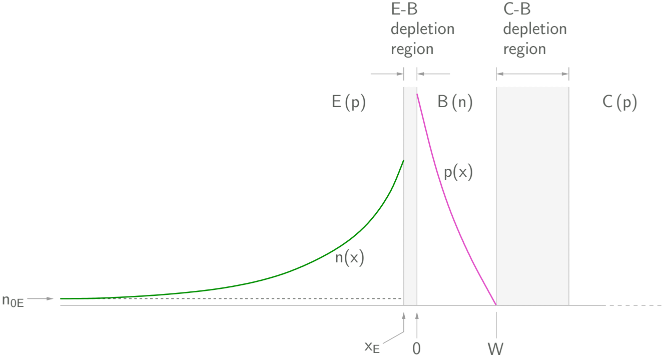
We assume that the emitter width is greater than 5 L_n.

Neglecting the drift components for minority carriers in the emitter and base neutral regions, we get

$$D_{nE}\,\frac{d^2\Delta n}{dx^2} - \frac{\Delta n}{\tau_{nE}} = 0, \quad x < x_E, \text{ with}$$

$$\Delta n(x_E) = n_{0E}\left[\exp\left(\frac{V_{EB}}{V_T}\right) - 1\right],$$

$$\Delta n(-\infty) = 0.$$



We assume that the emitter width is greater than $5\,L_n$.

Neglecting the drift components for minority carriers in the emitter and base neutral regions, we get

$$D_{nE}\,\frac{d^2\Delta n}{dx^2} - \frac{\Delta n}{\tau_{nE}} = 0, \quad x < x_E, \text{ with}$$

$$\Delta n(x_E) = n_{0E}\left[\exp\left(\frac{V_{EB}}{V_T}\right) - 1\right],$$

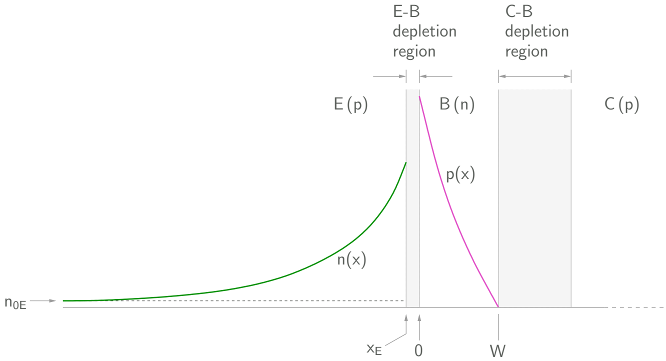
$$\Delta n(-\infty) = 0.$$

$$D_{pB}\,\frac{d^2\Delta p}{dx^2} - \frac{\Delta p}{\tau_{pB}} = 0, \quad 0 < x < W, \text{ with}$$

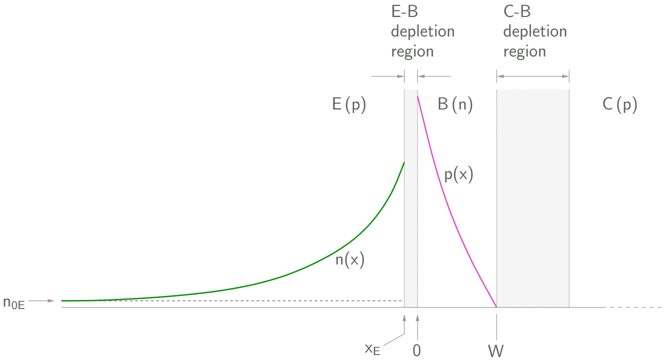
$$\Delta p(0) = p_{0B}\left[\exp\left(\frac{V_{EB}}{V_T}\right) - 1\right],$$

$$\Delta p(W) = p_{0B}\left[\exp\left(\frac{V_{CB}}{V_T}\right) - 1\right].$$

Dependence of α on device parameters



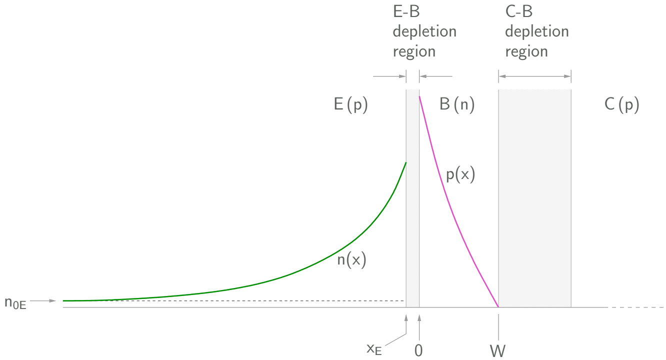
Dependence of α on device parameters



Solution: $\Delta n(x) = \Delta n(x_E)\, e^{-(x_E - x)/L_{nE}}, \quad x < x_E,$

$$\Delta p(x) = A\, e^{-x/L_{pB}} + B\, e^{+x/L_{pB}}, \quad 0 < x < W.$$

Dependence of α on device parameters



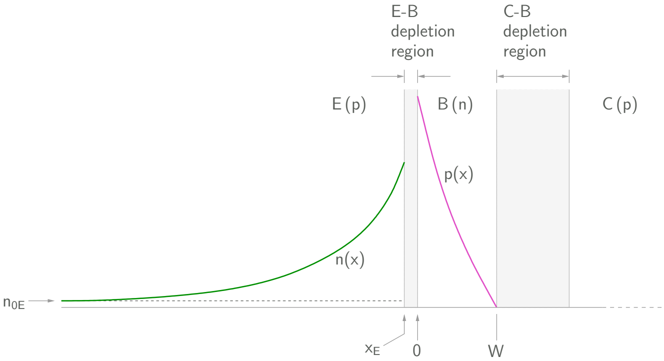
Solution: $\Delta n(x) = \Delta n(x_E)\, e^{-(x_E - x)/L_{nE}}, \qquad x < x_E,$

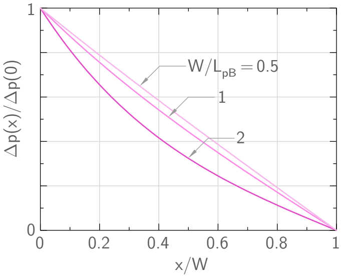
$$\Delta p(x) = A\, e^{-x/L_{pB}} + B\, e^{+x/L_{pB}}, \quad 0 < x < W.$$

Using the boundary conditions (last slide), we get

$$\Delta p(x) = \Delta p(0)\, \frac{\sinh\left(\dfrac{W-x}{L_{pB}}\right)}{\sinh\left(\dfrac{W}{L_{pB}}\right)} + \Delta p(W)\, \frac{\sinh\left(\dfrac{x}{L_{pB}}\right)}{\sinh\left(\dfrac{W}{L_{pB}}\right)}.$$

Dependence of α on device parameters



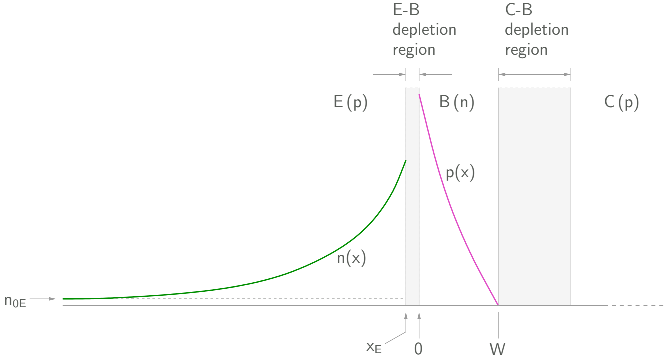


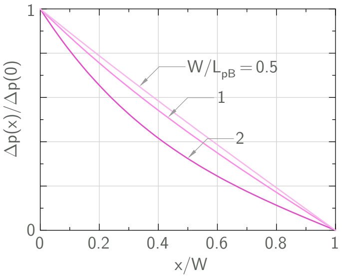
Solution: $\Delta n(x) = \Delta n(x_E)\, e^{-(x_E - x)/L_{nE}}, \quad x < x_E,$

$$\Delta p(x) = A\, e^{-x/L_{pB}} + B\, e^{+x/L_{pB}}, \quad 0 < x < W.$$

Using the boundary conditions (last slide), we get

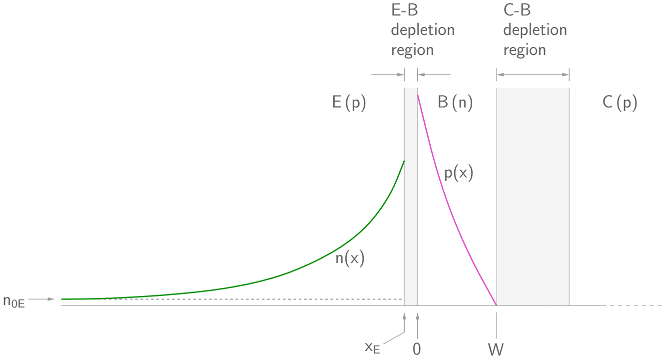
$$\Delta p(x) = \Delta p(0)\, \frac{\sinh\left(\dfrac{W-x}{L_{pB}}\right)}{\sinh\left(\dfrac{W}{L_{pB}}\right)} + \Delta p(W)\, \frac{\sinh\left(\dfrac{x}{L_{pB}}\right)}{\sinh\left(\dfrac{W}{L_{pB}}\right)}.$$

