References and bias

Carsten Wulff, carsten@wulff.no

Keywords: VREF, IREF, VD, BGAP, LVBGAP, VI, GMCELL

Status: 1.0

In our SPICE testbenches, and trial schematics, it's common to include voltage sources and current sources, like the symbols in Figure 1.

The ideal voltage source, or ideal current source, does not exist in the real world. There is no such thing.

We can come close to creating a voltage source, a known voltage, with a low source impedance, but not zero impedance. And it won't be infinitely fast either. If we suddenly decide to pull 1 kA from a lab supply I promise you the voltage will drop.

How do we create something that is a *good enough* voltage and current source on an IC? That's the goal of this chapter. To give you an introduction to "voltage sources" and "current sources" that we can make on an integrated circuit.

But before we take a take a look at the voltage and current source, I want you to think about how you would route a current, or a voltage on an IC.

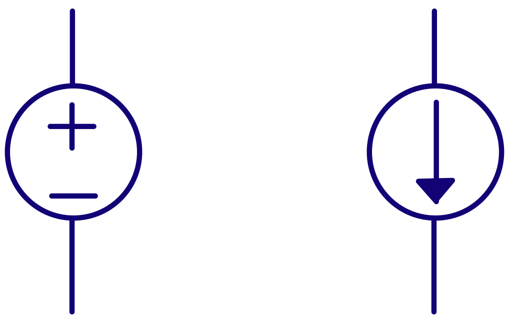


Figure 1: Symbols for voltage source and current source

I. Routing

Assume we have a known voltage on our IC, a reference voltage. How can we make sure we can share that voltage across an IC?

A voltage is only defined between two points. There is no such thing as the *voltage at a point on a wire*, nor *voltage in a node*. Yes, I know we say that, but it's not right. What we forget is that by *voltage in a node* we always, always mean *voltage in a node referred to ground*.

We've invented this magical place called *ground*, the final resting place of electrons, and we have agreed that voltages refer to that point.

As such, when we say "Voltage in node A is 1V", what we actually mean is "Voltage in node A is 1 V referred to ground".

Maybe you now understand why we can't just route a voltage across the IC, the *other side* might not have the same ground. The *other side* might have a different impedance to ground, and the impedance might be a function of time, voltage, frequency temperature, pressure and presence of gremlins.

Consider Figure 2. The ground impedance may depend on time, voltage, frequency, temperature, pressure (yes, stress in silicon can change the band structure, thus the conduction band energy levels, and thus the available charge carriers in the conduction band).

If there is no current flowing in the ground impedance at the destination we may be OK, but usually, there is some current flowing into ground at the destination. There is a circuit there.

If we choose to route a reference as a voltage we need to be careful with the ground.

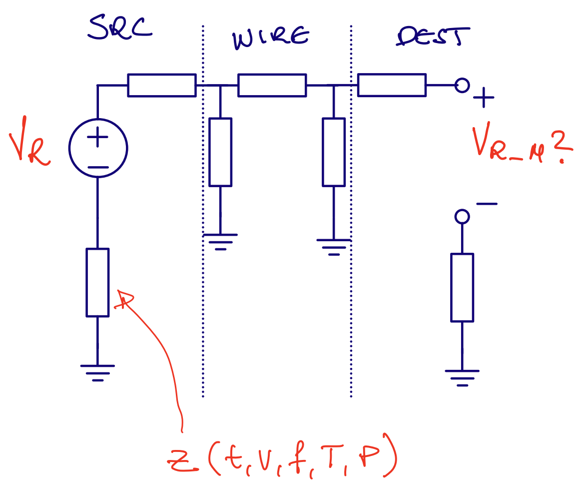


Figure 2: Voltage source with ground impedance. Routing long distances it's not possible to have guarantee we have the same ground impedance at the destination.

Most of the time, in order not to think about the ground impedance, we choose to route a known quantity, the reference, as a current instead of a voltage. That means, however, we must convert from a voltage to a current, but we can do that with a resistor (you'll see later), and as long as the resistor is the same on the other side of the IC, then we'll know what the voltage is.

