CAMBRIDGE
Cambridge University Press & Assessment
978-1-009-60223-5 — Analysis and Design of Data Converters
Behzad Razavi
Excerpt
More Information

1

Introduction to Data Conversion

Data conversion refers to the transformation of analog signals to their digital counterparts ("analog-to-digital" conversion) or vice versa ("digital-to-analog" conversion). In this chapter, we begin with the former, present some of its applications, and describe its basic implementation. We then repeat these steps for the latter.

1.1 Analog-to-Digital Conversion

Naturally-occurring signals are found in analog form. Our voice, an image captured by the pixels in a digital camera, and radiations received from far-away galaxies exemplify such signals. In most systems, we wish to process, carry, and store this information in digital form; hence the need for analog-to-digital converters (ADCs).

As an example, suppose a farmer facing severe drought uses a large number of sensors in an orchard that measure the water received by each tree and wirelessly transmit the information to the farmer's phone. It is much simpler in this case to digitize the sensors' outputs and deal with only ONEs and ZEROs beyond that point. Of course, the power consumed by each "node" (the sensor box) must be minimized because the farmer is not inclined to change the batteries frequently. The cost plays a critical role here as well.

There are also situations in which *digital* signals must be treated as analog information and must be *digitized*! Consider, for example, a USB cable connecting a high-resolution video camera to a computer (Fig. 1.1). The camera applies the image to an ADC and a digital signal processor (DSP) for various operations, e.g., automatic focus or brightness adjustment. The DSP binary output, $x(t)$, now travels over the cable, experiencing its limited bandwidth and emerging severely "distorted" as $y(t)$. Since it is difficult to distinguish the ONEs and ZEROs in this waveform, we may decide to digitize it, compensate for the cable's loss, and detect the bits reliably in the digital domain.

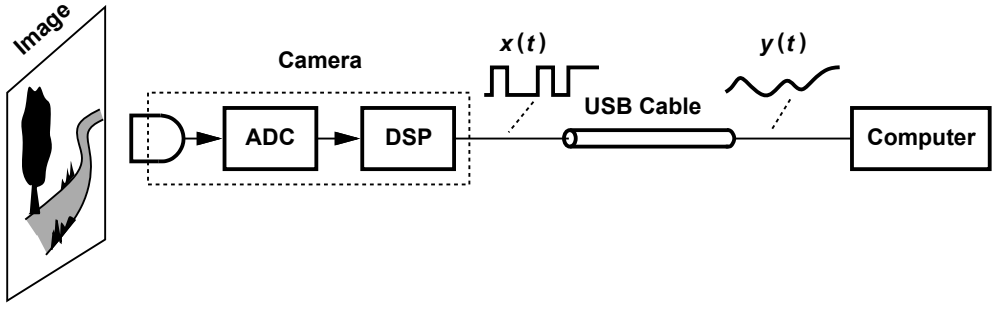


Figure 1.1 Distortion of binary data in a USB cable.

© in this web service Cambridge University Press & Assessment www.cambridge.org

CAMBRIDGE
Cambridge University Press & Assessment
978-1-009-60223-5 — Analysis and Design of Data Converters
Behzad Razavi
Excerpt
More Information

Figure 1.2 shows a conceptual ADC, where $x(t)$ is generally a "continuous-time" and "continuous-amplitude" signal. That is, it has a value at every point in time and the value can be any real (or complex) number between an upper and a lower bound. The output of a microphone is an example of such a signal. The ADC (binary) output, on the other hand, occurs only at discrete points in time, T_S, $2T_S$, etc., and assumes *discrete* amplitude values because of the finite number of bits available to represent the level. For example, a 4-bit ADC provides a binary output, $D_4D_3D_2D_1$, ranging from 0000 to 1111, and hence offers only 16 amplitude levels. We say an ADC converts a continuous-time, continuous-amplitude input to a discrete-time, discrete-amplitude output.

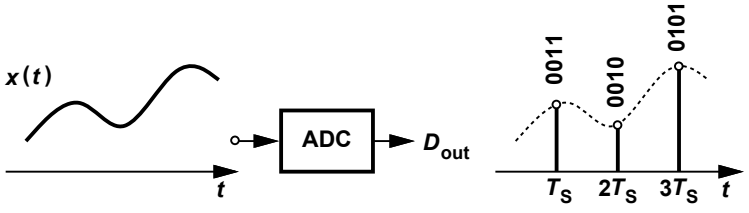


Figure 1.2 Conceptual operation of an ADC.