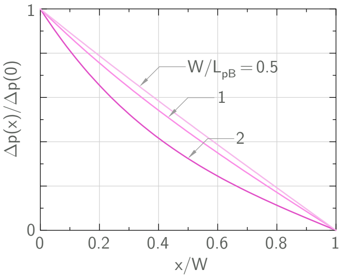




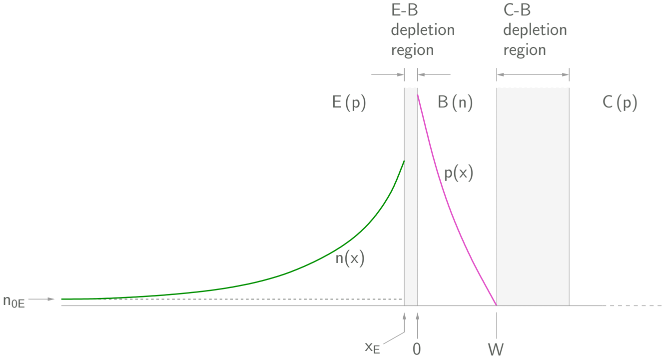
$$I_{nE} = qAD_{nE}\,\frac{dn}{dx}(x_E) = qAD_{nE}\,\frac{d\Delta n}{dx}(x_E)$$

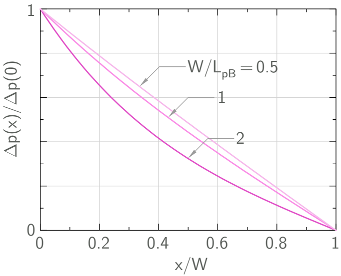
$$= \frac{qAD_{nE}}{L_{nE}}\,\Delta n(x_E) = \frac{qAD_{nE}}{L_{nE}}\,n_{0E}\left(e^{V_{EB}/V_T} - 1\right),$$





$$
\begin{aligned}
I_{nE} &= qAD_{nE}\,\frac{dn}{dx}(x_E) = qAD_{nE}\,\frac{d\Delta n}{dx}(x_E)\\
&= \frac{qAD_{nE}}{L_{nE}}\,\Delta n(x_E) = \frac{qAD_{nE}}{L_{nE}}\,n_{0E}\left(e^{V_{EB}/V_T}-1\right),\\
I_{pE} &= -qAD_{pB}\,\frac{dp}{dx}(0) = -qAD_{pB}\,\frac{d\Delta p}{dx}(0)\\
&= \frac{qAD_{pB}}{L_{pB}}\,p_{0B}\left(e^{V_{EB}/V_T}-1\right)\frac{\cosh(W/L_{pB})}{\sinh(W/L_{pB})}.
\end{aligned}
$$



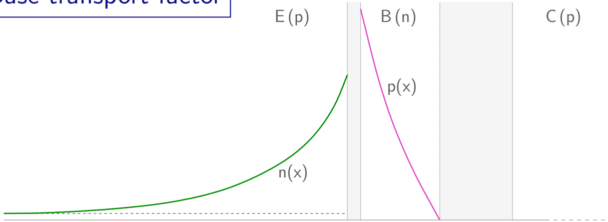


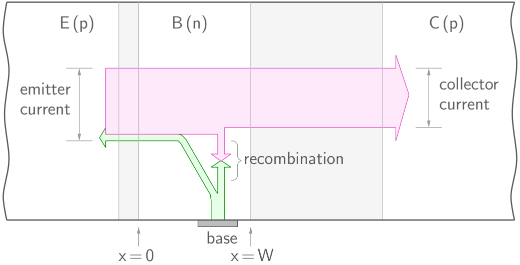
$$
\begin{aligned}
I_{nE} &= qAD_{nE}\,\frac{dn}{dx}(x_E) = qAD_{nE}\,\frac{d\Delta n}{dx}(x_E) \\
&= \frac{qAD_{nE}}{L_{nE}}\,\Delta n(x_E) = \frac{qAD_{nE}}{L_{nE}}\,n_{0E}\left(e^{V_{EB}/V_T}-1\right), \\
I_{pE} &= -qAD_{pB}\,\frac{dp}{dx}(0) = -qAD_{pB}\,\frac{d\Delta p}{dx}(0) \\
&= \frac{qAD_{pB}}{L_{pB}}\,p_{0B}\left(e^{V_{EB}/V_T}-1\right)\frac{\cosh(W/L_{pB})}{\sinh(W/L_{pB})}.
\end{aligned}
$$

$$
\begin{aligned}
\gamma &= \frac{I_{pE}}{I_{pE}+I_{nE}} = \frac{1}{1+(I_{nE}/I_{pE})} \\
&= \frac{1}{1+\left(\dfrac{D_{nE}}{D_{pB}}\,\dfrac{L_{pB}}{L_{nE}}\,\dfrac{N_{dB}}{N_{aE}}\right)\dfrac{\sinh(W/L_B)}{\cosh(W/L_B)}},
\end{aligned}
$$

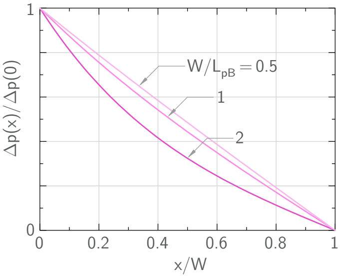
since $\dfrac{n_{0E}}{p_{0B}} = \dfrac{n_i^2}{N_{aE}} \times \dfrac{N_{dB}}{n_i^2} = \dfrac{N_{dB}}{N_{aE}}$.

Base transport factor

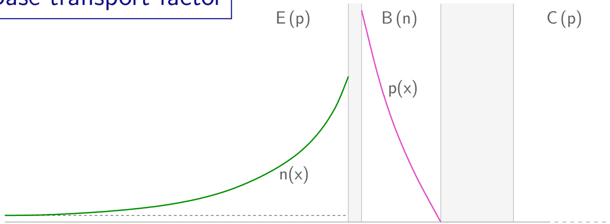


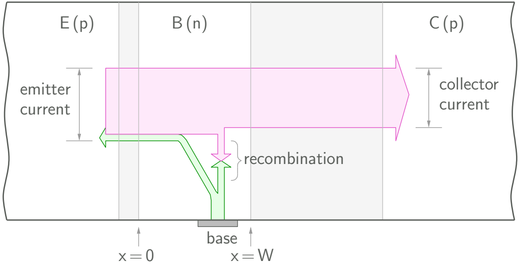


$$\Delta p(x) = \Delta p(0)\,\frac{\sinh\left(\dfrac{W-x}{L_{pB}}\right)}{\sinh\left(\dfrac{W}{L_{pB}}\right)} + \Delta p(W)\,\frac{\sinh\left(\dfrac{x}{L_{pB}}\right)}{\sinh\left(\dfrac{W}{L_{pB}}\right)}.$$



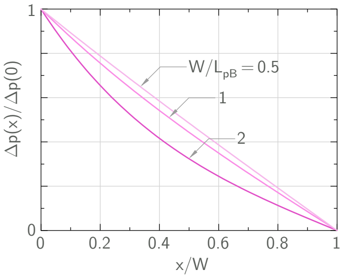
Base transport factor



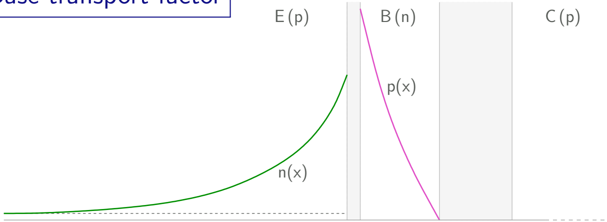


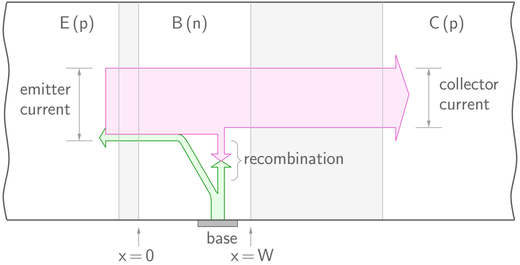
$$\Delta p(x) = \Delta p(0)\,\frac{\sinh\left(\dfrac{W-x}{L_{pB}}\right)}{\sinh\left(\dfrac{W}{L_{pB}}\right)} + \Delta p(W)\,\frac{\sinh\left(\dfrac{x}{L_{pB}}\right)}{\sinh\left(\dfrac{W}{L_{pB}}\right)}.$$

$$I_C \approx I_{pC} = -qAD_{pB}\,\frac{d\Delta p}{dx}(W)$$



Base transport factor

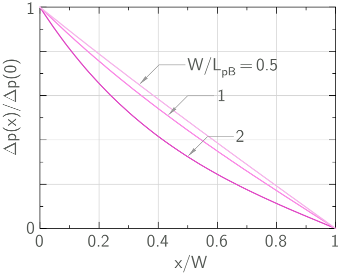




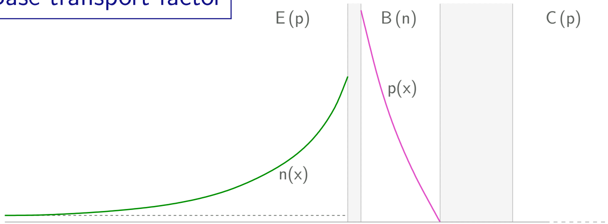
$$\Delta p(x) = \Delta p(0)\,\frac{\sinh\left(\dfrac{W-x}{L_{pB}}\right)}{\sinh\left(\dfrac{W}{L_{pB}}\right)} + \Delta p(W)\,\frac{\sinh\left(\dfrac{x}{L_{pB}}\right)}{\sinh\left(\dfrac{W}{L_{pB}}\right)}.$$

