



Figure 11.29. 12-bit dual DACs. (Courtesy of Analog Devices.)  
A. 7537 byte-wide loading.  
B. 7547 12-bit-wide loading.

would happen if a fixed voltage were applied to the board's analog inputs. The manufacturer's technical support department assured us that the result would be "at most two adjacent digital codes." In fact, the board actually bounces around among seven consecutive codes, which amounts to a 14-bit conversion. At their suggestion we brought the board back for tests, which confirmed the noise problem. When we asked about this, we were told that the particular employee no longer worked there. They told us all their boards work that way; and, to add insult to injury, they threatened to charge us for the "service" they had performed in testing the board.

### 11.12 Memory

In a commercial microcomputer it's easy to add memory – all you have to decide is how many megabytes you want to add, and from whom to buy it. You get to have more fun in a dedicated microprocessor application, where the assignment of memory is part of the design, with blocks of non-volatile (permanent) ROM being used for program storage, and read/write volatile RAM being used for temporary storage of data, stacks, and program workspace. Nonvolatile program storage is universally used in dedicated instruments so that it isn't necessary to load a program each time the instrument is turned on.

In this section we will discuss the various kinds of memory: static RAM, dynamic RAM, EPROM, and EEPROM. Once you get the hang of it, the choices are pretty clear. You may want to refer ahead to Figure 11.35, where we summarize memory families.

#### **Static and dynamic RAM**

Static RAM, or SRAM, stores bits in an array of flip-flops, whereas dynamic RAM, or DRAM, stores bits as charged

capacitors. A bit once written in a static RAM stays there until rewritten or until the power is turned off. In a dynamic RAM the data will disappear in less than a second, typically, unless "refreshed." In other words, a dynamic RAM is always busy forgetting data, and it is rescued only by periodic clocking through the "rows" of the two-dimensional pattern of bits in the chip. For example, you have to access each of 256 row addresses in a 256K-bit RAM every 4ms.

You might wonder what would possess anyone to choose a dynamic RAM. By not using flip-flops, the DRAM saves space, giving you more data on a chip, at lower cost. For example, the currently popular 32K×8 (256kbit) SRAMs cost about \$10, twice the current price of 1Mbit DRAMs. Thus, you can get four times as much memory on a large memory board, at half the price, by using dynamic RAM.

Now you might wonder why anyone would choose *static* RAM (fickle, aren't you?). The major virtue of SRAM is its simplicity, with no refresh clocks or timing complexity to worry about (the refresh cycle competes with normal memory access cycles and must be properly synchronized). Thus, for a small system with only a few memory chips, the natural choice is SRAM. Furthermore, most SRAM these days is CMOS, essential for battery-powered devices. In fact, CMOS static RAM, backed up with a battery when main power is off (using a power control chip like the MAX690), forms a good alternative to ROM for nonvolatile memory. A further advantage of SRAM is its availability in high-speed versions, down to 25ns or less, and convenient byte-wide packaging. Let's take a closer look at both SRAM and DRAM.

*Static RAM.* We saw SRAM in our microprocessor design, where we used a single 32K×8 SRAM for data, stack, and workspace (the program was in EPROM).

Using SRAM is like falling off a log: You assert the address, chip select ( $\text{CS}'$ ), and output enable ( $\text{OE}'$ ) for READ; the requested data appears on the three-state data lines a maximum of  $t_{aa}$  (address access time) later. For WRITE, you assert address, data, and  $\text{CS}'$ , then follow (after an address setup time  $t_{as}$ ) with a write enable ( $\text{WE}'$ ) pulse; valid data are written at the end of the  $\text{WE}'$  pulse. The actual timing constraints for 120ns SRAM are shown in Figure 11.30, from which you can see that the “speed” of the memory is the time from assertion of valid address to valid data (read) or to completion of the write cycle (write). For SRAM, the time from one memory access to the next (the “cycle time”) equals the access time; this isn’t true for DRAM, as we’ll see.

Static RAM is available in sizes from 1Kbit or less up to 1Mbit, organized as 1 bit, 4 bits, or 8 bits in width. Speeds range from a leisurely 150ns access time down to 10ns or so. Currently popular types include the inexpensive  $8\text{K} \times 8$  and  $32\text{K} \times 8$  CMOS SRAM with 80ns access times, and smaller high-speed ( $<30\text{ns}$ ) CMOS SRAM for cache memory. Variations include separate input/output pins, dual-port access, and SIP (single-row) packaging.

For whatever it’s worth, note that you don’t have to hook up the CPU’s data lines to the corresponding memory data pins – you can scramble them any way you want, since they get unscrambled when you read back what you wrote! The same goes for addresses. Don’t fool around with ROM in this way, though.

#### EXERCISE 11.18 Why not?

**Dynamic RAM.** By comparison with SRAM, DRAM is a royal headache. Figure 11.31 shows a normal cycle. The address (e.g., 20 bits for 1M DRAM) is split into two groups and multiplexed onto half the number of pins, first the “row address,”

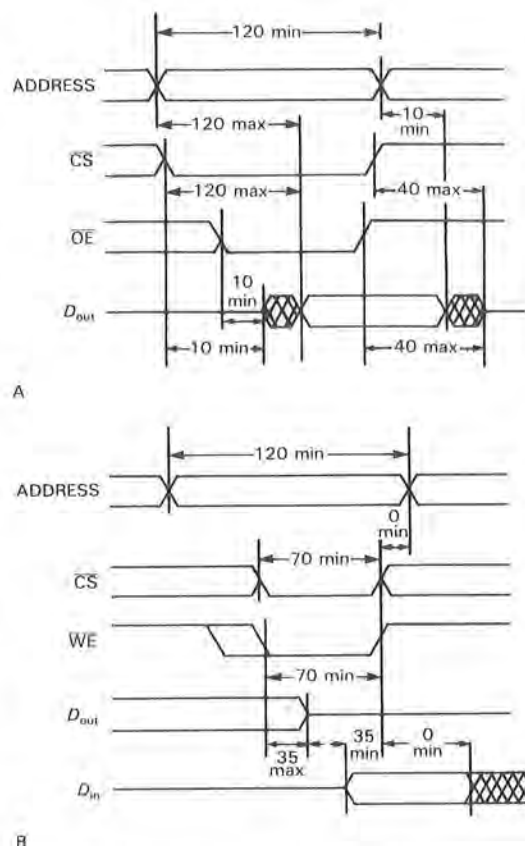


Figure 11.30. 120ns SRAM timing.  
A. READ cycle.  
B. WRITE cycle.

signaled by a “row address strobe” ( $\text{RAS}'$ ); then the “column address,” signaled by a “column address strobe” ( $\text{CAS}'$ ). The data are written (or read, as determined by the direction input  $\text{R/W}'$ ) following assertion of  $\text{CAS}'$ . Some additional “RAS-precharge” time must elapse before the next memory cycle, so the cycle time is greater than the access time; for example, a particular DRAM with 100ns access time has 200ns cycle time. A refresh cycle looks similar, but without the assertion of  $\text{CAS}'$ . Actually, normal memory accesses would refresh perfectly well, if you could count on accessing all required combinations of row addresses!

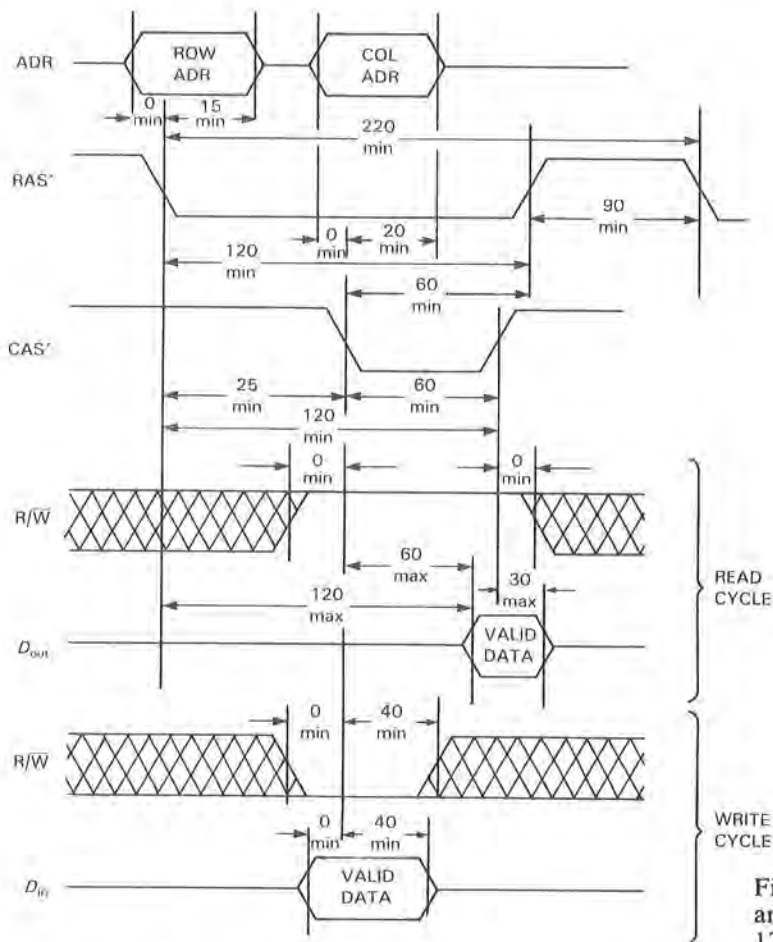


Figure 11.31. DRAM READ and WRITE cycles (Hitachi, 120ns).

DRAMs, like SRAMs, come in data widths of 1, 4, and 8 bits, in sizes from 64Kbit to 4Mbit, and in speeds from about 70ns to 150ns. The most popular are the large 1-bit types, which makes sense: If you are building a large memory array, say 16 bits wide and 4Mbyte total, and you have a choice of 1Mbit DRAMs, organized either as 1M $\times$ 1 or 128K $\times$ 8, it pays to use the 1-bit chips because (a) each data line will be tied to only two chips (rather than 16), which results in far less capacitive loading, and (b) the chips are smaller, since fewer data pins more than makes up for the extra address pins. Furthermore, the 1-bit chips will generally be cheaper. This argument assumes you're building a

large memory, of course, and, for example, doesn't apply for our simple 32K $\times$ 8 microprocessor. Note, however, that improved high-density packaging technologies ("ZIPs" and "SIMMs") are reducing the importance of minimizing pin count.

