CHAPTER 14.4
DIGITAL AND ANALOG SYSTEMS

Paul G. A. Jespers

INTEGRATED SYSTEMS

With the trend toward ever higher integration levels, an increasing number of *IC*s combine some of the circuits seen before in order to build large systems, digital as well as analog, or mixed, such as wave generators and A/D and D/A converters. Some of these are reviewed below.

COUNTERS[4,7]

To count any number N of events, at least k flip-flops are required, such that

$$2^k \geq N \tag{1}$$

Ripple Counters

JK flip-flops with *J* and *K* inputs equal to 1 are *divide-by-2* circuits. Hence, a cascade of k flip-flops with each output Q driving the clock of the next circuit forms a divide-by-2^k chain, or a binary counter (see Fig. 14.4.1). The main drawback of this circuit is its increasing propagation delay with k. When all the flip-flops switch, the clock signal must ripple through the entire counter. Hence, enough time must be allowed to obtain the correct count. Furthermore, the delays between the various stages of the counter may produce glitches, e.g., when parallel decoding is achieved.

Synchronous Binary Counters

Minimization of delay and glitches can be achieved by designing synchronous instead of asynchronous counters. In a synchronous counter all the clock inputs of the *JK* flip-flops are driven in parallel by a single clock signal. The control of the counter is achieved by driving the *J* input by means of the AND combination of all the preceding Q outputs, as shown in Fig. 14.4.2. In this manner, all state changes occur on the same trailing edge of the clock signal. The only remaining requirement is to allow enough time between clock pulses for the propagation through a single flip-flop and an AND gate. The drawback of course is increased complexity with the order k.

Divide-by-*N* Synchronous Counters

When the number N of counts cannot be expressed in binary form, auxiliary decoding circuitry is required. This is true also for counters that provide truncated and irregular count sequences. Their synthesis is based

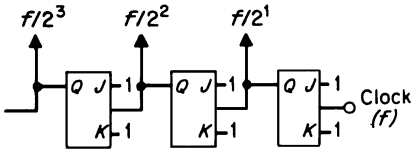


FIGURE 14.4.1 Ripple counter formed by cascading flip-flops.

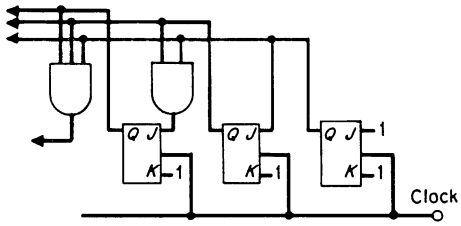


FIGURE 14.4.2 Synchronous counter.

on the so-called transition tables. The basic transition table of the *JK* flip-flop is derived easily from its truth table.

Q_n	Q_{n+1}	J	K	Line
0	0	0	×	1
0	1	1	×	2
1	0	×	1	3
1	1	×	0	4

Line 1, for instance, means that in order to maintain Q equal to 0 after a clock signal has occurred, the J input must be made equal to 0 whatever K may be (X stands for "don't care"). This can be easily verified with the truth table (lines 1 and 3). Hence, the synthesis of a synchronous counter consists simply of determining the J and K inputs of all flip-flops that are needed to obtain a given sequence of states.

Once the J and K truth tables have been obtained, classical minimization procedures can be used to synthesize the counter. For instance, consider the design of a divide-by-5 synchronous counter for which a minimum of three flip-flops is required. First, the present and next states of the flip-flops are listed. Then the required J and K inputs are found by means of the *JK* transition table:

	Present state			Next state			*JK* inputs					
	Q_3	Q_2	Q_1	Q_3	Q_2	Q_1	J_3	K_3	J_2	K_2	J_1	K_1
0	0	0	0	0	0	1	0	×	0	×	1	×
1	0	0	1	0	1	0	0	×	1	×	×	1
2	0	1	0	0	1	1	0	×	×	0	1	×
3	0	1	1	1	0	0	1	×	×	1	×	1
4	1	0	0	0	0	0	×	1	0	×	0	×

Using Karnaugh minimization techniques, one finds

$$J_3 = Q_1Q_2 \quad K_3 = 1$$
$$J_2 = Q_1 \quad K_2 = Q_1$$
$$J_1 = Q_3 \quad K_1 = 1$$

The corresponding counter is shown in Fig. 14.4.3. With this procedure it is quite simple to synthesize a decimal counter with four flip-flops. The same method also applies to the synthesis with D flip-flops.

Up-Down Counters

Upward and downward counters differ from the preceding ones only by the fact that an additional bit is provided to select the proper J and K controls for up or down count. The same synthesis methods are applicable.

Presettable Counters

Since a counter is basically a chain of flip-flops, parallel loading by any number within the count sequence is readily possible. This can be achieved by means of the set terminals. Hence, the actual count sequence can be initiated from any arbitrary number.

SHIFT REGISTERS[3]

Shift registers are chains of flip-flops connected so that the state of each can be transferred to its next left or right neighbor under control of the clock signal. Shift registers can be built with *JK* as well as *D* flip-flops. An example of a typical bidirectional shift register is shown in Fig. 14.4.4. Shift registers, such as counters, may be loaded in parallel or serial mode. This is also true for reading out.

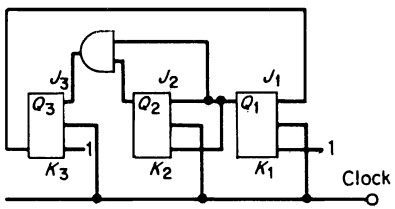
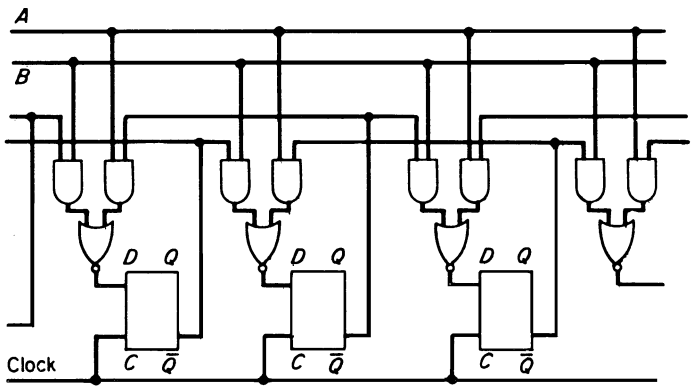


FIGURE 14.4.3 Synchronous divide-by-5 circuit.

In MOS technology, dynamic shift registers can be implemented in a simple manner. Memorization of states occurs electrostatically rather than by means of flip-flops. The information is stored in the form of charge on the gate of an MOS transistor. The main advantage of MOS dynamic shift registers is area saving resulting from the replacement of flip-flops by single MOS transistors. A typical dynamic 2-phase shift register (Fig. 14.4.5) consists of two cascaded MOS inverters connected by means of series switches. These switches, T_3 and T_6, are divided into two classes: *odd* switches controlled by the clock ϕ_1, and *even* switches controlled by ϕ_2.

The control signals determine two nonoverlapping phases. When ϕ_1 turns on the odd switches, the output signal of the first inverter controls the next inverter but T_6 prevents the information from going further. When ϕ_2 turns on, data are shifted one half cycle further. The signal jumps from one inverter to the next until it reaches the last stage.



Shift	A	B
Right	0	1
Left	1	0

FIGURE 14.4.4 Bidirectional shift register.

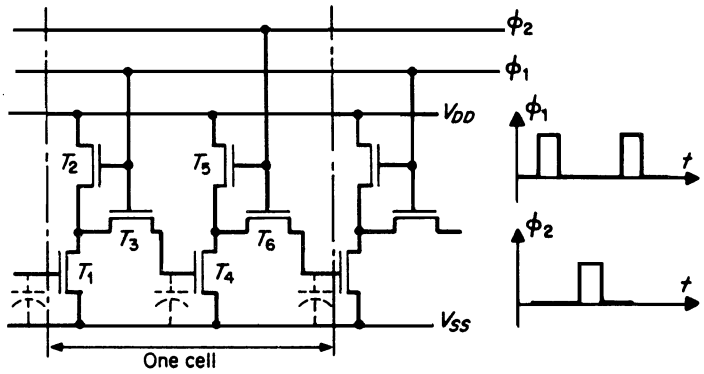


FIGURE 14.4.5 A 2-phase MOS dynamic shift register.

MULTIPLEXERS, DEMULTIPLEXERS, DECODERS, ROMS, AND PLAS[3–8]

A multiplexer is a combinatorial circuit that selects binary data from multiple input lines and directs them to a single output line. The selected input line is chosen by means of an address word. A representation of a 4-input multiplexer (MUX) is shown in Fig. 14.4.6 as well as a possible implementations. MOS technology lends itself to the implementation of multiplexers based on pass transistors. In the example illustrated by Fig. 14.4.7 only a single input is connected to the output through two conducting series transistors according to the S_1 S_0 code address.

Multiplexers may be used to implement canonical logical equations. For instance, the 4-input MUX considered above corresponds to the equation

$$y = x_0\bar{S}_0\bar{S}_1 + x_1S_0\bar{S}_1 + x_2\bar{S}_0S_1 + x_3S_0S_1 \tag{3}$$

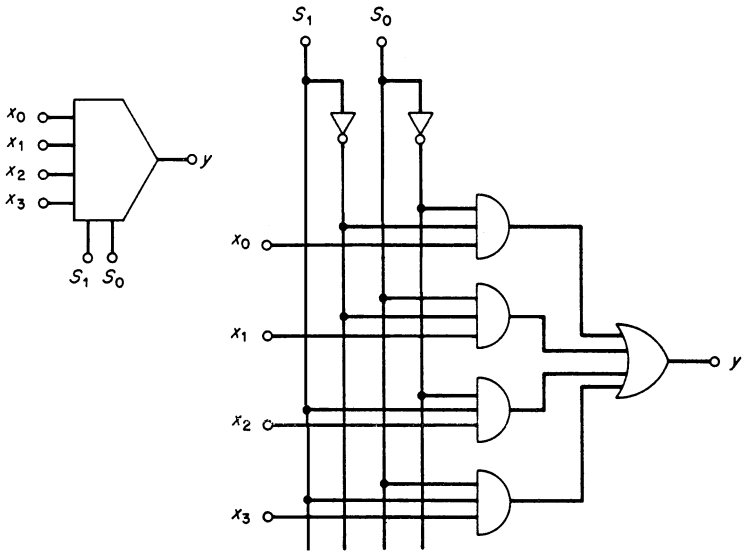


FIGURE 14.4.6 A 4-input multiplexer (MUX) with a 2-bit (S_1S_0) address.