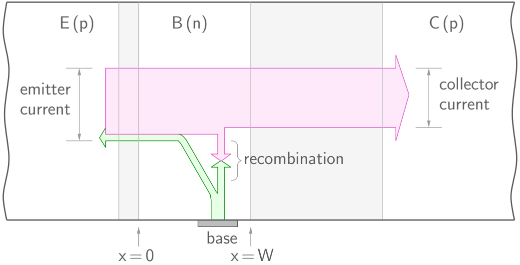
$$I_C \approx I_{pC} = -qAD_{pB}\,\frac{d\Delta p}{dx}(W)$$

$$= \frac{qAD_{pB}}{L_{pB}}\,p_{0B}\left(e^{V_{EB}/V_T} - 1\right)\frac{1}{\sinh(W/L_B)}.$$



Base transport factor





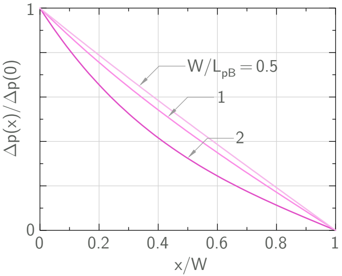
$$\Delta p(x) = \Delta p(0)\,\frac{\sinh\left(\dfrac{W-x}{L_{pB}}\right)}{\sinh\left(\dfrac{W}{L_{pB}}\right)} + \Delta p(W)\,\frac{\sinh\left(\dfrac{x}{L_{pB}}\right)}{\sinh\left(\dfrac{W}{L_{pB}}\right)}.$$

$$I_C \approx I_{pC} = -qAD_{pB}\,\frac{d\Delta p}{dx}(W)$$

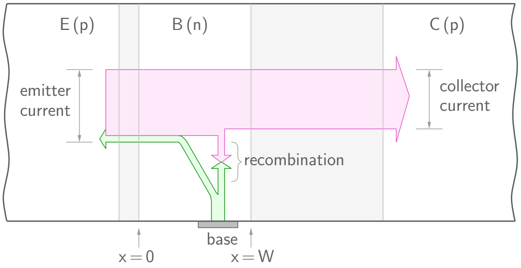
$$= \frac{qAD_{pB}}{L_{pB}}\,p_{0B}\left(e^{V_{EB}/V_T}-1\right)\frac{1}{\sinh(W/L_B)}.$$

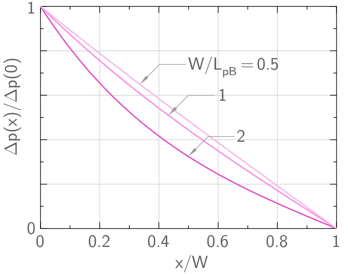
The base transport factor is (using I_{pE} from the last slide),

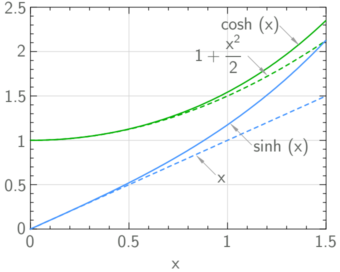
$$\alpha_T = \frac{I_{pC}}{I_{pE}} = \frac{1}{\cosh(W/L_{pB})}.$$



α_T with $W \ll L_{pB}$

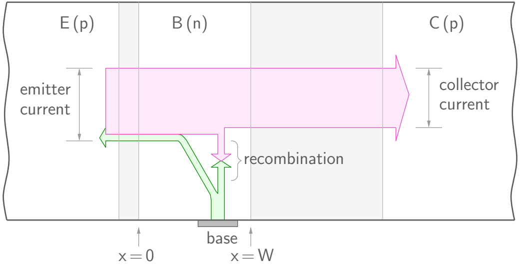


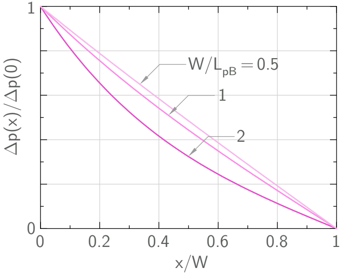




$$\Delta p(x) = \Delta p(0)\,\frac{\sinh\left(\dfrac{W-x}{L_{pB}}\right)}{\sinh\left(\dfrac{W}{L_{pB}}\right)} + \Delta p(W)\,\frac{\sinh\left(\dfrac{x}{L_{pB}}\right)}{\sinh\left(\dfrac{W}{L_{pB}}\right)}.$$

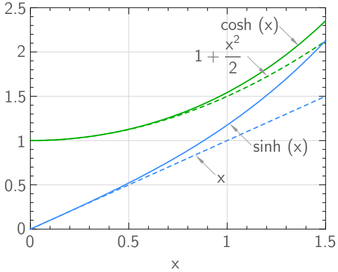
α_T with $W \ll L_{pB}$



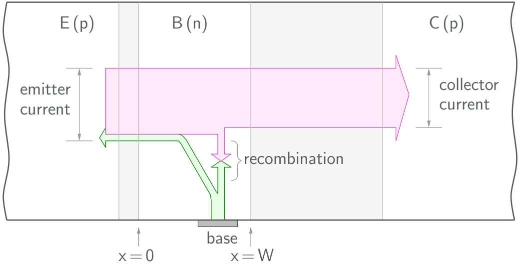


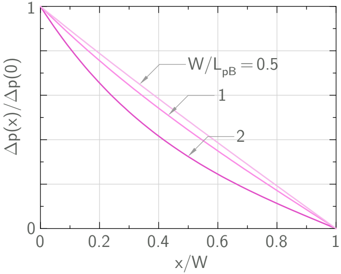
$$\Delta p(x) = \Delta p(0)\,\frac{\sinh\left(\dfrac{W-x}{L_{pB}}\right)}{\sinh\left(\dfrac{W}{L_{pB}}\right)} + \Delta p(W)\,\frac{\sinh\left(\dfrac{x}{L_{pB}}\right)}{\sinh\left(\dfrac{W}{L_{pB}}\right)}.$$

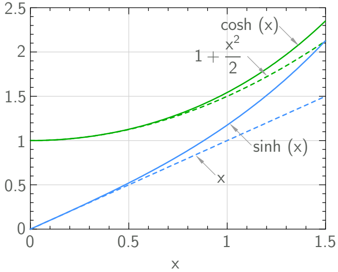
$$\alpha_T = \frac{I_{pC}}{I_{pE}} = \frac{1}{\cosh(W/L_{pB})} \approx \frac{1}{1+\dfrac{1}{2}\left(\dfrac{W}{L_{pB}}\right)^2}.$$



α_T with $W \ll L_{pB}$





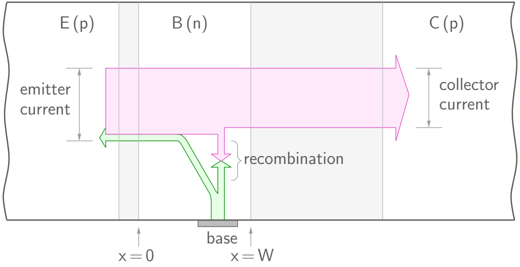


$$\Delta p(x) = \Delta p(0)\,\frac{\sinh\left(\dfrac{W-x}{L_{pB}}\right)}{\sinh\left(\dfrac{W}{L_{pB}}\right)} + \Delta p(W)\,\frac{\sinh\left(\dfrac{x}{L_{pB}}\right)}{\sinh\left(\dfrac{W}{L_{pB}}\right)}.$$

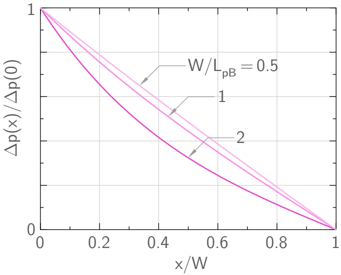
$$\alpha_T = \frac{I_{pC}}{I_{pE}} = \frac{1}{\cosh(W/L_{pB})} \approx \frac{1}{1+\dfrac{1}{2}\left(\dfrac{W}{L_{pB}}\right)^2}.$$

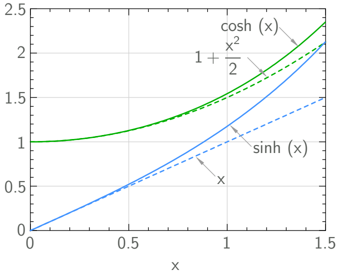
Remark: $\alpha_T \rightarrow 1$ if the base width W is made small compared to L_{pB}.