Example 1.1

Why should the digital output of an ADC appear only at discrete points in time?

Solution

From one perspective, we recognize that the DSP following the ADC can operate up to some speed and can therefore accept only up to a certain number of ADC output samples per second. But, more fundamentally, to represent the information carried by $x(t)$ in Fig. 1.2, the ADC need only provide a certain number of samples per second. Nyquist's sampling theorem tells us that the sampling rate, $f_S = 1/T_S$, must be only slightly greater than twice the bandwidth of $x(t)$ to avoid aliasing. That is, it's not necessary to sample the signal any faster.

Example 1.2

Must the digital output samples in Fig. 1.2 appear at integer multiples of T_S? This scheme is called "uniform sampling."

Solution

In general, sampling need not be uniform. Depending on how "busy" the signal is, we can adjust the sampling rate. For example, when a microphone output is quiet, the ADC sampling frequency can be lowered. Nonetheless, uniform sampling is by far the most common.

From these thoughts emerges the basic implementation depicted in Fig. 1.3. Here, a sampler driven by a clock frequency of f_S takes uniform samples of $x(t)$, generating $y(t)$. The latter is a discrete-time but continuous-amplitude signal (why?). Next, $y(t)$ is applied to a digitizer, more precisely called a "quantizer," which approximates the input analog values by a finite number of levels. The following example illustrates the concept of (amplitude) quantization.

Example 1.3

A 2-bit ADC receives a voltage varying from 0 V to 400 mV. Sketch the input-output characteristic of the circuit.

© in this web service Cambridge University Press & Assessment www.cambridge.org

CAMBRIDGE
Cambridge University Press & Assessment
978-1-009-60223-5 — Analysis and Design of Data Converters
Behzad Razavi
Excerpt
More Information

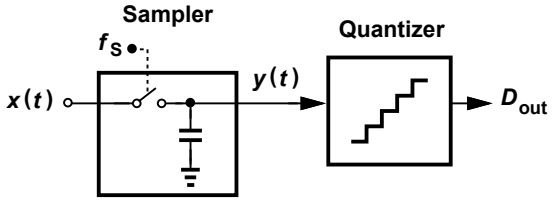


Figure 1.3 Sampling and quantization as fundamental operations for A/D conversion.

Solution

As shown in Fig. 1.4, we first identify the input range on the horizontal axis and divide it by $2^2 = 4$. For $0 < V_{in} < 100$ mV, we have $D_{out} = 00$. As V_{in} exceeds 100 mV, $D_{out} = 01$. Similarly, as V_{in} crosses 200 mV and 300 mV, D_{out} increments to 10 and 11, respectively.

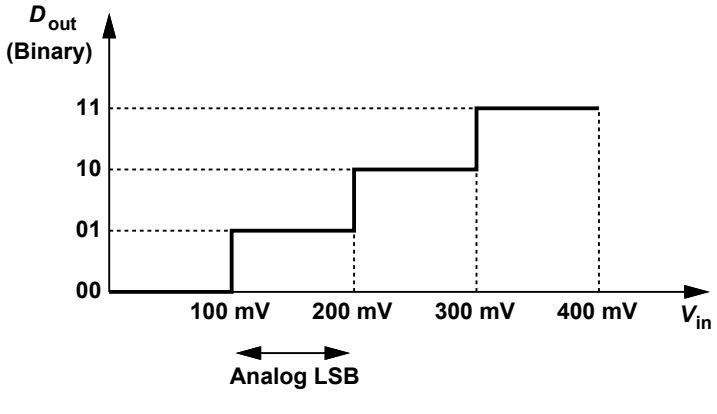


Figure 1.4 Input-output characteristic of a 2-bit quantizer.

The foregoing example indicates that quantization is similar to "truncation" rather than to "rounding." An input voltage of 160 mV, for example, is mapped to $D_{out} = 01$, as if it were 101 mV. We could also write $D_{out} \equiv [V_{in}/100\ \text{mV}]$, where the square brackets denote the "integer part" of the quantity. (Recall $y = [x]$ from calculus.) We also say the ADC has an analog least-significant bit (LSB) of 100 mV. We deal with sampling in Chapter 2 and with quantization in Chapter 12.

