Behzad Razavi

The Delta-Sigma Modulator

Delta-Sigma modulators (DSMs) are a class of oversampling analog-to-digital converters (ADCs) that perform "quantization noise shaping," thus achieving a high signal-to-noise ratio (SNR). An efficient solution for resolutions above approximately 12 b, DSMs are extensively used in analog and RF applications. In this article, we study the fundamentals of this vast field.

The Delta Modulator

It is helpful to first study the predecessor of DSMs, namely, the delta modulator. Shown in Figure 1(a), the latter consists of a 1-b quantizer (e.g., a single comparator) and an integrator, both placed in a negative-feedback loop. The high loop gain ensures that $V_F \approx V_{in}$ and hence the digital output is a representation of the analog input.

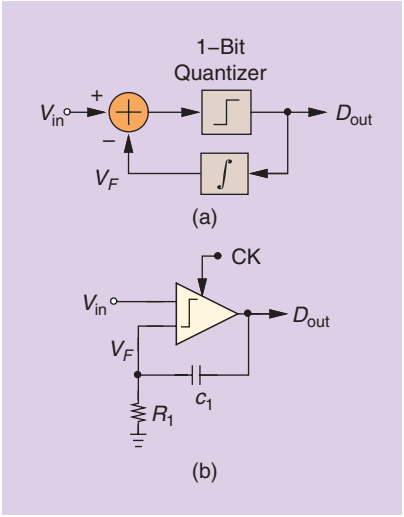

FIGURE 1: (a) A delta modulator and (b) its simple implementation.

Digital Object Identifier 10.1109/MSSC.2016.2543061
Date of publication: 21 June 2016

Figure 1(b) depicts a simple implementation where the integrator is approximated by a low-pass filter.

The principal difficulty with the delta modulator is that the output digital representation in fact contains only the *derivative* of the input, as can be seen by noting that $V_F = \int D_{out}\,dt$ in Figure 1(a). This differentiation alters the signal spectrum, attenuates the low-frequency content of the signal, and amplifies high-frequency noise.

Delta-Sigma modulators are a class of oversampling analog-to-digital converters that perform "quantization noise shaping," thus achieving a high signal-to-noise ratio.

Brief History

To avoid the differentiation effect in Figure 1(a), Inose et al. [1] cleverly moved, in 1962, the integrator from the feedback path to the forward path, introducing the "$\Delta\Sigma$ modulator" shown in Figure 2. Here, the high-loop gain forces the running average of D_{out} to follow V_{in}. Of course, D_{out} also contains the quantization noise created by the quantizer, but with certain interesting and useful alterations.

The $\Delta\Sigma$ modulator structure actually predates the work by Inose et al. In a patent filed in 1961 [2], Brahm discloses the system shown in Figure 3, where the loop contains an integrator, a multibit quantizer (an ADC), and a multibit digital-to-analog converter (DAC).

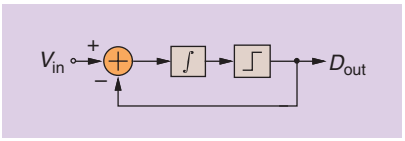

FIGURE 2: A DSM.

In the 1970s, the potential of DSMs was further explored. Candy proposed the use of the structure for robust analog-to-digital conversion in 1974 [3] and, along with Ching and Alexander, in 1976 demonstrated a resolution of 13 b with a 1-b quantizer in the loop [4]. These two papers pointed out that the overall resolution increases as the quantizer is clocked faster and its output circulates around the loop more frequently. An important observation made by Candy was that the overall output noise is the "first difference" of the quantizer's additive noise, exhibiting a spectrum of the form $\sin^2(\omega T_{CK}/2)$, where T_{CK} is the quantizer clock period [3]. That is, the noise is suppressed at low frequencies. Candy also recognized that the performance negligibly degrades with the imperfections of the analog components within the loop.

The first integrated DSM was evidently reported by van de Plassche in 1977 [5]. Using a continuous-time (CT) integrator, the ADC achieved a resolution of about 17 b in bipolar technology.