There are several ways to generate the proper sequence of multiplexed addresses, RAS, CAS, and R/W' signals that you need to use dynamic RAM. Since the DRAM is invariably connected to a microprocessor bus, you start things going when you see an AS' (or equivalent) signaling a valid address in DRAM memory space (as determined by the high-order address lines). The traditional method uses discrete MSI components to do the address multiplexing



(a few '257 quad 2-input MUXs) and generates the RAS, CAS, and MUX control signals either with a shift register clocked from a multiple of the processor clock or, better, from a tapped delay line. You need some logic to interpose RAS-only refresh cycles periodically, with a counter to generate successive row addresses. Figure on at least 10 chips if you design your circuit this way.

PALs offer an attractive alternative to "discrete" DRAM control circuitry, and you can usually get most of the logic you need in one or two chips this way. An even easier way is to use one of the special-purpose "DRAM support" chips, for example the AM2968. These chips handle not only the address MUX and RAS/CAS generation but also the refresh arbitration and row address generation; furthermore, they even include the robust drivers and damping resistors you need to drive a large array of memory chips, as we'll explain in a moment. These DRAM controller chips usually include companion chips for timing and error detection/correction; thus, a small chip set forms a complete solution to your DRAM design.

Well, *almost* a complete solution! The real agony of DRAM memory comes when

you try to keep all those bused strobes and address lines noise-free. The basic problem is that you typically have a few dozen MOS ICs spread over a substantial area of PC board, with the address lines and other control signals bused to all the chips. To drive a few dozen chips you need high-current Schottky output stages; but the long lines and distributed input capacitances, in combination with fast output-stage edge times, result in severe ringing. It's not unusual to see -2 volts undershoot at DRAM address pins! The usual solution (not always completely successful) is to put series damping resistors, typically about 33 ohms, at the output of each driver. A related problem is caused by the enormous transient currents, which can easily reach 100mA per line. Imagine an octal driver chip in which most of the outputs happen to make a HIGH-to-LOW transition. This causes a transient current of nearly an ampere, which makes the ground pin rise momentarily above ground, and with it any outputs that were supposed to stay LOW. This problem is far from academic - we once had memory errors because just such a ground transient, caused by CAS-driver current spikes, let the RAS drivers on the same chip spike HIGH enough to terminate the memory cycle!

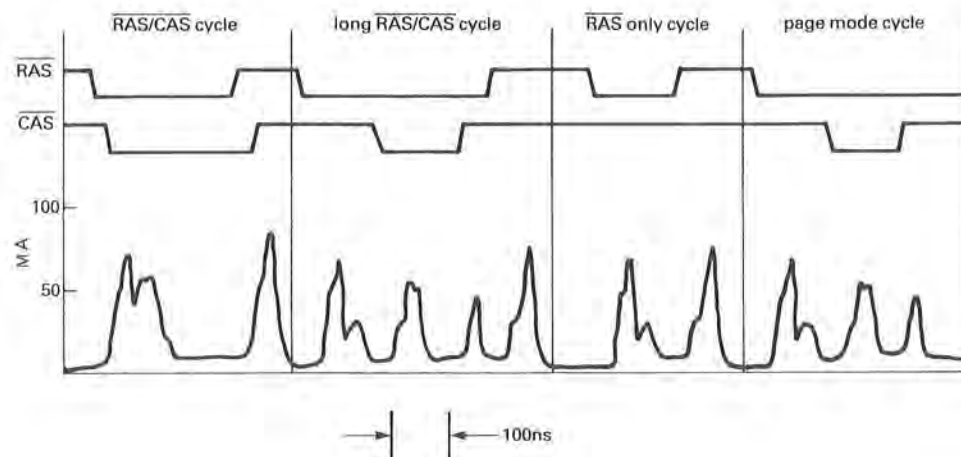


Figure 11.32. DRAM current transients.

An additional noise source in DRAM memories is the large rail-to-rail transient current produced by the memory chips themselves, which some manufacturers are even honest enough to show in their data sheets (Fig. 11.32). The usual solution is to use extensive bypassing to a low-inductance groundplane; current practice favors one  $0.1\mu\text{F}$  ceramic capacitor per memory chip.

We have found that 74F logic drivers with external resistors work well with DRAM, as do special drivers, such as the Am2966, that include integral damping resistors. The Am2968 DRAM controller mentioned above claims to drive up to 88 memory chips without external components, with  $-0.5$  volts maximum undershoot. Perhaps more important than the particular driver is the use of a low-inductance groundplane and liberal  $V_+$  bypassing. The best PC-board memories we've seen are invariably 4-layer (or more), with layers devoted entirely to ground and  $V_+$ . Two-sided boards with skimpy grounds are certain invitations to disaster; Wire-Wrap panels often don't do much better.

It's important to realize that noise-induced memory errors tend to be highly pattern-sensitive and may not show up in a simple read/write memory test. Conservative design and exhaustive memory testing (with oscilloscope examination of waveforms) are the best ways to ensure reliable memory operation.

### **Read-only memory (ROM)**

These are nonvolatile memories, which you need in nearly all computer systems. In microcomputers, for example, you need as a minimum some nonvolatile memory to hold the boot-up sequence of instructions, which includes not only the stack allocation and port and interrupt initializations but also enough code to get the operating system read from disk storage.

When your PC does its memory test and then comes up in DOS, it's following orders from some ROM. In addition, microcomputers usually keep a portion of the operating system (usually the more hardware-specific portions) in ROM; this is called the "ROM-BIOS" (basic I/O system), and it makes the operating system feel at home by providing universal "hooks" to its particular ports. ROMs are widely used for lookup tables, for example for character generators for a display device. In the most extreme case, a microcomputer can put the entire operating system, and even compilers and graphics routines, in ROM. The Macintosh, for example, provides almost all system software in ROM, liberating nearly 256Kbyte of RAM for user programs. However, this ROM-based approach has enjoyed only limited popularity in microcomputers, because it limits flexibility; note, however, that software "fixes" and modest upgrades can usually be accommodated by placing patches in RAM.

In dedicated microcontroller applications, ROM is used more extensively. As in our signal-averaging example, the entire stand-alone program lives in ROM, and volatile RAM is used only for arrays and temporary storage. Finally, ROM is sometimes useful in discrete digital hardware, for example to construct arbitrary sequential state machines, or as lookup tables for linearizing the response of some measuring system. Let's take a look at non-volatile memory choices: UV-erasable EPROMs, fusible-link ROMs, mask-programmed ROMs, and electrically erasable EEPROMs.

**EPROM.** Those large chips with the quartz windows are EPROMs, erasable programmable read-only memories. They are by far the most popular form of non-volatile memory for computers. They are available in CMOS and NMOS and consist of a large array of MOSFETs with floating

gates that can be charged by “avalanche injection,” a fancy name for breakdown of the gate insulating layer by an applied pulse of more than 20 volts. These memories store data by retaining indefinitely a tiny charge (about  $10^6$  electrons) on these insulated “buried” gates, which you can think of as capacitors with time constants of centuries. You read out the state of an individual capacitor by allowing it to be the gate of an associated MOSFET channel. Since the gate is not electrically accessible, it can be erased only by exposing it to intense ultraviolet radiation for 10 to 30 minutes, which causes the stored charge to leak off by photoconduction. Thus, the individual stored bytes cannot be erased selectively.

In the first edition of this book we mentioned the “classic 2716”  $2K \times 8$  EPROM, with its price of \$25. It’s now so classic that you can’t even get it! The current crop of EPROMs range from  $8K \times 8$  up to  $128K \times 8$ , with prices of a few dollars. Typical access times are in the range of 150–300ns, though companies like Cypress offer parts as fast as 25ns in the small sizes. To program EPROMs, you have to apply elevated programming voltages (usually 12.5V or 21V) while asserting the desired bytes. The original algorithms specified 50ms programming time per byte (which is 100 seconds for the 2716, but would add up to a half hour for a modest-sized  $32K \times 8$  EPROM). Thus, the larger ROMs forced the designers to concoct newer “smart” algorithms, in which for each byte you try a succession of 1ms programming pulses, checking the result by reading after each write attempt; when the byte reads back correctly, you hit it with a final write pulse equal to three times the previous total. Typically most of the bytes program the first time, so you spend about 4ms per byte, or 2 minutes for a  $32K \times 8$  ROM.

EPROMs are great for prototyping, since you can re-use them after erasure.

They’re fine also for small production runs. You can get them in a cheaper no-window variant, sometimes called “one-time-programmable” (OTP) EPROM. Although the latter shouldn’t be called EPROM, engineers persist in this misnomer. EPROM manufacturers are conservative, only guaranteeing a 10 year data retention. That figure assumes worst-case conditions (especially high temperature, which causes charge leakage); in reality they don’t seem to lose data, unless you happen to get a defective part.

EPROMs have a limited *endurance*, the number of times the memory can be erased and reprogrammed. The manufacturers are notoriously shy about revealing these numbers, but you can generally expect to get 100 or so erase/program cycles before the chip degrades seriously.

□ **Mask ROM and fusible-link ROM.** Mask-programmed ROMs are essentially custom chips born with your bit pattern built in. The semiconductor house converts your bit specification into a custom metallization mask used to process the chip. These are for large production jobs, and you wouldn’t dream of having a mask-programmed ROM designed for prototyping. Typical costs are \$1k to \$3k setup charge, with the manufacturer strongly discouraging you from buying fewer than a thousand ROMs at a time. In those quantities the chip might cost a few dollars.

Many single-chip microcontrollers include a few kilobytes of on-chip RAM and ROM, so that a finished instrument needn’t have any external memory chips. In most cases the micro-controller family includes versions that take external ROM and, in some cases, versions with on-chip EPROM (Fig. 11.33). The idea is that you develop the instrument and write the code using the EPROM (or external ROM) version, then go into production with the cheaper mask-programmed controller.

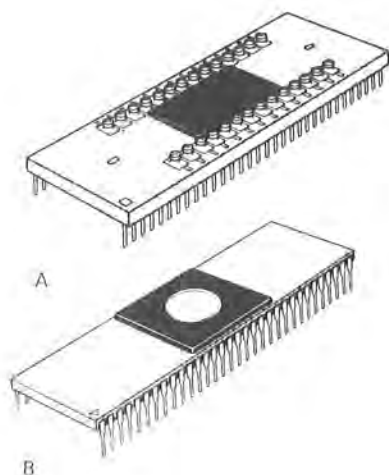


Figure 11.33. Microcontroller/EPROM combinations.  
A. 8-bit microcontroller with "piggyback" EPROM.  
B. 8-bit microcontroller with on-chip EPROM.