A quantizer with 2^N levels has a "resolution" of N bits. We expect that a greater N makes the quantized signal a more accurate copy of the analog input. That is, the *error* introduced by the quantizer decreases. Called "quantization noise," this error leads to a finite signal-to-noise ratio (SNR) at the output. We derive this SNR in terms of N for a sinusoidal input in Chapter 12, but the result proves useful even at the beginning of our studies:

$$\text{SNR} = 6.02N + 1.74\ \text{dB}. \tag{1.1}$$

For example, a 10-bit ADC cannot provide an SNR greater than 62 dB.

Depending on the application, N lies between 4 bits and 20 bits, with $6 \leq N \leq 12$ the most common range. For example, the camera in a mobile device employs 10-bit or 12-bit ADCs to digitize images. Lower resolutions would translate to a lower SNR, yielding a "noisy" image.

© in this web service Cambridge University Press & Assessment www.cambridge.org

CAMBRIDGE
Cambridge University Press & Assessment
978-1-009-60223-5 — Analysis and Design of Data Converters
Behzad Razavi
Excerpt
More Information

1.2 Digital-to-Analog Conversion

Digital-to-analog converters (DACs) serve two purposes. First, they interface digital processors with an analog "environment." For example, the pixels on a laptop display must receive an analog voltage level so as to produce a certain light intensity. Thus, the digital information delivered by the graphics card must be translated to analog form before reaching the display (Fig. 1.5). Since our eyes are sensitive to any imperfections on the display, the DAC must reproduce the information with high "fidelity," i.e., with minimal noise and distortion.

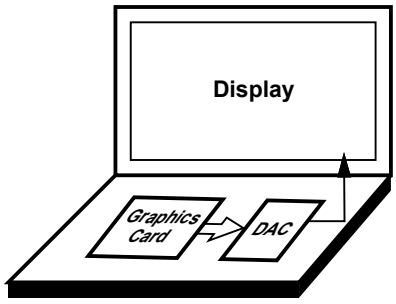


Figure 1.5 Use of a DAC to drive the pixels in a laptop display.

The second application of DACs is within ADCs. As explained in Chapters 14 and 17, some ADC architectures incorporate DACs.

Shown conceptually in Fig. 1.6, a DAC receives a digital input, D_{in}, and generates an analog output, A_{out}. We note from the input-output characteristic example that A_{out} rises in steps equal to Δ as D_{in} goes from 00 to 11, exhibiting four analog levels. The output can be any analog quantity, including voltage, current, or even time.

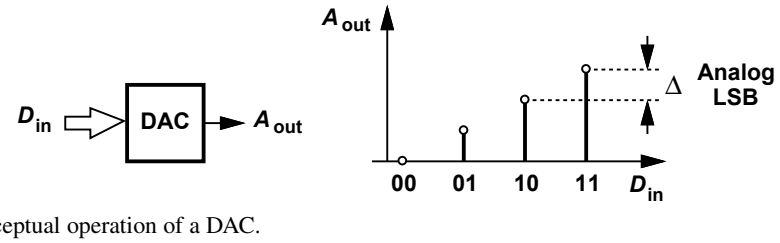


Figure 1.6 Conceptual operation of a DAC.

Example 1.4

A sinusoidal digital input is applied to a 2-bit DAC. Sketch the output as a function of time.

Solution

To construct the output, we first note that both D_{in} and A_{out} can assume only one of four values. The input sinusoid thus appears as shown in Fig. 1.7(a).[1] For each input sample, the DAC produces an integer multiple of the analog step, Δ, and *holds* it until the next value arrives [Fig. 1.7(b)]. The resulting waveform barely resembles a sinusoid.

[1] The reader may wonder why the digital input values are not evenly spaced in time. We ignore this point here.

© in this web service Cambridge University Press & Assessment www.cambridge.org

CAMBRIDGE
Cambridge University Press & Assessment
978-1-009-60223-5 — Analysis and Design of Data Converters
Behzad Razavi
Excerpt
More Information