In 1978, Tewksbury and Hallock described higher-order DSMs, presenting the architecture shown in Figure 4 (but attributing it to G.R. Ritchie) [6]. They also showed that the quantization noise spectrum is attenuated

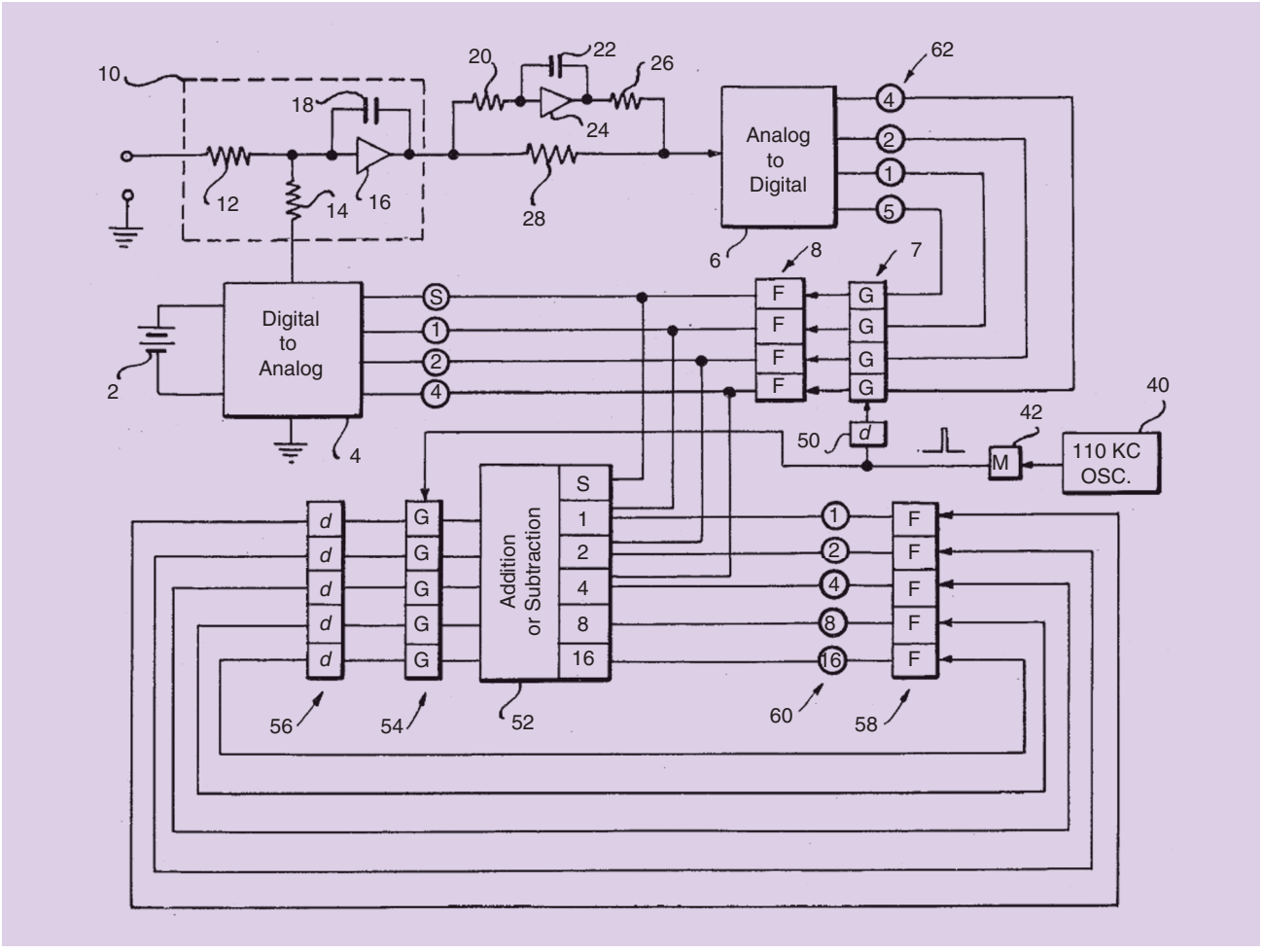

FIGURE 3: The DSM proposed by Brahm in 1961.

according to the shaping function $(1 - z^{-1})^N$, where N denotes the order (the number of integrators).

In 1981, an n-type metal-oxide-semiconductor (NMOS) implementation using a single passive discrete-time integrator was reported [7], and in 1982, a patent was filed disclosing a loop with two active switched-capacitor integrators [8]. CMOS realizations followed in 1986 [9] and 1988 [10].

It is interesting that some authors use the term "$\Delta\Sigma$ modulator" and others use the term "$\Sigma\Delta$ modulator" to refer to the circuit. One argument in favor of the former is that the loop first subtracts and then accumulates.