γ with $W \ll L_{pB}$

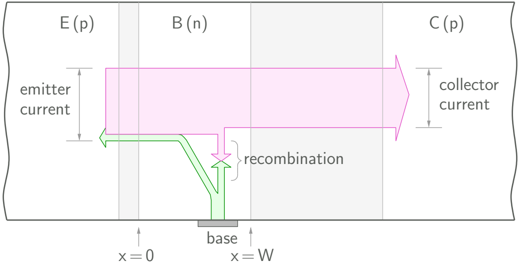


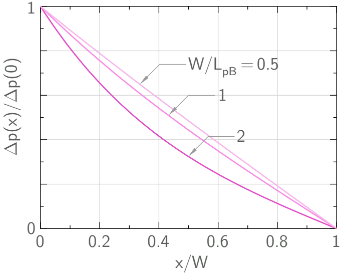
$$\gamma = \frac{I_{pE}}{I_{pE}+I_{nE}} = \frac{1}{1+(I_{nE}/I_{pE})} = \frac{1}{1+\left(\dfrac{D_{nE}}{D_{pB}}\,\dfrac{L_{pB}}{L_{nE}}\,\dfrac{N_{dB}}{N_{aE}}\right)\dfrac{\sinh(W/L_B)}{\cosh(W/L_B)}}$$





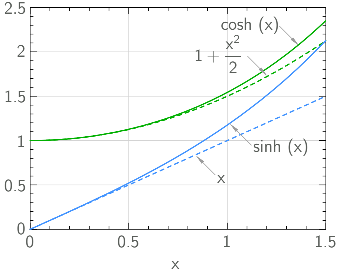
γ with $W \ll L_{pB}$



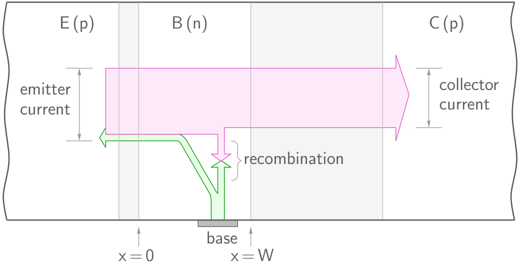


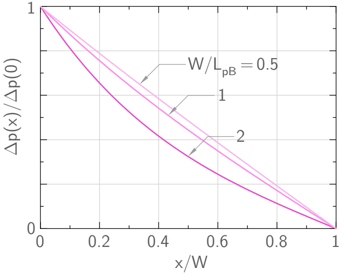
$$\gamma = \frac{I_{pE}}{I_{pE}+I_{nE}} = \frac{1}{1+(I_{nE}/I_{pE})} = \frac{1}{1+\left(\dfrac{D_{nE}}{D_{pB}}\,\dfrac{L_{pB}}{L_{nE}}\,\dfrac{N_{dB}}{N_{aE}}\right)\dfrac{\sinh(W/L_B)}{\cosh(W/L_B)}}$$

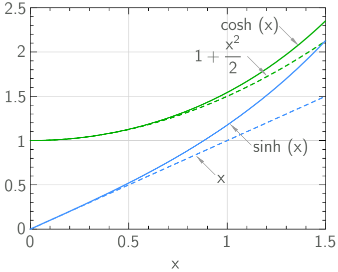
$$\approx \frac{1}{1+\left(\dfrac{D_{nE}}{D_{pB}}\,\dfrac{L_{pB}}{L_{nE}}\,\dfrac{N_{dB}}{N_{aE}}\right)\dfrac{W/L_{pB}}{1+\frac{1}{2}(W/L_{pB})^2}}$$



γ with $W \ll L_{pB}$





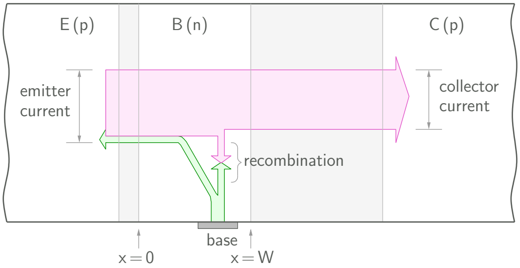


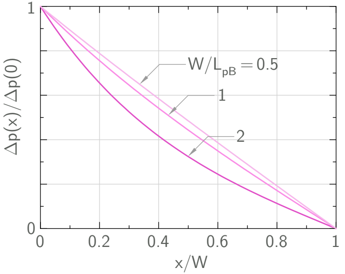
$$\gamma = \frac{I_{pE}}{I_{pE}+I_{nE}} = \frac{1}{1+(I_{nE}/I_{pE})} = \frac{1}{1+\left(\dfrac{D_{nE}}{D_{pB}}\,\dfrac{L_{pB}}{L_{nE}}\,\dfrac{N_{dB}}{N_{aE}}\right)\dfrac{\sinh(W/L_B)}{\cosh(W/L_B)}}$$

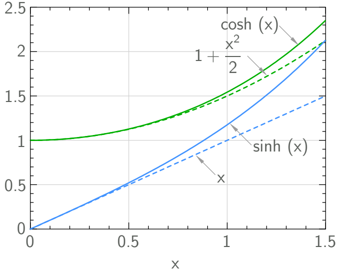
$$\approx \frac{1}{1+\left(\dfrac{D_{nE}}{D_{pB}}\,\dfrac{L_{pB}}{L_{nE}}\,\dfrac{N_{dB}}{N_{aE}}\right)\dfrac{W/L_{pB}}{1+\frac{1}{2}(W/L_{pB})^2}}$$

$$\approx \frac{1}{1+\left(\dfrac{D_{nE}}{D_{pB}}\,\dfrac{W}{L_{nE}}\,\dfrac{N_{dB}}{N_{aE}}\right)}$$

γ with $W \ll L_{pB}$

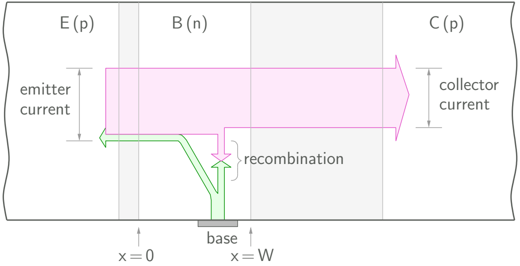


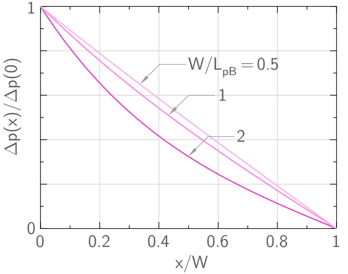




$$\gamma \approx \frac{1}{1 + \left(\dfrac{D_{nE}}{D_{pB}} \, \dfrac{W}{L_{nE}} \, \dfrac{N_{dB}}{N_{aE}} \right)}.$$

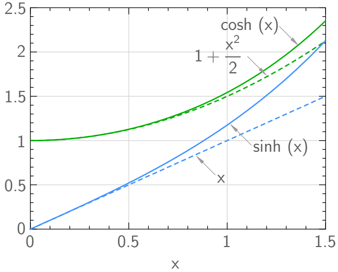
γ with $W \ll L_{pB}$



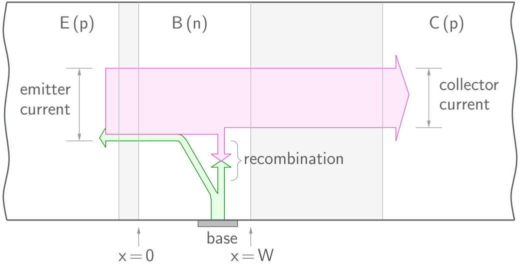


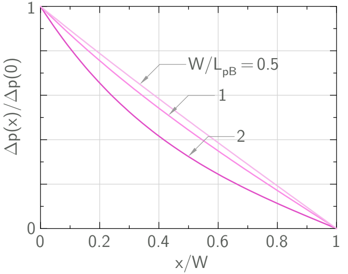
$$\gamma \approx \frac{1}{1 + \left(\frac{D_{nE}}{D_{pB}} \, \frac{W}{L_{nE}} \, \frac{N_{dB}}{N_{aE}} \right)}.$$

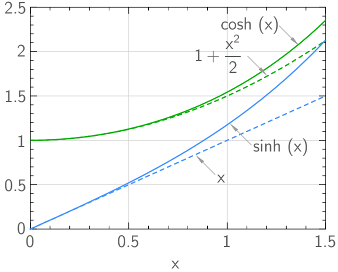
$*$ $\gamma \rightarrow 1$ if $N_{aE} \gg N_{dB}$.



