1.7 Digital Logic Inverters

Reading Assignment: *pp. 40-48*

Consider the **ideal** digital logic inverter.

Q:

A: HO: The Ideal Inverter

Q:

A:

HO: Noise Margins

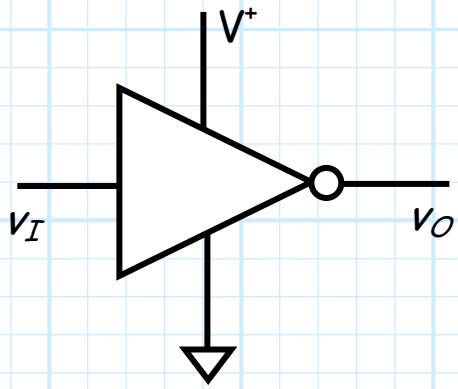
HO: Power Dissipation

HO: Propagation Delay

HO: The Delay-Power Product

The Ideal Inverter

Q: *How would the "**perfect**" inverter behave?*



A: Clearly, if $v_I = 0$, then $v_O = V^+$, and if $v_I = V^+$, then $v_O = 0$.

v_I	v_O
0	V^+
V^+	0

But what happens if v_I is **not** equal to precisely 0.0 or V^+ ??

In other words, what is the **ideal transfer function** $v_O = f(v_I)$ of a digital inverter?

For example, say $V^+ = 5V$. How should the inverter respond to $v_I = 1V$, or $v_I = 2V$, or $v_I = 4V$??

Since v_I = 1V and v_I = 2V are closer to 0.0 V (low level) than they are to 5.0 V (high level), the inverter **should** interpret them as **low** inputs and the output should then be placed precisely at the high state v_O = 5V.

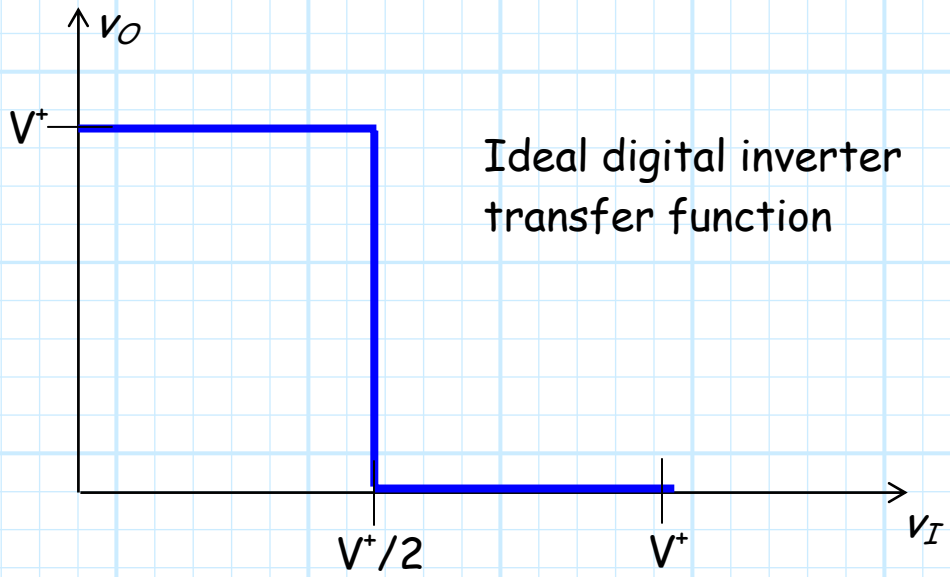
Similarly, for v_I = 4V, the inverter **should** interpret it as **high** input and thus the output should be placed precisely at the low state v_O = 0V.

Therefore, we can say that an **ideal** digital inverter will interpret **input** values **less** than $V^+/2$ (i.e., < 2.5 V) as **low** inputs, and thus produce an **ideal output** of V^+ (i.e., 5.0 V).

Likewise, any **input** values **greater** than $V^+/2$ (i.e., >2.5 V) will be interpreted as a **high** input, and thus an **ideal low** value of 0.0 V will be placed at the output.

Thus, the **ideal transfer function** for a digital inverter is:

$$v_O = \begin{cases} V^+ & \text{if} \quad v_I < V^+/2 \\ 0 & \text{if} \quad v_I > V^+/2 \end{cases}$$

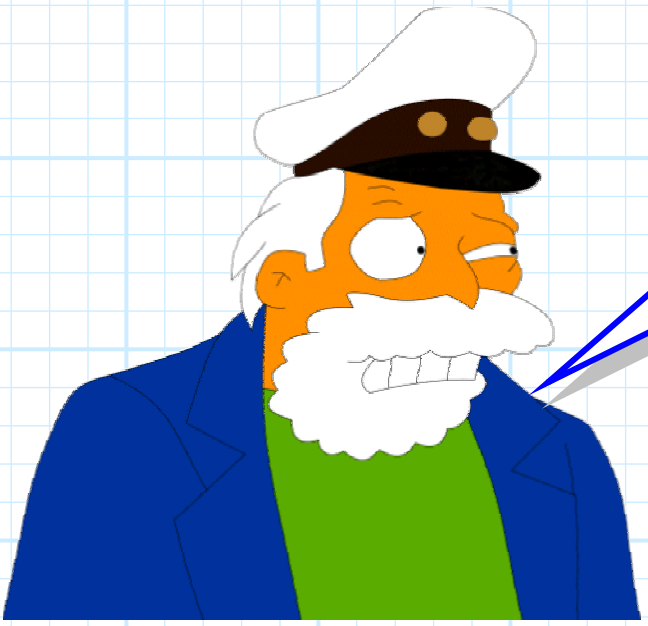


Ideal digital inverter transfer function

By the way, the ideal inverter has **noise margins** (NM) of:

$$NM = \frac{V^+}{2}$$

This is the **largest possible** noise margin, therefore the **ideal** noise margin!