Another kind of one-time-programmable ROM is the *fusible-link* ROM. These begin life with all bits set, and you give them electrical shock treatments until you've blown out the offending bits. A typical part is the Harris HM6617, a 2K×8 CMOS PROM. Fusible-link PROMs are also available in bipolar (TTL) technology.

**EEPROM.** These ROMs can be selectively erased and reprogrammed electrically, while in-circuit. They are therefore ideal for holding configuration information, calibration parameters, etc., that cannot be frozen before the computer or instrument is used. They use the same MOS floating-gate technique as EPROMS.

First-generation EEPROMs were difficult to use and required elevated voltages and lengthy programming procedures, just like EPROMs. More recent EEPROMs run from a single +5 volt supply and work nearly the same as SRAMs – in other words, you can reprogram any byte with a single bus WRITE cycle. The EEPROM uses internal circuitry to generate the

higher programming voltages, and internal logic latches the data and generates the several-millisecond programming sequence, either setting a BUSY' flag during the process, or producing complemented DATA on a READ to indicate WRITE-in-progress. Some EEPROMs implement *both* protocols, which are commonly called "RDY/BUSY'" and "DATA'-polling." It's easy to interface these ROMs – just hook them up like SRAMs, and use the BUSY' line to make an interrupt (or read it or the DATA as a status flag); see Figure 11.34. DATA'-polling has the advantage of working in a standard SRAM socket without any circuit changes (you must, of course, alter your programs to test the read-back data, waiting until it agrees with what you wrote). Since you don't write often to EEPROM, you don't really need the interrupt capability of RDY/BUSY'.

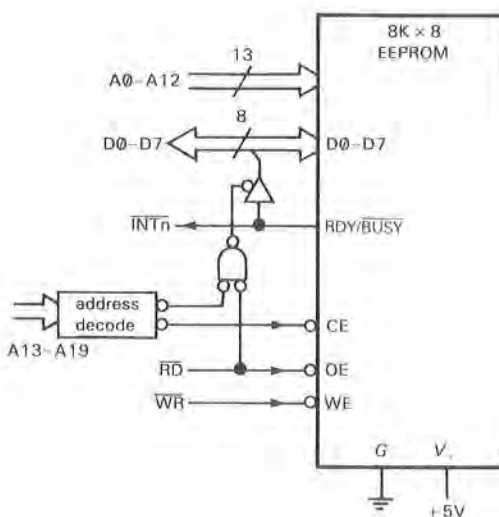


Figure 11.34. EEPROM.

CMOS EEPROMs are currently available in 2K×8, 8K×8, and 32K×8 sizes, with prices around \$10–\$50. Access times (200ns–300ns) and programming speeds (~2ms/byte, with internal "smart" algorithm) are comparable with standard



EPROMs. EEPROMs, like EPROMs, have limited read/write endurance. Although the manufacturers avoid specifying actual endurance figures, you see numbers like 100,000 read/write cycles at 25°C operating temperature.

*Note:* Although EEPROMs are unique in allowing *in-circuit* reprogramming, they can perfectly well be programmed externally, in an EPROM-type programmer. This makes them ideal during firmware development, because you don't have to wait around while your old EPROMs cook for a half hour in the UV hot seat.

There are two interesting variations on these EEPROMs. From National, Xicor, etc., you can get little 8-pin mini-DIP EEPROMs, with memory sizes from 16×16 to 2K×8 bits; they use serial access, with a clock input and single data line. You can't easily use these memories without a microprocessor around to access them; but in microprocessor instrumentation applications they're great for storing a few setup parameters, etc. And from Xicor comes the "EEPOT," a clever use of electrically erasable memory to hold the position of a "digital wiper": These chips have a string of 99 equal resistors, with a tap position you can set digitally and store in on-chip nonvolatile memory. You can imagine applications where it would be nice to have automatic or remote calibration of some instrument, without the need for adjusting a mechanical trimmer pot.

A recent variety of EEPROM known as *flash* combines the high density of EPROM with the in-circuit reprogrammability of EEPROM. However, with flash EEPROMs you generally are not able to erase single bytes, as you can with conventional EEPROM: Intel's current flash EEPROM permits only *complete* erasure (like EPROM), whereas Seeq's product permits erasure either by sector (512 bytes) or in bulk. Furthermore, most available flash EEPROMs require an additional switched +12 volt supply during erase/write, a

significant price to pay when compared with the single +5 volt supply of conventional EEPROMs. Flash EEPROMs claim endurences of 100 to 10,000 programming cycles.

These PROM technologies are still evolving, and we're watching to see what happens; you should, too!

### **Nonvolatile RAM**

EPROM is fine for nonvolatile read-only storage, but you often want nonvolatile read/write memory. EEPROM achieves that, but with sluggish (10ms) write cycle times (and limited total number of read/write cycles). There are two ways to get full RAM read/write speeds (~100ns) with unlimited read/write endurance: Use either battery backed-up CMOS static RAM or the unusual "NOVRAM" from Xicor, which combines SRAM and EEPROM on the same chip.

We talked about battery backed-up SRAM earlier. In many ways this gives you the best of both worlds, namely the low price and full read/write speed of RAM combined with the nonvolatility of ROM. You must, of course, use CMOS RAM with specified micropower shut-down current. Some manufacturers are producing "nonvolatile RAM" by packaging a lithium battery and shut-down logic together with the CMOS RAM chip in an ordinary DIP package. An example is Dallas Semiconductor's 8K×8 DS1225 and 32K×8 DS1230; they also make a line of "smart-sockets" containing battery and logic, which magically render your ordinary CMOS RAM nonvolatile. Note that nonvolatile RAM constructed this way is not exactly immortal; the battery, and therefore the data, has a life of about 10 years. As with ordinary SRAM, there is no limit on the number of read/write cycles the memory can sustain.

NOVRAM (nonvolatile RAM) from Xicor combines ordinary static RAM with



some on-chip “shadow” EEPROM. A STORE<sup>1</sup> input preserves the contents of SRAM on the EEPROM, with a total write cycle time of 10ms; the data are retrieved more quickly – about 1 $\mu$ s. With a power supervisory chip like the MAX690 series, you get enough warning to save the SRAM contents before the +5 volt power drops too far. The NOVRAM is specified to withstand 10,000 storage operations, and, like ordinary SRAM, it has unlimited RAM read/write endurance.

Among these two options in nonvolatile RAM, the battery-backed SRAM seems generally preferable, because it lets you use any SRAM you can find, as long as it has a zero current shut-down mode. That means you can use the latest large RAMs, and you can, for example, shop for memory with quick access time, if that is what you need. Although battery life is finite, it’s long enough for most applications. For short retention times (a day or less), you can substitute a high-capacitance double-layer capacitor for the lithium battery; these are available from Panasonic, Sohio, and others, and they offer capacitances up to a farad or more in a very small package.

### Memory wrap-up

Figure 11.35 summarizes the important attributes of the various memory types.

|                         | Volatile         | Nonvolatile                          |
|-------------------------|------------------|--------------------------------------|
| fast READ<br>fast WRITE | DRAM<br>SRAM     | SRAM<br>w/backup<br>NOVRAM           |
| fast READ<br>slow WRITE |                  | EEPROM<br>EPROM                      |
| fast READ<br>once WRITE |                  | OTP “EPROM”<br>fuse PROM<br>mask ROM |
|                         | In-circuit write | External write                       |

Figure 11.35. Memory types.

Among the types shown, we would recommend 1-bit-wide DRAM for the largest read/write memory arrays, byte-wide SRAM for smaller microprocessor memory arrays, EPROM for read-only parameters and programs, and either EEPROM (if write speed is unimportant) or battery-backed SRAM (for full-speed read/write) for nonvolatile read/write storage.

### 11.13 Other microprocessors

Like any Darwinian process, the evolution of microprocessors has proceeded in several divergent directions. In the contest for survival, some of the less fit have become endangered species. As an example of different evolutionary paths, there is the division between processors with separate I/O instructions and those that require “memory-mapped I/O,” in which the peripheral device registers simply look like individual locations in memory (examples are the 8086 and 68000 families, respectively). Then there is the division created by machines that use memory instead of registers for most arithmetic operations. Another choice involves the use of available pinouts: Some CPUs make multiple use of pins, to allow more flexibility within the package constraint. Then there is the question of word size (4, 8, 16, or 32 bits), stacks, and elegance (or

richness) of the instruction set. Each microprocessor family has its own assembly language, another stumbling block for the beginner. Microprocessors are usually fabricated with CMOS technology, but you'll see some built with NMOS or bipolar logic.

There are even greater differences among processors that have to do with their intended applications. There are the "single-chip" processors with on-chip RAM and ROM, parallel ports, UARTs, timers, and even analog/digital converters. At the other extreme, powerful 32-bit CPUs like the 80486, 68040, and AM29000 overlap the computational finesse of large computers, but they require extensive hardware and software support to take full advantage of their advanced features. "High-integration" chips try to strike a balance; for example, the H16 is a CMOS 68000 with 2 UARTs, timers, and DMA channels on-chip.

We have used the 68008 for all our examples in this chapter, but we don't want to leave the impression that other microprocessors are less useful. Table 11.8 lists a selection of the most popular microprocessors now available. This table is not intended to be a comprehensive listing of everything that is available.

#### 11.14 Emulators, development systems, logic analyzers, and evaluation boards

How are you going to get a program written, assembled, debugged, and loaded into a ROM for use in some instrument you've designed? This is a major problem, especially for someone beginning to work with microprocessors. There are a few techniques available, involving methods that range from the simple stand-alone "burn and crash" up to elaborate development systems and high-level language emulators. In this section we will try to describe what is available and how it can be useful in designing instruments with microprocessors.

##### **"Burn and crash"**

This colorful term describes the simplest method of code development for microprocessor-based hardware. It goes as follows. You first use an assembler or compiler in some general-purpose computer (perhaps a desktop personal computer) to generate executable code that can be burned in EPROM. If the target processor – the one in your custom instrument – is different from the processor in the development computer (which we'll call a PC), you'll need a "cross-assembler"; otherwise just use the PC's "native assembler." Now you program an EPROM (that's the "burn") and try it in your instrument (that's the "crash"). Debugging now consists of looking at the symptoms created by the faulty code (or faulty hardware), making corrections or inserting diagnostic tests, and trying again. There's plenty of room for cleverness as you go about finding bugs. For example, if you put the right statements in your trial programs, you can make diagnostic use of LED indicators or other ports in the instrument to see what is happening in the program. Don't forget about the traditional tools of the electronic trade – logic probe, oscilloscope, and (if you're desperate) logic analyzer (see below).