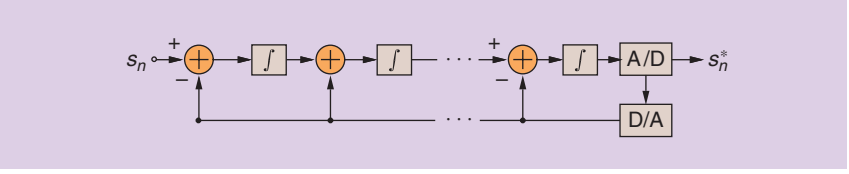

FIGURE 4: A high-order DSM attributed to Ritchie.

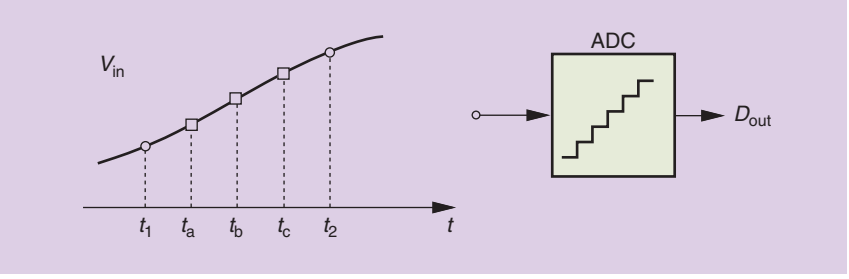

FIGURE 5: Oversampling to create correlation between consecutive samples.

Basic Operation

Suppose we wish to digitize the analog waveform shown in Figure 5. A Nyquist-rate ADC would sample and quantize V_{in} at t_1 and t_2, with $f_s = 1/(t_2 - t_1)$ slightly greater than twice the signal bandwidth. In this case, the samples at t_1 and t_2 exhibit little correlation, and so do their quantization errors. On the other hand, if we additionally sample and digitize V_{in} at t_a, t_b, and t_c, we create correlated quantization errors between consecutive samples. From another perspective, if the signal changes slowly enough from t_1 to t_a, then

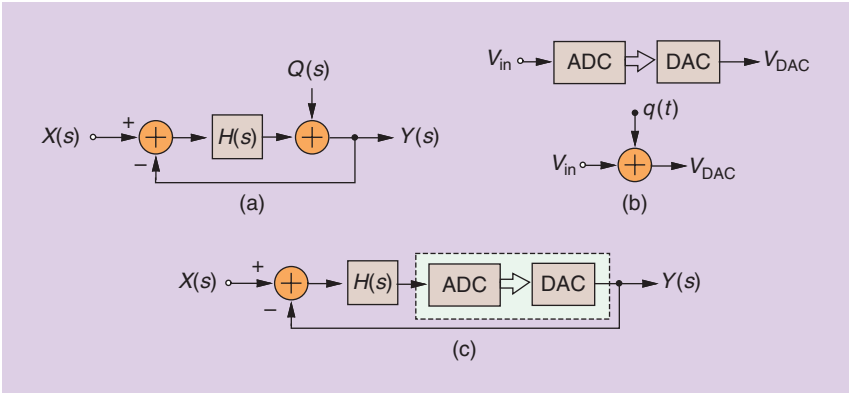

FIGURE 6: (a) A negative-feedback system with noise injected near the output, (b) an ADC/DAC cascade modeled in terms of additive noise, and (c) the rejection of quantization noise by negative feedback.

the quantization errors incurred by these two samples are almost equal. We then surmise that subtracting the quantization error of one sample from the next can reduce the overall quantization noise.

This method of noise suppression can also be explained in the frequency domain, culminating in the concept of noise shaping. First, consider the negative-feedback system shown in Figure 6(a), where an unwanted signal $Q(s)$ is injected "near" the output but inside the loop. The transfer function from Q to Y can be chosen to provide a high-pass behavior. For example, if $H(s)$ is an integrator, then $H \times (s) = 1/s$ and hence $Y/Q = s/(s+1)$. We say the spectrum of Q is "shaped" by the feedback loop, an effect also observed for the phase noise of oscillators in phase-locked loops.

This method of noise suppression can also be explained in the frequency domain.