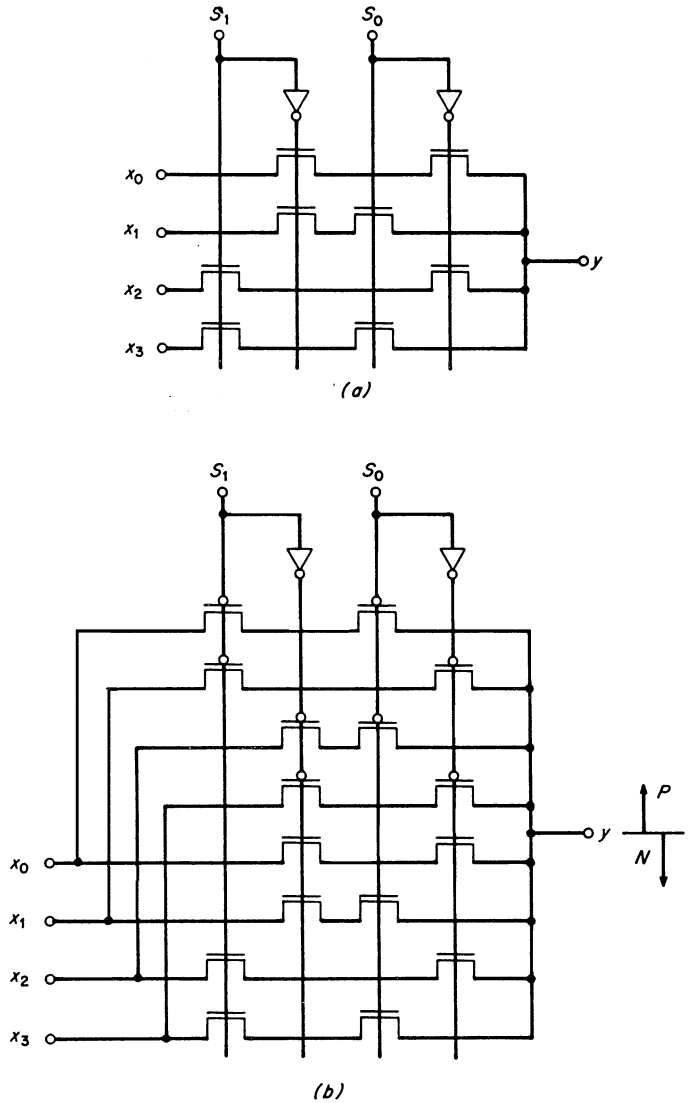


FIGURE 14.4.7 (*a*) NMOS and (*b*) CMOS implementations of multiplexers.

Hence, logical functions of the variables S_0 and S_1 can be synthesized by means of multiplexers. For instance, an EXOR circuit corresponds to $x_0 = x_3 = 0$ and $x_1 = x_2 = 1$.

Demultiplexers (DEMUX) perform the inverse passage from a single input line to several output lines under the control of an address word. They implement logic functions that are less general than those performed by multiplexers because only miniterms are involved. The symbolic representation of a 4-input DEMUX is shown in Fig. 14.4.8 with a possible implementation. The MUX represented in Fig. 14.4.7 may seem attractive for the purpose to achieve demultiplexation but one should not forget that when a given channel is disconnected, the data across the corresponding output remain unchanged. In a DEMUX, unselected channels should take a well-defined state, whether 0 or 1.

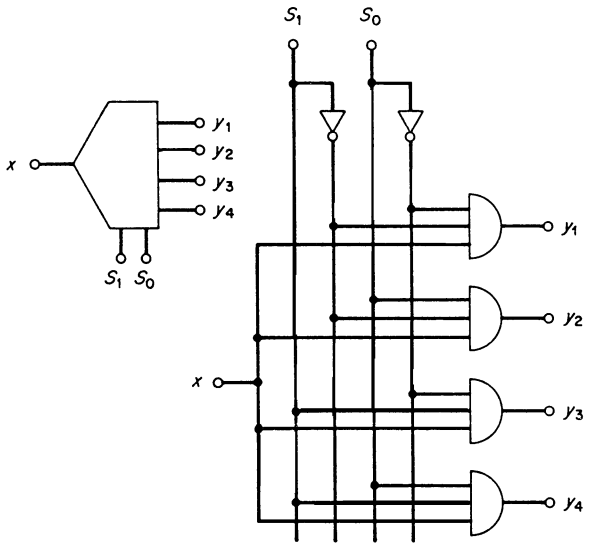


FIGURE 14.4.8 A 4-output demultiplexer (DEMUX).

A decoder is a DEMUX with a constant input. Decoders are currently used to select data stored in memories. Figure 14.4.9 shows a simple NMOS decoder. In this circuit, all unselected outputs are grounded while the selected row is at the logical 1. Any row may be viewed as a circuit implementing a miniterm. For instance, the first row corresponds to

$$y_0 = \overline{S_1 + S_0} = \overline{S}_1 \overline{S}_0 \tag{4}$$

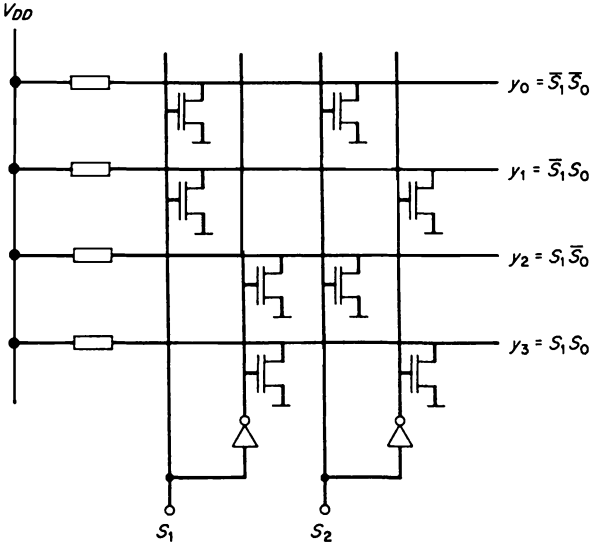


FIGURE 14.4.9 An NMOS decoder circuit.

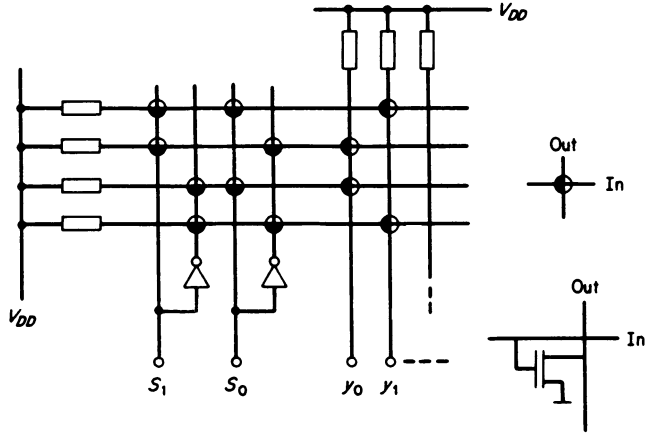


FIGURE 14.4.10 A ROM memory.

When a decoder drives a NOR circuit, a canonical equation is obtained again. Figure 14.4.10 shows a decoder driving several NOR circuits which are implemented along vertical columns. This circuit is called a read-only memory (ROM). One can easily recognize an EXOR function (y_1) and its complement (y_0) in the example. The actual silicon implementation strictly follows the pattern illustrated by Fig. 14.4.10. Notice that the decoder block and ROM column block both have the same structure after turning them by 90°.

When a ROM is used to implement combinatorial logic, the number of outputs usually is restricted to only those miniterms which are necessary to achieve the desired purpose. The ROM is then called a programmable logic array (PLA)

ROMs as well as PLAs are extensively used in integrated circuits because they provide the means to implement logic in a very regular manner. They contribute substantially to area minimization and lend themselves to automatic layout, reducing design time (silicon compilers).

A large number of functions can be implemented by means of the circuits described above: circuits converting the format of digital data, circuits coding pure binary into binary-coded decimal (BCD) or decimal data, and so forth. The conversion from a 4-bit binary-coded number into a 7-segment display by means of a PLA is illustrated by Fig. 14.4.11.

All the required OR functions are obtained by means of the right-plane decoder, while the left one determines the AND functions (they are called, respectively, OR-AND planes). In applications where multiple-digit displays are required, rather than repeating the circuit of Fig. 14.4.11 as many times as there are digits, a combination of MUX and DEMUX circuit and a single PLA can be used. An example is shown in Fig. 14.4.12.

The 4-input codes representing 4 digits are multiplexed in a quadruple MUX in order to drive a PLA 7-segment generator. Data appear sequentially on the seven common output lines connected to the four displays. Selection of a given display is achieved by a decoder driven by the same address as the MUX circuit. If the cyclic switching is done at a speed of 60 Hz or more, the human eye cannot detect the flicker.

MEMORIES[3]

Memories provide storage for large quantities of binary data. Individual bits are stored in minimum-size memory cells that can be accessed within two- or three-dimensional arrays. Read-only memories (ROMs) provide access only to permanently stored data. Programmable ROMs (PROMs) allow data modifications only during special write cycles, which occur much less frequently than readout cycles. Random access memories (RAMs) provide equal opportunities for read and write operations. RAMs store data in bistable circuits (flip-flops, SRAMs), or in the form of charge across a capacitor (DRAM).

Decimal display	Binary coded input x_3	x_2	x_1	x_0	7 segment code A	B	C	D	E	F	G
0	0	0	0	0	1	1	1	1	1	1	0
1	0	0	0	1	0	1	1	0	0	0	0
2	0	0	1	0	1	1	0	1	1	0	1
3	0	0	1	1	1	1	1	1	0	0	1
4	0	1	0	0	0	1	1	0	0	1	1
5	0	1	0	1	1	0	1	1	0	1	1
6	0	1	1	0	1	0	1	1	1	1	1
7	0	1	1	1	1	1	1	0	0	0	0
8	1	0	0	0	1	1	1	1	1	1	1
9	1	0	0	1	1	1	1	1	0	1	1
10–15	—				x						

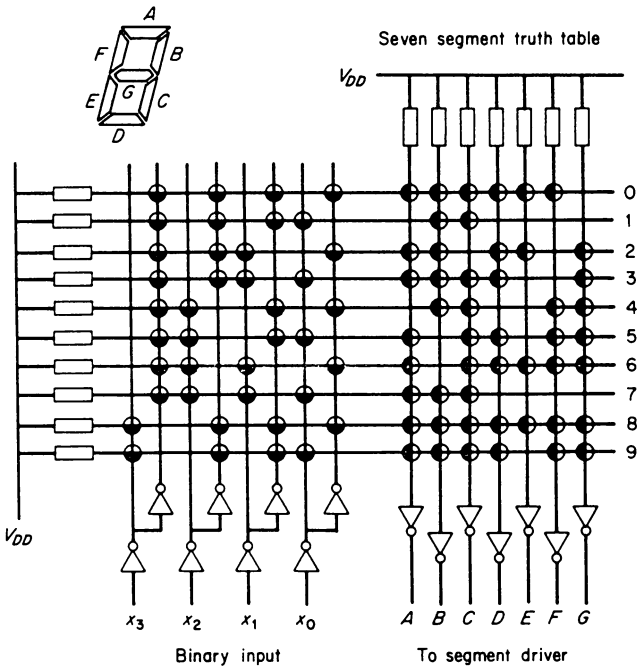


FIGURE 14.4.11 A 7-segment PLA driver.

Static Memories

Static memory cells are arrays of integrated flip-flops. The state of each flip-flop represents a stored bit. Read-out is achieved by selecting a row (word-line: WL) and a column (but or bit-lines: BL) crossing each other over the chosen cell. Writing occurs in the same manner. A typical six-transistor MOS memory cell is shown in Fig. 14.4.13. The word-line controls the two transistors that connect the flip-flop to the read-write bus (BL and BLbar). In order to read out nondestructively the data stored in the cell, the read-write bus must be precharged. The inverter that is in the low-state discharges the corresponding bit-line. Writing data either confirms or changes the state of the flip-flop. When a change must happen, the high bit-line must overrule the corresponding low-state inverter, while the low bit-line simply discharges the output node of the high inverter. The second event takes less time than the first for the big differences between bus and inverter node capacitances. Bus capacitances are at least 10 times larger than cell node capacitances. The rapid discharge of the high inverter output node blocks the other half of the flip-flop before it has a chance to discharge the high bus. In order to prevent a read cycle from

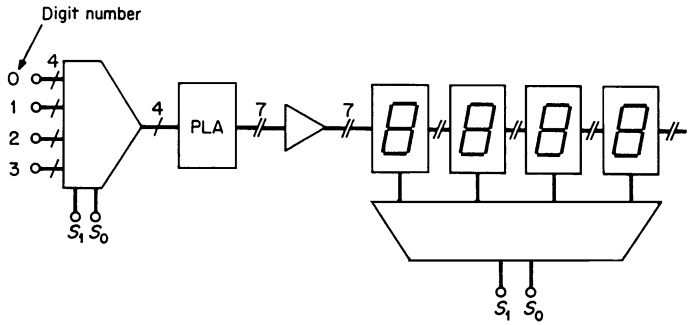


FIGURE 14.4.12 A 4-digit, 7-segment multiplexed display.

becoming a write-cycle, it is important to equalize the voltages of the bit-lines prior to any attempt to read out. This is achieved by means of a single equalizing transistor on top of the read-write bus. This transistor shorts the two bit-lines for just long enough time to neutralize any bit-line voltage difference whatever the mean voltage may be. The pull-up circuitry precharging the read-write bus accommodates conflicting requirements. It must load the bit-lines as fast as possible but not counteract the discharge of the bit-line, which is tied to the low data during read out. Usually, the pull-up circuitry is clocked and driven by the same clock that controls the equalizing transistor.