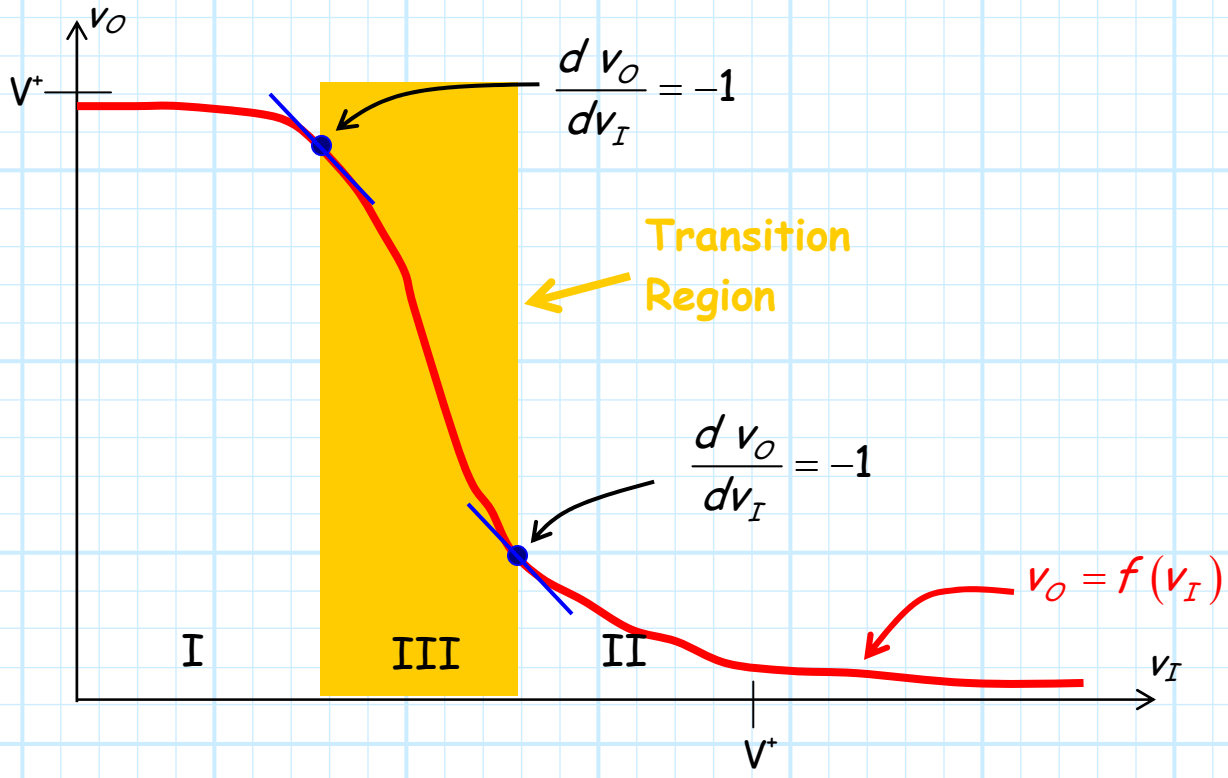


Q: *Arrgh! Just what is exactly is a "noise margin"?*

A: The subject of the **next** handout!

Noise Margins

The **transfer function** of a **digital inverter** will typically look something like this:



Note that there are essentially **three** regions to this curve:

I. The region where v_I is relatively **low**, so that the output voltage v_O is **high**.

II. The region where v_I is relatively **high**, so that the output voltage v_O is **low**.

III. The **transition region**, where the input/output voltage is in an **indeterminate state** (i.e, an **ambiguous** region between high and low.

Note that the **transition region** is rather **arbitrarily** defined by the points on the transfer function where the magnitude of the **slope** is **greater than one** (i.e., where $\left|dv_O/dv_I\right| > 1.0$).

Although this transfer function **looks** rather simple, there are actually **several parameters** that we use to **characterize** this transfer function—and thus characterize the digital inverter **as well**!

1. **First** of all, let's consider the case when v_I=**0**. The **output** of the digital inverter in this condition is **defined** as V_{OH} (i.e., OH → "output high"), i.e.:

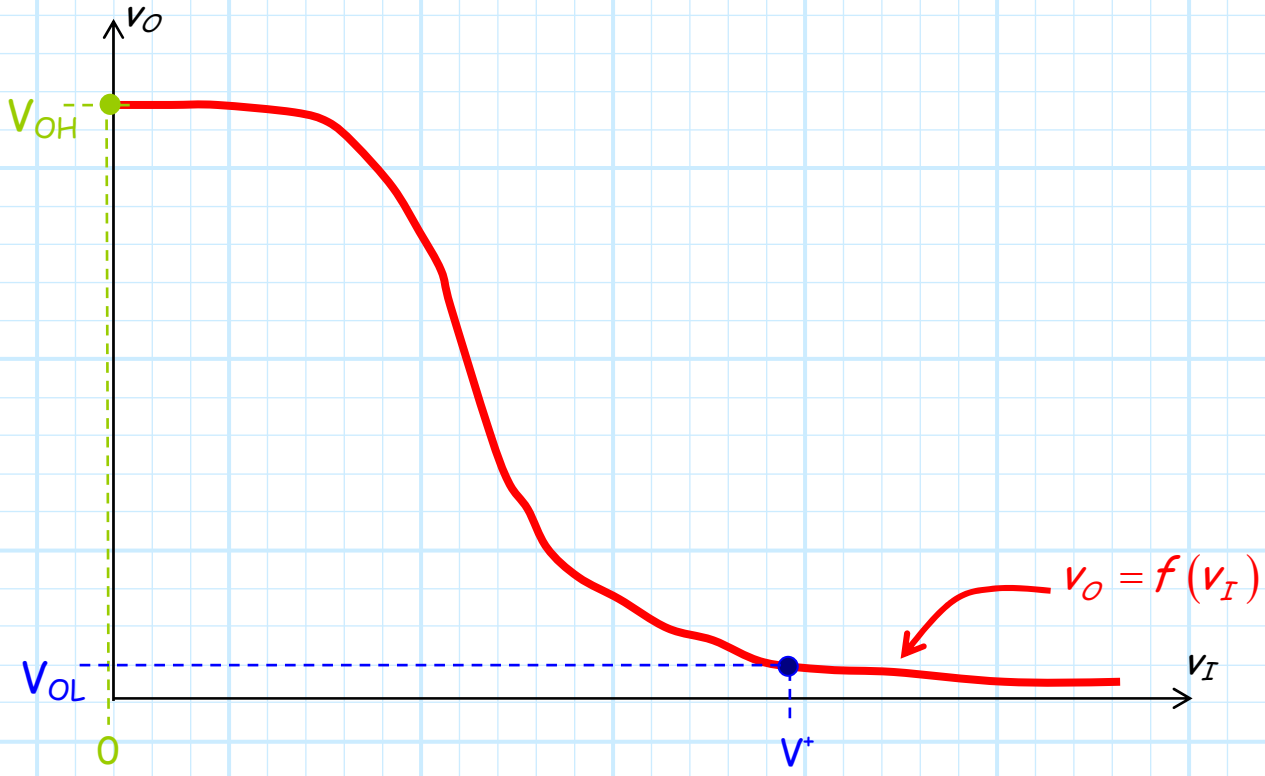
$$V_{OH} \doteq v_O \text{ when } v_I = 0$$

Thus, V_{OH} is essentially the "**ideal**" inverter **high** output, as it is the output voltage when the inverter input is at its ideal low input value v_I=0. Typically, V_{OH} is a value just **slightly** less than supply voltage V^+.

2. Now, let's consider the case when $v_I = V^+$. The output of the digital inverter in this condition is **defined** as V_{OL} (i.e., OL → "output low"), i.e.:

$$V_{OL} \doteq v_O \text{ when } v_I = V^+$$

Thus, V_{OL} is essentially the "**ideal**" inverter **low** output, as it is the output voltage when the inverter input is at its ideal high input value $v_I = V^+$. Typically, V_{OL} is a value just **slightly** greater than 0.



3. The "**boundary**" between region I and the transition region of the transfer function is denoted as $\mathbf{V_{IL}}$ (i.e., IL → "input low"). Specifically, this is the value of the **input** voltage that corresponds to the **first** point on the transfer function where the **slope** is equal to -1.0 (i.e., where $dv_O/dv_I = -1.0$).

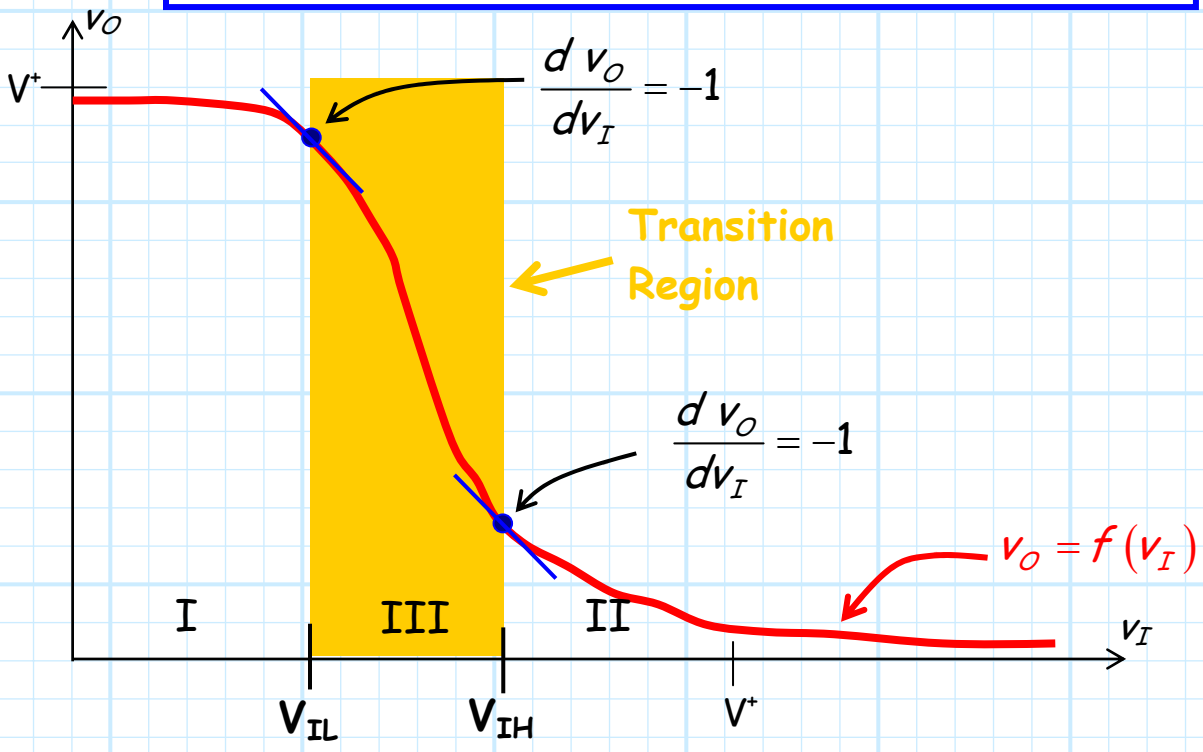
Effectively, the value V_{IL} places an **upper bound** on an acceptably "**low**" value of input v_I—any v_I **greater** than V_{IL} is **not** considered to be a "low" input value. I.E.:

v_I considered "low" only if $v_I < V_{IL}$

4. Likewise, the "**boundary**" between region II and the transition region of the transfer function is denoted as $\mathbf{V_{IH}}$ (i.e., IH → "input high"). Specifically, this is the value of the **input** voltage that corresponds to the **second** point on the transfer function where the **slope** is equal to -1.0 (i.e., where $dv_O/dv_I = -1.0$).