In the next step, consider the ADC–DAC cascade depicted in Figure 6(b), noting that V_{DAC} is equal to the ideal analog input plus the ADC's quantization noise, $q(t)$, if the DAC is ideal. It is therefore expected that placing this cascade within the feedback loop of Figure 6(a) can reduce the overall quantization noise for some frequency range [11]. Illustrated in Figure 6(c), such an arrangement guarantees that $Y(s)$ tracks $X(s)$—and that $Q(s)$ is suppressed—so long as $H(s)$ provides a high loop gain. For $H(s) = 1/s$, this occurs at low frequencies.

The foregoing observations lead to the first-order $\Delta\Sigma$ modulator shown in Figure 7(a), where the ADC is realized as a 1-b quantizer (a single comparator) and the DAC as two switches producing $\pm V_{REF}$. By virtue of its high gain, the comparator enforces a virtual ground at node X, but due to the discrete-time nature of the loop, only the average value of V_X remains close to zero. This in turn means that the average difference between V_{in} and V_{DAC}, and hence between V_{in} and D_{out}, is nulled. For example, if V_X crosses from negative to positive, the comparator and the DAC apply a pulse to the integrator so as to return V_X toward zero. Figure 7(b) illustrates how the running average of the digital output tracks the analog input [10].

The 1-b quantizer in Figure 7(a) suffers from enormous quantization noise, $q(t)$. One might wonder, then, whether a more resolute quantizer can be used instead. This question leads to two different architectures, namely, loops containing a multibit quantizer or a greater number of integrators. Before describing these solutions, we need to derive the noise-shaping properties of the first-order modulator.

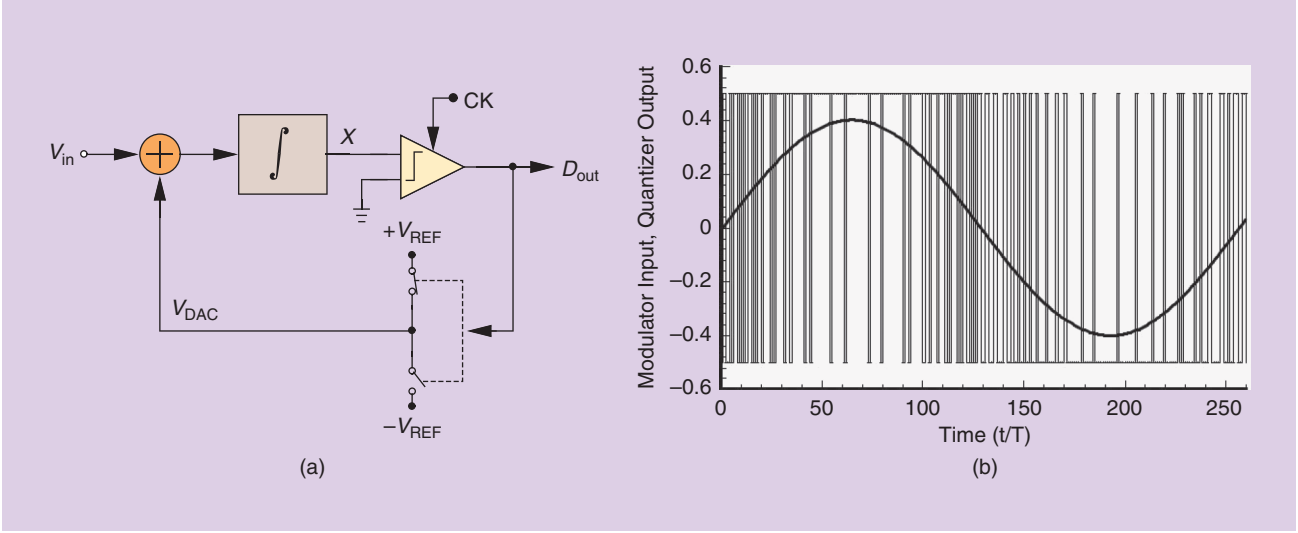

FIGURE 7: (a) A simple first-order DSM with a 1-b quantizer, and (b) input and output waveforms.