γ with $W \ll L_{pB}$



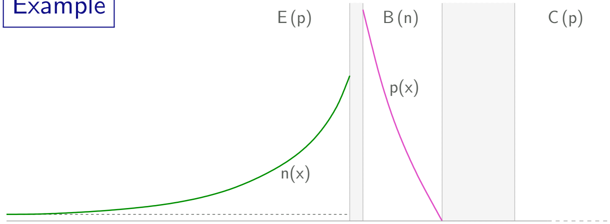


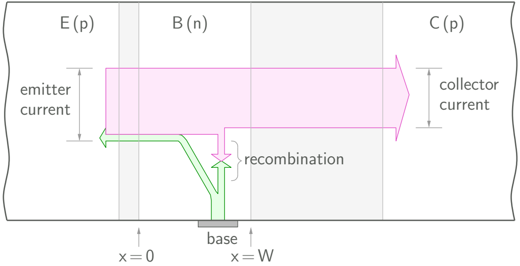


$$\gamma \approx \frac{1}{1 + \left(\frac{D_{nE}}{D_{pB}} \frac{W}{L_{nE}} \frac{N_{dB}}{N_{aE}} \right)}.$$

- $*$ $\gamma \rightarrow 1$ if $N_{aE} \gg N_{dB}$.
- $*$ It is now clear why a higher doping density in the emitter region (compared to the base doping density) is desirable.

Example





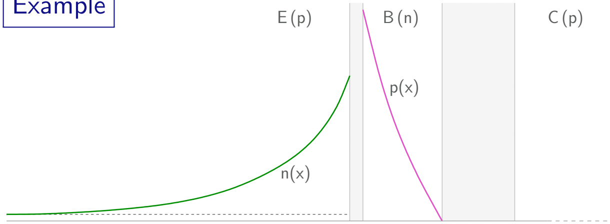
Consider a *pnp* BJT with $N_{aE} = 10^{18}\,\text{cm}^{-3}$, $N_{dB} = 5 \times 10^{16}\,\text{cm}^{-3}$, $N_{aC} = 10^{15}\,\text{cm}^{-3}$, and with a base width $W = 2\,\mu\text{m}$ ($T = 300\,\text{K}$).

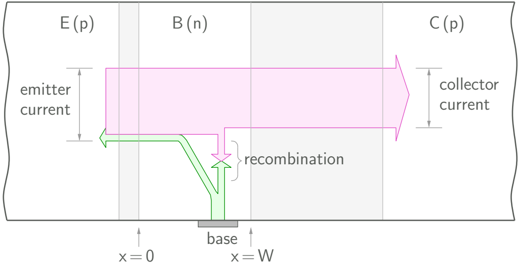
(a) Calculate α_T, γ, α, and β, using the following parameters.

$\mu_{nE} = 250\,\text{cm}^2/\text{V-s}$, $\mu_{pB} = 500\,\text{cm}^2/\text{V-s}$, $\mu_{nC} = 1500\,\text{cm}^2/\text{V-s}$,

$\tau_{nE} = 0.2\,\mu\text{s}$, $\tau_{pB} = 1\,\mu\text{s}$, $\tau_{nC} = 1\,\mu\text{s}$.

Example





Consider a *pnp* BJT with $N_{aE} = 10^{18}\,\text{cm}^{-3}$, $N_{dB} = 5 \times 10^{16}\,\text{cm}^{-3}$, $N_{aC} = 10^{15}\,\text{cm}^{-3}$, and with a base width $W = 2\,\mu\text{m}$ ($T = 300\,\text{K}$).

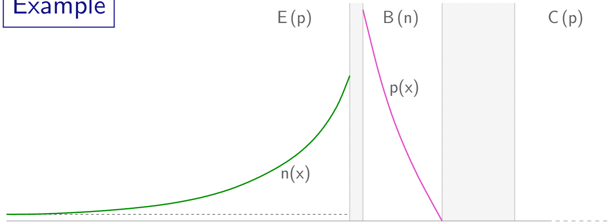
(a) Calculate α_T, γ, α, and β, using the following parameters.

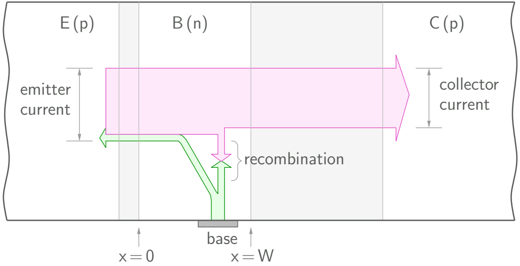
$\mu_{nE} = 250\,\text{cm}^2/\text{V-s}$, $\mu_{pB} = 500\,\text{cm}^2/\text{V-s}$, $\mu_{nC} = 1500\,\text{cm}^2/\text{V-s}$,

$\tau_{nE} = 0.2\,\mu\text{s}$, $\tau_{pB} = 1\,\mu\text{s}$, $\tau_{nC} = 1\,\mu\text{s}$.

(b) Repeat (a) for the BJT operating in the reverse active mode.

Example





Solution:

The minority carrier diffusion lengths are

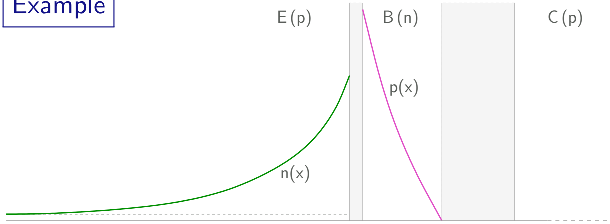
$$L_{nE} = \sqrt{D_{nE}\tau_{nE}} = \sqrt{V_T\,\mu_{nE}\tau_{nE}} = \sqrt{0.0258 \times 250 \times 0.2 \times 10^{-6}} = 1.14 \times 10^{-3}\,\text{cm} = 11.4\,\mu\text{m}.$$

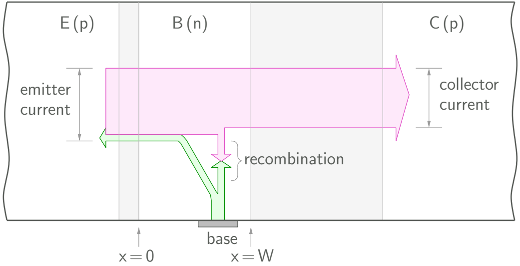
$$L_{pB} = \sqrt{D_{pB}\tau_{pB}} = \sqrt{V_T\,\mu_{pB}\tau_{pB}} = \sqrt{0.0258 \times 500 \times 1 \times 10^{-6}} = 3.59 \times 10^{-3}\,\text{cm} = 35.9\,\mu\text{m}.$$

$$L_{nC} = \sqrt{D_{nC}\tau_{nC}} = \sqrt{V_T\,\mu_{nC}\tau_{nC}} = \sqrt{0.0258 \times 1500 \times 1 \times 10^{-6}} = 6.22 \times 10^{-3}\,\text{cm} = 62.2\,\mu\text{m}.$$

Note that $L_{pB} \gg W$ ($2\,\mu$m).

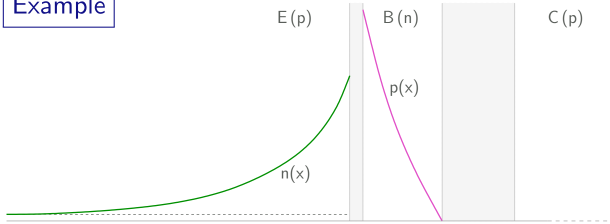
Example

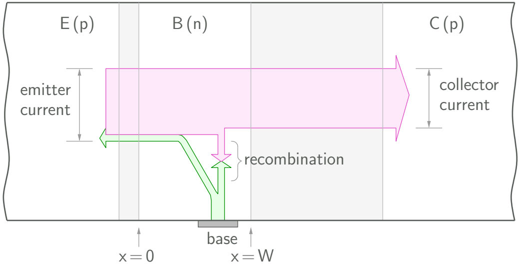




(a) $$\frac{D_{nE}}{D_{pB}}\,\frac{W}{L_{nE}}\,\frac{N_{dB}}{N_{aE}}=\frac{\mu_{nE}}{\mu_{pB}}\,\frac{W}{L_{nE}}\,\frac{N_{dB}}{N_{aE}}=\frac{250}{500}\,\frac{2\times10^{-4}}{1.14\times10^{-3}}\,\frac{5\times10^{16}}{10^{18}}=4.386\times10^{-3}.$$