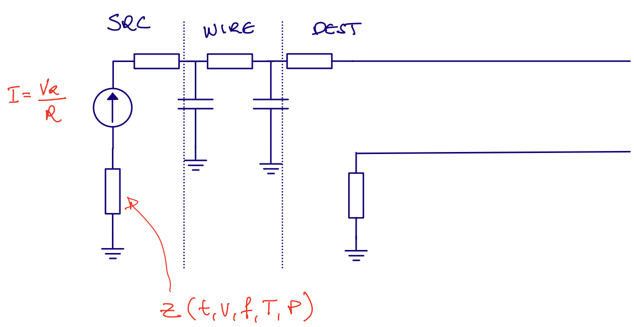


Figure 3: Routing a reference as a current.

Resistors have finite matching across die, let's say 2 % 3-sigma variation. A limitation on how accurate we can distribute reference across the IC with current method.

For most voltage regulators (think about the circuit that delivers the digital voltage for an MCU) 2 % percent may be an acceptable portion of the error budget. For a battery charger, however, the termination voltage of Li-ion batteries need to be precise, more accurate than 1 %.

For that application we cannot distribute current, we must distribute voltage, but we need to care deeply about ground.

But how can "It's better to distribute a voltage as a current across the IC, it's more accurate" and "If you need something really accurate, you must distribute voltage" both be true?

Imagine I have a 0.5 % 3-sigma accurate voltage reference at 1.22 V, that's a sigma of 2 mV. I need this reference voltage on a block on the other side of the IC, I don't want to distribute voltage, because I don't know that the ground is the same on the other side, at least not to a precision of 2 mV. I convert the voltage into a current, however, I know the R has a 2 % 3-sigma across die, so my error budget immediately increases to 2.06%.

But what if I must have 0.5 % 3-sigma voltage in the block? For example in a battery charger, where the 4.3 V termination voltage must be 1 % accurate? I have no choice but to go with voltage directly from the reference, but the key point, is then the receiving block **cannot** be on the other side of the IC. The reference must be right next to my block.

I could use two references on my IC, one for the ADC and one for the battery charger. Ask yourself, "Why do we care if there is two references?" And the answer is "Silicon area is expensive, to make things cheep, we must make things small", in other words, we should not duplicate features unless we absolutely have to.

II. Bandgap voltage reference

One of the ways to create a known reference on an integrated circuit is the "bandgap voltage reference". There are flavors of bandgaps, but all rely on the bandgap of silicon, which is about 1.12 eV.

We can't access the bandgap voltage directly, but we can use the fact that diodes, and BJTs all have a voltage across the PN junction of about 1.12 V at absolute zero (actually, slightly higher, maybe 1.2 V), and that they have a well known temperature dependence from that point.

A. A voltage complementary to temperature (CTAT)

A diode connected bipolar transistor, shown in Figure 4, or indeed a PN diode, assuming a fixed current, will have a voltage across that is temperature dependent

$$I_D = I_S \left(e^{\frac{V_{BE}}{V_T}} - 1 \right) + I_B \approx I_S e^{\frac{V_{BE}}{V_T}}$$

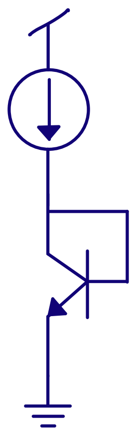


Figure 4: Diode connected bipolar transistor

As I_S is much smaller that I_D we can ignore the -1, and we assume that the base current is much smaller than the drain current.

Re-arranging for V_{BE} and inserting for

$$V_T = \frac{kT}{q}$$

$$V_{BE} = \frac{kT}{q} \ln \frac{I_C}{I_S}$$

$$I_S = qAn_i^2 \left[\frac{D_n}{L_n N_A} + \frac{D_p}{L_p N_D} \right]$$

From this equation, it looks like the voltage V_{BE} is proportional to temperature, however, it turns out that the V_{BE} decreases with temperature due to the temperature dependence of I_S.

The V_{BE} is almost linear with temperature with a property that if you extrapolate the V_{BE} line to zero Kelvin, then all diode voltages seem to meet at the bandgap voltage of silicon (approx 1.12 eV).

To see the temperature coefficient, I find it easier to re-arrange the equation above.