Formulation of Noise Shaping

Let us examine how the quantization noise introduced by the quantizer in Figure 8(a) propagates to the output. We assume a discrete-time integrator and express its output as $u(kT_s) = u[(k-1)T_s] + g[(k-1)T_s]$, where $g[(k-1)T_s] = x[(k-1)T_s] - y[(k-1)T_s]$. The quantizer output is given by $y(kT_s) = u(kT_s) + q(kT_s)$, and reaches the DAC output unchanged if the DAC is ideal. Substituting for $g[(k-1)T_s]$ and for $y[(k-1)T_s]$, we obtain

$$y(kT_s) = u[(k-1)T_s] + x[(k-1)T_s] - y[(k-1)T_s] + q(kT_s). \quad (1)$$

Since $u[(k-1)T_s] - y[(k-1)T_s] = -q[(k-1)T_s]$, we have

$$y(kT_s) = x[(k-1)T_s] + q(kT_s) - q[(k-1)T_s]. \quad (2)$$

As expected, the output quantization noise is equal to the difference between the quantization errors incurred by two consecutive samples. Taking the z transform of both sides yields

$$Y(z) = z^{-1}X(z) + (1 - z^{-1})Q(z). \quad (3)$$

The output thus contains the input with no change but just a delay. The quantization noise experiences a $1 - z^{-1}$ transfer function. We say the system provides a "signal transfer function" (STF) equal to z^{-1} and a "noise transfer function" (NTF) equal to $1 - z^{-1}$.

To determine the output noise spectrum, we replace z in $1 - z^{-1}$ with $\exp(j\omega T_s)$ and multiply the spectrum of $q(t)$, $S_Q(f)$, by $|1 - \exp(-j\omega T_s)|^2 = (2\sin \pi f T_s)^2$. Figure 8(b) plots this noise-shaping function, revealing that integrated quantization noise can be small if the input signal bandwidth $f_B \ll f_s/2$. The quantity $M = (f_s/2)/f_B$ is called the "oversampling ratio" (OSR) and signifies how far above the Nyquist rate the system operates. The area under the curve in Figure 8(b) from zero to f_B is proportional to $1/M^3$, revealing the strong dependence of the performance upon the oversampling ratio.

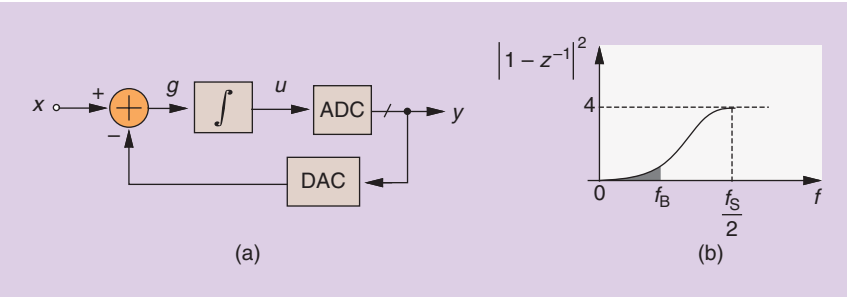

FIGURE 8: (a) First-order DSM for noise shaping calculations and (b) its noise-shaping function.

In addition to noise shaping, DSMs provide two other advantages over Nyquist-rate ADCs. First, for a given amount of kT/C noise, the sampling capacitors in the former can be smaller than those in the latter by a factor of M. This can be intuitively explained by noting that the extra samples taken in Figure 5 are eventually combined with those at t_1 and t_2 (by means of a "decimator"), benefiting from kT/C noise (and op amp noise) averaging. Second, the antialiasing filter in the former has a more relaxed selectivity than in the latter.

DSMs with Multibit Quantizers

In the spirit of Brahm's patent (Figure 3) and to lower the quantization noise, we can digitize the integrator output with more than one bit of resolution and feed the result to a multibit DAC. Typically realized as a flash stage, the quantizer injects proportionally less noise as its resolution increases. The performance of the system, however, is limited by the DAC nonlinearity, as pointed out by van de Plassche in 1979 [5]. In contrast to the two-level DAC in Figure 7(a), a multibit DAC exhibits nonlinearity in its input–output characteristic if its constituent components (resistors, capacitors, or current sources) have mismatches. This phenomenon can be viewed in Figure 8(a) as an undesirable term subtracted by the DAC from x and hence indistinguishable from nonlinearity in the input path.

The problem of DAC nonlinearity proves serious because DSMs typically target high resolutions, at which the "raw" device mismatches produce considerable distortion. For this reason, loops containing multibit DACs employ "dynamic element matching" techniques to reduce this nonlinearity [5].

Higher-Order DSMs

Another approach to reducing the noise of the quantizer is to replace it with another $\Delta\Sigma$ modulator [Figure 9(a)]. Here, the outer loop further shapes the quantization noise of the inner loop, yielding a shaping function of the form $(1 - z^{-1})^2$ for the 1-b quantizer's noise. The area under $|1 - z^{-1}|^2$ from zero to f_B is now proportional to $1/M^5$, a marked reduction compared to that of the first-order loop.