*ROM substitutes.* Burn-and-crash, as described above, is a slow procedure. Although it's sometimes adequate to the task, you quickly become impatient with the need to cycle EPROMs through the ultraviolet eraser and ROM programmer. There are some shortcuts: (a) Use EEPROMs instead of EPROMs. They don't program any faster, but at least you don't have to fool around erasing them. (b) Use a battery-backed-up (nonvolatile) CMOS RAM, instead of EPROM, during the development cycle. These are as fast as conventional RAM (i.e., they "program" instantly) and have EPROM-compatible pinouts so that you can plug them into

TABLE 11.8. MICROPROCESSORS

| Type                           | Reg size (bits) | Bus width (bits) | Adr space | Fastest instr (μs) | Clock @ freq (MHz) | CMOS? | Queue? | Segments? | MMU? | Virtual mem? | Cache? | FPU? | Pipelined? | Comments                              |
|--------------------------------|-----------------|------------------|-----------|--------------------|--------------------|-------|--------|-----------|------|--------------|--------|------|------------|---------------------------------------|
| <b>Traditional 8-bit</b>       |                 |                  |           |                    |                    |       |        |           |      |              |        |      |            |                                       |
| 6502                           | 8               | 8                | 64kB      |                    |                    | A     | —      | —         | —    | —            | —      | —    | —          | Apple II; 65C802=upgrade              |
| 6800                           | 8               | 8                | 64kB      |                    |                    | —     | —      | —         | —    | —            | —      | —    | —          |                                       |
| 8085                           | 8               | 8                | 64kB      | 0.8                | 10                 | A     | —      | —         | —    | —            | —      | —    | —          | replaced 8080                         |
| Z80                            | 8/16            | 8                | 64kB      | 0.5                | 8                  | A     | —      | —         | —    | —            | —      | —    | —          | popular CP/M μP                       |
| 6809                           | 8/16            | 8                | 64kB      |                    |                    | A     | —      | *         | —    | —            | —      | —    | —          |                                       |
| <b>High Integration 8-bit</b>  |                 |                  |           |                    |                    |       |        |           |      |              |        |      |            |                                       |
| 50740, 65C124                  |                 |                  | 64kB      |                    |                    | *     | —      | —         | —    | —            | —      | —    | —          | 6502 inst set                         |
| 6801, 6301, 68HC11             |                 |                  | 64kB      |                    |                    | A     | *      | *         | —    | —            | —      | —    | —          | 6800 inst set                         |
| 64180, Z180                    | 8               | 8                | 1MB       |                    | 10                 | *     | *      | *         | —    | —            | —      | —    | —          | Z80 inst set, on-chip peripherals     |
| Z280                           | 8/16            | 16               | 16MB      |                    | 25                 | *     | —      | *         | *    | —            | *      | E    | *          | Z80 inst set                          |
| <b>Traditional 16-bit</b>      |                 |                  |           |                    |                    |       |        |           |      |              |        |      |            |                                       |
| 8088                           | 16              | 8                | 1MB       |                    |                    | A     | *      | *         | —    | —            | —      | E    | —          |                                       |
| 8086                           | 16              | 16               | 1MB       | 0.3                | 10                 | A     | *      | *         | —    | —            | —      | E    | —          |                                       |
| Z8000                          | 16              | 16               | 8MB       | 0.3                | 10                 | —     | *      | *         | E    | —            | —      | E    | —          | fast interrupts                       |
| F9450                          | 16              | 16               | 4MB       | 0.2                | 20                 | —     | *      | —         | *    | *            | —      | *    | *          | MIL std 1750A                         |
| <b>Advanced 16-bit</b>         |                 |                  |           |                    |                    |       |        |           |      |              |        |      |            |                                       |
| 65C816                         | 16              | 16               | 16MB      | 0.13               | 8                  | *     | —      | —         | —    | —            | E      | —    | —          | Apple IIGS, has 6502 subset           |
| 80286                          | 16/32           | 16               | 16MB      | 0.16               | 12.5               | —     | *      | *         | *    | *            | —      | E    | *          | Harris 80C286 to 20MHz                |
| V20,V30                        | 16              | 8/16             | 1MB       |                    | 10                 | *     | *      | *         | —    | —            | —      | E    | —          | has 8080 subset; CMOS                 |
| V60                            | 32              | 16               | 4GB       |                    | 16                 | *     | *      | *         | *    | —            | *      | *    | *          | 1.5 MIPS average; CMOS                |
| RTX2000                        | 16              | 16               | 1MB       | 0.1                | 10                 | *     | —      | *         | —    | —            | —      | —    | —          | Forth at 10 MIPS                      |
| <b>High-integration 16-bit</b> |                 |                  |           |                    |                    |       |        |           |      |              |        |      |            |                                       |
| 80188                          | 16              | 8                | 1MB       |                    |                    | A     | *      | *         | —    | —            | —      | E    | —          |                                       |
| 80186                          | 16              | 16               | 1MB       |                    |                    | A     | *      | *         | —    | —            | —      | E    | —          | popular as a controller               |
| V40,V50                        | 16              | 8/16             | 1MB       |                    | 8                  | *     | *      | *         | —    | —            | —      | E    | —          | 8086 inst set controller              |
| <b>Traditional 32-bit</b>      |                 |                  |           |                    |                    |       |        |           |      |              |        |      |            |                                       |
| 68008                          | 32              | 8                | 4MB       |                    |                    | —     | *      | —         | —    | —            | —      | —    | —          |                                       |
| 68000                          | 32              | 16               | 8MB       |                    |                    | —     | *      | —         | —    | —            | —      | —    | —          |                                       |
| 68010                          | 32              | 16               | 16MB      |                    |                    | —     | *      | —         | —    | —            | —      | —    | *          |                                       |
| 68012                          | 32              | 16               | 4GB       |                    |                    | —     | *      | —         | *    | *            | —      | —    | *          |                                       |
| 32008                          | 32              | 8                | 16MB      |                    | 10                 | *     | *      | —         | *    | *            | —      | *    | *          |                                       |
| 32C016                         | 32              | 16               | 16MB      |                    | 10                 | *     | *      | —         | *    | *            | —      | E    | *          |                                       |
| 32C032                         | 32              | 32               | 16MB      |                    | 10                 | *     | —      | —         | E    | *            | —      | E    | *          |                                       |
| <b>Advanced 32-bit</b>         |                 |                  |           |                    |                    |       |        |           |      |              |        |      |            |                                       |
| V70                            | 32              | 32               | 4GB       |                    | 20                 | *     | *      | *         | *    | *            | *      | *    | *          | 2.5 MIPS avg; CMOS                    |
| 32332                          | 32              | 32               | 4GB       | 0.2                | 15                 | *     | *      | —         | E    | *            | *      | E    | *          | 1st 32-bit data+adr; NMOS             |
| 68020                          | 32              | 32               | 4GB       | 0.08               | 25                 | *     | —      | —         | E    | *            | *      | E    | *          | 5 MIPS sustained                      |
| 80386                          | 32              | 32               | 4GB       | 0.03               | 33                 | *     | *      | *         | *    | *            | E      | E    | *          | has 8086 subset (SX=16-bit bus)       |
| WE32100                        | 32              | 32               | 4GB       | 0.13               | 18                 | *     | *      | —         | E    | *            | *      | E    | *          | AT&T "3B" computer; 4 MIPS avg        |
| Clipper                        | 32              | 32               | 4GB       | 0.03               | 33                 | *     | *      | —         | E    | *            | *      | *    | *          | RISC (Harv arch), 5 MIPS avg (C100)   |
| Z80000                         | 32              | 32               | 4GB       | 0.08               | 25                 | —     | *      | *         | *    | *            | *      | E    | *          | 5 MIPS avg (has Z8000 subset)         |
| 68030                          | 32              | 32               | 4GB       | 0.06               | 30                 | *     | —      | —         | *    | *            | *      | E    | *          | has data cache, twice 68020 speed     |
| 32C532                         | 32              | 32               | 4GB       | 0.06               | 30                 | *     | *      | —         | *    | *            | *      | E    | *          | UNIX top performer, 10 MIPS           |
| T414                           | 32              | 32               | 4GB       | 0.05               | 20                 | *     | *      | —         | —    | —            | E      | E    | *          | transputer; 4 multiprocessing links   |
| 86C010                         | 32              | 32               | 64MB      | 0.08               | 12                 | *     | *      | —         | E    | *            | *      | E    | —          | Acorn RISC (44 instructions)          |
| 80960                          | 32              | 32               | 4GB       | 0.06               | 16                 | *     | —      | —         | *    | *            | *      | *    | *          | RISC, multitasking, 8MIPS, 0.5K cache |
| 68040                          | 32              | 32               | 4GB       | 0.06               | 33                 | *     | —      | —         | *    | *            | *      | *    | *          | 17 MIPS, 3MIPS, 8K cache              |
| 80486                          | 32              | 32               | 4GB       | 0.03               | 33                 | *     | *      | *         | *    | *            | *      | *    | *          | 20 MIPS, 8MIPS, 8K cache              |
| 29000                          | 32              | 32               | 4GB       | 0.04               | 25                 | *     | *      | —         | *    | *            | *      | E    | *          | RISC, 17 MIPS sustained               |
| 88000                          | 32              | 32               | 4GB       |                    |                    | —     | —      | —         | E    | E            | E      | *    | *          | RISC                                  |
| T800                           | 32              | 32               | 4GB       | 0.05               | 20                 | *     | —      | —         | —    | —            | E      | *    | *          | Transputer                            |
| 86900                          | 32              | 32               | 4GB       | 0.07               | 17                 | —     | —      | —         | —    | —            | E      | E    | *          | Sun μP (107 instructions) SPARC       |
| WE32200 <sup>a</sup>           | 32              | 32               | 4GB       | 0.13               | 30                 | *     | *      | —         | E    | *            | *      | E    | *          | 8 MIPS avg (28MHz)                    |
| SPARC                          | 32              | 32               | 4GB       | 0.05               | 33                 | *     | —      | —         | *    | *            | —      | *    | *          | Cypress, Fujitsu, etc, >20 MIPS       |

(a) AT&T may discourage OEM sales. A - CMOS version available. E - external chip can provide this function.

the target instrument carrying trial ROM code. You can get battery backed-up RAM from Dallas, Thomson-Mostek, and others, or you can make your own. (c) Use a "ROM-emulator": This is a clever little box you buy, with a cable and 28-pin DIP plug at the end. You plug the latter into the EPROM socket of your instrument, and it emulates a ROM. However, the little box actually contains a dual-ported RAM, which you download from your PC via a serial port. The ROM-emulator technique is the fastest of all, because you don't have to plug and unplug memory chips each time you change the software – you just download the new code to the emulator, which stays plugged into the target instrument. ROM emulators go by names like "Memulator" or "Romulator" (the latter from Onset Computer Corp., N. Falmouth, MA).