Cell size and power consumption are the two key items that determine the performances and size of present memory chips which may count as much as several million transistors. Cell sizes have shrunken continuously until they are a few tens of microns square. Static power is minimized by using CMOS instead of resistively loaded inverters. Most of the power needed is to provide fast load and discharge cycles of the line capacitances. Therefore a distinction is generally made between standby and active conditions. The load-discharge processes imply short but very large currents. The design of memories (static as well as dynamic) has always been at the forefront of the most advanced technologies.

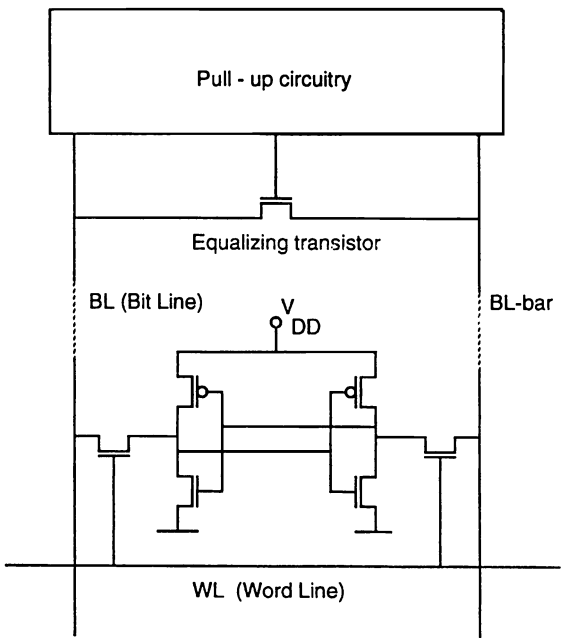


FIGURE 14.4.13 The basic six-transistor circuit of a static MOS memory cell.

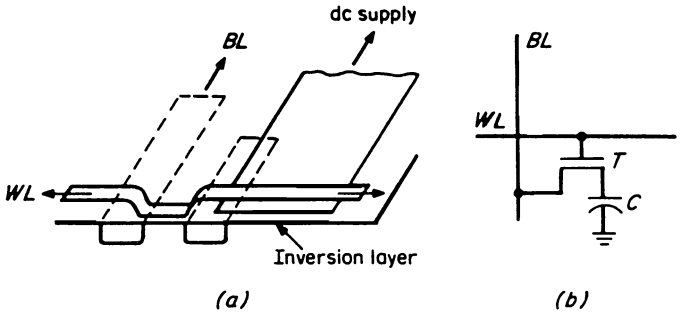


FIGURE 14.4.14 The one-transistor MOS memory cell; (*a*) IC implementation; (*b*) equivalent circuit.

Dynamic Memories[8,9]

The trend toward ever larger memory chips has led to the introduction of single transistor memory cells. Here the word-line controls the gate of the MOS switch that connects the single bit line to a storage capacitor (Fig. 14.1.14). Charge is pumped from (or fed back to) the bit-line by properly choosing the voltage of the bit-line. The actual data consist of charge (or absence of charge) stored in the capacitor in the inversion layer below a field plate. Typical storage capacitances are 100 to 50 fF. The useful charge is quite small, around 0.1 or 0.2 pC. In order to read data, this charge is transferred to the bit-line capacitance, which is usually one order of magnitude larger. The actual voltage change of the bit-line is thus very small, typically 100 to 200 mV. In order to restore the correct binary data, amplification is required. The amplifier must fulfill a series of requirements: It must fit in the small space between cells, be sensitive and respond very rapidly, but also distinguish data from inevitable switching noise injected by the overlap capacitance of the transistor in series with the storage capacitor. To solve this problem, dynamic memories generally use two bit-line output lines instead of a single one. Read-out of a cell always occurs in the same time as the selection of a single identical dummy cell. This compensates switching noises as far as tracking of parasitic capacitances of both switching transistors is achieved. The key idea behind the detection amplifier is the metastable state of a flip-flop. The circuit is illustrated in Fig. 14.4.15. Before read-out, the common source terminal is left open. Hence, assuming the bit-line voltages are the same, the flip-flop behaves like two cross-coupled diode-connected transistors.

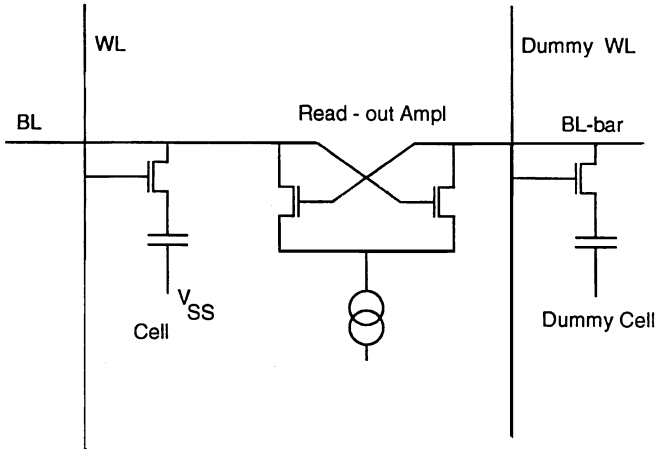


FIGURE 14.4.15 The readout amplifier of dynamic one-transistor-per-cell MOS memories consists of a flip-flop in the metastable state in order to enhance sensitivity and speed and to minimize silicon area.

The common source voltage adjusts itself to the pinch-off voltage of the transistors. After read-out has occurred, a small voltage imbalance is experienced between bit-lines. The sign of this imbalance is determined by the actual data previously stored in the cell. If then the common source node voltage is somewhat reduced, one of the two transistors becomes slightly conducting while the other remains blocked. The conducting transistor slightly discharges the bit-line with the lowest voltage, increasing the voltage imbalance. The larger this difference, the faster the common source node can be lowered without risk to switch-on of the second transistor. After a short time, the imbalance is large enough to connect the common source to ground. The final state of the flip-flop reproduces the content of the selected cell. Notice that data read during the read-out cycle are now available for rewriting in the cell. Permitting the transistor in series with the storage capacitor to conduct longer performs automatic rewriting. Because storage is likely to be corrupted by leakage current, this same procedure is repeated regularly, even when data are not demanded, in order to keep the memory active. Dynamic memories are the densest circuits designed. Memory sizes currently attain 16 Mbits, and 256 Mbits memories are being developed.

Commercially Available Memory Chips

Memory arrays usually consist of several square arrays packed on a single chip. In order to access data, the user provides an address. The larger the memory size, the longer the address. In order to reduce pin count, address words are divided in fields controlling less pins in a sequential manner. Many memories are accessed by a row address strobe (RAS) followed by a column address strobe (CAS), each one requiring only half of the total number of address bits. Memories represent a very substantial part of the semiconductor market. Besides RAMs (random access memories), a large share of nonvolatile memories is available comprising ROMs (read-only memories), programmable read-only memories (PROMs), and electrically alterable read-only memories (EAROMs), which are considered below.

RAM. Random access memories (RAMs) are either static or dynamic memory arrays like those described above. They are used to read and write binary data. Storage is warranted as long as the power supply remains on.

ROM. In read-only memories, the information is frozen. Information can only be read out. Access is the same as in RAMs. Memory elements are single transistors. Whether a cell is conducting or not has been determined during the fabrication process. Some MOS transistors have a thin gate oxide layer; others have a thick oxide layer, or some are connected to the access lines, while others are not.

PROM. The requirements for committing to a fixed particular memory content, inherent in the structure of ROMs, is a serious disadvantage in many applications. PROMs (programmable ROMs) allow the manufacturer or the user to program ROMs. Various principles exist:

- With mask programmable ROMs, a single mask is all the manufacturer needs to implement the client code.
- With fuse link programmable ROMs, the actual content is written in the memory by blowing a number of microfuses. This allows the customer to program the ROM, but the final stage is irreversible as in mask programmable ROMs.

EPROM. In electrically programmable ROMs the data are stored as shifts of the threshold voltages of the memory transistors. Floating-gate transistors are currently used for this purpose. Loading the gate occurs by hot electrons tunneling through the oxide from an avalanching junction toward the isolated gate. The charges trapped in the gate can only be removed by ultraviolet light, which provides the energy required to cross the oxide barrier.

EEPROM. Electrically erasable PROMs use metal-nitride-oxide-silicon transistors in which the charges are trapped at the nitride oxide interface. The memory principle is based on Fowler-Nordheim tunneling to move charges from the substrate in the oxide or vice versa. The electric field is produced by a control gate. EAROMs (electrically alterable ROMs) use the same principle to charge or discharge a floating gate with some additional features providing better yields.

DIGITAL-TO-ANALOG CONVERTERS (D/A OR DAC)[10]

Converting data from the digital to the analog domain may be achieved in a variety of ways. One of the most obvious is to add electrical quantities such as voltages, currents or charges following a binary weighted scale. Circuits illustrating this principle follow. They are of two kinds: first, converters aiming at integral linearity and, second, converters exchanging integral linearity for differential linearity.

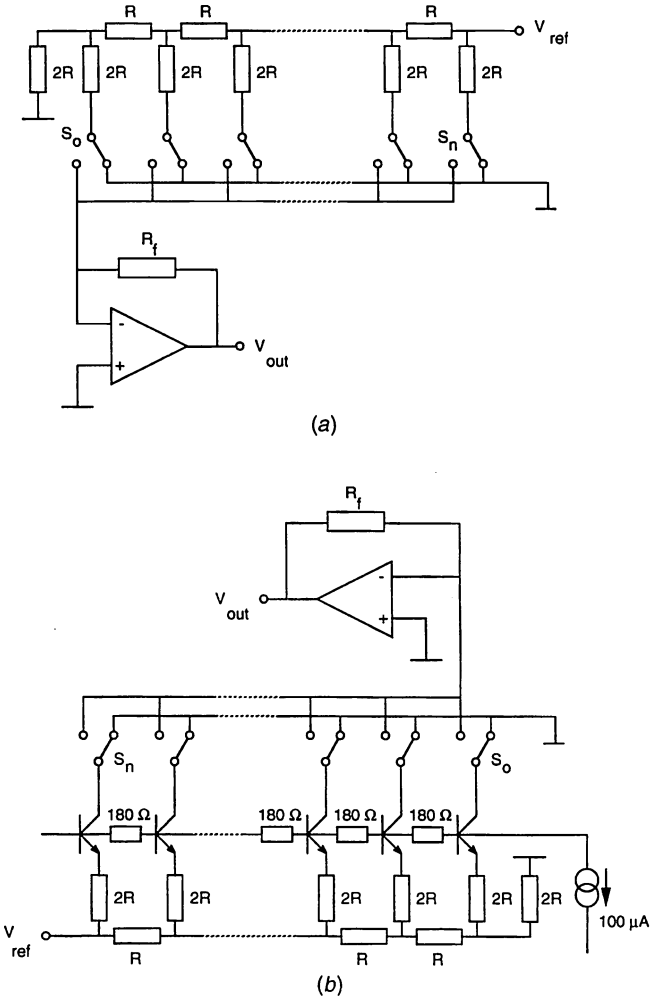


FIGURE 14.4.16 Integral linearity converters. Switching is implemented by (*a*) MOS transistors and (*b*) bipolar circuits.