Effectively, the value V_{IH} places a **lower bound** on an acceptably "**high**" value of input v_I—any v_I **lower** than V_{IH} is **not** considered to be a "high" input value. I.E.:

v_I considered "high" only if $v_I > V_{IH}$

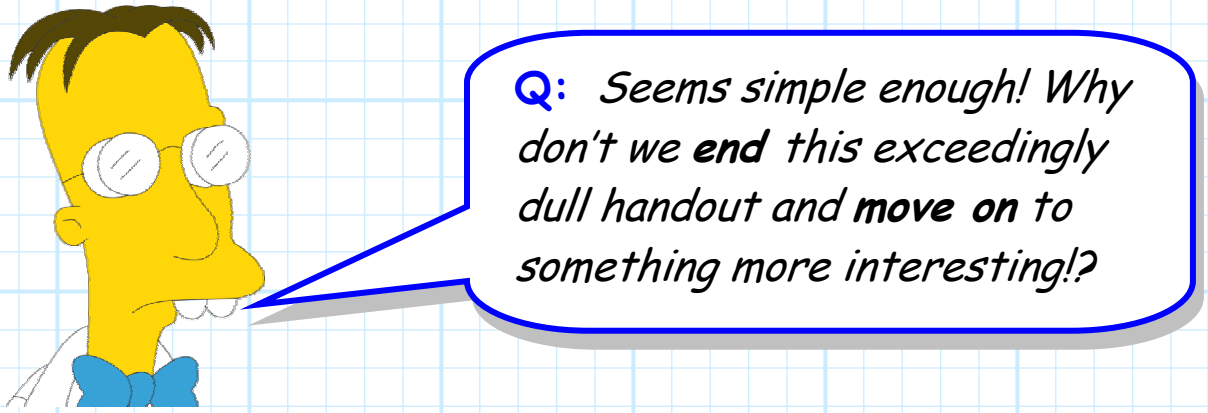


Note then that the **input** voltages of the **transition region** (i.e., $V_{IL} < v_I < V_{IH}$) are **ambiguous** values—we **cannot** classify them as either a digital "low" value or a digital "high" value.

Accordingly, the **output** voltages in the transition region are both significantly less that V_{OH} and significantly larger then V_{OL}. Thus, the **output** voltages that occur in the transition region are **likewise ambiguous** (cannot be assigned a logical state).

Lesson learned → Stay away from the **transition region**!

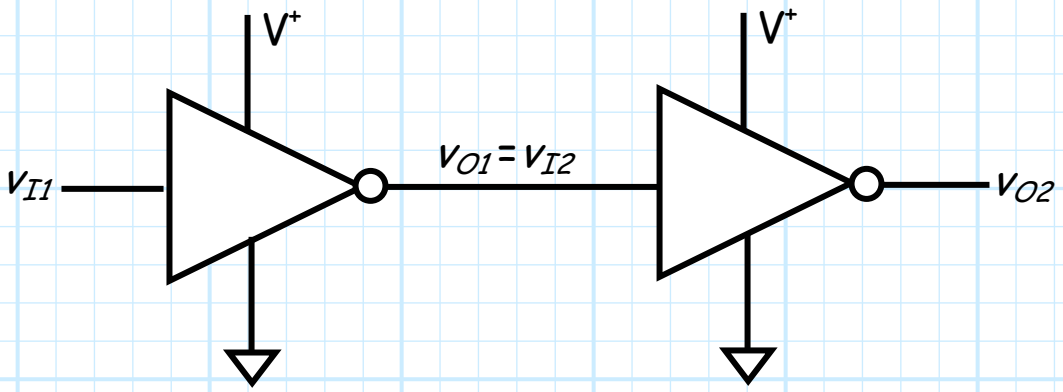
In other words, we must ensure that an **input** voltage representing a logical "**low**" value is **significantly lower** than V_{IL}, and an **input** voltage representing a logical "**high**" value is **significantly higher** than V_{IH}.



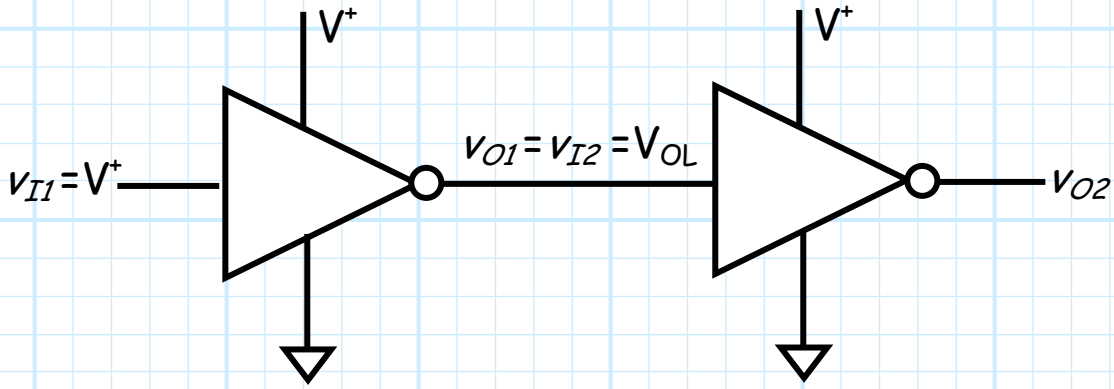
A: Actually, staying **out** of the transition region is sometimes **more difficult** than you might first imagine!