**Monitor ROM.** If your target instrument has a serial port in it, you can make the software development job easier by coding a simple "monitor" EPROM whose task is not to run the instrument but only to serve as a communication link between the port and the target instrument's memory. The simplest monitor lets you load code into RAM, then begin program execution. That lets you speed up code development, because you can load trial programs into the target instrument's RAM directly from your PC. With just a little more effort you can add additional features to the monitor, for example, the ability to look at the contents of specified locations in RAM. With that addition, your trial programs can tell you what's happening, for example, by depositing numbers in specified locations in RAM, then jumping back to the monitor (through which the PC can examine those RAM locations). You might exploit this capability by coding diagnostic "software breakpoints" into your trial code: At such a breakpoint the contents of designated registers or memory locations would be

copied into an unused area of RAM, there to be subsequently read by your PC via the monitor. This sort of tool really shortens the development cycle, because it lets you quickly locate the source of trouble in the program.

#### ***In-circuit emulator***

The burn-and-crash technique, and its elaborations above, can usually succeed in getting the job done. But it's not an ideal method. For one thing, it requires the use of resources on the target instrument, such as serial ports. And, more seriously, it doesn't provide a way to get at faulty operation at the *hardware* level. To understand this last statement, imagine you have an instrument that is crashing because it is improperly trying to write to EPROM. You've figured that out, somehow, but you're stuck because there's no easy way to code a *software* breakpoint to trap a fault known only from its hardware symptom. In this case, for example, an address register may have been overwritten earlier. It's a subtle error: The actual crash occurs long after the initiating event, and there's just no way to find the bug by looking at the software code. What you need is a way to set a "hardware breakpoint."

The solution is an *in-circuit emulator* (ICE): This is a box of hardware (or plug-in card) you buy, this time emulating the *microprocessor* in the target system via a cable with CPU-like plug on the end. The in-circuit emulator can execute code in the target system's memory (EPROM or RAM), or it can execute code you have downloaded to the emulator. Either way, it knows everything that is happening in the CPU – it can monitor the contents of registers, and it can set *hardware* breakpoints. For example, to solve our hypothetical problem we could ask it to check for a write cycle to the EPROM's area of memory space, providing us with a register dump and listing of the last 100



instructions executed before the doomed write-to-EPROM.

In-circuit emulation is the best way to develop code – it's the fastest, and it's the most powerful. The disadvantages are cost (several thousand dollars, sometimes much more) and the need to buy a new emulator for each type of microprocessor you use. The ROM-emulator, by comparison, is completely general, but not as powerful. It's the poor man's in-circuit emulator.

### Development systems

"Development system" is the collective term for the paraphernalia of cross-assembler, EPROM burner, and hardware in-circuit emulator. Development systems traditionally were packaged as impressive stand-alone systems, but recent trends are favoring plug-in cards that use a PC as host, perhaps controlling an external box containing additional circuitry. Whatever the format, you should invest in a development system if you are in the business of developing instruments based on a particular microprocessor. Development systems are generally offered by the manufacturer of the microprocessor family you're using. There are, in addition, companies making "universal emulators" that accept plug-in boards for additional processor types. Hewlett-Packard, Tektronix, Microcosm, and Applied Microsystems are some companies offering development systems.

### Logic analyzers

Logic analyzers are the "super-oscilloscopes" of digital hardware development. The logic analyzer was a major character in Tracy Kidder's book *Soul of a New Machine*. These marvellous devices look like fancy oscilloscopes, but they have dozens of channels, lots of memory, complex

"word-recognizer" triggers, and the ability to disassemble executing instructions and display them on the screen. Logic analyzers work in two modes, *state analysis* and *timing analysis*. Here's how you use them.

**State analysis.** In state-analysis mode, you provide the logic analyzer with a synchronous clock signal from your circuit (usually the CPU clock), and you clip clusters of leads (which come from "pods") onto data and address bus lines, and any other signals you're interested in. Good logic analyzers can handle 60 or 80 channels, and clock rates to 25MHz or more. You then set the triggering to catch the software event you're interested in: There are usually several *word recognizers*, which you set up with the pattern of 0's, 1's, and X's (don't care) you want to trigger on; these are, of course, address and data bits. Good logic analyzers let you combine the outputs of the word recognizers in Boolean and state expressions; you could, for example, trigger on the 10th pass through a certain subroutine.

The logic analyzer waits for the trigger, then records successive states on all input lines. You can display them as digital waveforms, or as lists of 1's and 0's (or hex or octal), with disassembled code shown alongside. You can walk through the array of recorded states (typically 4K or more), and, most important, you can look "backward in time" before the triggering event. This lets you see the few thousand states that preceded the triggering event, which usually provides the cause of the trouble.

**Timing analysis.** In timing-analysis mode, the logic analyzer runs from a fast *asynchronous* clock, typically 100MHz, logging the logic states of a smaller number of input lines (typically 16). The triggering

logic is necessarily simpler, usually a single word recognizer. The analyzer waits for the trigger condition, then fills up its memory with the fast samples. In timing mode you can see fast glitches and other waveform aberrations that might be missed in state-analysis mode. You can also turn on a "glitch-detection" mode, in which the analyzer looks for two edges occurring during a 10ns sample period.

**Cross-triggering.** You can combine state analysis and timing analysis with a powerful technique called *cross-triggering*. In this combined mode, state analysis can arm timing analysis, or vice versa. Thus, you might set the state triggering logic to catch a problem that occurs within some particular software loop, arming the timing logic to store a fast burst on the next occurrence of its trigger word. In that way you could discover a short logic glitch that happens only rarely.

Logic analyzers with cross-triggering capability provide split-screen display, so you can watch the fast timing waveform go by as you scroll through the state display. Some of the bigger names in logic analyzers are Gould, Hewlett-Packard, Philips, and Tektronix.

### **Evaluation boards**

During the 1970s, when 8-bit microprocessors like the 6800 and Z80 were becoming popular, there was an *evaluation board* offered by the manufacturer for each new microprocessor. These boards had a small keypad, hex display, RAM, monitor EPROM, some parallel and serial ports, and a breadboard area for adding custom circuitry of your own choosing. You could hand-assemble little programs, enter them through the keypad, and delight in the results. For a world feeling its way into microprocessors, it was an easy way to learn the ropes.

It's a more sophisticated world now, and these breadboards have become almost extinct. However, you'll still see evaluation boards offered for specialized processors, for example, large signal-processing ICs, or complex video processors. These boards contain the processor itself, surrounded by "glue" logic, analog signal components, and often a conventional microprocessor used for control. These evaluation boards are frequently made in the form of plug-in cards for personal computers, complete with driver software. Given the complexity of many of these new special-purpose processors, the evaluation board saves lots of time and makes good sense.



# ELECTRONIC CONSTRUCTION TECHNIQUES

CHAPTER

12

Between completing a circuit design and testing a finished product, you've got a number of decisions to make: Is the instrument going in a benchtop case, a "relay rack" enclosure, or perhaps some sort of modular "bin" chassis? Should the circuit itself be constructed with point-to-point wiring, on a "breadboard" card, with Wire-Wrap connections, or on a printed-circuit board? Are connections to the circuit board made with solder lugs, flat ribbon-cable connectors, or edge connectors? Should individual circuit cards be housed in a card "cage," plug into a motherboard, or what? Does it pay to design a printed-circuit motherboard or use hand-wired backplane connections? Which adjustments should be on the circuit board, and which on the front (or rear) panel? Decisions like these are important for the appearance, reliability, and serviceability of the finished product, not to mention cost and ease of construction and testing. In this chapter we will try to give some information and guidance on this important subject, one that tends to be overlooked in electronics courses where circuit

work is usually done on plug-in breadboarding gadgets. We will begin with the circuit construction itself, treating interconnections, controls, and enclosures last.

Because this chapter does not deal with circuit design, it could be skipped in an abbreviated reading.

## PROTOTYPING METHODS

### 12.01 Breadboards

The unusual name "breadboard" seems to have arisen from the early practice of building radios on handsome slabs of varnished wood, with tubes, coils, condensers, etc., and the interconnecting wires all fastened to the topside of the board. Later, radios of greater refinement and elegance (to be used by stiff-skirted ladies in the parlor) were built with holes near each component so that the wiring would be hidden from view underneath the board. The practice of testing circuits by constructing trial versions on some sort of board or jig is still called breadboarding.



Wooden breadboards are no longer used (except in the kitchen). Instead, you can get handy plastic blocks with rows of holes spaced to accommodate ICs or other components and (usually) some extra rows for distributing the power-supply voltages. These breadboards are typified by those manufactured by AP and Global Specialties, and more elaborate breadboarding boxes are made by these companies and by E&L Instruments, among others. These are intended for *testing* circuits, not for constructing permanent (even semipermanent) versions.

### 12.02 PC prototyping boards

To construct one-of-a-kind circuits of some permanence, probably the best approach is to use a different kind of breadboard, one of the many printed-circuit (PC) prototyping cards available with predrilled pads for ICs and other components, but with no interconnections laid out on the board itself. Instead, each component pad is connected to two or three uncommitted pads nearby, and you wire the circuit by soldering lengths of insulated wire from pad to pad. There are usually some additional lines running around the board for power-supply distribution and ground. These boards, made by companies such as Douglas Electronics, Artronics, Vector, Triad, Radio Shack, and many others, usually include a card-edge connector – gold plated “fingers” of copper board material aligned at one end of the card for easy plug-in connection to a mating PC card-edge socket.

There are several standard connector configurations, a very common one being 22 connections on each side of the card, with 0.156 inch spacing (spacings of 0.125 and 0.1 inch are also common). A connector to mate with such a card is called a 44-pin “dual-readout” PC card-edge connector. Prototyping cards are available in various sizes, accommodating from 12 to 36 or more IC packages, with some larger

computer-compatible cards available that will accept 100 or more ICs and will plug directly into small-computer mainframes. Some of these boards are single-sided, and others are double-sided with plated-through holes, a subject we will discuss shortly in connection with custom PC design. Figure 12.1 shows a small PC prototyping card (Douglas Electronics 11-DE-3) plugged into a 44-pin dual-readout socket that includes card-supporting guides (Elco 6022).