Integral Linearity Converters

A binary scale of voltages and currents is easily obtained by means of the well-known *R*-2*R* network. In the circuit shown in Fig. 14.4.16*a*, binary weighted currents flowing in the vertical resistances are either dumped to ground or injected into the summing node of an operational amplifier. The positions of the switches reproduce the code of the digital word to be converted. Practical switches are implemented by means of MOS transistors. Their widths double going from left to right to keep the voltage drops constant between drains and sources. The resulting constant offset voltage can be compensated easily. Another approach, better suited for bipolar circuits, is illustrated in Fig. 14.4.16*b*. Here the currents are injected into the emitters of transistors whose outputs feed bipolar current switches performing the digital code conversion. To keep all emitters at the same potential, small resistors are placed between bases of the transistors. They introduce voltage drops of 18 mV (UT.ln2) to compensate for the small base to emitter voltages changes resulting from the systematic halving of collector current going from left to right.

DIGITAL AND ANALOG SYSTEMS

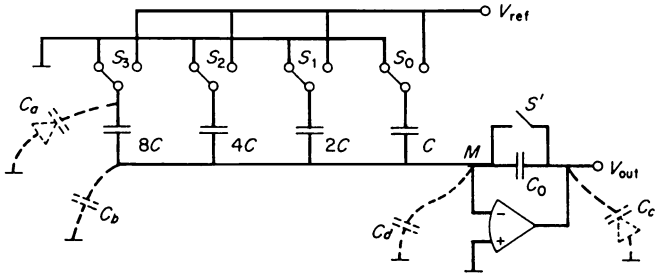


FIGURE 14.4.17 Capacitive D/A converter with charge integration op-amp.

The accuracy of *R*-2*R* circuits is based on the matching of the resistances. It does not exceed 10 bits in practice, and may reach 12 bits when laser trimming is used.

Switched capacitors are used also to perform D/A conversion. Capacitors are easy to integrate and offer superior temperature characteristics compared to thin-film resistors. However, they suffer from several drawbacks and require special care in order to provide good matching. Integrated capacitors are sandwiches of metal-oxide-silicon layers (MOS technology) or polysilicon-oxide-polysilicon (double poly technology). The latter exhibit a highly linear behavior and extremely small temperature coefficient (typically 10 to 20 ppm/°C). In order to improve geometrical tracking, any capacitor must be a combination of "unit" capacitors that represent the minimum-sized element available (typically 100 fF). MOS capacitors have relatively large stray capacitances to the substrate through their inversion layer, about 10 to 30 percent of their nominal value, depending on the oxide thickness. Doubly poly capacitors offer better figures, but their technology is more elaborate. Whatever choice is made, the circuits must be designed to be inherently insensitive to stray capacitance. This is the case for the circuit shown in Fig. 14.4.17.

In this circuit, a change of the position of a switch results in a charge redistribution among a weighted capacitor and the feedback capacitor C_0. The transferred charge is stored on the feedback capacitor so that the output voltage stays constant, provided leakage currents are small (typically 10^{-15} A). The stray capacitors are illustrated in Fig. 14.4.17. They have no influence on the converter accuracy. Indeed, C_a and C_c are loaded and discharged according to position changes of switch S_3, but the currents through those capacitors do not flow through the summing junction. On the other hand, C_b as well as C_d is always discharged and therefore do not influence the accuracy.

Another version of a capacitive D/A converter is shown in Fig. 14.4.18. In this network, changing the position of switches S_3 and S' produces a charge redistribution between the capacitor at the extreme left and the parallel combination of all other capacitors that represent a total of 8 *C*. Hence the upper node undergoes a voltage swing equal to $V_{ref}/2$. A binary scale of voltages consequently may be produced according to the various switches. Switch S' fixes the initial potential of the upper node and allows the discharge of all capacitors prior to conversion. There is no specific requirement that the upper node be connected to ground or to any fixed potential V_0, provided V_{out} is evaluated against V_0.

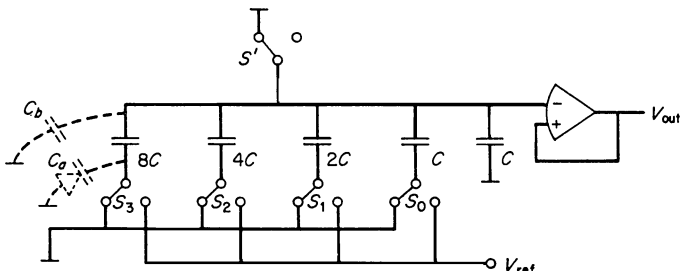


FIGURE 14.4.18 Capacitive D/A converter with unity-gain buffer.

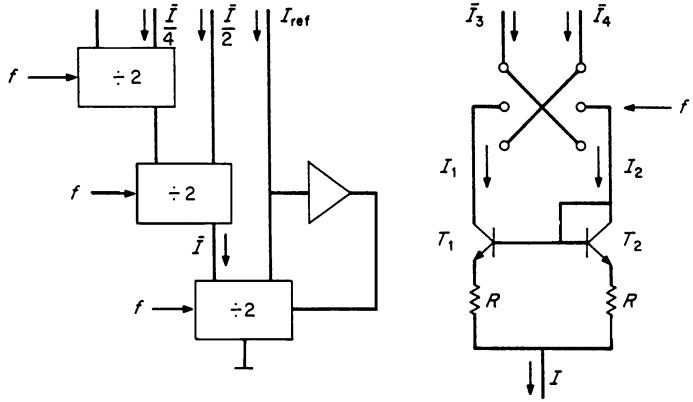


FIGURE 14.4.19 Dynamic element matching converter principle.

Notice that the stray capacitance C_b degrades the converter accuracy, while C_a has no effect at all. Careful layout therefore is required to minimize C_b, unless this stray capacitance is included in the evaluation of the capacitor C at the extreme right. For this reason, only small numbers of bits should be considered. The unavoidable exponential growth of capacitances with the number of bits is relieved to some extent because the capacitance area increases as the square of dimensions.

Another widely used technique to perform D/A conversion is paralleling of many identical transistors controlled by a single base-to-emitter or gate-to-source voltage source. The output terminals are tied together in order to implement banks of binary weighted current sources. This requires $2^{(N+1)}$ transistors to make an N-bit converter, which is very efficient for transistors are the smallest on-chip devices available. Usually all the transistors are placed in an array and controlled by a row and column decoder, as in memories. One of the interesting features of this type of converter is its inherent monotonicity. Furthermore, if access to the individual transistors follows a pseudorandom pattern, many imperfections resulting from processing, such as an oxide gradient in the MOS process, have counterbalancing effects. Hence, an accuracy of 10 bits may be achieved with transistors whose standard deviation is much worse.

A converter[11] with excellent absolute accuracy is illustrated in Fig. 14.4.19. It is based on high-accuracy divide-by-2 blocks like those shown in the right part of the illustration. Each block consists of a Widlar circuit with equal resistances in the emitters of transistors T_1 and T_2 in order to split the current I approximately into two equal currents: I_1 and I_2 (1 percent tolerance). A double-throw switch is placed in series with the Widlar circuit in order to alternate I_1 and I_2 at the rate of the clock f. Provided the half clock periods t_1 and t_2 can be made equal within 0.1 percent, a condition that can easily be met, one obtains two currents whose averages I_3 and I_4 represent $I/2$ with an accuracy approaching 10^{-5}. This technique, known as the dynamic element matching method, has been successfully integrated in bipolar technology offering an accuracy of 14 bits without expensive element trimming. A band-gap current generator provides the reference current I_{ref}, which is compared to the right-side output current of the first divide-by-2 block. A high current-gain amplifier closes the feedback loop controlling the base terminals of T_1 and T_2 of the first block.

Segment Converters

In some applications, accuracy does not imply absolute linearity but rather differential linearity, which is very critical because errors between two successive steps larger than a half LSB cannot be tolerated. The difficulty occurs when changes of MSBs occur. To overcome the problem, segment converters were designed. The idea is to divide the full conversion scale into segments. A D/A converter with a maximum of 10 bits is used within segments. Passing from one segment to another is achieved in a smooth manner as in the circuit illustrated by Fig. 14.4.20.[12] In this circuit, the 10-bit segment converter is powered by one of the four bottom current sources. A 2-bit decoder,

DIGITAL AND ANALOG SYSTEMS

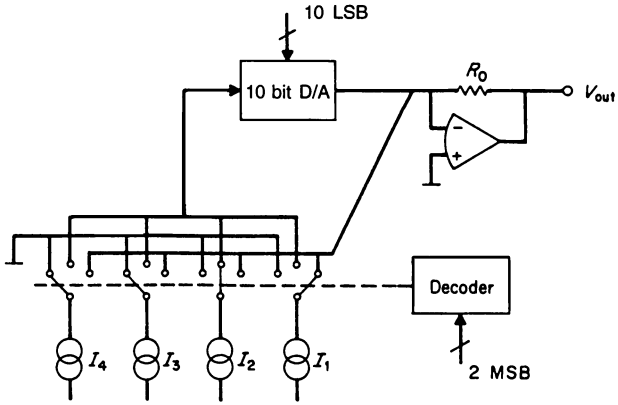


FIGURE 14.4.20 Segment D/A converter.

which is controlled by the two MSBs, makes the appropriate connections. When the two MSBs are 00, current source I_1 is chosen and the three others are short-circuited to ground. When the MSB pattern changes to 01, the switches take the positions illustrated in the figure. Then I_2 supplies the 10-bit converter while I_1 is injected directly into the summing node of the op amp. Hence, no change occurs in the main current component. The differential linearity is entirely determined by the 10-bit converter. The price paid for this improvement is a degradation of absolute linearity. Small discrepancies among the four current sources introduce slope or gain changes between segments. The absolute linearity specifications thus do not reach the differential linearity specs of the converter.

Voltage segment converters can also be integrated. Instead of parallel current sources, series voltage sources are required. This is achieved by means of a chain of equal resistors dividing the reference voltage into segments. The D/A converter is connected along the chain by means of two unity-gain buffers in order not to affect input voltages. Passing from one segment to the next implies that the buffers are interchanged in order to avoid the degradation of the differential linearity, which could result from different offsets in the buffers. An integrated 16-bit D/A converter has been built along these lines.[13]

When a switched-capacitor D/A converter is embedded in a voltage segment converter, no buffer amplifiers are needed because no dc current drain is required. This property simplifies greatly the design of moderate accuracy converters and has been extensively used in the design of integrated codecs.

Codecs. Codecs are nonlinear converters used in telephony. To cover the wide dynamic range of speech satisfactorily with only 8 bits, steps of variable size must be considered: small steps for weak signals, and large steps when the upper limit of the dynamic range is approached. The four LSBs of a codec-coded word define a step size within a segment whose number is given by the three following bits. The polarity of the sample is given by the MSB. Going from one segment to the next implies doubling the step size in order to generate the nonlinear conversion law (μlaw). In the codec circuit of Fig. 14.4.21,[14] bottom plates are connected either in the middle of the resistive divider considered as an ac ground (position 1), or to V_{low} or V_{high} (position 2 plus switch S_1 or S_2) representing, respectively, the negative and positive reference voltage sources, or to a third position (position 3) connecting a single capacitor to an appropriate node of the resistive voltage divider. Hence, the resistive divider determines the step size (four LSBs), the capacitive divider defines the segment (three following bits), and switches S_1 or S_2 control the sign (MSB).