Some algebra (see Diodes)

$$V_{BE} = \frac{kT}{q}(\ell - 3\ln T) + V_G$$

The ℓ is a temperature independent constant given by

$$\ell = \ln I_C - \ln qA - \ln\left[\frac{D_n}{L_n N_A} + \frac{D_p}{L_p N_D}\right]$$
$$-2\ln 2 - \frac{3}{2}\ln m_n^* - \frac{3}{2}\ln m_p^* - 3\ln\frac{2\pi k}{h^2}$$

And if we plot the diode voltage, we can see that the voltage decreases as a function of temperature.

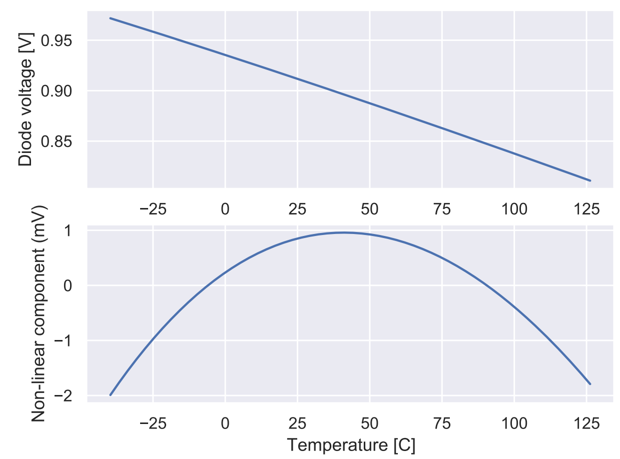


Figure 5: Diode voltage versus temperature. Bottom plot shows deviation from a straight line.

B. A current proportional to temperature (PTAT)

If we take two diodes, or bipolars, biased at different current densities, as shown in Figure 6, then

$$V_{D1} = V_T \ln\frac{I_D}{I_{S1}}$$

$$V_{D2} = V_T \ln\frac{I_D}{I_{S2}}$$

The OTA will force the voltage on top of the resistor to be equal to V_{D1}, thus the voltage across the resistor R_1 is

$$V_{D1} - V_{D2} = V_T \ln\frac{I_D}{I_{S1}} - V_T \ln\frac{I_D}{I_{S2}} = V_T \ln\frac{I_{S2}}{I_{S1}} = V_T \ln N$$

This is a remarkable result. The difference between two voltages is only defined by Boltzmann's constant, temperature, charge, and a know size difference.

This differential voltage can be used to read out directly the temperature on an IC, provided we can compare to a known voltage.

We often call this voltage ΔV_D or ΔV_{BE}, and we can see it's proportional to absolute temperature.

We know that the V_D decreases linearly with temperature, so if we combined a multi-plum of the ΔV_{BE} with a V_D voltage, then we should get a constant voltage.

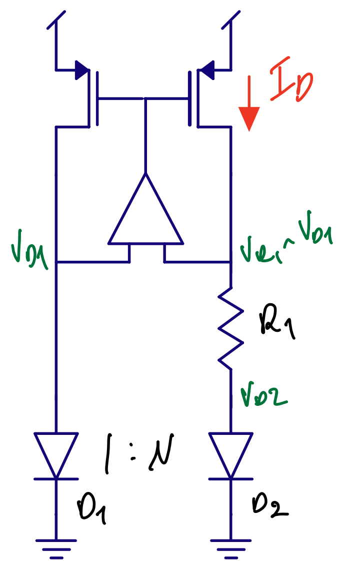


Figure 6: Circuit to create a PTAT current controlled by the resistor and ΔV_{BE}

C. How to combine a CTAT with a PTAT ?

One method is Figure 7. The voltage across resistor R_2 would compensate for the decrease in V_{D3}, as such, R_2 would be bigger than R_1.

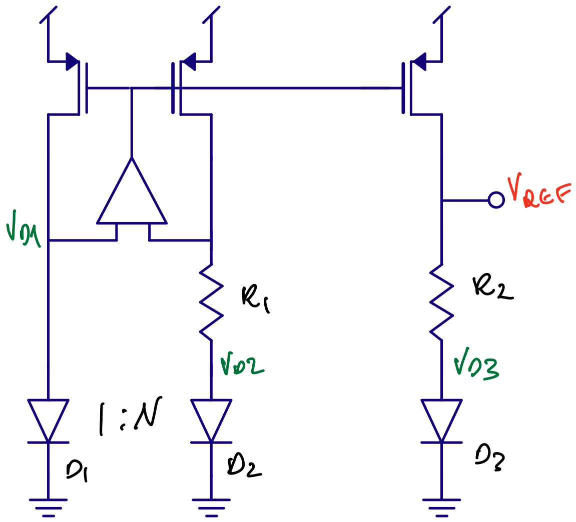


Figure 7: A bandgap voltage reference with a constant output voltage.