Another form of breadboard that has enjoyed considerable popularity is the so-called perfboard, a thin sheet of laminated insulating material manufactured with regularly spaced holes ( $3/16$  inch is a common spacing) designed to accept little metal pins. To wire up a circuit, you shove in dozens of the little pins wherever you want, then stick in the components. You then solder wires from pin to pin to complete the circuit. Perfboard works OK, but it is awkward to use with the tight spacings of IC packages (0.1 inch between pins). Figure 12.2 shows an example.

### 12.03 Wire-Wrap panels

A variation on the PC breadboard is the Wire-Wrap (a registered trademark of Gardner-Denver) panel, a circuit card festooned with IC sockets (or pads), with a pin 0.3 to 0.6 inch long protruding from each socket connection (Fig. 12.3). These pins are square in cross section, typically 0.025 inch on a side, and made of a hard metal with sharp corners, plated with gold or tin. Instead of soldering to the pin, you wrap an inch of bare wire tightly around it, using an electric wire-wrapping tool (there’s an inexpensive variation known euphemistically as a “hand-operated Wire-Wrap tool”). Wire-wrapping is very fast. You just stick the stripped end of the wire into the Wire-Wrap tool, put the tool over the Wire-Wrap pin, and zip, you’re done. The standard wire used for this purpose is

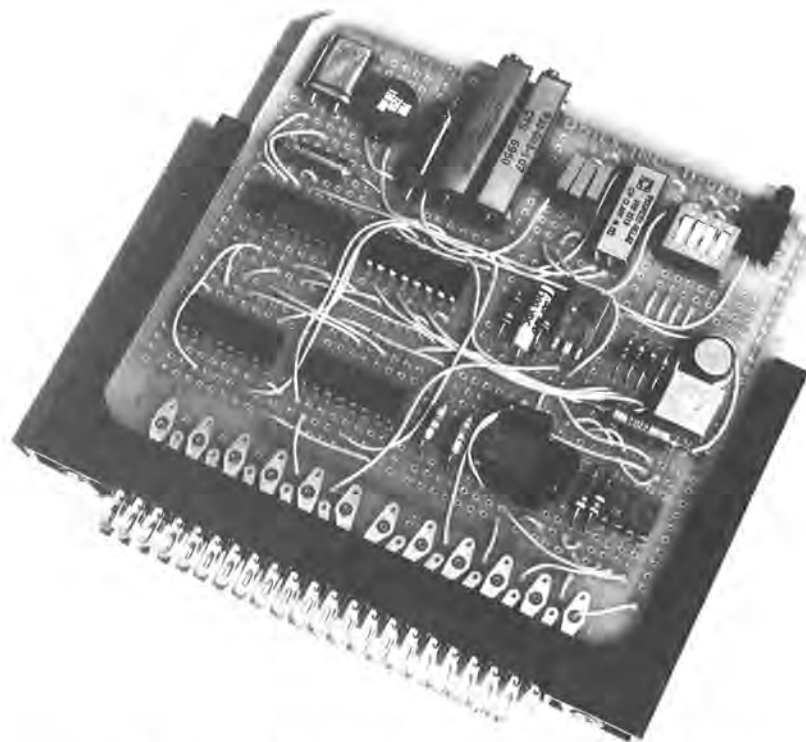


Figure 12.1. A "solder breadboard" prototyping card is useful for wiring up small circuits, especially those involving both discrete parts and ICs. This particular specimen will accommodate 12 dual in-line (DIP) packages, and it includes common lines for ground and power supply voltages. An "edge-connector" foil pattern is standard, so the card can be plugged into a card cage or supporting connector, as shown. This particular circuit uses a variety of components, including single-turn and multiturn trimmers, inductors, a crystal, a DIP switch, a miniature relay, and a logic-state indicator, in addition to both transistor and IC circuitry.

silver-plated copper wire of 26 or 30 gauge, with Kynar insulation. There are special tools available for stripping the insulation from the thin wire without nicking it. The wire is stretched tightly around the sharp corners during the wrapping process, forming a few dozen gas-tight cold welds. As a result, wire-wrapped connections are as reliable as well-soldered connections, and it is extremely easy to do them rapidly. For logic circuits, where you have few discrete components, wire-wrapping is probably the best technique for constructing one or two custom circuits of reasonable complexity.

Because Wire-Wrap panels are laid out primarily for IC packages, the technique is less convenient for linear circuits with many resistors, capacitors, etc., and the soldered prototype breadboard technique described earlier will usually be better.

It is possible to use discrete components with a Wire-Wrap panel. You just mount them on little "headers" that plug into IC sockets, then do the wire-wrapping as usual from the socket pins. Some Wire-Wrap panels have extra solderable pads (rather than IC sockets) available for discrete components. A very nice kind

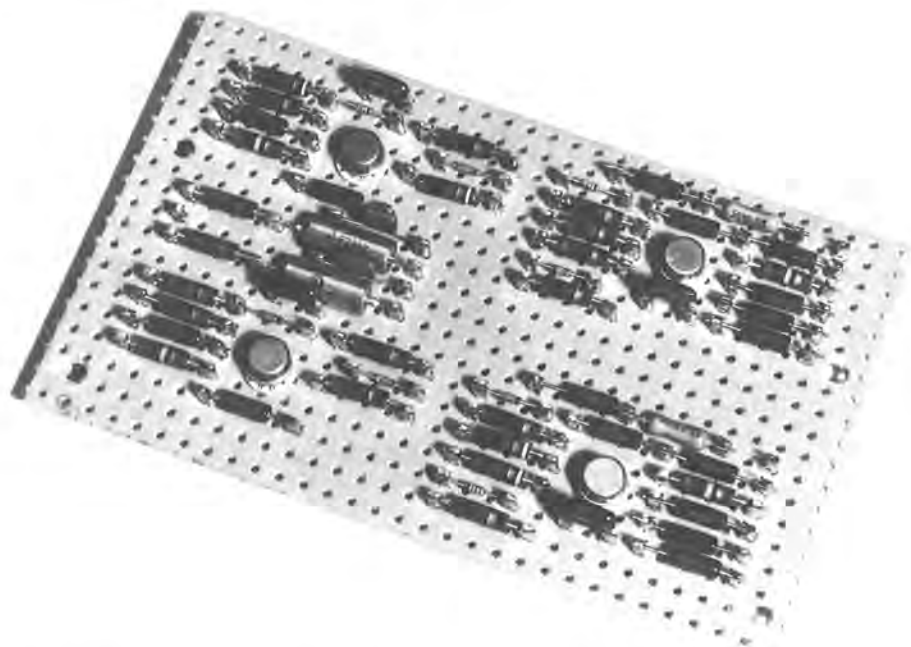


Figure 12.2. “Perfboard” can be handy for prototype circuits constructed with discrete components, although it is not particularly good for ICs. The terminals are press-fit into the holes (or flared with a special tool) and wired underneath.

of Wire-Wrap board is available with the pins on the *component* side of the board (the usual procedure is to have the pins stick out the other side). Although this type of board is less dense (fewer ICs per square inch), it is easier to use with discrete components, since you can see them while wrapping, and it allows closer spacing between adjacent circuit boards, since both components and Wire-Wrap pins take space on the same side. This kind of board without sockets is actually quite convenient for construction of linear or digital circuits. Figure 12.4 shows an example. In Figure 12.5 we have compared a prototype circuit built on a Wire-Wrap panel with the final printed-circuit version used in production. Printed circuits are much easier to produce in quantity; they are superior electrically and less cluttered than Wire-Wrap panels. We will talk about PC cards next.

## PRINTED CIRCUITS

### 12.04 PC board fabrication

The best method of constructing any electronic circuit in quantity is to use a printed circuit, a stable insulating sheet of material with thin plated copper lines bonded to the sheet forming the circuit paths. Although early printed circuits were associated with poor reliability (Remember the advertisements stressing the superior quality that only handcrafted television sets without printed circuits could provide?), the process of manufacturing board material and producing finished boards has been perfected to the point that printed-circuit boards now have very few problems. In fact, PC boards offer the most reliable fabrication technique. They are routinely used in computers, spacecraft, and military electronics where high reliability is essential.