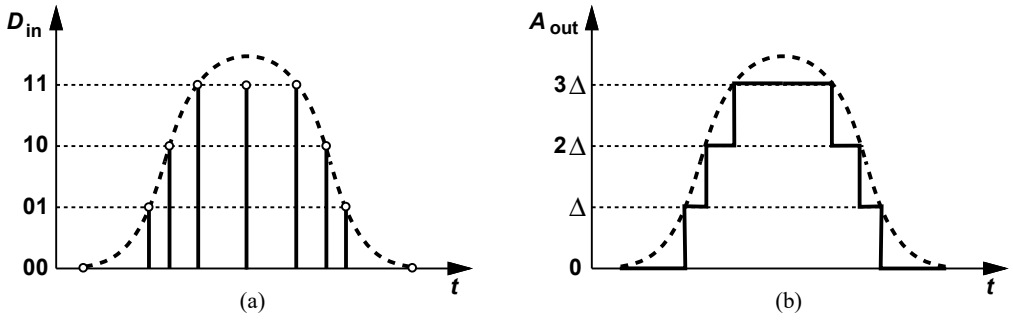


Figure 1.7 (a) A digital sinusoid applied to a DAC, and (b) the DAC output.

Example 1.5

We expect a DAC to perform the inverse of an ADC's operation. Does a DAC convert a discrete-time, discrete-amplitude signal to a continuous-time, continuous-amplitude output?

Solution

The DAC output is continuous in time but not in amplitude. The waveform observed in Fig. 1.7(b) exhibits jumps between discrete amplitude levels, creating unwanted high-frequency components (Chapter 8). For this reason, stand-alone DACs are followed by a low-pass filter so as to produce a smoother waveform (Fig. 1.8).

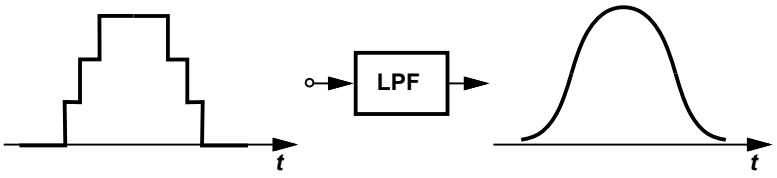


Figure 1.8 Use of a low-pass filter to remove high-frequency components from a DAC output.

© in this web service Cambridge University Press & Assessment www.cambridge.org

CAMBRIDGE
Cambridge University Press & Assessment
978-1-009-60223-5 — Analysis and Design of Data Converters
Behzad Razavi
Excerpt
More Information

2

Basic Circuit Concepts

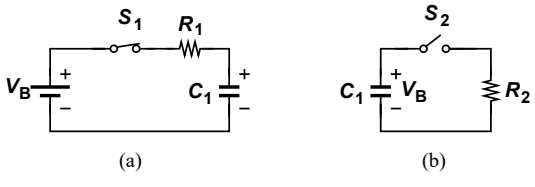
Many data converters rely heavily on switched-capacitor (SC) circuits. While the reader is assumed to have a solid understanding of circuit theory, there are a number of basic concepts that should be revisited before we delve into data converter design. This chapter briefly reviews these concepts, beginning with charge conservation and charge sharing. We then study integrated capacitor structures and three general circuit imperfections: offset, gain error, and nonlinearity.

We should make two remarks. If one plate of a capacitor receives charge in the amount of $+Q$, the other *must* receive $-Q$. Also, we say a switch is "on" when it conducts and "off" when it does not; these correspond to the "closed" and "open" states, respectively.

2.1 Charge Conservation

We know from physics that electric charge must be conserved. Applied to the charge arriving at a node, this principle translates to Kirchhoff's current law (KCL). While the exact definition of charge conservation is beyond our scope, we wish to verbalize it in a particular language suited to switched-capacitor circuits.

Let us begin with the circuit of Fig. 2.1(a), where the capacitor starts with a zero initial condition and charges to V_B after S_1 turns on, drawing a total charge of $Q = C_1 V_B$ from the battery. This positive charge



(a) (b)

Figure 2.1 (a) A capacitor charging through a resistor, and (b) the capacitor discharging through a resistor.

resides on the top plate of C_1, attracting a negative charge equal to $-C_1 V_B$ to its bottom plate. We can also say that positive charge flows from the top terminal of the battery to the top plate of C_1 and *from* the bottom plate of C_1 to the bottom terminal of V_B. Thus, it appears that positive current flows clockwise around the circuit.

In the next step, we turn S_1 off and, as shown in Fig. 2.1(b), connect another resistor across C_1. Now, positive charge flows from the top plate of C_1, through R_2, to its bottom plate. As time progresses, the positive and negative charge cancel, yielding zero volts across C_1.