Another method would be to stack the R_2 on top of R_1 as shown in Figure 8.

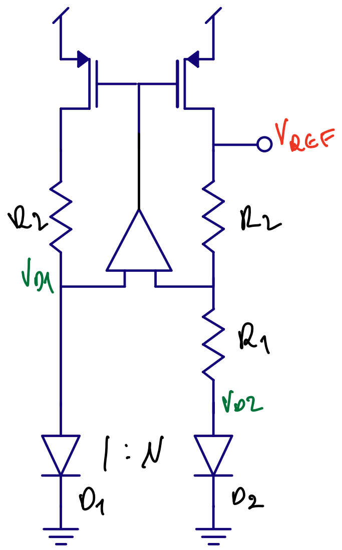


Figure 8: Another bandgap voltage reference with a constant output voltage.

D. Brokaw reference

Paul Brokaw was a pioneer within reference circuits (I met him once in the restroom queue in Tropisueno behind the Marriot hotel in SF during ISSCC). Below is the Brokaw reference, which I think was first published in [A simple three-terminal IC bandgap reference](#).

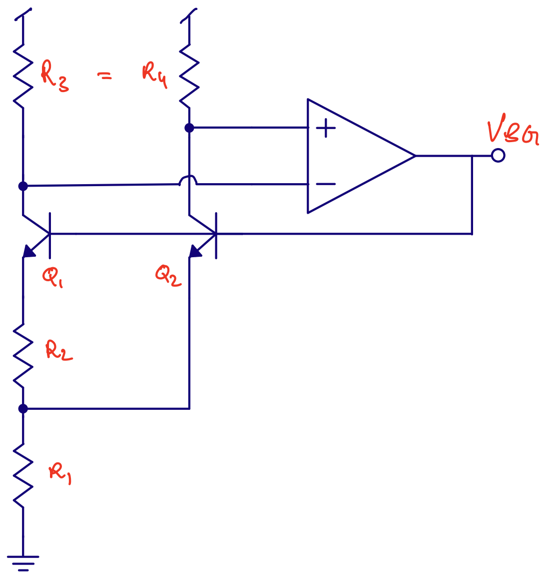


Figure 9: Brokaw bandgap voltage reference

The opamp ensures the two bipolars have the same current. Q_1 is larger than Q_2. The ΔV_{BE} is across the R_2, so we know the current I. We know that R_1 must then have $2I$.

The voltage at the output will then be.

$$V_{BG} = V_{G0} + (m-1)\frac{kT}{q}\ln\frac{T_0}{T} + T\left[\frac{k}{q}\ln\frac{J_2}{J_1}\frac{2R2}{R1} - \frac{V_{G0} - V_{be0}}{T_0}\right]$$

where V_{G0} is the bandgap, V_{be0} is the base emitter measured at a temperature T_0 and the J's are the current densities.

To get a constant output voltage, the relationship between the resistors should be approximately

$$\frac{R2}{R1} = \frac{V_{G0} - V_{be0}}{2T_0\frac{k}{q}\ln(\frac{J_2}{J_1})}$$

In typical simulations, the variation can be low over the temperature range. The second order error is the remaining error from

$$V_{BG} = V_{G0} + (m-1)\frac{kT}{q}\ln\frac{T_0}{T} + T\left[\frac{k}{q}\ln\frac{J_2}{J_1}\frac{2R2}{R1} - \frac{V_{G0} - V_{be0}}{T_0}\right]$$