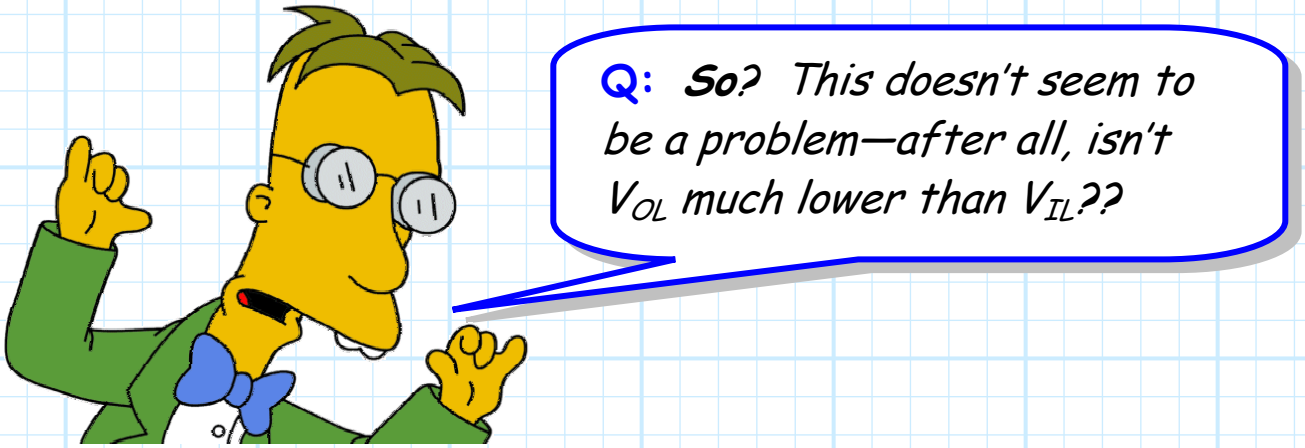
The reason for this is that in a **digital system**, the devices are **connected** together—the input of one device is the output of the other, and vice versa.

For **example**:



Say that the **input** to the **first** digital inverter is $v_{I1} = V^+$. The **output** of that inverter is therefore $v_{O1} = V_{OL}$. Thus, the **input** to the **second** inverter is **likewise** equal to V_{OL} (i.e., $v_{I2} = v_{O2} = V_{OL}$).



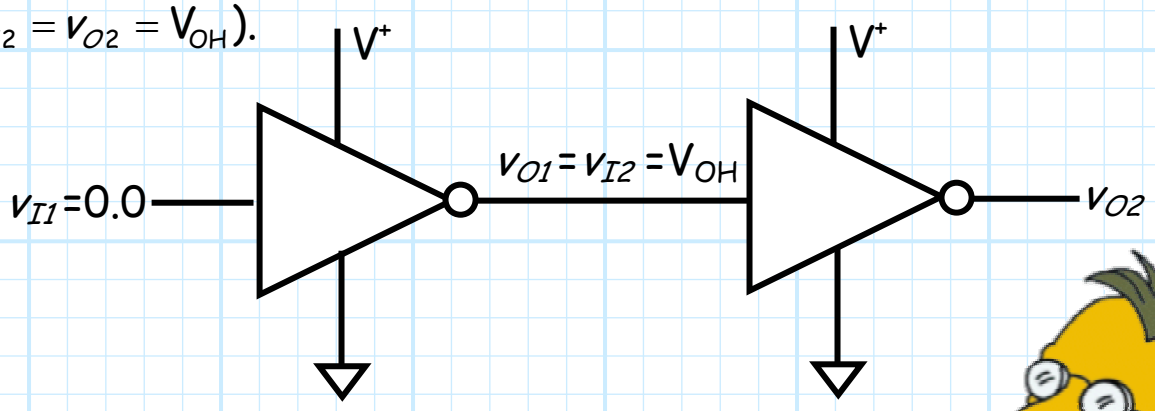


A: True enough! The input $v_{I2}=V_{OL}$ is typically **well below** the maximum acceptable value V_{IL}. In fact, we have a specific name for the **difference** between V_{IL} and V_{OL}—we call this value **Noise Margin** (NM):

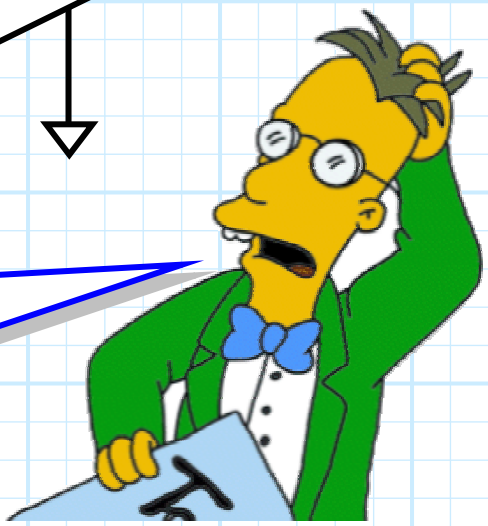
$$NM_L = V_{IL} - V_{OL} \qquad [Volts]$$

The noise margin essentially tells us **how close** we are to the **ambiguous** transition region for a typical case where $v_I = V_{OL}$. Of course, we do **not** wish to be close to this transition region at all, so **ideally** this noise margin is **very large**!