ANALOG-TO-DIGITAL (A/D) CONVERTERS (ADC)[10]

Some analog-to-digital converters are in fact D/A converters embedded in a negative feedback loop with appropriate logic (see Fig. 14.4.22). The analog voltage to be converted is sensed against the analog output of a D/A converter, which in turn is controlled by a logic block so as to minimize the difference sensed by the

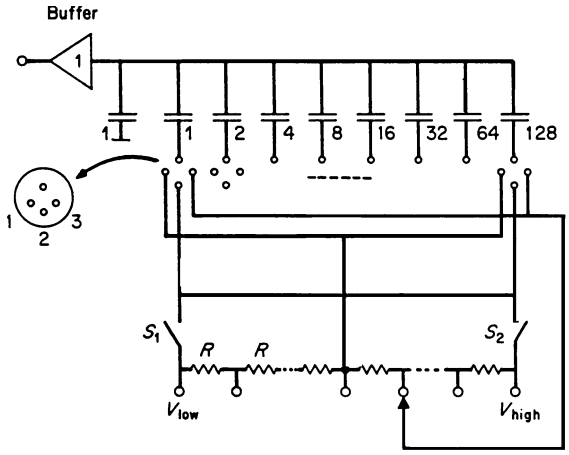


FIGURE 14.4.21 Segment type codec.

comparator. Various strategies have been proposed for the logic block, depending on the type of D/A converter used. Most of the converters described previously have been implemented as ADCs in this manner. In particular, ladder and capacitive D/A converters are used in the so-called successive approximation D/A converter. In those devices, the logic block first sets the MSB; then the comparator determines whether the analog input voltage is larger or smaller than the half-reference voltage. Depending on the result, the next bit is introduced, leaving the MSB unchanged if the analog signal is larger and resetting the MSB in the opposite state. The algorithm is repeated until the LSB is reached.

Successive approximation ADCs are moderately fast, for conversion time is strictly proportional to the number of bits and is determined by the settling time of the op-amp. The algorithm implies that the analog input voltage remains unaltered during the conversion time; otherwise it may fail. To avoid this, a sample-and-hold (SH) circuit must be placed in front of the converter.

A capacitive A/D converter is shown in Fig. 14.4.23.[15] The conversion is accomplished by a sequence of three operations. During the first, the unknown analog input voltage is sampled. The corresponding positions of the switches are as illustrated in Fig. 14.4.23*a*. The total stored charge is proportional to the input voltage V_{in}. In the second step, switch S_A is opened and the positions of all S_i switches are changed (Fig. 14.4.23*b*). The bottom plates of all capacitors are grounded, and consequently the voltage at the input node of the comparator

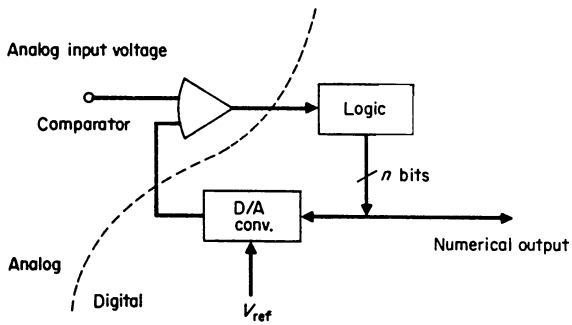


FIGURE 14.4.22 A D/A converter within a feedback loop forms an A/D converter.

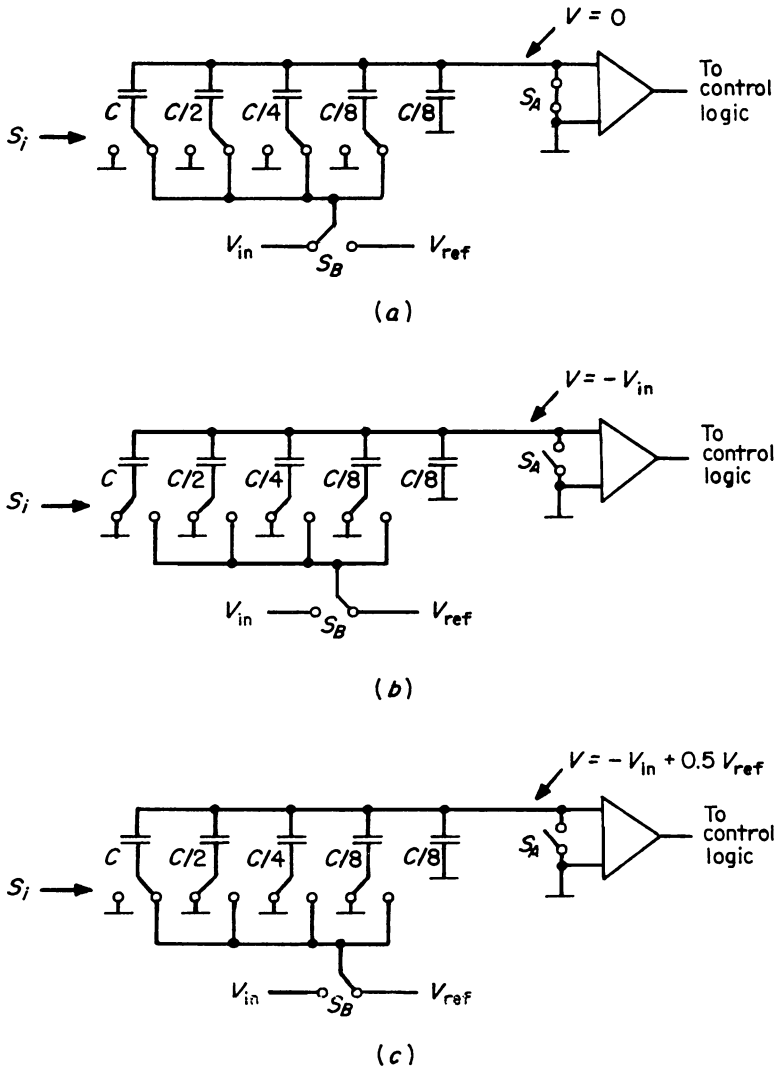


FIGURE 14.4.23 An integrated switched-capacitor A/D converter: (*a*) sample mode; (*b*) hold mode; and (*c*) redistribution mode.

equals $-V_{in}$. The third step is initiated when raising the bottom plate of the MSB capacitor from ground to the reference voltage V_{ref} (Fig. 14.4.23*c*). This is done by changing again the position of the MSB switch and connecting S_b to V_{ref} instead of V_{in}. The voltage at the input node of the comparator is thus increased by $V_{ref}/2$, so that the comparator's output is a logic 1 or 0, according to the sign of $(V_{ref}/2 - V_{in})$.

The circuit operates similarly to any successive approximations converter. That is, V is compared to $(V_{ref}/2 + V_{ref}/4)$ when the result of the previous operation is negative, or the MSB switch returns to its initial position and only the comparison with $V_{ref}/4$ is considered. After having carried out the same test until the LSB is reached, the conversion is achieved and the digital output may be determined from the position of the switches. The voltage of the common capacitive node is approximately equal to, or smaller than, the smallest incremental step.

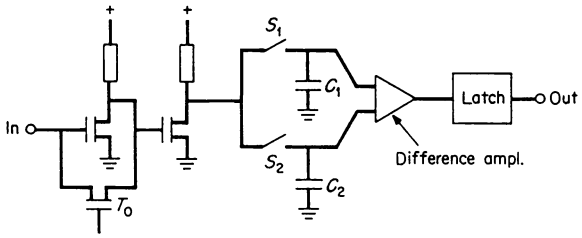


FIGURE 14.4.24 A typical MOS comparator for an A/D converter.

Since this is almost negligible, the stray capacitance of the upper common node is practically discharged. It thus has no effect on the overall accuracy of the converter. Hence the problem in the circuit of Fig. 14.4.18 does not occur. The name "charge redistribution A/D converter" was given to this circuit because the charge stored during the first sequence is redistributed among only those capacitors whose bottom plates are connected to V_{ref} after the conversion algorithm is completed.

The design of comparators able to discriminate steps well below the offset voltage of MOS amplifiers is a problem deserving careful attention. An illustration of an MOS comparator is shown in Fig. 14.4.24.[15] The amplifier comprises three parts: a double inverter, a differential amplifier, and a latch. The input-output nodes of the first inverter may be short-circuited by means of transistor T_0 to bias the gates and drains of the identical two first inverters at approximately half the supply voltage. This procedure occurs during the "sampling" phase of the converter and corresponds to the grounding of switch S_A in Fig. 14.4.23. As stated above in the comments concerning the circuit of Fig. 14.4.17, there is no necessity to impose zero voltage on this node during sampling; any voltage is suitable as long as the impedance of the voltage source is kept small. This is what occurs when T_0 is on, since the input impedance of the converter is then minimum because of the negative-feedback loop across the first stage. Once the double inverter is preset, T_0 opens and the output voltage of the second inverter is stored across the capacitor C_1, momentarily closing switch S_1. Then the floating input node of the comparator senses the voltage change resulting from charge redistribution during a test cycle and reflects the amplified signal at the output of the second inverter. Switch S_2 in turn samples the new output so that the differential amplifier sees only the difference between the two events. Any feed through noise from T_0 is ignored since it only affects the common mode of the voltages sampled on C_1 and C_2. Hence, the overall offset voltage of this comparator is equal to the offset voltage of the differential amplifier divided by the gain of the double inverter input stage. Signals in the millivolt range can be sensed correctly regardless of the poor behavior of MOS transistors' offset voltages.

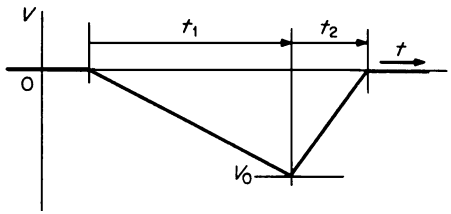


FIGURE 14.4.25 The dual-slope A/D converter principle.

Another integrated converter widely accepted in the field of electrical measurements is the dual-ramp A/D converter.[13] The principle of this device is first to integrate the unknown voltage V_x during a fixed duration t_1 (see Fig. 14.4.25). Then the input signal V_x is replaced by a reference voltage V_{ref}. Since the polarities of both signals are opposite, the integrator provides an output ramp with opposite slope. It takes a time t_2 for the integrator output voltage to return from V_0 to zero. Hence

$$V_0 = -\frac{1}{RC}\int_0^{t1} V_x\,dt \tag{5}$$

and

$$V_0 = (V_{ref}/RC)t_2 \tag{6}$$

Consequently, if V_x is a constant

$$V_x = -V_{\text{ref}} t_2/t_1 \tag{7}$$

the time t_1 is determined by a counter. At the moment the counter overflows, switching of the input signal from V_x to V_{ref} occurs. The counter is automatically reset, and a new counting sequence initiated. When the integrator output voltage has returned to zero, the comparator stops the counting sequence. Thus the actual count is a direct measure of t_2.

The dual-ramp converter has a number of interesting features. Its accuracy is not influenced by the value of the integration time constant *RC*; neither is it sensitive to the long-term stability of the clock-frequency generator. The comparator offset can easily be compensated by autozeroing techniques. The only signal that actually controls the accuracy of the A/D converter is the reference voltage V_{ref}. Excellent thermal stability of V_{ref} can be achieved by means of band-gap reference sources.

Another interesting feature of the dual-ramp A/D converter is the fact that since V_x is integrated during a constant time t_1, any spurious periodic signal with zero mean whose period is a submultiple of t_1 is automatically canceled. Hence, by making t_1 equal to an entire number of periods of the power supply one obtains excellent hum rejection.