In some circuits, as in the foregoing examples, the total charge on the capacitor(s) changes after a switching event. We thus say the charge on the capacitor(s) is not conserved. To appreciate the significance of this point, let us consider a simple circuit in which the capacitor charge *is* conserved. As shown in Fig. 2.2(a), a voltage step is applied to the left plate of C_1 while its right plate is not attached to any other device. Assume the initial voltage on C_1 is zero. How does V_{out} behave? Since no current can flow through C_1, the voltage across C_1

© in this web service Cambridge University Press & Assessment www.cambridge.org

CAMBRIDGE
Cambridge University Press & Assessment
978-1-009-60223-5 — Analysis and Design of Data Converters
Behzad Razavi
Excerpt
More Information

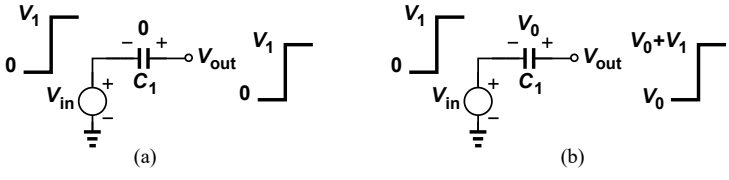


Figure 2.2 (a) A step applied to a floating capacitor, and (b) step response with an initial condition.

remains zero, demanding that $V_{out} = V_{in} = V_1 u(t)$, where $u(t)$ denotes the unit step. The reader may recall from circuit theory that we can view C_1 as a short circuit at $t = 0^+$ and arrive at the same conclusion. The charge on C_1 is zero before and after the step. This is because the right plate of C_1 has no path through which it can receive charge. In other words, even though V_{in} is capable of delivering charge to the left plate of C_1, it does not.

We now repeat the above experiment but with an initial voltage of V_0 on C_1 [Fig. 2.2(b)]. Since the charge on the right plate is conserved, so is the charge on the left plate, yielding a constant voltage on C_1 equal to V_0. The output therefore jumps from V_0 to $V_0 + V_1$. We say the capacitor passes *changes* in the input voltage in both examples.

Example 2.1

The voltage source shown in Fig. 2.3 applies a random signal to the left plate of C_1 with a zero initial

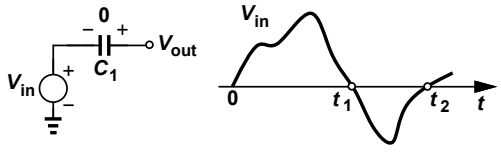


Figure 2.3 A random signal applied to a floating capacitor.

condition. Determine V_{out} at $t = t_1$ and $t = t_2$.

Solution

In this case, $V_{out} = V_{in}$ at all times; thus, $V_{out} = 0$ at $t = t_1$ and $t = t_2$. From another perspective, the net change in V_{in} is zero from $t = 0$ to $t = t_1$, and so is the change in V_{out}. The same point applies to $t = t_2$.

Let us turn to the slightly more complex arrangement depicted in Fig. 2.4(a), assuming zero initial condi-

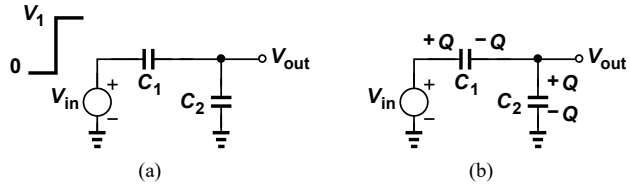


Figure 2.4 (a) A step applied to a capacitive voltage divider, and (b) charge distribution on the capacitors.

tions. We analyze the circuit in three steps. First, we ask, is the charge on the capacitors conserved here? We recognize that the right plate of C_1 can receive charge only from the top plate of C_2. That is, the sum of the charges on these two plates is conserved and equal to zero.

© in this web service Cambridge University Press & Assessment www.cambridge.org

CAMBRIDGE
Cambridge University Press & Assessment
978-1-009-60223-5 — Analysis and Design of Data Converters
Behzad Razavi
Excerpt
More Information

Second, we surmise that the positive step generated by V_{in} may deliver positive charge to the left plate of C_1. Denoting this charge by $+Q$, we observe that the right plate of C_1 *must* receive $-Q$, which can only come from the top plate of C_2 [Fig. 2.4(b)]. Upon losing $-Q$, this plate now carries an amount of charge equal to $+Q$, requiring that the bottom plate of C_2 absorb $-Q$ from the ground. We say a charge equal to Q flows clockwise around the loop. We also note that two capacitors in series must carry equal charges if they begin with zero initial conditions.