Example

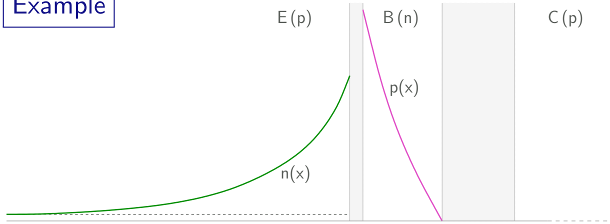


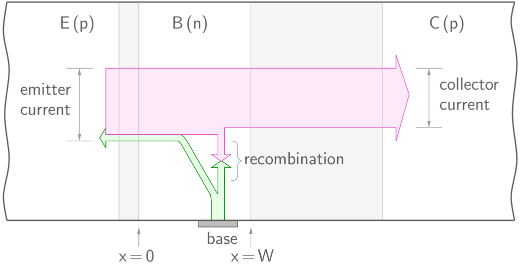


(a) $$\frac{D_{nE}}{D_{pB}}\,\frac{W}{L_{nE}}\,\frac{N_{dB}}{N_{aE}} = \frac{\mu_{nE}}{\mu_{pB}}\,\frac{W}{L_{nE}}\,\frac{N_{dB}}{N_{aE}} = \frac{250}{500}\,\frac{2\times10^{-4}}{1.14\times10^{-3}}\,\frac{5\times10^{16}}{10^{18}} = 4.386\times10^{-3}.$$

$$\gamma = \frac{1}{1+4.386\times10^{-3}} = 0.9956.$$

Example



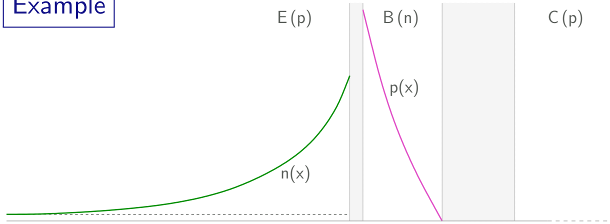


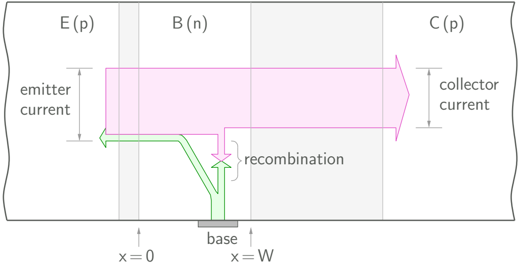
(a) $$\frac{D_{nE}}{D_{pB}}\,\frac{W}{L_{nE}}\,\frac{N_{dB}}{N_{aE}} = \frac{\mu_{nE}}{\mu_{pB}}\,\frac{W}{L_{nE}}\,\frac{N_{dB}}{N_{aE}} = \frac{250}{500}\,\frac{2\times 10^{-4}}{1.14\times 10^{-3}}\,\frac{5\times 10^{16}}{10^{18}} = 4.386\times 10^{-3}.$$

$$\gamma = \frac{1}{1 + 4.386\times 10^{-3}} = 0.9956.$$

$$\alpha_T = \frac{1}{1+\frac{1}{2}(W/L_{pB})^2} = \frac{1}{1+\frac{1}{2}(2.0/35.9)^2} = 0.9985.$$

Example





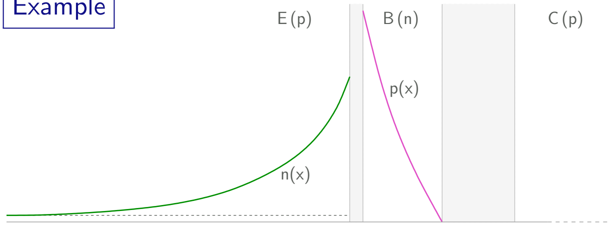
(a) $$\frac{D_{nE}}{D_{pB}}\,\frac{W}{L_{nE}}\,\frac{N_{dB}}{N_{aE}} = \frac{\mu_{nE}}{\mu_{pB}}\,\frac{W}{L_{nE}}\,\frac{N_{dB}}{N_{aE}} = \frac{250}{500}\,\frac{2\times10^{-4}}{1.14\times10^{-3}}\,\frac{5\times10^{16}}{10^{18}} = 4.386\times10^{-3}.$$

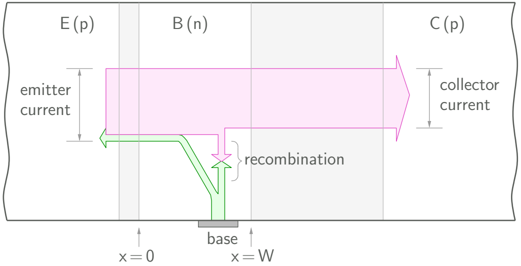
$$\gamma = \frac{1}{1+4.386\times10^{-3}} = 0.9956.$$

$$\alpha_T = \frac{1}{1+\frac{1}{2}(W/L_{pB})^2} = \frac{1}{1+\frac{1}{2}(2.0/35.9)^2} = 0.9985.$$

$$\alpha = \gamma\alpha_T = 0.9940 \rightarrow \beta = \frac{\alpha}{1-\alpha} = 166.$$

Example





(b) With $E \leftrightarrow C$,

$$\frac{D_{nE}}{D_{pB}}\,\frac{W}{L_{nE}}\,\frac{N_{dB}}{N_{aE}} \rightarrow \frac{D_{nC}}{D_{pB}}\,\frac{W}{L_{nC}}\,\frac{N_{dB}}{N_{aC}} = \frac{\mu_{nC}}{\mu_{pB}}\,\frac{W}{L_{nC}}\,\frac{N_{dB}}{N_{aC}} = \frac{1500}{500}\,\frac{2\times10^{-4}}{6.22\times10^{-3}}\,\frac{5\times10^{16}}{10^{15}} = 4.823.$$

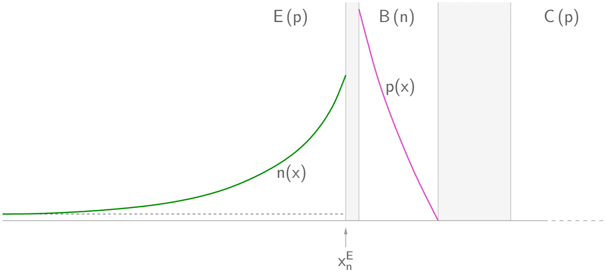
$$\gamma = \frac{1}{1+4.823} = 0.1717,\ \alpha_T = \frac{1}{1+\frac{1}{2}(2/35.9)^2} = 0.9985.$$

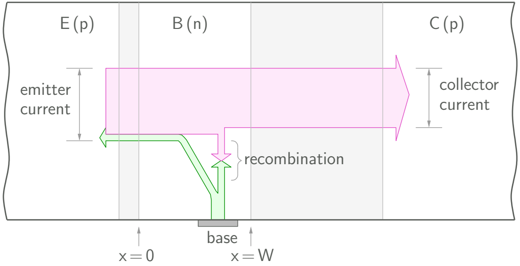
$$\rightarrow \alpha = \gamma\,\alpha_T = 0.1714 \rightarrow \beta = 0.2, \text{ a disaster.}$$

Conclusion: $N_{aE} \gg N_{dB}$ is crucial.

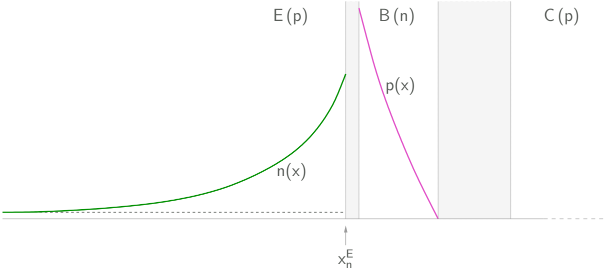
(Note that we have treated W as a constant, but it would vary with bias conditions.)

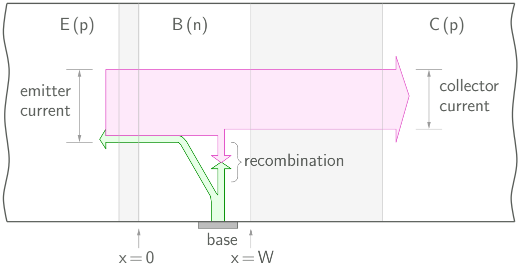
γ with $W << L_{pB}$





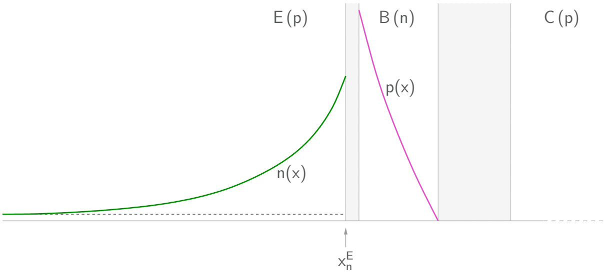
γ with $W << L_{pB}$

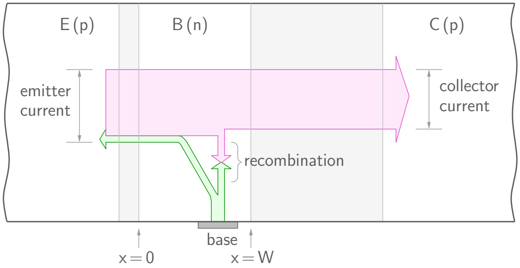




$$I_{nE} = -q\,A\,D_{nE}\,\frac{d}{dx}\left[n_{0E}\exp\left(\frac{V_{EB}}{V_T}\right)\exp\left(-\frac{x_n^E - x}{L_{nE}}\right)\right] \;\text{ at } x = x_n^E$$