Providing identical outputs, the two DACs in Figure 9(a) can be merged, resulting in the more compact architecture shown in Figure 9(b). Exemplified by the implementation in Figure 9(c) [10], this simple, robust topology is the most commonly used DSM for moderate-performance applications. It can be shown that such imperfections as capacitor mismatch, op amp offset, op amp gain error, and comparator offset have much less impact here than in, for example, pipelined ADCs. The order of the loop can be increased further by adding more integrators, but instability becomes problematic, requiring other measures.

Problem of Tones

As explained earlier, the average output of a DSM tracks the input signal.

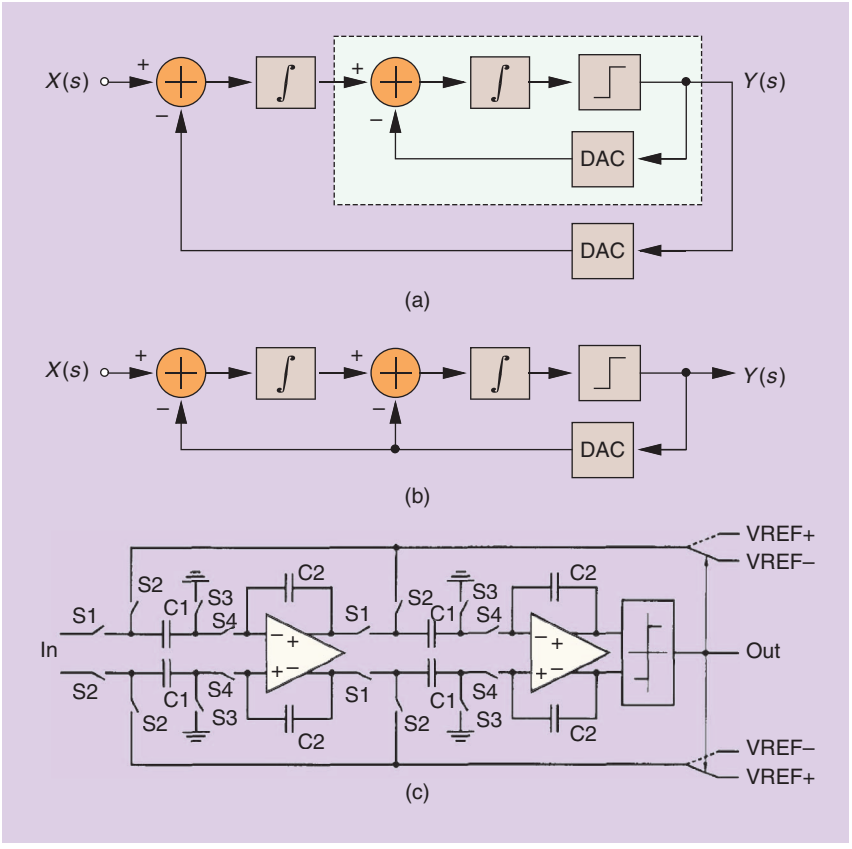

FIGURE 9: (a) A second-order DSM, (b) the simplified architecture, and (c) a discrete-time implementation.

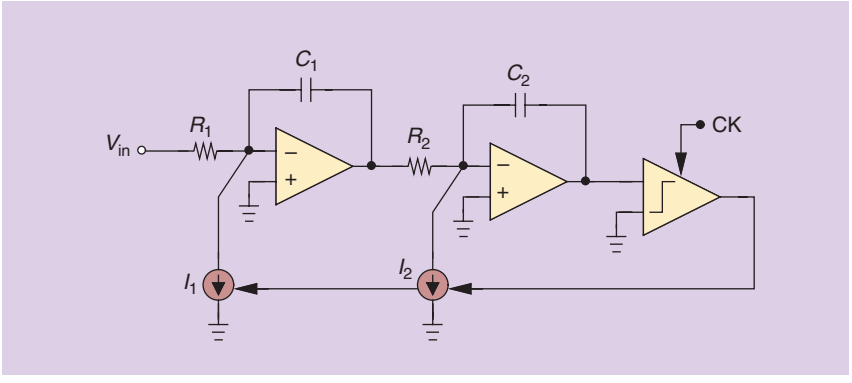

FIGURE 10: A simple second-order CTDSM.