Now, consider the **alternate** case where v_{I1}=0.0 V. The **output** of the **first** inverter is therefore $v_{O1} = V_{OH}$. Thus, the **input** to the **second** inverter is **likewise** equal to V_{OH} (i.e., $v_{I2} = v_{O2} = V_{OH}$).



Q: *This **still** doesn't seem to be a problem—after all, isn't V_{OH} much larger than V_{IH}??*

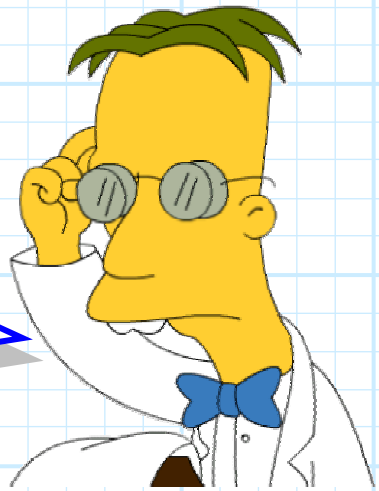


A: Again, this is **true** enough! The input $v_{I2}=V_{OH}$ is typically **well** above the minimum acceptable value V_{IH}. We can **again** specify the **difference** between V_{IH} and V_{OH} as a **noise margin** (NM):

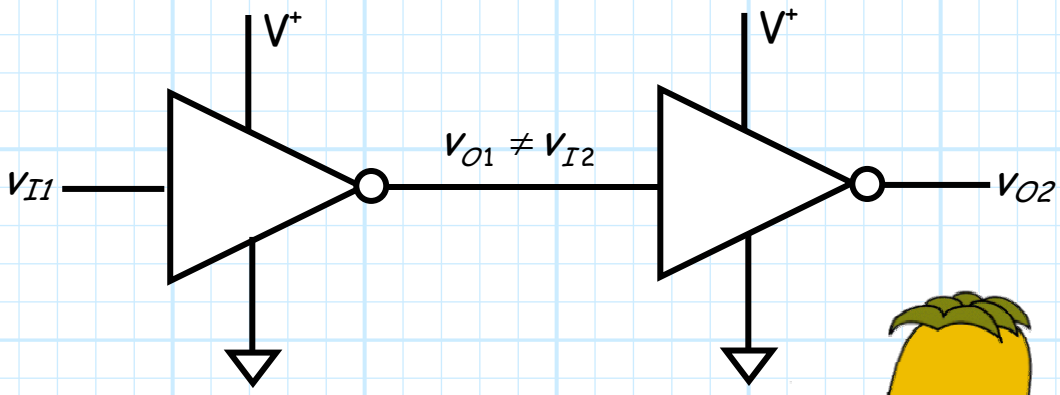
$$NM_H = V_{OH} - V_{IH} \qquad [Volts]$$

This **noise margin** essentially tells us how **close** we are to the ambiguous **transition region** for a typical case where $v_I = 0.0\ V$. Of course, we do **not** wish to be close to this transition region at all, so ideally this noise margin is **very large**!

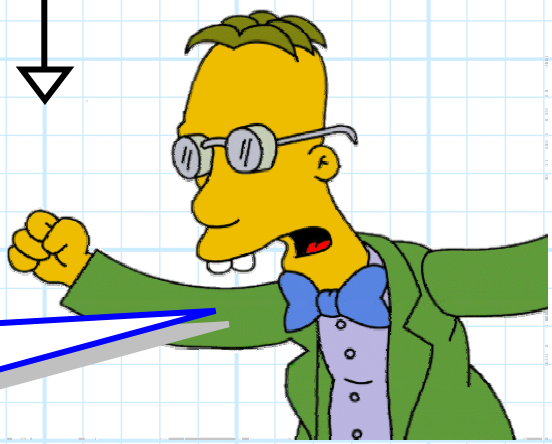
Q: *I don't see **why** we care about the values of these "noise margins". Isn't the simple fact that $V_{OL}<V_{IL}$ and $V_{OH}>V_{IH}$ **sufficient?***



A: **Ideally** yes. However, in our example we have made one important **assumption** that in fact may **not** be true! It turns out that in a **real** digital circuit, v_{I2} may **not** be equal to v_{O1} !!



Q: *What! How can this be possible? It appears to* ***me*** *that* v_{I2} ***must*** *be equal to* v_{O1} *!?*