γ with $W << L_{pB}$

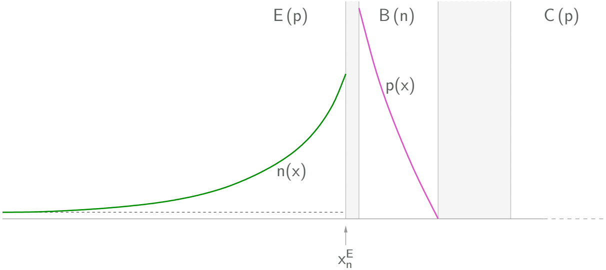


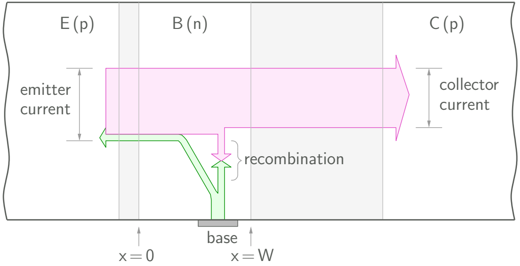


$$I_{nE} = -q\,A\,D_{nE}\,\frac{d}{dx}\left[n_{0E}\exp\left(\frac{V_{EB}}{V_T}\right)\exp\left(-\frac{x_n^E - x}{L_{nE}}\right)\right] \quad \text{at } x = x_n^E$$

$$= q\,A\,D_{nE}\,n_{0E}\exp\left(\frac{V_{EB}}{V_T}\right) \times \frac{1}{L_{nE}}$$

γ with $W << L_{pB}$



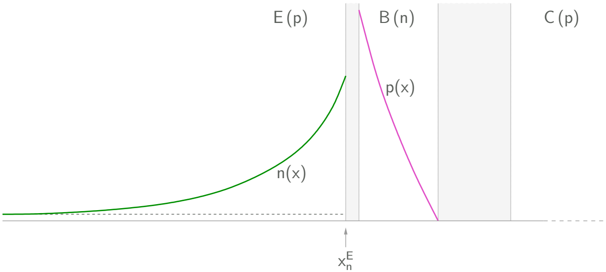


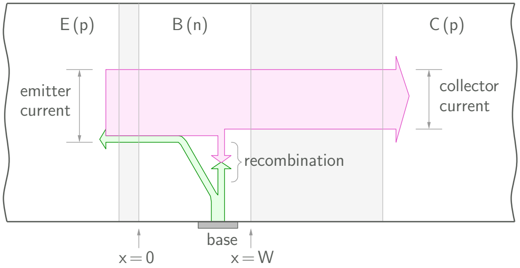
$$I_{nE} = -q\,A\,D_{nE}\,\frac{d}{dx}\left[n_{0E}\exp\left(\frac{V_{EB}}{V_T}\right)\exp\left(-\frac{x_n^E - x}{L_{nE}}\right)\right] \quad \text{at } x = x_n^E$$

$$= q\,A\,D_{nE}\,n_{0E}\exp\left(\frac{V_{EB}}{V_T}\right) \times \frac{1}{L_{nE}}$$

$$I_{pE} \approx q\,A\,D_{pB}\,p_{0B}\exp\left(\frac{V_{EB}}{V_T}\right) \times \frac{1}{W}$$

γ with $W << L_{pB}$





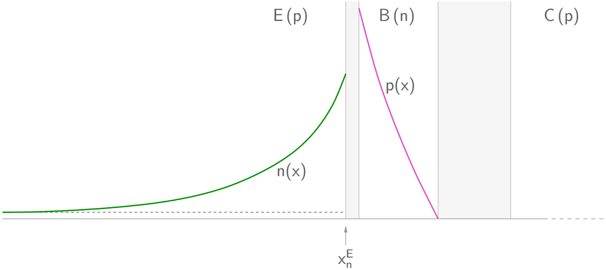
$$I_{nE} = -q\,A\,D_{nE}\,\frac{d}{dx}\left[n_{0E}\exp\left(\frac{V_{EB}}{V_T}\right)\exp\left(-\frac{x_n^E - x}{L_{nE}}\right)\right] \quad \text{at } x = x_n^E$$

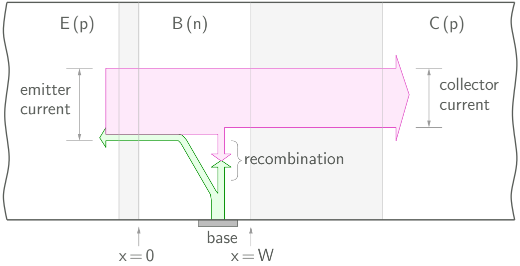
$$= q\,A\,D_{nE}\,n_{0E}\exp\left(\frac{V_{EB}}{V_T}\right)\times\frac{1}{L_{nE}}$$

$$I_{pE} \approx q\,A\,D_{pB}\,p_{0B}\exp\left(\frac{V_{EB}}{V_T}\right)\times\frac{1}{W}$$

$$\frac{I_{nE}}{I_{pE}} \approx \frac{D_{nE}}{D_{pB}}\,\frac{n_{0E}}{p_{0B}}\,\frac{W}{L_{nE}}$$

γ with $W << L_{pB}$





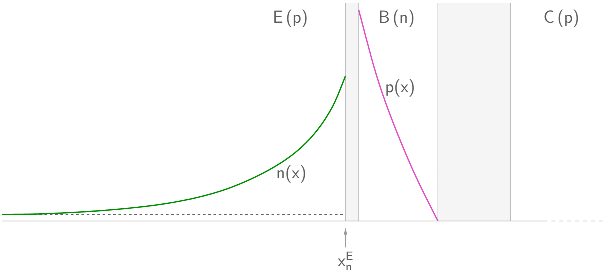
$$I_{nE} = -q\,A\,D_{nE}\,\frac{d}{dx}\left[n_{0E}\exp\left(\frac{V_{EB}}{V_T}\right)\exp\left(-\frac{x_n^E - x}{L_{nE}}\right)\right] \quad \text{at } x = x_n^E$$

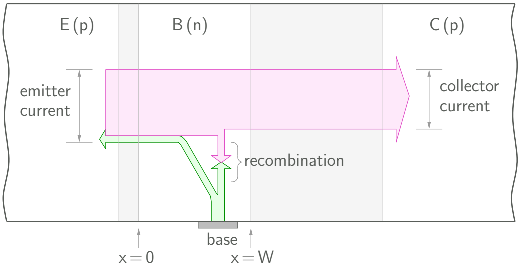
$$= q\,A\,D_{nE}\,n_{0E}\exp\left(\frac{V_{EB}}{V_T}\right) \times \frac{1}{L_{nE}}$$

$$I_{pE} \approx q\,A\,D_{pB}\,p_{0B}\exp\left(\frac{V_{EB}}{V_T}\right) \times \frac{1}{W}$$

$$\frac{I_{nE}}{I_{pE}} \approx \frac{D_{nE}}{D_{pB}}\,\frac{n_{0E}}{p_{0B}}\,\frac{W}{L_{nE}} = \frac{D_{nE}}{D_{pB}}\,\frac{n_i^2}{N_{aE}}\,\frac{N_{dB}}{n_i^2}\,\frac{W}{L_{nE}}$$

γ with $W << L_{pB}$





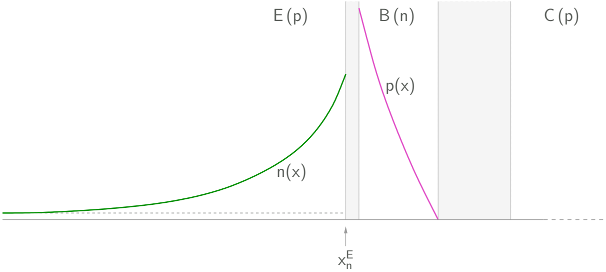
$$I_{nE} = -q\,A\,D_{nE}\,\frac{d}{dx}\left[n_{0E}\exp\left(\frac{V_{EB}}{V_T}\right)\exp\left(-\,\frac{x_n^E - x}{L_{nE}}\right)\right] \quad \text{at } x = x_n^E$$

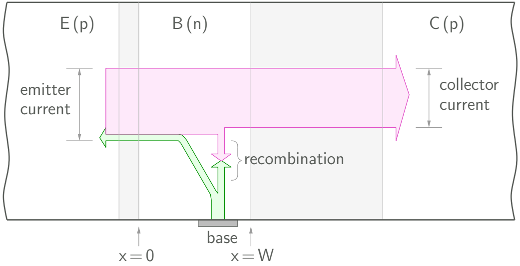
$$= q\,A\,D_{nE}\,n_{0E}\exp\left(\frac{V_{EB}}{V_T}\right)\times\frac{1}{L_{nE}}$$

$$I_{pE} \approx q\,A\,D_{pB}\,p_{0B}\exp\left(\frac{V_{EB}}{V_T}\right)\times\frac{1}{W}$$