DELTA-SIGMA CONVERTERS[10,20]

Because the development of integrated systems is driven mainly by digital applications, mixed analog-digital circuits, especially converters, should be implementable equally well without loss of accuracy in digital technologies, even the most advanced ones. Short channels, however, do not just improve bandwidth, they also negatively affect other features, such as dynamic range and 1/*f* noise, because both supply voltage and gate area of the MOS transistors get even smaller. Therefore, a trade-off of speed and digital complexity for resolution in signal amplitude is needed.

Delta-Sigma converters illustrate this trend. Their object is to abate the quantization noise to enhance the signal-to-noise ratio (SNR) of the output data (quantization noise is the difference between the continuous analog data and their discrete digital counterpart). When the number of steps of a converter increases the quantization noise decreases. Similarly, if we enhance the SNR the resolution increases. Delta-Sigma converters take advantage of *oversampling* and *noise shaping* techniques to improve the SNR.

Oversampling refers to the Nyquist criterion, which requires a sampling frequency twice the baseband to keep the integrity of sampled signals. Oversampling does not improve the signal. Neither does it increase the quantization noise power, nor is it a way to spread the quantization noise power density over a large spectrum, thereby lessening its magnitude to keep the noise power unchanged. If we restrict the bandwidth of the oversampled signal to the signal baseband in compliance with the Nyquist criterion, the amount of noise power left in the baseband is divided automatically by the ratio of the actual sampling frequency over the Nyquist frequency. We improve thus the SNR and consequently increase the resolution.

Noise shaping is the other essential feature of Delta-Sigma converters. It refers to the filtering step oversampled data must undergo. The purpose is to further decrease the amount of noise left in the baseband by shifting a large part of the remaining quantization noise outside the baseband. During the step, the SNR is substantially improved.

Both A/D and D/A converters lend themselves to Delta-Sigma architectures. The same principles prevail, but their implementation differs. We consider both separately.

A/D Delta-Sigma Converters[16,26]

In A/D converters, oversampling is done simply by increasing the sampling frequency of the analog input signal. Noise shaping is achieved by means of a nonlinear filter like the one illustrated in the upper part of Fig. 14.4.26. As stated already above, the goal is to shift most quantization noise out of the baseband. The noise shaper consists of a feedback loop similar to the linear circuit shown in the lower part. In the upper figure, the quantization noise is produced by the A/D converter. This converter is followed by a D/A converter, which is required because the difference between the continuous analog input signal and the discrete digital output signal delivered by the

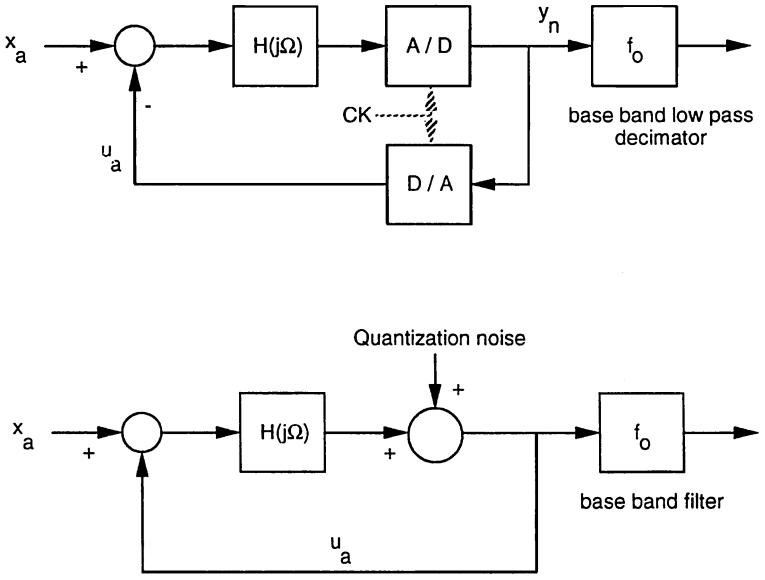


FIGURE 14.4.26 Principle of sigma-delta converters.

A/D converter must be sensed in the analog domain. This difference is amplified and low-passed by means of the analog loop filter *H*. The idea is truly to minimize the signal fed to the amplifier as in the continuous circuit shown below. Consequently, the output signal becomes a replica of the input. This requires some caution, however, because matching continuous input data and discrete output data is not feasible. The situation is very different from the one illustrated in the lower part of Fig. 14.4.26 where the quantization noise is simulated by means of an independent analog continuous noise source. Clearly, if the loop gain is high, the signal outputted by the amplifier should be the sum of the input signal minus the noise in order to make u_a look like x_a. In the upper circuit, however, the amplifier senses steps whose magnitude is determined by the A/D and D/A quantizer. The signal delivered by the amplifier is an averaged image over time of the difference between the input and the discrete signal fed back. The latter tracks the input the best it can—thanks to the feedback loop. A look at the signals displayed in Fig. 14.4.27 confirms this statement. The figure represents output data of a third-order noise shaper with a 3-bit quantizer. The loop filter consists of three integrators in cascade. Their outputs are respectively illustrated in the three upper plots. The signal delivered by the third integrator is applied to the quantizer whose output is shown below. It is obvious that notwithstanding the poor resolution of the A/D converter, the noise shaper tracks the input sine wave pretty well by correcting its output data continuously.

Once the bandwidth of the signal outputted by the noise shaper has been restricted to the baseband, we get a signal whose SNR may be very high. To illustrate this, consider Fig. 14.4.28, which shows a plot of the signal-to-noise improvement that can be obtained with the linear circuit shown in Fig. 14.4.26. Although the actual noise shaper differs from its analog counterpart because the quantization noise is correlated with the input and the quantizer is a nonlinear device, results are comparable. Large SNR figures, 60 or 80 dB and even better, are readily achievable. It is obvious that large OSRs are not the only way to get high SNRs; these can also be obtained by increasing the order of the loop filters (assuming of course stability is achieved, which is not always easy in a nonlinear system). Another interesting feature that stems from the above figure is that A/D and D/A converters with few bits resolution only do not impair the resolution of the Delta-Sigma converter given the large SNRs noise shapers can achieve. Even a single-bit quantizer, notwithstanding its large amount of quantization noise, suffices.

In practice the above converters are parallel devices in order not to slow down needlessly the conversion rate. The A/D converter is a flash converter and the D/A converter a unit-elements converter. The accuracy of the A/D converter is not critical as it is a part of the forward branch and it does not control the overall accuracy like

DIGITAL AND ANALOG SYSTEMS

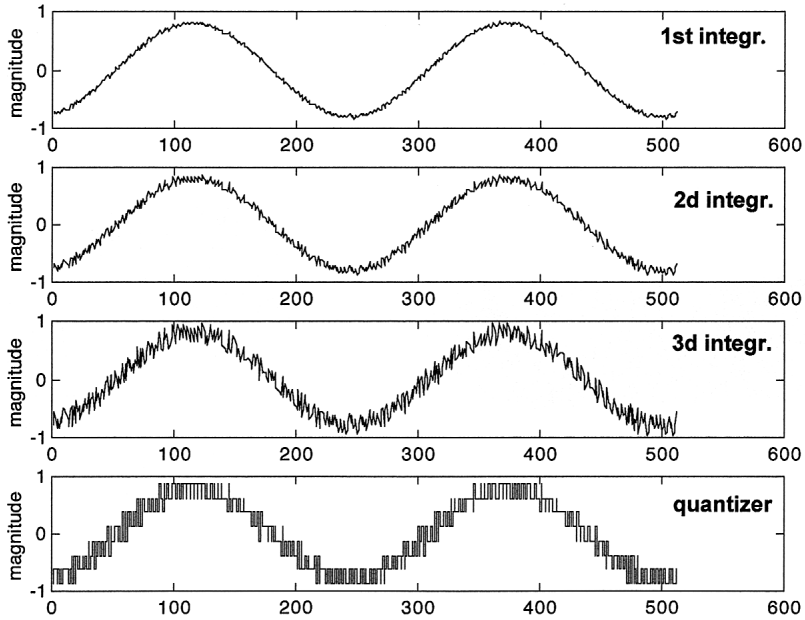


FIGURE 14.4.27 Waveforms observed in a third-order noise shaper making use of a 3-bit quantizer.

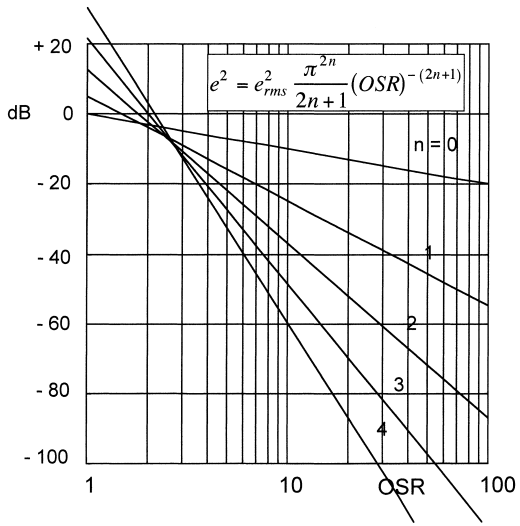


FIGURE 14.4.28 Plot of the quantization noise attenuation vs. the oversampling rate (OSR) and the loop filter order. The case $n = 0$ corresponds to oversampling without noise shaping.

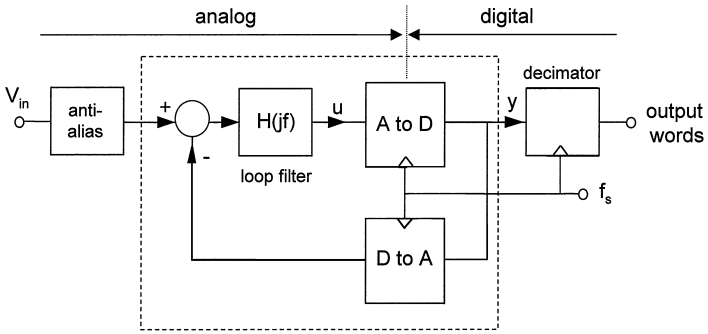


FIGURE 14.4.29 The generic Delta-Sigma A/D converter.

in any feedback loops. The D/A accuracy is more critical because it is located in the feedback branch and impairs the overall accuracy since the feedback loop minimizes only the difference between the oversampled input signal and the signal delivered by the D/A converter. This puts, of course, a heavy toll on the D/A converter unless we use a single-bit architecture. In single-bit quantizers, the A/D converter resumes to a comparator and the D/A converter to a device whose output can take only two states. No possibility exists that steps of unequal heights are found as in multi-bit D/As. This explains why single-bit quantizers are preferred generally to multi-bit. Multi-bit quantizers, however, are not evaded; they offer better alternatives for large bandwidth converters where the sampling frequency can be bound by the technology and the OSRs may take large values. To improve the performances of the D/A converter, randomization of their unit elements is generally recommended. This spreads the impairments of the unit elements mismatch over a large bandwidth. Without randomization, the same impairments produce harmonic distortion, which is a lot more annoying than white noise in terms of SNRs.