In the third step, we write the voltages on C_1 and C_2 as Q/C_1 and Q/C_2, respectively, and enforce Kirchhoff's voltage law (KVL):

$$V_1 = \frac{Q}{C_1} + \frac{Q}{C_2}. \tag{2.1}$$

That is,

$$Q = \frac{C_1 C_2}{C_1 + C_2} V_1, \tag{2.2}$$

as expected for a series combination. It follows that

$$V_{out} = \frac{Q}{C_2} \tag{2.3}$$

$$= \frac{C_1}{C_1 + C_2} V_1. \tag{2.4}$$

The voltage across C_1 is equal to $[C_2/(C_1 + C_2)]V_1$ (why?). The capacitive divider attenuates the input voltage change. If undesirable, this effect is minimized by choosing C_1 much greater than C_2.

We have seen that the charge on C_1 is not conserved in Fig. 2.1(b) but that on C_1 and C_2 in Fig. 2.4 is. Why is that? The principal difference is that the former contains a *resistive* path through which charge can flow while the latter does not. We articulate the latter's property by drawing a closed surface (also called a "Gaussian surface") around the right plate of C_1 and the top plate of C_2 (Fig. 2.5), recognizing that the total charge within this surface must remain constant. We also say X is a "high-impedance" node.

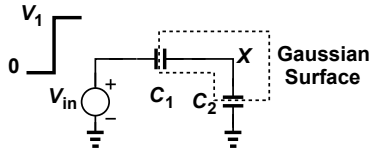


Figure 2.5 Concept of Gaussian surface for two capacitors.

2.2 Charge Sharing

Charge sharing commonly occurs in many data converter circuits—intentionally or not—and merits an in-depth study.

In the circuit of Fig. 2.6(a), the initial voltages on C_1 and C_2 are equal to V_0 and zero, respectively. We draw a Gaussian surface around the capacitors' top plates and note that its total charge must be constant. Thus, after S_1 turns on [Fig. 2.6(b)], the initial charge on C_1, C_1V_0, is shared between the two capacitors. Denoting the new voltage by V_X, we write

$$C_1 V_0 = C_1 V_X + C_2 V_X \tag{2.5}$$

© in this web service Cambridge University Press & Assessment www.cambridge.org

CAMBRIDGE
Cambridge University Press & Assessment
978-1-009-60223-5 — Analysis and Design of Data Converters
Behzad Razavi
Excerpt
More Information

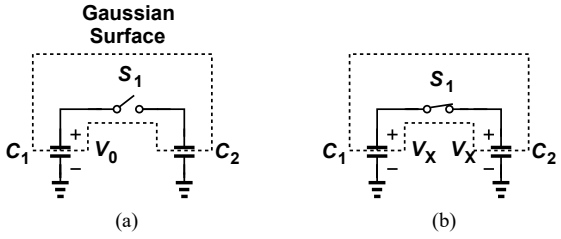


Figure 2.6 (a) Two capacitors within a Gaussian surface, and (b) conservation of charge before and after the switch turns on.

and hence

$$V_X = \frac{C_1}{C_1 + C_2} V_0. \tag{2.6}$$

If undesirable, charge sharing can be minimized by selecting $C_1 \gg C_2$.

Example 2.2

Calculate the final capacitor voltages in Fig. 2.7(a) if the initial conditions are V_0 for C_1 and zero for C_2 and C_3. The switch turns on at $t = 0$.

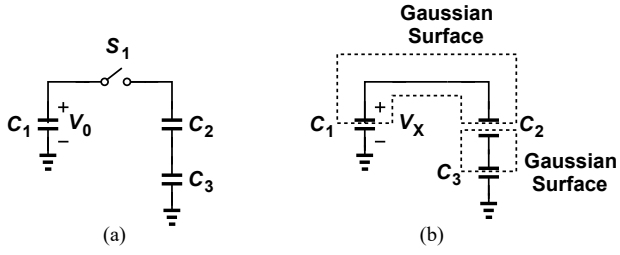


Figure 2.7 (a) Two series capacitors sharing charge with another, and (b) use of Gaussian surfaces for analysis.