Where the last term is zero, so

$$V_{BG} = V_{G0} + (m-1)\frac{kT}{q}\ln\frac{T_0}{T}$$

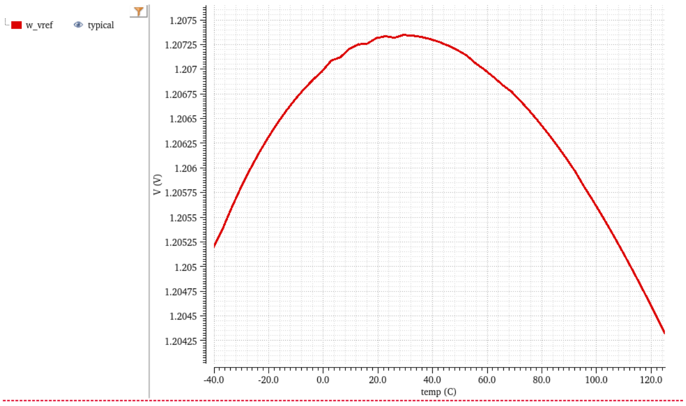


Figure 10: Simulation of a Brokaw reference in GF 130 nm.

Over corners, I do expect that there is variation, as we can see from Figure 11. It may be that the V_D modeling is not perfect, which means the cancellation of the last term is incomplete.

We could include trimming of PTAT to calibrate for the remaining error, however, if we wanted to remove the linear gradient, we would need a two point temperature test of every IC, which too expensive for low-cost devices.

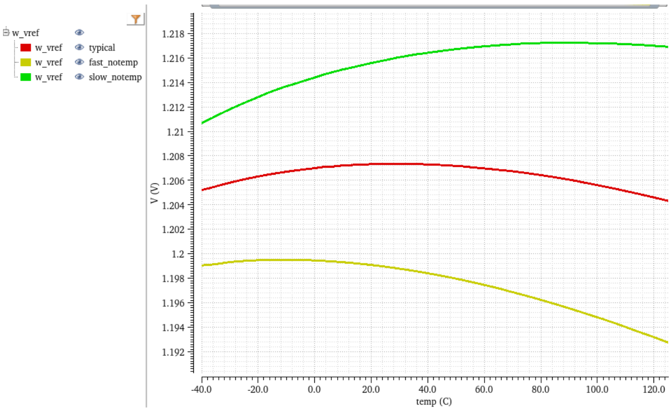


Figure 11: Typical, Slow, Fast simulation of the Brokaw bandgap

E. Low voltage bandgap

The Brokaw reference, and others, have a 1.2 V output voltage, which is hard to make if your supply is below about 1.4 V. As such, people have investigated lower voltage references. The original circuit was presented by Banba A CMOS bandgap reference circuit with sub-1-V operation

In real ICs though, you should ask yourself long and hard whether you really need these low-voltage references. Most ICs today still have a high voltage, either 1.8 V or 3.0 V.

If you do need them, consider the circuit in Figure 12. We have two diodes at different current densities. The ΔV_D will be across R_1. The voltage at the input of the OTA will be V_D and the OTA will ensure the both inputs are equal.

The current will then be

$$I_1 = \frac{\Delta V_D}{R_1}$$

and we know the current increases with temperature, since ΔV_D increases with temperature.

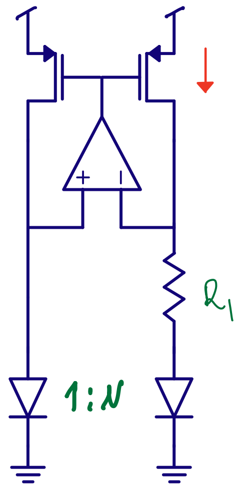


Figure 12: PTAT current generator

I use ΔV_{BE} and ΔV_D interchangeably, appologies.

In Figure 13 we copy the V_D to another node, and place it across a second resistor R_2.

The current in this second resistor is then

$$I_2 = \frac{V_D}{R_2}$$

and we know the current decreases with temperature, since V_D decreases with temperature.

From before, we know the current in R_1 is proportional to temperature. As such, if we combine the two current with the correct proportions, then we can get a current that does not change with temperature.