A generic A/D Delta-Sigma converter is shown in Fig. 14.4.29. Besides the noise shaper, two additional items are visible—an anti-alias filter at the input and a decimator at the output. The anti-alias filter is not specific to Delta-Sigma converters, but its implementation is much less demanding than in Nyquist converters. The reason therefore stems from the fact that the sampling frequency is much larger than the signal baseband so that a third- or second-order analog filter already is enough. At the output of the noise shaper, the decimator restricts the bandwidth to the baseband and lowers the sampling rate from oversampled to the Nyquist frequency. Since the signal inputted in the decimator is taken after the A/D converter, the decimator is in fact a digital filter. This is important for decimation is a demanding operation that cannot be done easily in the analog domain. The decimator is indeed supposed not only to restrict the bandwidth of the converted data but also to get rid of the large amounts of high-frequency noise lying outside the baseband. Therefore the decimator consists of two or three cascaded stages, an finite impulse response (FIR) filter and one or two accumulate-and-dump filters. The FIR filter takes care of the steep low-pass frequency characteristic that fixes the baseband, while the accumulate-and-dump filters attenuate the high-frequency noise. An illustration of the input-output signals of a fourth-order decimator fed by the third-order noise shaper considered in Fig. 14.4.27 is shown in Fig. 14.4.30. The signal from the noise shaper is shown in Fig. 14.4.27. The output of the decimator is illustrated below together with continuous analog input sine wave. The down-sampling rate in the example is equal to 8 and the effective number of bits (ENOB) of the decimated signal reaches 9.6 bits. In practice, resolutions of 16, even 24 bits, are obtained.

The resolution of Delta-Sigma converters is determined by the ENOB, which is derived from the relation linking quantization noise to number of bits:

$$\text{ENOB} = \frac{\text{SNR}_{\text{dB}} - 1.8}{6} \tag{8}$$

The SNR is measured by comparing the power of a full-scale pure sine wave to the quantization noise power taking advantage of the fast Fourier transform (FFT) algorithm. The measured noise consists of two

DIGITAL AND ANALOG SYSTEMS

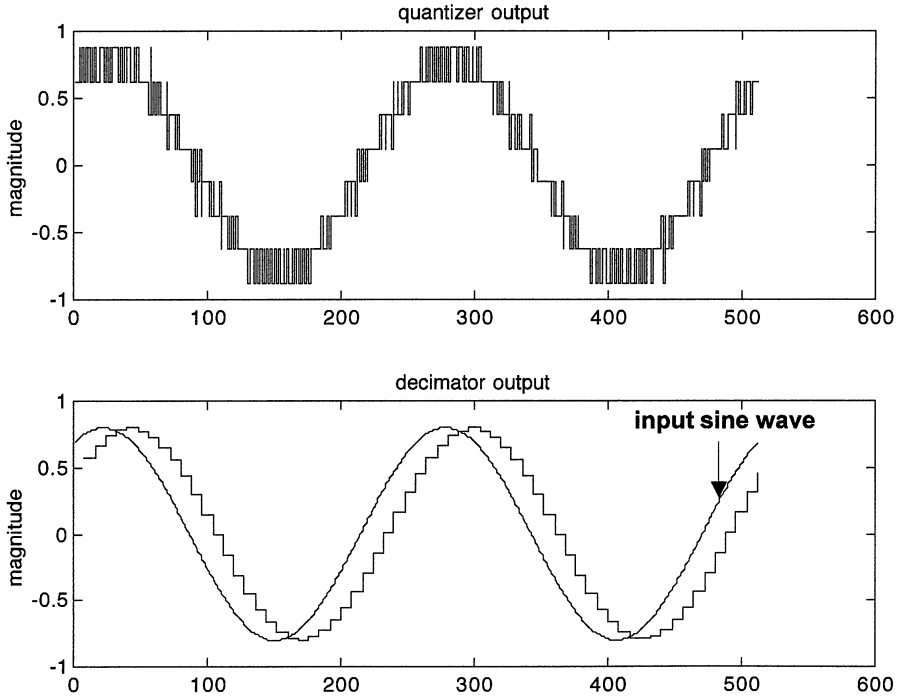


FIGURE 14.4.30 (Above) The quantizer output signal of Fig. 14.4.28. (Below) The same after decimation (order 4) compared to the input signal.

components—the actual quantization noise, which is a measure of the resolution, and the noise caused by the converter impairments. The first does not represent a defect; the second is unwanted and inevitable. The magnitude of this extra noise generally increases very rapidly when the input signal gets large because more and more nonlinear distortion tends to enter the picture. In order to avoid the impact of distortion on the ENOB evaluation, the SNR is evaluated generally as follows. The magnitude of the input signal is varied from a very low level, for instance, 60 dB below full scale, until full scale and the measured SNRs are plotted versus the magnitude of the input signal. As long as distortion does not prevail, the SNR varies like the power of the input signal, but beyond this it departs from ideal. The figure that must be considered in the above equation is the SNR obtained after extrapolating the linear portion of the SNR plot until full scale.

D/A Delta-Sigma Converters[10,20]

The principles underlying A/D converters can be transposed in D/A Delta-Sigma converters (Fig. 14.4.31). Oversampling and noise shaping are applied concurrently but in a different way since the input data are digital words and the output is an analog signal. An interpolation step is needed first in order to generate additional samples interleaved with the input data. The output of the interpolator is then noise shaped before being applied to a low-resolution D/A converter whose analog output is filtered by means of a low-order analog low-pass filter. The quantization noise is evaluated and fed back to the loop filter before closing the feedback loop of the noise shaper. One of the differences with respect to A/D Delta-Sigma converters is the manner quantization noise is measured. All that is needed is to split the words delivered by the noise shaper in two fields—an MSB field and an LSB field. The MSB field, which consists of a few bits, or even a single-bit, controls the D/A converter, while the LSB field closes the feedback loop after the digital loop filter.

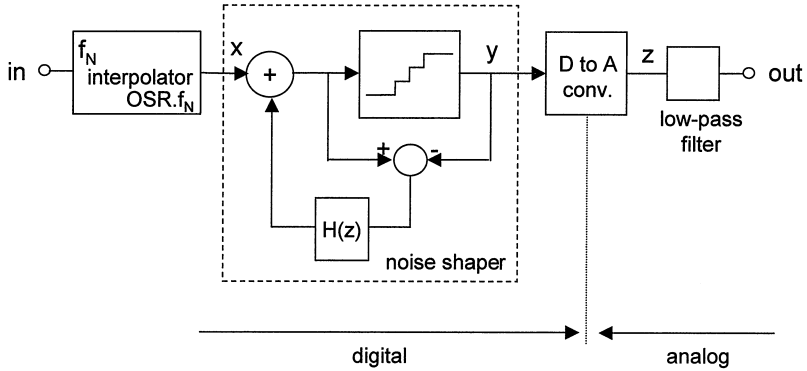


FIGURE 14.4.31 Block diagram of a Delta-Sigma D/A converter.

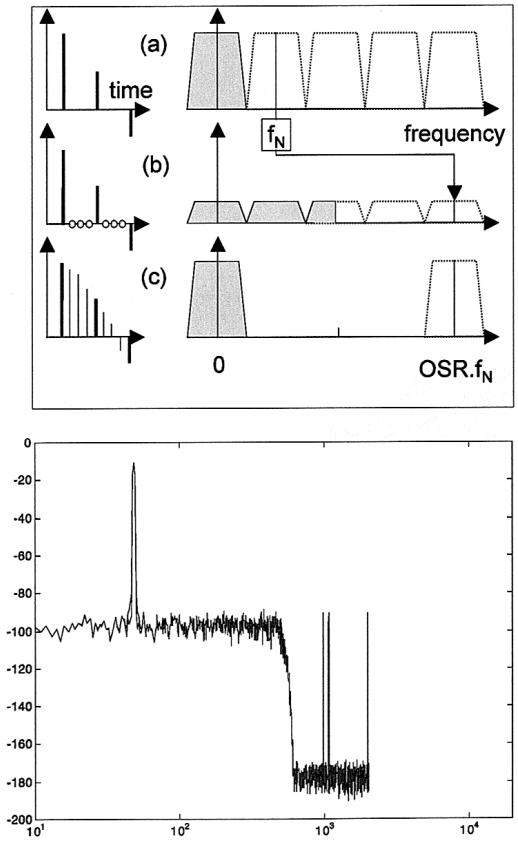


FIGURE 14.4.32 Interpolation (above) the principle (below) spectrum of a fourfold interpolated sine wave.

DIGITAL AND ANALOG SYSTEMS

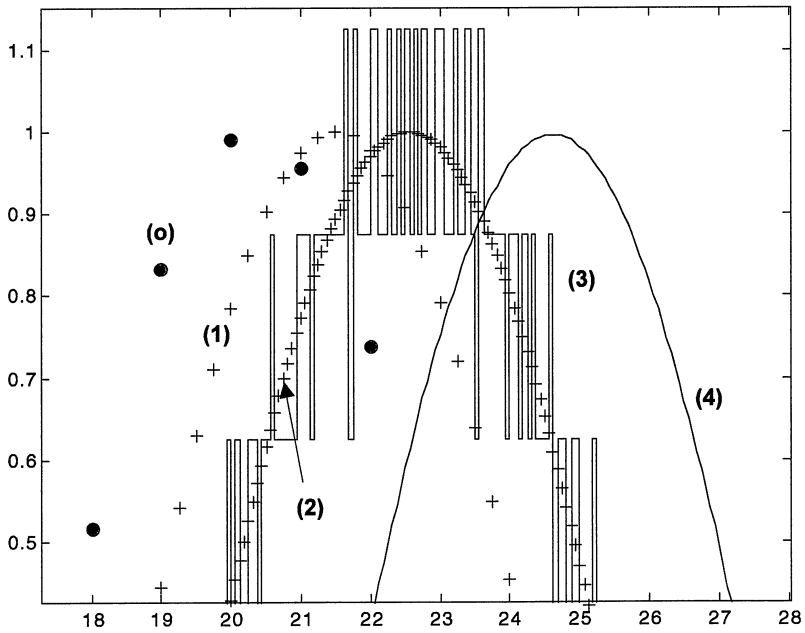


FIGURE 14.4.33 Timing of signals in a D/A Delta-Sigma converter: (0) the 13-bit digital input sine wave, (1) after a first fourfold interpolation, (2) after the second interpolation, (3) the noise shaped 3-bit signal, and (4) the analog output signal.

Interpolators are the counterpart of decimators. Instead of down-sampling they add data between samples as shown in the upper part of Fig. 14.4.32. In the figure, three zeros are placed between every sample. This multiplies already the sampling frequency by 4 but does not suffice because the spectrum of the resulting oversampled signal is like the one shown in Fig. 14.4.32*b*. One must erase part of the spectrum and multiply the signal by 4 in order to get the correct spectrum shown under Fig. 14.4.32*c*. This is done by means of a digital filter with a sharp cutoff frequency at the edge of the baseband. In practice this filter consists of several cascaded filters like in the decimator. An FIR filter takes care of the sharp cutoff frequency, whereas one or two low-pass filters perform the rest. The FFT of a sine wave after a fourfold interpolator with a single FIR filter is shown in the lower part of Fig. 14.4.32. The fundamental ray that is reproduced three times lies approximately 80 dB below the quantization noise floor and does no harm.

A second more elaborate example is shown in Fig. 14.4.33. It illustrates the changes signals undergo versus time. The input is a 13-bit sine wave represented by large black dots. After a first interpolation, which multiplies the sampling frequency by 4, we get the data marked (1). A second interpolation by 4 yields data (2). The signal is then noise shaped. The three MSBs control the D/A converter whose analog output (3) consists of large steps that approximate the signal from the second interpolator. When the signal outputted by the D/A converter is filtered by means of a third-order low-pass, the continuous sine wave (4) is obtained, which is the actual output signal of the converter. With an SNR of nearly 80 dB, the 13-bit accuracy of the digital input signal is still met.

VIDEO A/D CONVERTERS[10]

Video applications require extremely short conversion times. Few of the previous converters meet the required speed. An obvious solution is full parallel conversion with as many comparators and reference levels as there are quantization steps. Fortunately, most video applications do not require accuracies higher than 8 bits, so 256 identical channels are sufficient. Only integrated circuits offer an economical solution in this respect.