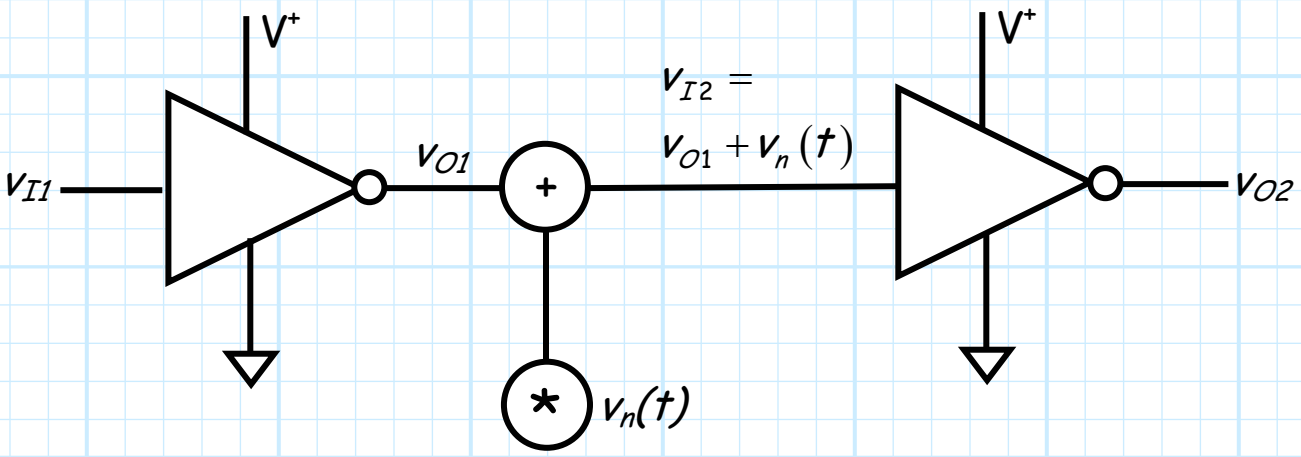


A: It turns out that for a **real** digital circuit, a lot can happen **between** the output of one device and the input to another. The voltage at the input of a device might be affected by **many** sources—**only one** of which is the output device connected to it!

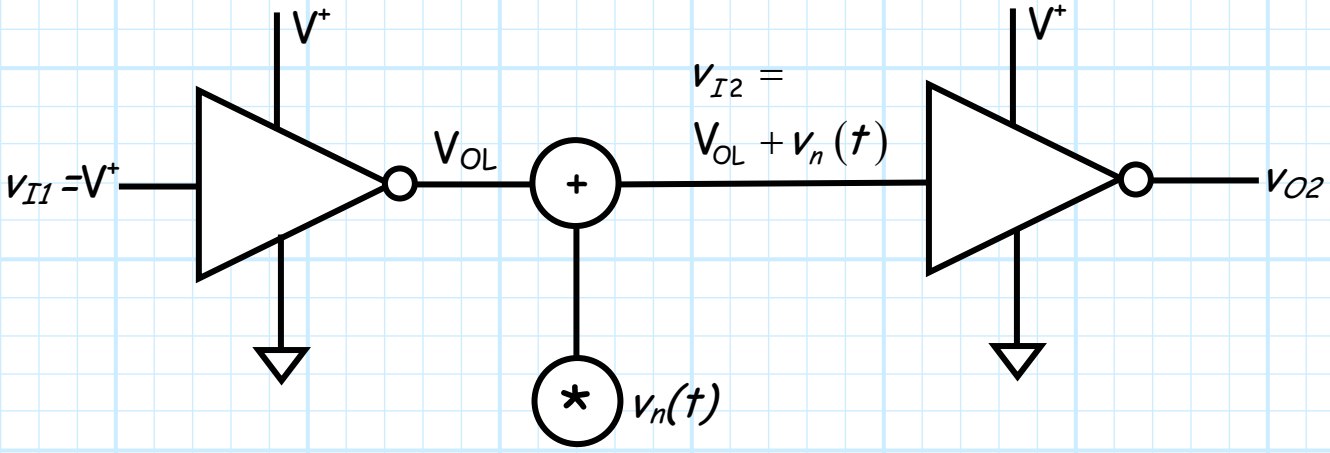
Examples of these "**extra**" sources include:

1. Thermal **noise**
2. Coupled signals
3. Power supply **transients**

We will **combine** the effect of **all** of these sources together into one "**noise**" source $v_n(t)$. Thus, a **better model** for our digital circuit example is:



Now, let's **reconsider** the case where $v_{I1} = V^+$. We find that the **input** to the **second** digital inverter is then $v_{I2} = V_{OL} + v_n(t)$:



Now we see the **problem**! If the **noise** voltage is **too large**, then the **input** to the second inverter will **exceed** the maximum low input level of V_{IL}—we will have entered the dreaded **transition region**!!!!

To **avoid** the transition region, we find that the **input** to the **second** inverter must be less than V_{IL}:

$$V_{OL} + v_n(t) < V_{IL}$$
$$v_n(t) < V_{IL} - V_{OL}$$
$$v_n(t) < NM_L$$

Look at what this means! It says to avoid the transition region (i.e., for the input voltage to have an unambiguously "low" digital level), the **noise** must be **less** than **noise margin** NM_L for **all time** t!

Thus, if the **noise margin** NM_L is **large**, the noise $v_n(t)$ can be large **without** causing any deleterious effect (deleterious effect→transition region). Conversely, if the noise margin NM_L is **small**, then the noise **must** be small to avoid **ambiguous** voltage levels.

Lesson learned → **Large noise margins are very desirable**!

Considering **again** the example circuit, only this time with v_I=0.0 V, we find that to **avoid** the transition region (verify this for yourself!):

$$V_{OH} + v_n(t) > V_{IH}$$
$$v_n(t) > V_{IH} - V_{OH}$$
$$v_n(t) > -NM_H$$
$$-v_n(t) < NM_H$$

Note that the noise $v_n(t)$ is **as likely** to be positive as negative—it is in fact **negative** valued noise that will send v_{I2} to a value **less** than V_{IH}!

Thus, we can make the statement that the **magnitude** of the **noise** $v_n(t)$ must be **less** than the **noise margins** to avoid the ambiguous values of the disturbing **transition region**! I.E., make sure that:

$$|v_n(t)| < NM \qquad \text{for all time } t$$

Gate Power Dissipation

Every digital gate will require some amount of **power**. It must **dissipate** this power in the form of heat. We consider **two** types of power:

Static P_D - Power dissipated when gate **is not** changing state.

Dynamic P_D - Power dissipated when gate **is** changing states.