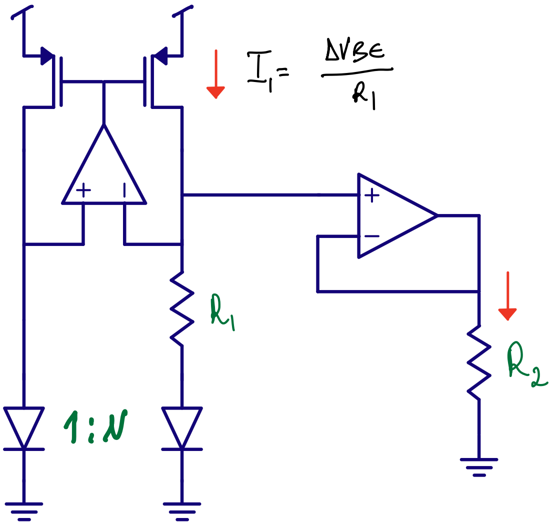


Figure 13: Extending the PTAT current generator

Let's remove the OTA, and connect R_2 directly to V_D nodes, as shown in Figure 14.

You should convince yourself of the fact that this does not change I_1.

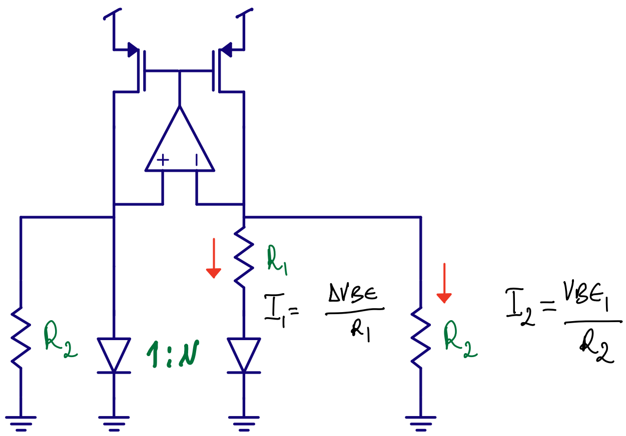


Figure 14: The Banba bandgap voltage reference core

It does, however, change the current in the PMOS. Provided we scale R_2 correctly, then the PTAT I_1 can compensate for CTAT I_2, and we have a current that is independent of temperature.

$$I_{PMOS} = \frac{V_D}{R_2} + \frac{\Delta V_D}{R_1}$$

Assuming we copy the current into another resistor R_3, as shown in Figure 15, we can get a voltage that is

$$V_{OUT} = R_3 \left[\frac{V_D}{R_2} + \frac{\Delta V_D}{R_1}\right]$$

We can choose the output voltage freely, and it be lower than 1.2 V.

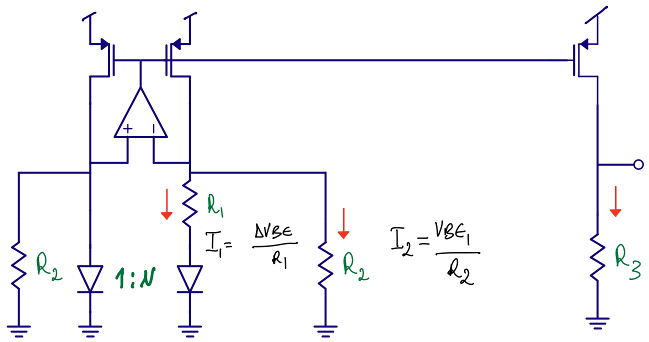


Figure 15: The Banba bandgap voltage reference

III. Bias

Sometimes we just need a current

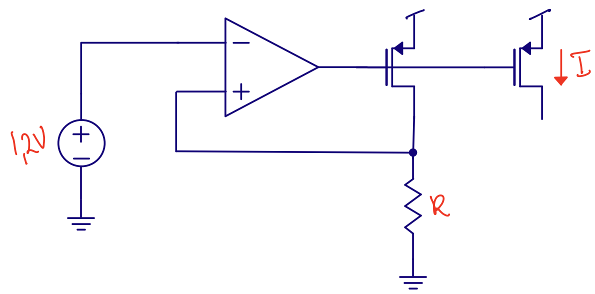
A. Voltage to current conversion

With a known voltage, we can convert to a known current with the circuit in Figure 16.

On-chip we don't have accurate resistors, but for bias currents, it's usually ok with $+-20$ variation (the variation of R).

Across a IC, we can expect the resistors to match within 2 % percent, as such, we can recreate a voltage with a accuracy of about 2 %percent difference from the original if we have a second resistor on the other side of the IC.

If we wanted to create an accurate current, then we'd trim the R in production test until the current is what we want.



B. GM Cell

Sometimes we don't need a full bandgap reference. In those cases, we can use a GM cell, as shown in Figure 17.