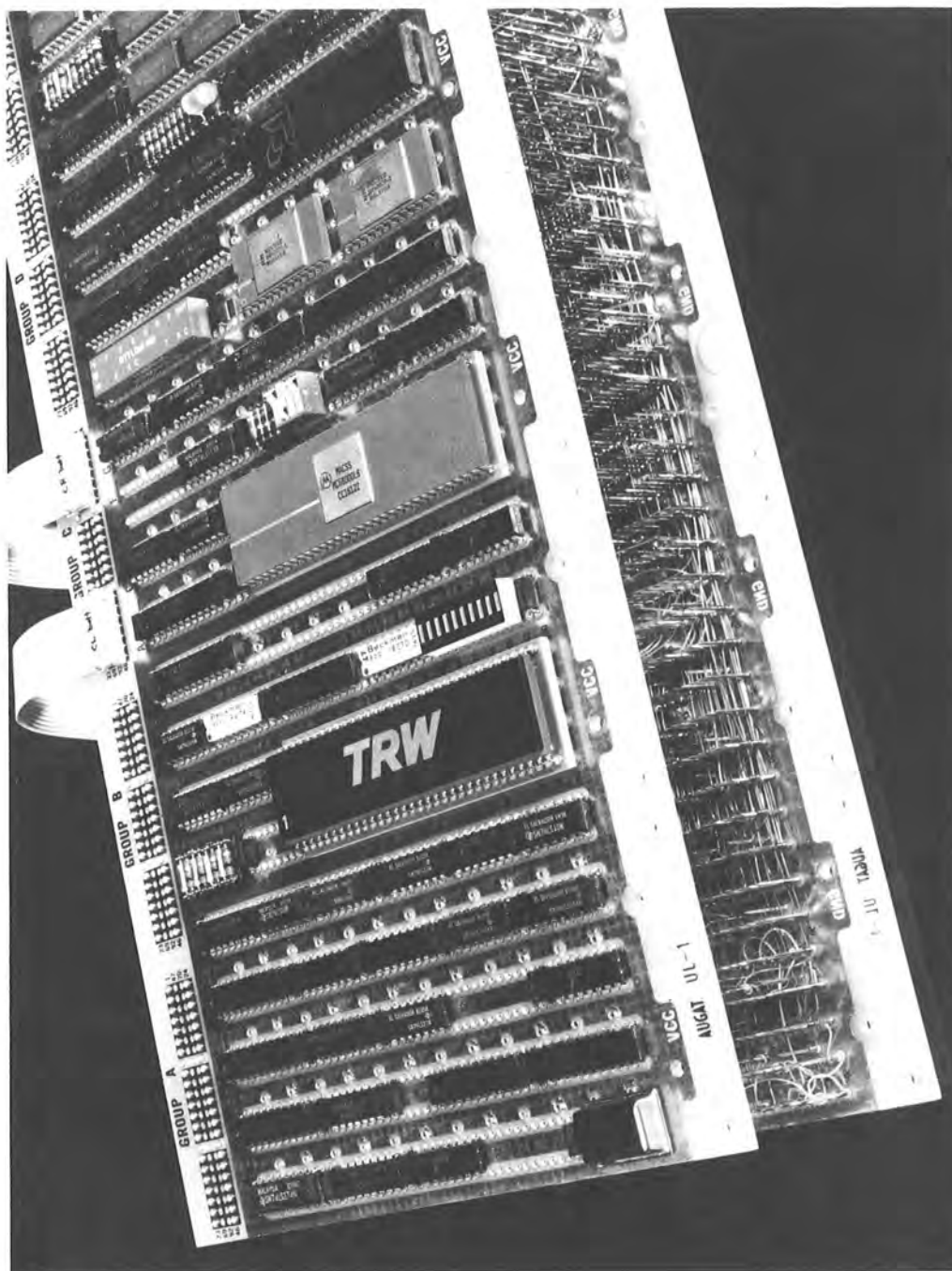


Figure 12.3. A large Wire-Wrap® panel and its underside (visible in a mirror). This microcomputer circuit was wired by machine ("CAD/CAM," see Section 12.08). Chips of different sizes, as well as component-carrying "headers," are accommodated by this general-purpose panel, which also includes an area for off-board connections via flat ribbon cable with 14-pin DIP plugs.

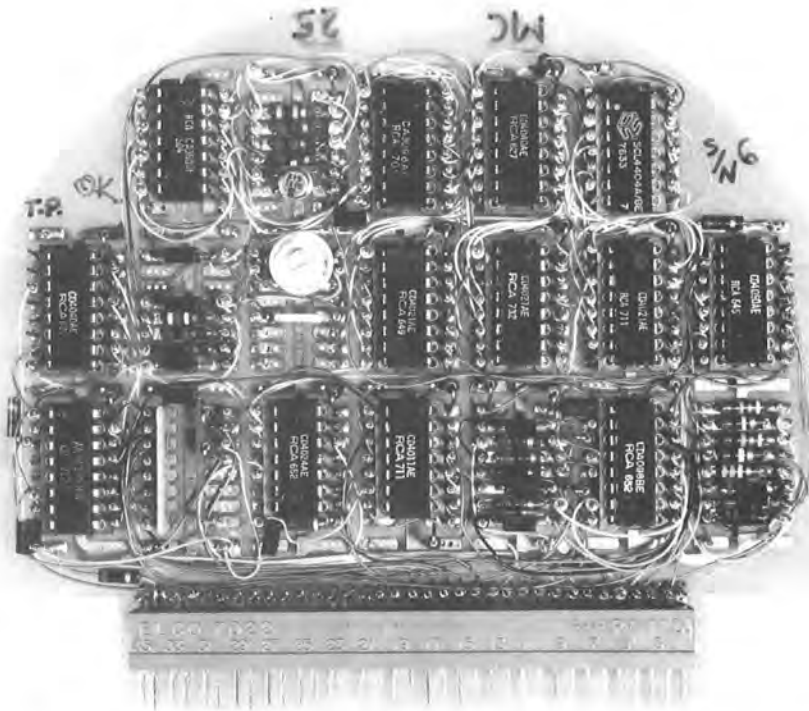


Figure 12.4. Wire-Wrap boards provide a neat and fast construction method particularly good for circuits made with digital ICs. This board uses a printed-circuit pattern to bring out the Wire-Wrap pins on the component side, an alternative to the usual underside pin configuration. Its peculiar shape is dictated by the interior of the oceanographic pressure cell into which it fits.

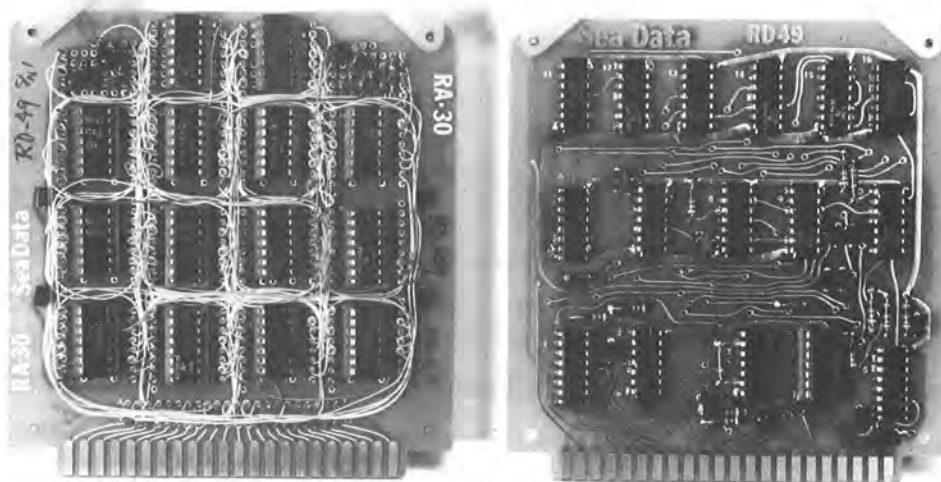


Figure 12.5. A Wire-Wrap prototype board and its printed-circuit successor. PC boards are less cluttered and far easier to fabricate in quantity. They eliminate wiring errors, too.



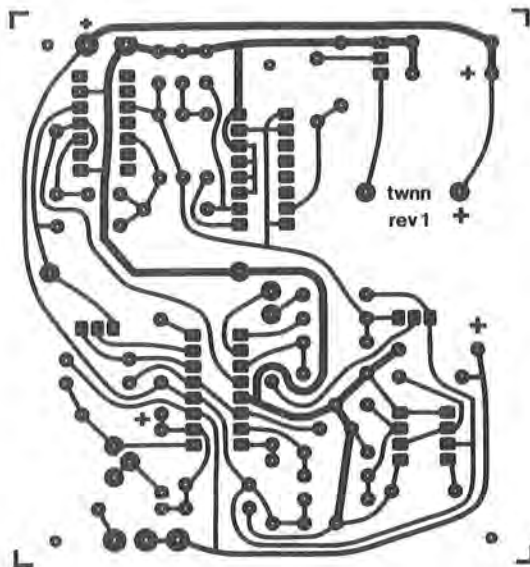


Figure 12.6. Foil pattern for a simple single-sided PC board. This “layer positive” was photographically reduced from a double-sized hand-taped Mylar pattern. The circuit board is 3.25×3.5 inches and has 4 ICs, 24 resistors, 11 capacitors, 5 diodes, 1 trimmer, and 1 piezo buzzer.

### The “Mylar” or photoplot

Printed-circuit-board production begins with a set of actual-size transparencies on which an opaque pattern delineates the desired circuit traces and “pads” (Fig. 12.6). There are lots of rules and tricks in this business, but the basic idea is to figure out how to make all the interconnections the circuit demands by running lines around a board. As we will describe shortly (Section 12.08), these transparencies are usually produced directly on film by computer-driven photoplotters or laser plotters, working from a description of the circuit that you produce on a CAD (computer-aided design) system. However, for simple circuits, you may choose to do the layout manually, sticking opaque tape and patterns onto clear Mylar film. In the latter case you usually make the taped Mylar pattern twice actual

size, from which you photographically produce an actual-size transparency.

No matter what the parentage of the final 1:1 transparency, the result is a set of conductor patterns on film. For the simplest boards you may be able to make all the needed connections (perhaps aided by a few wire “jumpers”) with a *single-sided* board, which has all its traces on the underside (officially called the “solder” side; the top is called the “component” side). Most often, however, you need traces on both sides of the board. Such *double-sided* boards invariably use *plated-through* holes (the holes in the board are lined with foil, connecting corresponding pads on both sides of the board). This really makes a difference when you are attempting to route traces, because you can always switch sides (using a “via” hole) when a trace runs into a dead end, whereas with a single-sided board you often get hopelessly boxed in. As an important side benefit, the use of plated-through holes ensures a superior solder joint to the component leads, because the solder wicks up through the hole.

For complex digital circuits you often see laminated multilayer PC boards, in which interior layers (called the “core”) are used to carry ground and power-supply voltages, and sometimes signal lines as well. Four-layer and six-layer boards are pretty routine these days, with occasional use of more layers (up to 40!) in desperate situations.

### Manufacturing

The board material (usually 1/16 inch of so-called FR-4 board, a fire-resistant epoxy-bonded fiberglass) comes clad on both sides with copper (“2 ounce” thickness is standard; the copper is 0.0027 inch thick). The first step is to drill the holes, using a template or automated drilling machine keyed to the full-size photopositive from the photoplotter or the Mylar

pattern. The holes are then “plated through” by a tricky multistep copper plating process, creating continuous conducting paths from one side of the board to the other.

The next step is to create a tough “resist” material, adhering to both sides of the board everywhere except where the foil for the circuit is to remain. This is done by (a) coating the board with a light-sensitive film (usually a thin adhesive “dry film”), then (b) exposing the board to light with the full-size photopositive accurately sandwiched on top, and (c) chemically “developing” the film (as in conventional photography) to make the exposed areas permanent. A step analogous to photographic “fixing” then removes the unexposed film, in precisely the pattern that will ultimately become the circuit traces. Then the board, with the pattern of resist now masking the areas where the copper foil is ultimately to be removed, is immersed into a solder-plating bath. The result is to plate solder (a tin/lead alloy) everywhere that the foil pattern is to remain, including the insides of the holes.

Next the resist is removed chemically, exposing the copper that is to be removed, and the board is treated with a copper-etching compound. That leaves the desired pattern of solder-plated copper, complete with plated-through holes. At this point it is important to carry out a step known as “reflow soldering,” which consists of heating the board to make the thin solder plating flow. This prevents the formation of tiny slivers of metal (from the undercutting action of the etching bath) that could otherwise cause conductive bridges. Reflow soldering also improves the solderability of the finished board; a reflow-soldered board is a delight to “stuff” with components.