What happens if V_{in} in Figure 2 is *constant*? Since the loop is periodically clocked and V_{in} does not change with time, we surmise that the output is also periodic. For example, if $V_{in} = 0.001\,V_{REF}$, then D_{out} consists of one ONE and another 999 ZEROs so as to produce such an average. Repeating with a period of $1000T_s$, the output therefore exhibits harmonics given by $mf_s/1000$, many of which can fall within the signal band. These "tones" corrupt the digitized signal. The tones tend to be smaller in magnitude in higher-order loops or at higher oversampling ratios, but one must often incorporate "dithering" to break their periodicity and convert them to noise.

Continuous-Time DSMs

The evolution of DSMs has made a 360° turn over the years. The earliest designs, those by Brahm and Inose et al., for example, employed continuous-time integrators, but, as switched-capacitor techniques matured in CMOS technology, discrete-time integrators became more common. In the late 1990s, it was recognized that continuous-time integrators offer certain advantages, and continuous-time DSMs (CTDSMs) rapidly rose as a formidable contender. It is important to note, however, that even CTDSMs are discrete-time feedback loops, still facing tone and stability issues.

Depicted in Figure 10 is a simple CTDSM realization of a second-order loop, where the current sources act as 1-b DACs. This arrangement provides three advantages over its discrete-time counterparts: 1) the sampling is performed by the comparator, obviating the need for highly linear front-end (bootstrapped) samplers, 2) the DSM presents less input capacitance and kickback noise, easing the demand on the preceding circuit, and 3) the two integrators naturally provide antialiasing filtering, simplifying the other filter stages in the signal path.

CTDSMs entail their own drawbacks. First, the jitter in the comparator clock modulates the amount of charge delivered by the feedback DACs to the integrators. This issue has been addressed by various techniques, e.g., the use of switched-capacitor DACs [12]. Second, the integrator op amps must have enough bandwidth to avoid slewing, a difficult issue because the comparator quantization noise traveling through the DACs and arriving at the integrators presents fast changes. This translates to a greater power consumption than that of op amps in discrete-time DSMs. Third, the signal-dependent delay of the comparator, each time it approaches metastability, also modulates the DACs' outputs, leading to distortion. The comparator must therefore be designed for a short regeneration time so that metastable states occur infrequently enough to negligibly affect the signal. Fourth, the thermal noise of R_1 and I_1 in Figure 10 limits the performance.

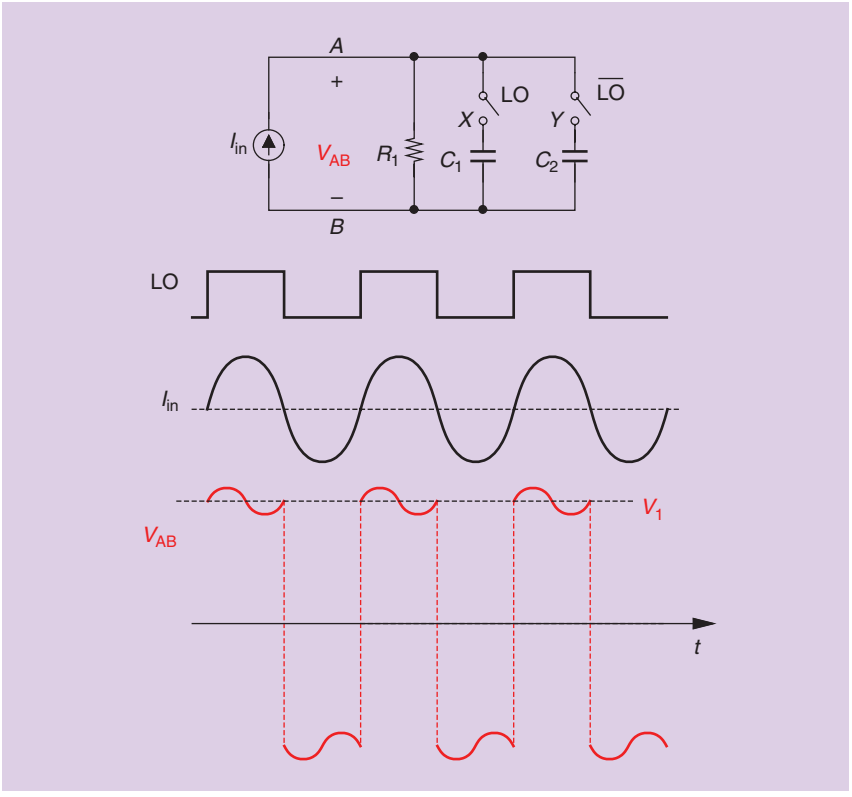

FIGURE 11: Steady-state waveforms in a commutated circuit.