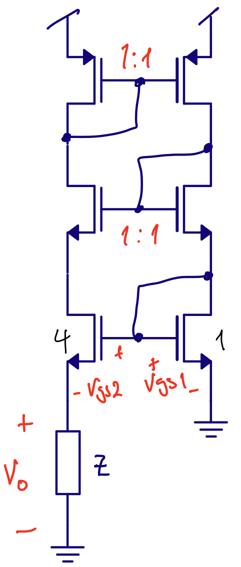


Figure 17: GM cell.

The top PMOS current mirror ensures that both branches have the same current. The middle NMOS current mirror copies the drain voltage on top of the diode connected bottom NMOS to the left NMOS. Consider the bottom transitors, those marked with "1" and "4". The V_o voltage is

$$V_o = V_{GS1} - V_{GS2} = V_{eff1} + V_{tn} - V_{eff2} - V_{tn} = V_{eff1} - V_{eff2}$$

Assuming transistors in strong inversion, then

$$I_{D1} = \frac{1}{2}\mu_n C_{ox}\frac{W_1}{L_1}V_{eff1}^2$$

$$I_{D2} = \frac{1}{2}\mu_n C_{ox}4\frac{W_1}{L_1}V_{eff2}^2$$

$$I_{D1} = I_{D2}$$

$$\frac{1}{2}\mu_n C_{ox}\frac{W_1}{L_1}V_{eff1}^2 = \frac{1}{2}\mu_n C_{ox}4\frac{W_1}{L_1}V_{eff2}^2$$

$$V_{eff1} = 2V_{eff2}$$

Inserted into above

$$V_o = V_{eff1} - \frac{1}{2}V_{eff1} = \frac{1}{2}V_{eff1}$$

Still assuming transistors in strong inversion, such that

$$g_m = \frac{2I_d}{V_{eff}}$$

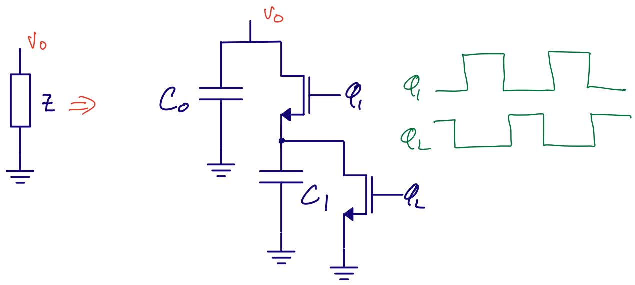
we find that

$$I = \frac{V_{eff1}}{2Z}$$

$$Z \Rightarrow \frac{1}{g_m}$$

If we use a resistor for Z, then we can get a transconductance that is proportional to a resistor, or a constant g_m bias.

We can use other things for Z, like a switched capacitor



IV. Want to learn more?

A simple three-terminal IC bandgap reference

A CMOS bandgap reference circuit with sub-1-V operation

A sub-1-V 15-ppm//spl deg/C CMOS bandgap voltage reference without requiring low threshold voltage device

The Bandgap Reference

The Design of a Low-Voltage Bandgap Reference

Carsten Wulff received the M.Sc. and Ph.D. degrees in electrical engineering from the Department of Electronics and Telecommunication, Norwegian University of Science and Technology (NTNU), in 2002 and 2008, respectively. During his Ph.D. work at NTNU, he worked on open-loop sigma-delta modulators and analog-to-digital converters in nanoscale CMOS technologies. In 2006-2007, he was a Visiting Researcher with the Department of Electrical and Computer Engineering, University of Toronto, Toronto, ON, Canada. Since 2008 he's been with Nordic Semiconductor in various roles, from analog designer, to Wireless Group Manager, to currently Principle IC Scientist. From 2014-2017 he did a part time Post.Doc focusing on compiled, ultra low power, SAR ADCs in nanoscale technologies. He's also an Adjunct Associate Professor at NTNU. His present research interests includes analog and mixed-signal CMOS design, design of high-efficiency analog-to-digital converters and low-power wireless transceivers. He is the developer of Custom IC Compiler, a general purpose integrated circuit compiler, and makes the occational video on analog integrated circuits at https://www.youtube.com/@analogicus. For full CV see https://analogicus.com/markdown-cv/.