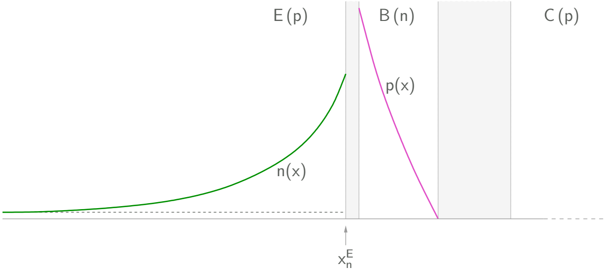
$$\frac{I_{nE}}{I_{pE}} \approx \frac{D_{nE}}{D_{pB}}\,\frac{n_{0E}}{p_{0B}}\,\frac{W}{L_{nE}} = \frac{D_{nE}}{D_{pB}}\,\frac{n_i^2}{N_{aE}}\,\frac{N_{dB}}{n_i^2}\,\frac{W}{L_{nE}} = \frac{D_{nE}}{D_{pB}}\,\frac{N_{dB}}{N_{aE}}\,\frac{W}{L_{nE}}\,.$$

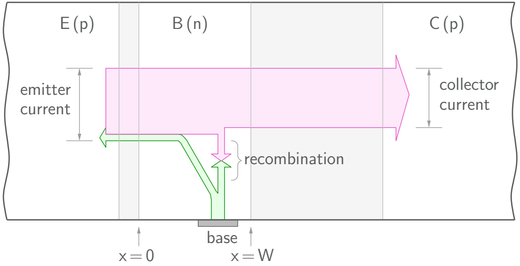
γ with $W << L_{pB}$





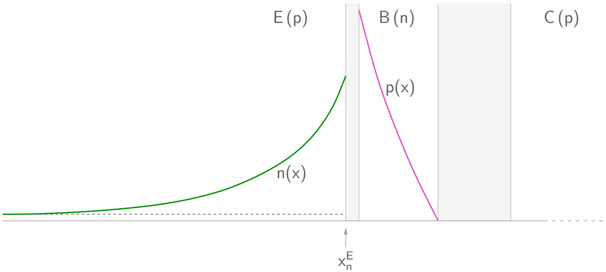
γ with $W << L_{pB}$

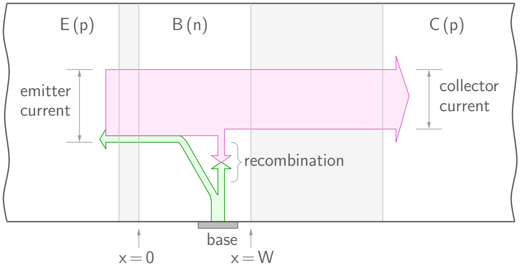




$$\frac{I_{nE}}{I_{pE}} \approx \frac{D_{nE}}{D_{pB}}\,\frac{N_{dB}}{N_{aE}}\,\frac{W}{L_{nE}}$$

γ with $W << L_{pB}$

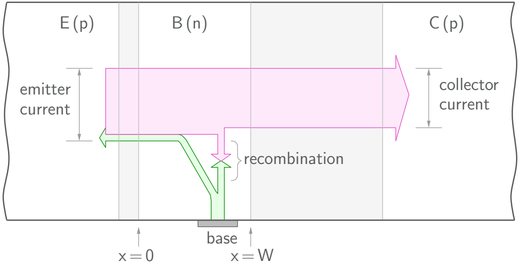


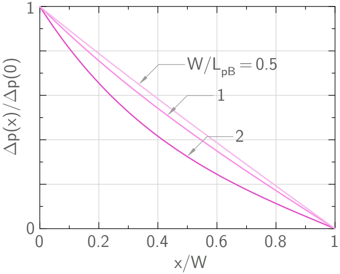


$$\frac{I_{nE}}{I_{pE}} \approx \frac{D_{nE}}{D_{pB}} \frac{N_{dB}}{N_{aE}} \frac{W}{L_{nE}}$$

$$\rightarrow \gamma = \frac{I_{pE}}{I_{pE} + I_{nE}} = \frac{1}{1 + \dfrac{I_{nE}}{I_{pE}}} \approx \frac{1}{1 + \left(\dfrac{D_{nE}}{D_{pB}} \dfrac{W}{L_{nE}} \dfrac{N_{dB}}{N_{aE}} \right)}.$$

α_T with $W \ll L_{pB}$

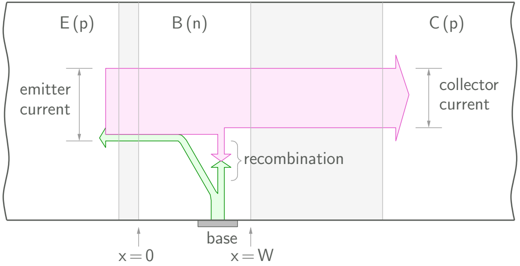


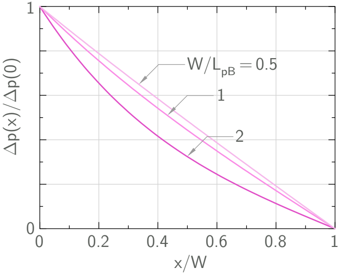


When $W \ll L_{pB}$, $\Delta p(x)$ is linear.

$$\alpha_T = \frac{I_{pC}}{I_{pE}} = \frac{I_{pC}}{I_{pC} + I_{pB}} = \frac{1}{1 + \dfrac{I_{pB}}{I_{pC}}}.$$

α_T with $W \ll L_{pB}$



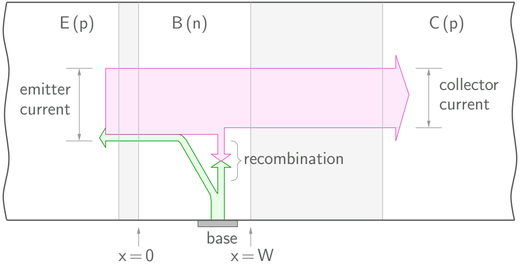


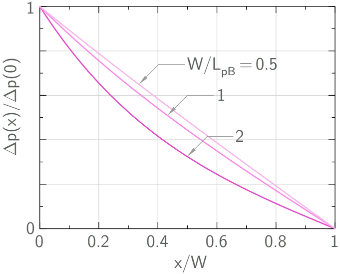
When $W \ll L_{pB}$, $\Delta p(x)$ is linear.

$$\alpha_T = \frac{I_{pC}}{I_{pE}} = \frac{I_{pC}}{I_{pC} + I_{pB}} = \frac{1}{1 + \dfrac{I_{pB}}{I_{pC}}}.$$

$$I_{pC} = -q\,A\,D_{pB}\,\frac{dp}{dx}(W) \approx q\,A\,D_{pB}\,\frac{\Delta p(0)}{W}.$$

α_T with $W \ll L_{pB}$





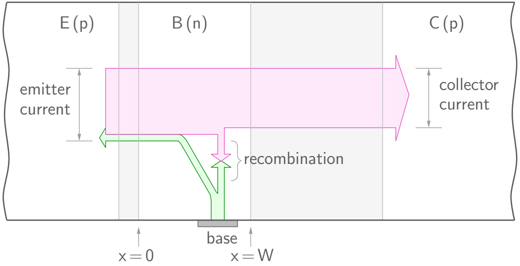
When $W \ll L_{pB}$, $\Delta p(x)$ is linear.

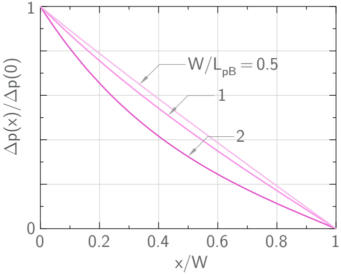
$$\alpha_T = \frac{I_{pC}}{I_{pE}} = \frac{I_{pC}}{I_{pC} + I_{pB}} = \frac{1}{1 + \dfrac{I_{pB}}{I_{pC}}}.$$

$$I_{pC} = -q\,A\,D_{pB}\,\frac{dp}{dx}(W) \approx q\,A\,D_{pB}\,\frac{\Delta p(0)}{W}.$$

$$I_{pB} = \frac{Q_p}{\tau_{pB}} = \frac{q\,A\,\frac{1}{2}\,\Delta p(0)\,W}{\tau_{pB}}$$

α_T with $W \ll L_{pB}$





When $W \ll L_{pB}$, $\Delta p(x)$ is linear.

$$\alpha_T = \frac{I_{pC}}{I_{pE}} = \frac{I_{pC}}{I_{pC} + I_{pB}} = \frac{1}{1 + \dfrac{I_{pB}}{I_{pC}}}.$$

$$I_{pC} = -q\,A\,D_{pB}\,\frac{dp}{dx}(W) \approx q\,A\,D_{pB}\,\frac{\Delta p(0)}{W}.$$

$$I_{pB} = \frac{Q_p}{\tau_{pB}} = \frac{q\,A\,\frac{1}{2}\,\Delta p(0)\,W}{\tau_{pB}} \rightarrow \alpha_T \approx \frac{1}{1 + \dfrac{1}{2}\left(\dfrac{W}{L_{pB}}\right)^2}.$$