Typically we find that:

$$\text{Dynamic } P_D \geq \text{Static } P_D$$

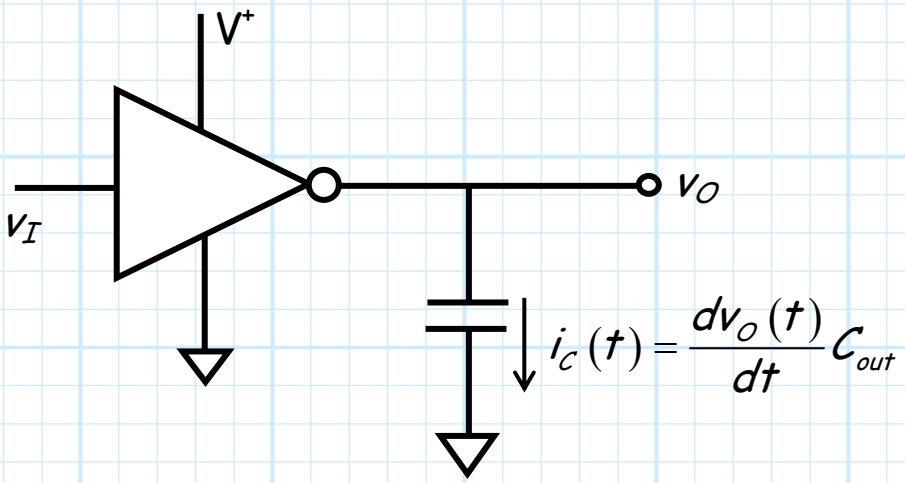
In fact, for CMOS logic gates (e.g., a CMOS digital inverter), we will find that the **static** P_D is nearly **zero**!

However, we will find that it **always** takes some power to **change** the output state of a digital logic gate.

Q: *Why is that??*

A: **Capacitance**! When we change the **voltage** at the output, we must charge or discharge the **capacitance** associated with the output. This requires **current**!

Although engineers work very hard to **minimize** the output capacitance of digital circuits, we **cannot** fully eliminate it.



Note that the **displacement current** $i_C(t)$ flowing "through" this capacitor is proportional to the **time derivative** of the output voltage.

* Thus, if the output state is **static** (i.e., it is not changing), this derivative is **zero**, as is the current.

* However, if the output is **dynamic** (i.e. it is changing with time), the derivative is **non-zero**, and thus displacement current **flows** in the capacitor.

* Note that the **faster** we change the output, the **more** displacement current is produced, meaning **more power** is required!

* Thus, we come to the (correct) conclusion that **dynamic** power dissipation is **dependent** on the speed (e.g., clock frequency) of the digital logic—the "**faster**" the logic, the **higher** the dynamic power dissipation!

* As a result, dynamic P_D is typically **specified** as a function of **frequency**.

Gate Propagation Delay

Say the **input** to a logic gate changes its state (e.g., 0 to V^+, or V^+ to 0). The **output** of the gate will likely change state as a result.

However, the output will **not change instantaneously** when the input changes. Instead, the output will change **after** a small **delay**.

We call this delay the **propagation delay**. Ideally, this delay is as **small** as possible; typically, it is on the order of a few **nanoseconds** or less.

Often, the delay when the output changes from **low to high** is a **different value** than the delay when the output changes from **high to low**. Therefore, we can define:

t_{pHL} = delay for output changing from high to low

t_{pLH} = delay for output changing from low to high

We can therefore define the **propagation delay** t_p as the **average** of these values:

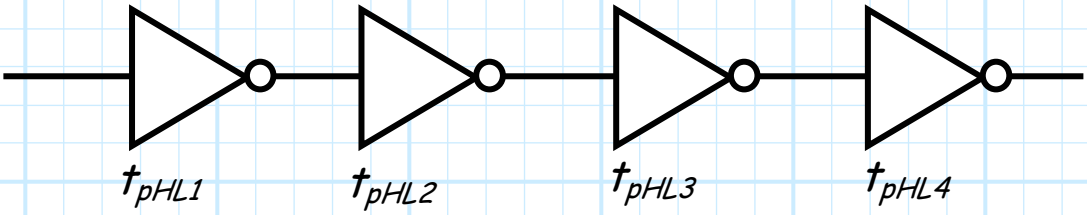
$$t_p = \frac{t_{pHL} + t_{pHL}}{2}$$

Q: *Why does this delay occur?*

A: Again, the reason is **output capacitance**!

It takes a **non-zero** amount of time to **charge** or **discharge** a capacitor. In other words, the output voltage **cannot** change instantaneously to a change in the input.

Propagation delay is a particularly disturbing **problem** when we construct a **complex** digital circuit consisting of **many** interconnecting stages. For example:



The **total** propagation delay for this complex digital circuit is therefore:

$$t_p = t_{pHL1} + t_{pHL2} + t_{pHL3} + t_{pHL4}$$

Thus, although the propagation delay of **one** individual logic gate may be insignificant, the **total delay** through a complex digital circuit consisting of many stages and gates can be quite **large**!

This can cause **big problems** in the precise timing required of sophisticated and complex digital systems!

Delay-Power Product

Q: *So, an ideal digital technology would have BOTH very small propagation delay t_p, AND very small power dissipation P_D, right ?*

A: True ! But, there is a problem. Designing a "faster" (e.g., lower t_p) digital gate usually requires **greater power**. And designing a gate to minimize power consumption usually **slows down** the digital device.

Propagation delay and power dissipation generally form a **design trade off** - improve one and you degrade the other !

To quantify how effective, or efficient a digital design technology is in terms of delay and power, we use the product of the **propagation delay** and the **power dissipation**:

The **delay-power product** (DP) !

The delay-power product is therefore **defined** as:

$$DP = P_D \; t_p$$

* Note we could define either a **static or dynamic** delay-power product, depending on P_D.

* Note also the **unit** of the delay-power product—**Joules** !!