Solution

As shown in Fig. 2.7(b), we identify two Gaussian surfaces whose internal charge must be constant. But the circuit is simple enough to allow us to write charge conservation as

$$C_1 V_0 = C_1 V_X + \frac{C_2 C_3}{C_2 + C_3} V_X. \tag{2.7}$$

That is,

$$V_X = \frac{C_1}{C_1 + \dfrac{C_2 C_3}{C_2 + C_3}} V_0. \tag{2.8}$$

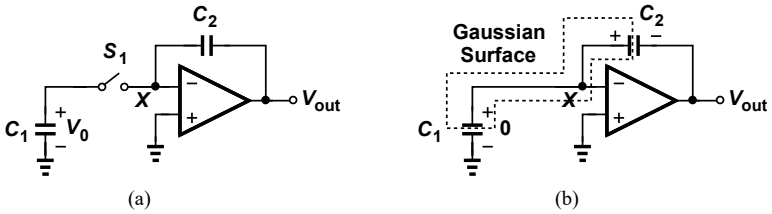
This voltage is attenuated by C_2 and C_3, yielding $[C_3/(C_2+C_3)]V_X$ across the former and $[C_2/(C_2+C_3)]V_X$ across the latter (why?).

© in this web service Cambridge University Press & Assessment www.cambridge.org

CAMBRIDGE
Cambridge University Press & Assessment
978-1-009-60223-5 — Analysis and Design of Data Converters
Behzad Razavi
Excerpt
More Information

Example 2.3

In the circuit of Fig. 2.8(a), the initial voltages on C_1 and C_2 are equal to V_0 and zero, respectively. Explain what happens after the switch turns on, assuming that the op amp has a high gain.



(a) (b)

Figure 2.8 (a) A charged capacitor switched to a virtual ground, and (b) use of a Gaussian surface for analysis.

Solution

After S_1 turns on, positive charge flows from the top plate of C_1 to the left plate of C_2. The high op amp gain ensures that node X is a virtual ground, forcing the voltage across C_1 to fall to zero [Fig. 2.8(b)]. Consequently, all of the initial charge on C_1 is transferred to C_2, creating a voltage across it equal to C_1V_0/C_2. In this case, the Gaussian surface encloses the top plate of C_1 and the left plate of C_2. Since V_{out} is equal to the negative of the voltage across C_2 (why?),

$$V_{out} = -\frac{C_1}{C_2}V_0. \tag{2.9}$$

We say the circuit amplifies the initial voltage by a factor of $-C_1/C_2$.

It is important to recognize the different charge sharing mechanisms in Figs. 2.6(b) and 2.8(b): in the former, C_1 still retains some charge because its final voltage is *not* zero, whereas in the latter, C_1 loses all of its charge owing to the virtual ground. This distinction proves useful in our studies of data converter circuits.

2.3 Integrated Capacitors

Capacitors can be realized in different forms in CMOS technology. The simplest candidate is the MOS capacitor, obtained by tying the source and drain of a device together [Fig. 2.9(a)]. Offering $C_{AB} \approx WLC_{ox}$, this topology typically provides the *densest* structure (i.e., one with the greatest capacitance per unit area). For example, in 28-nm technology, the gate capacitance is around 2 fF/μm^2. However, such a value is obtained only if the MOS device is *turned on*. In fact, the capacitance is a nonlinear function of V_{AB} [Fig. 2.9(b)], proving unsuitable for most switched-capacitor circuits.

A much more linear capacitor can be realized by creating a vertical stack of metal plates, as shown in Fig. 2.9(c). We recognize that metal-9 and metal-8 plates form a capacitor, and so do metal-8 and metal-7 plates, etc. The connections shown in the figure place these capacitors in parallel, yielding

$$C_{AB} = C_9 + C_8 + C_7. \tag{2.10}$$

The stack can include lower metal layers as well to provide a greater capacitance density, but it is still less dense than the MOS structure.

The parallel-plate geometry of Fig. 2.9(c) suffers from a parasitic capacitance, C_p, to the substrate, leading to the model shown in Fig. 2.9(d). This parasitic is given by the capacitance between the bottom-most layer—M_6 in our example—and the grounded substrate. Of course, we can decide on which side of C_{AB} the parasitic component should be placed. Typically, C_p is around 5% to 10% of C_{AB}. To appreciate the effect of C_p, let us study the following two examples.

© in this web service Cambridge University Press & Assessment www.cambridge.org