Note on StrongArm Latch

In my article on the StrongArm latch [13], I had traced the circuit to a 1992 paper by Kobayashi et al. The idea was in fact filed for a patent by Madden and Bowhill on 27 June 1988 in the United States and by Kobayashi's coauthor, Nogami, in Japan on 13 July 1988.

Questions for the Reader

1) Explain in the time domain why $1-z^{-1}$ represents a high-pass function.
2) Explain why the comparator clock jitter in a discrete-time SDM such as that in Figure 9(c) is not critical.

Answers to Last Issue's Questions

1) The commutated capacitors of Figure 11 are placed at the antenna port of a Global System for Mobile Communication (GSM) receiver so as to attenuate by 20 dB a 0-dBm blocker at 20-MHz offset. What issues does such a circuit face?

   Such an approach faces three issues. First, from Smith's equation, the array must employ a large capacitance to provide a small bandwidth with a 50-Ω source resistance. Second, the on-resistance of the switches must be about 5 Ω, demanding a high power in the LO drive circuitry. Third, the switches experience a large voltage swing in the presence of a 0-dBm blocker, exhibiting considerable nonlinearity.
2) Does V_1 in Figure 11 change if the circuit contains four capacitive branches that are driven by 25%-duty-cycle local oscillator phases?

   No, it does not. The steady-state swing remains the same.

References

[1] H. Inose, Y. Yasuda, and J. Murakami, "A telemetering system by code modulation—Δ-Σ modulation," *IEEE Trans. Space Electron. Telemetry*, vol. 3, pp. 204–209, Sept. 1962.

[2] C. B. Brahm, "Feedback integrated system," U.S. Patent 3,192,371, Sept. 1961.

[3] J. C. Candy, "A use of limit cycle oscillations to obtain robust analog-to-digital conversion," *IEEE Trans. Commun.*, vol. COM-22, pp. 298–305, Mar. 1974.

[4] J. C. Candy, C. Y. Ching, and D. S. Alexander, "Using triangularly weighted interpolation to get 13-bit PCM from a sigma-delta modulator," *IEEE Trans. Commun.*, vol. COM-24, pp. 1268–1275, Nov. 1976.

[5] R. van de Plassche, "A five-digit analog-digital converter," *IEEE J. Solid-State Circuits*, vol. SSC-12, pp. 656–662, Dec. 1977.

[6] S. K. Tewksbury and R. W. Hallock, "Oversampled, linear and predictive noise-shaping coders of order N > 1," *IEEE Trans. Circuits Syst.*, vol. CAS-25, pp. 436–447, July 1978.

[7] T. Misawa, J. E. Iwersen, L. J. Loporcaro, and J. G. Ruch, "Single-chip per channel codec with filters utilizing Δ-Σ modulation," *IEEE J. Solid-State Circuits*, vol. 16, pp. 333–342, Aug. 1981.

[8] K. Shenoi and B. Agrawal, "Delta-sigma modulator with switch capacitor implementation," U.S. Patent 4,439,756, Jan. 1982.

[9] T. Hayashi, Y. Inabe, K. Uchimura, and T. *Kimura*, "A multistage delta-sigma modulator without double integration loop," in *ISSCC Dig. Tech. Papers*, Feb. 1986, pp. 182–183.

[10] B. Boser and B. A. Wooley, "The design of sigma-delta modulation analog-to-digital converters," *IEEE J. Solid-State Circuits*, vol. 23, pp. 1298–1308, Dec. 1988.

[11] H. A. Spang and P. M. Schultheiss, "Reduction of quantizing noise by use of feedback," *IRE Trans. Commun. Syst.*, vol. 10, pp. 373–380, Dec. 1962.

[12] M. Ortmann, F. Gerfers, and Y. Manoli, "Jitter insensitive feedback DAC for continuous-time ΣΔ modulators," in *Proc. Int. Conf. Electronics, Circuits, and Systems*, 2001, pp. 1049–1052.

[13] B. Razavi, "The StrongArm latch," *IEEE Solid-State Circuits Mag.*, vol. 7, no. 2, pp. 12–17, Spring 2015.