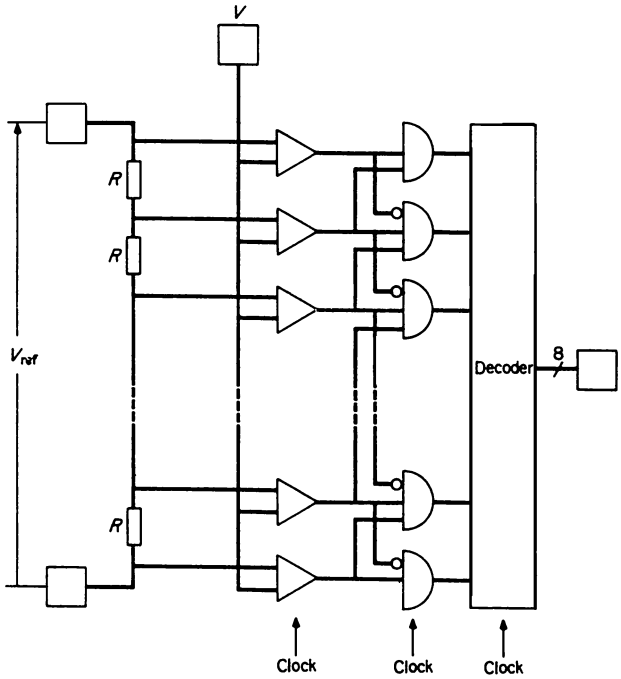


FIGURE 14.4.34 A flash converter.

A typical parallel video converter, called a flash converter, is shown in Fig. 14.4.34.[17] The architecture of the converter is simple. A string of 255 identical thin-film aluminum or polysilicon resistors is used in order to produce 2^8 reference levels. The input voltage V produces at the outputs of the comparators a vector which may be divided into two fields of 1s and 0s. An additional layer of logic transforms this vector into a new vector with 0s everywhere except at the boundary between the two fields. The new vector is further decoded to produce an 8-bit coded output. No sample-and-hold (SH) circuit is required since all comparators are synchronously driven by the same clock, but extremely fast comparators are required. An 8-bit resolution and a 5-MHz bandwidth imply binary decisions in a time as short as 250 ps. This is achieved by means of ECL logic drivers and storage flip-flops. Signal propagation across the chip also requires careful layout. The resistive divider impedance may not exceed 50 to 100 Ω. Another particular feature is the inevitable high-input capacitance, which results from the paralleling of many comparators.

The huge area and power consumption inherent to flash devices has stimulated the design of less greedy converters. Although these cannot compete with flash converters as far as speed is concerned, they offer conversion times short enough to comply with the requirements of a number of "fast" applications. Their principle comes to fragment the conversion process into several cycles, eventually two, during which strings of bits are evaluated by means of a subconverter. The process starts with the bits representing the MSB field. The other sets are evaluated one after another until the correct output word can be reconfigured by concatenating the individual segment codes. Each conversion step requires a single clock. This means, of course, that all subconverters must be fast devices, like flash converters. A 9-bit converter, for instance, operating by means of 3- bit segments requires three cycles for full conversion. Although only three clock cycles are required, the real conversion time is always much longer than three times what is needed for the 9-bit flash converter. Segmentation supposes indeed that some kind of analog memory be used in order to store intermediate results. Op-amps are thus required, which is a drawback with respect to flash converters, for the latter don't need op-amps, and consequently ignore dominant poles.

The most stringent difference between flash converters and segmented A/D converters is the number of comparators. In a flash converter, the number of comparators increases exponentially with the number of bits.

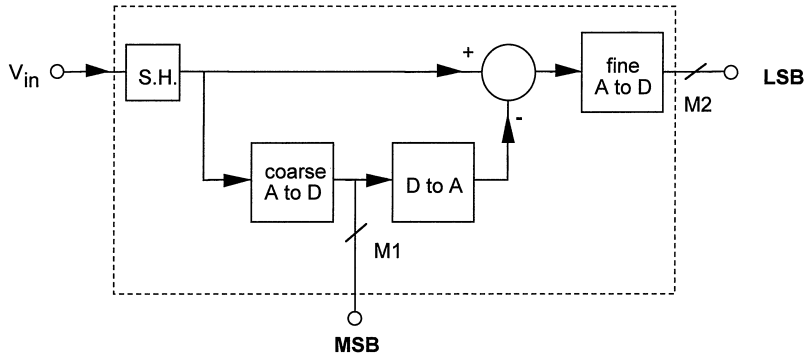


FIGURE 14.4.35 A subranging converter.

In a segmented converter, less comparators are required. Not only is the area much smaller, but the power consumption is drastically reduced.

Figure 14.4.35 shows the implementation of a two-cycle segmented A/D converter called a *subranging converter*. The coarse A/D converter evaluates first the M1 MSBs. For the remaining M2 LSBs, an analog replica of the MSBs is subtracted from the analog input signal. This is done by means of the M1-bit wide D/A converter whose output is subtracted from the input signal. The difference, which is nothing but coarse conversion quantization noise, is fed to the fine A/D flash converter, which evaluates the LSBs. The output code word is obtained by concatenation of the M1 and M2 segment codes. The total number of comparators in this type of converter is drastically reduced. For instance, in a 10-bit converter taking advantage of two 5-bit segments, each flash converter requires 31 comparators, that is, a total of 62 comparators, which is very little compared to the 1023 comparators required by a true 10-bit flash converter. Naturally the coarse D/A converter, which is a unit-elements parallel device, and the sample-and-hold, which holds the input while A/D and D/A conversions take place, must be brought into the picture also. In any case, segmented converters save a lot of area and power.

The number of comparators can be decreased further if we consider the *recycling converter*[24] shown in Fig. 14.4.36. In this circuit the difference between the input and its coarse quantized approximation is fed back to the input, stored in the sample-and-hold device, and recycled through the A/D converter. Eventually a second cycle may be envisaged. Each cycle, a new set of lower rank bits is generated. The fine flash converter is not needed, but an interstage amplifier is required. Its purpose is to amplify the quantization noise so that it spans exactly over the full dynamic range of the A/D and D/A converters. If this were not the case, we would be forced to adapt continuously the resolution of the converters to cope with the decreasing magnitude of the difference signal.

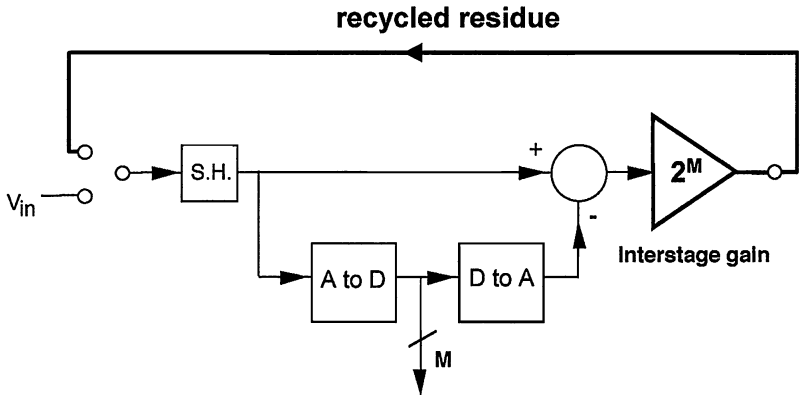


FIGURE 14.4.36 A recycling converter.

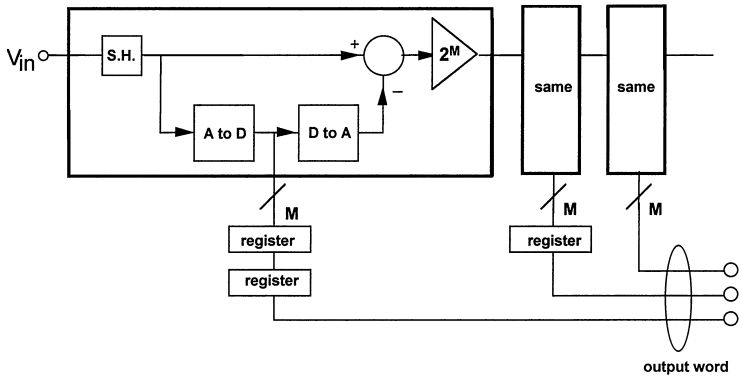


FIGURE 14.4.37 A pipelined converter.

Of course, it is useless to try to repeat this procedure over and over again because the errors generated in every new cycle pile up and corrupt the difference signal. It is very important to determine which sources of errors are important. The main ones are the interstage gain error, the D/A and the A/D nonlinearities. The A/D errors can easily be corrected by extending the dynamic range of both converters. Errors from the flash converter are corrected automatically during recycling. The other errors are more difficult to correct but one should not overlook the fact that they affect only bits generated after the first cycle.

The conversion time of recycling converters varies, of course, like the number of cycles. Pipelined converters offer a good means to keep the conversion time equal to a single cycle time, however, at the expense of area. Such a converter is shown in Fig. 14.4.37. The idea is simply to exchange time for space. In other words, we cascade blocks identical to the circuit of Fig. 14.4.36, but each circuit feeds its neighbor instead of recycling its own data. Every bloc thus deals with data that belong to different time samples. In order to reconfigure the correct output words, one must reshuffle the code segments in order to recover time consistency. This is done by means of registers.

Recycling and pipelined converters that operate with segments only 1-bit wide are currently designated *algorithmic converters*. They are not fast devices since they require as many cycles as number of bits to

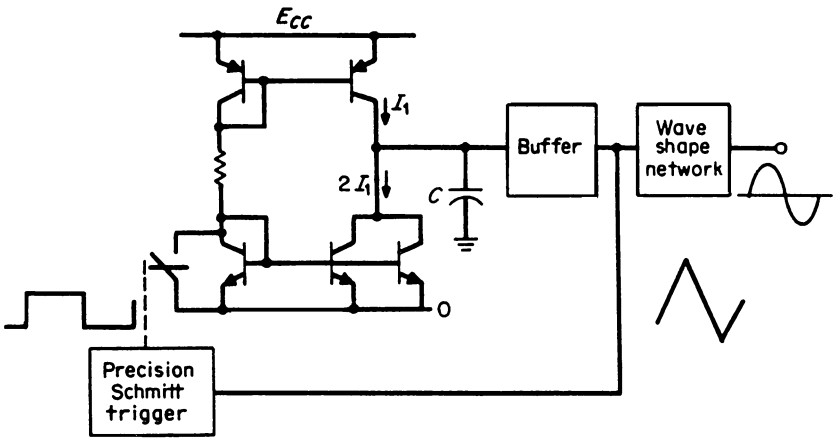


FIGURE 14.4.38 A typical integrated wave generator that can deliver a square wave, a triangular wave, and a sine wave.

DIGITAL AND ANALOG SYSTEMS

output code words. Algorithmic converters are similar to successive approximation converters although they operate differently.

FUNCTION GENERATORS

Integrated function generators consist generally of a free-running relaxation oscillator controlling a nonlinear shaping circuit. A typical block diagram of a function generator is shown in Fig. 14.4.38. The relaxation oscillator is a combination of a time-base generator and a Schmitt trigger. The time base in the present case is obtained by successive loading and discharging of a capacitor using two current sources. One is a constant current source I_1 and the other a controlled current source delivering a current step equal to $-2I_1$ or zero. Hence, the voltage across the capacitor C is a sawtooth.

The switching of the controlled current source is monitored by the logical output signal of the Schmitt trigger. This last circuit is in fact a precision Schmitt trigger. The oscillating voltage across C is obtained in the same manner. The output sawtooth signal is buffered and drives a network consisting of resistors and diodes which changes the sawtooth into a more or less sinusoidal voltage.

The advantage of function generators over RC or op-amp oscillators is their excellent amplitude stability versus frequency. Also frequency modulation can easily be achieved by changing the current delivered by the two current sources. This type of function generator can be frequency-swept over a wide dynamic range without spurious amplitude transients since no selective network is involved.