The next step in board manufacture is to electroplate the edge-connector fingers with gold. The final process in board manufacture consists of applying a tough

“soldermask” coating over the entire board, covering all of the foil except for the pad areas. This greatly reduces the tendency to form “solder bridges” between closely spaced traces during subsequent soldering operations. It also makes the board moisture- and scratch-resistant. Soldermask materials can be applied by silk-screen methods (“wet mask”) or by the same photographic “resist” methods used to create the foil circuit pattern (“dry mask”). You can recognize the soldermask by its dark green color and by the observed fact that it is practically impossible to remove. In industrial board manufacture, the board might then be stuffed by automatic machines, with all joints soldered in a few seconds in a “wave-soldering” machine. The alternative is to stuff and solder by hand.

There is a simpler process of board manufacture that is sometimes used, especially in small production situations or for single-sided boards, where plated-through holes aren’t needed. In this method you begin by coating the board with photo-resist, then expose it through a full-size *negative* (rather than positive) of the desired pattern, i.e., a photographic film that is transparent wherever you want foil to *remain*. The resist is developed, and then the unexposed resist is dissolved away. This board then has a layer of tough resist covering the copper that is to remain, so you simply expose it directly to the etching compound (omitting the solder-plating step, above). After the superfluous copper has been etched away, the remaining resist is washed off with solvent, leaving the desired pattern in copper. At this point it is best to treat the board with an “electroless” tin-plating bath in order to cover the copper with a metal less susceptible to corrosion. As before, the edge-connector fingers will then be gold-plated. The final step in this process consists of drilling the holes by hand, using the actual conductive pattern as a guide (each “donut” pad has a small

opening in the center to aid in drilling the finished board).

#### □ 12.05 PC board design

There are several important decisions you have to make during PC board design, during component “stuffing,” and finally when the board is used in an instrument. In this section we will try to touch on the most important of these.

#### □ PC board layout

To make a printed-circuit board, you must ultimately convert your schematic diagram into a corresponding pattern of desired copper-foil traces that will compose the finished board. There are basically two ways: (a) Working from the manually drawn schematics, you can use pencil and paper (and lots of erasers!) to figure out a set of interconnection paths (“routes”) that does the job, then manually apply opaque tape and preformed connection patterns accurately onto clear polyester film to make the finished “Mylar” masters, or (b) you can convert your hand-drawn schematic to a connection “netlist,” then use a CAD (computer-aided design) program to figure the trace routing, producing a set of precision machine-drawn “photoplots” directly; better still, you can replace the manual schematic drafting with CAD-based “schematic capture,” in which you draw the diagrams directly on a graphics workstation, using a graphic tablet or mouse.

CAD-based circuit drawing and layout (option b, above) has many advantages, including automatic netlist extraction, painless documentation, the ability to check for design and layout rule errors, the ability to make changes relatively easily, and the ability to produce complex multilayer boards with precise alignment of pads and traces. It is the method of choice for nearly all complex and high-density PC board production. However, we would like

to begin simply, by briefly describing manual methods of PC board layout (option a, above). Once you know how to do a PC board layout by hand, you will understand what you’re trying to make with the complex computer-based CAD tools. Furthermore, for simple jobs the manual methods may be all you need, and they are cheaper (and sometimes quicker) than CAD-based methods. They are particularly appropriate for the uncrowded boards you might use in simple unsophisticated instruments, of the sort you might build in small quantities for research laboratory use. They are also well adapted to boards containing parts with unusual shapes and lead spacings. Later, in Section 12.08, we will describe the CAD methods that are mandatory for high-density digital multilayer boards intended for large production.

There are several stages along the way from a schematic diagram to a final printed circuit. Beginning with the diagram, you generally work out trial pencil sketches of component layouts and interconnections, eventually working these together into a final pencil layout drawing. From this you make the “Mylar,” consisting of accurately aligned “pads” (terminal areas for component connections) and taped interconnections. Precut patterns are used for IC and transistor pads and for ribbon and edge connectors, since these have standard spacings and dimensions. The pencil sketch and Mylar are usually made double size to allow greater accuracy (and to keep your eyes from popping out!). When the Mylar (two Mylars for double-sided boards) is completed, it is photographically reduced to an actual-size negative, from which a trial board is made as described previously. You generally “stuff” the prototype board with components, turn on the power, and then hunt down the errors; this lets you correct the Mylar artwork to produce final boards. The following subsections provide some further details and hints.

### □ *Initial sketch*

We recommend doing the initial layout with pencil on grid paper (5 lines/inch), with two colors to indicate foil patterns on the top and bottom (assuming it is a double-sided board). We usually use black pencil for runs on the bottom and green or red for the top (component) side. Since you're likely to do plenty of erasing, it is best to use vellum graph paper. The 0.2 inch gridding corresponds to 0.1 inch final size, the universal measure for IC pin spacings, transistor pinouts, edge connectors, etc. Your drawing should be the view from the component side; i.e., the sketch of the component-side (top-side) foil pattern looks like the final pattern, and the sketch of the bottom foil pattern is what you would see looking down through the finished board with x-ray eyes. While working on the layout, indicate component outlines with a pencil of a third color. All this work should be freehand. Don't waste time with an outline template; just use the grid lines as a guide to draw IC and component pinouts.

It is generally best to work up some trial layouts on a piece of scratch paper, particularly for sections of the circuit that may require special layout to minimize long lines or capacitive coupling. It may take some experimentation to arrive at good component arrangements. A trial layout might consist of a block of the circuit with two or three op-amps, or perhaps the input or output section of the circuit. These blocks should then be worked together onto the large gridded vellum, with adjustments being made as you go. Don't hesitate to do lots of erasing!

### □ *Layout dimensions and hints*

Try to have all ICs pointing in the same direction, preferably in straight rows. Likewise, resistors should be in even rows, not staggered. We use 0.030 or 0.040 black tape for signal runs, with wider tape for

power supplies (0.05 or 0.062 inch) and very wide ground runs (0.1 to 0.2 inch, or even wider; it's common to broaden the ground runs with lots of tape). Be sure to include plenty of bypass capacitors, one 0.1 $\mu$ F for every two to four ICs. As you scratch your head, trying to juggle the tangled maze of interconnections, don't forget that components act as "jumpers" – they can hop over runs on the board.

**Dimensions and spacings:** On the actual-size PC board, we recommend holes spaced 0.4 inch for resistors ( $\frac{1}{4}$ W size), with spacing of 0.1 or 0.15 inch between resistors (with 0.15 inch spacing you can get a tape run between adjacent pads). We favor CK05 and CK06 types of ceramic capacitors, with their controlled 0.2 inch lead spacing, or the "DIP" 0.3 inch types (AVX type MD01, or Kemet C630C104M5U); they can also be spaced 0.1 inch from other capacitors or resistors. Leave some room around ICs for logic clips: a minimum of 0.2 inch to the next IC pads and a minimum of 0.15 inch to the nearest resistor or capacitor pads. Leave 0.030 inch spacing between tape runs, and don't run anything closer than 0.25 inch from the edge of the board, to allow room for card lifters, guides, standoffs, etc. Avoid running lines between the 0.1 inch spaced pads of an IC, unless necessary. You can fit up to six tape runs lengthwise between the pads of a standard DIP IC pattern (they're spaced 0.3 inch).

Recent PC design practice generally favors much higher line densities (both thinner lines and closer spacing) than the values recommended above; the latter would be called "15-15" design rules, signifying 0.015 inch minimum line width with 0.015 inch minimum conductor clearance. In the PC board industry, 15-15 rules are considered quite relaxed, with 12-12 pretty standard; 12-12 rules let you run one trace between adjacent IC pads on 0.1 inch centers (check for yourself that this is permitted, providing the pads aren't larger



than 0.064 inch in diameter). High-density boards often use 10-10 or 8-8 rules, with which you can squeeze *two* traces between adjacent IC pads (the maximum pad diameters are then 0.050 and 0.060 inch, respectively). Occasionally you see daring boards with 0.006 inch or thinner traces; the designers of such boards are trading the increased circuit density (*three* traces between adjacent pads, 20 down the center channel!) against smaller plated-through holes and other compromises that result in poorer production yield and board robustness.

#### □ **Connections to the board**

For the majority of boards it is probably best to bring out all connections through "edge-connector" contacts, which mate directly with sockets available in a variety of contact configurations. The most commonly used spacings are 0.156 inch, 0.125 inch, and 0.100 inch between fingers. Generally you'll put an edge-connector pattern at one end of the card, bringing power-supply voltages and signals through that connector. The card is mechanically supported, and it plugs in at that end (more on that shortly).

Often you see an edge-connector pattern at the other end of the card also, used instead of a flat ribbon connector to bring some other signals off the board or to other boards. Another method for bringing out signals is to use flat ribbon cable terminated in DIP plugs; such cables plug right into IC sockets on the board. You can buy these cables prefabricated in various lengths, or you can make them yourself with a kit consisting of flat cable, unassembled DIP plugs, and a crimping tool. Ribbon cables can also connect to the board via in-line or "mass-termination" connectors, which use one or two rows of pins on 0.1 inch centers.

For simple boards the best method of connection may be to use swage-solder

terminals or PC-type barrier strips with screw terminals. Avoid the use of large pads alone for connection of external wires to PC boards.

Figure 12.7 illustrates a variety of PC board connection techniques.

#### □ **Odds and ends**

With plated-through boards, use several holes to join ground foils on opposite sides of the board. Try to avoid using multiple passes through the board to reach your destination, since plated-through connections where no component is mounted are more likely to give trouble. The layout of a double-sided board generally winds up with most tape runs going horizontally on one side, vertically on the other.

General philosophy: Use smooth curves or 45° turns, rather than right-angle turns, for hand-taped layouts. Bring lines into pads as if heading for the center of the pad, rather than coming in at an oblique angle. Don't mount heavy components on boards (a couple of ounces ought to be the limit); assume that the instrument will be dropped 6 feet onto a hard surface sometime during its life! Put polarity markings on the component side for diodes and electrolytic capacitors, and label IC numbers and pin 1 location (if there's room). It is always nice to label test points, trimmer functions (e.g., "ZERO ADJ"), inputs and outputs, indicator light functions, etc., if you have room.

#### □ **Taping the Mylar**

General advice: Use an illuminated "light table" with a piece of precision gridded Mylar taped to it. Don't confuse this with the inexpensive gridded plastic films that are neither accurate nor dimensionally stable; a piece of precision gridded film will set you back at least \$20. Put your clear Mylar over, and stick down the IC pads accurately on it. Use the pencil sketch for guidance while taping. Wash



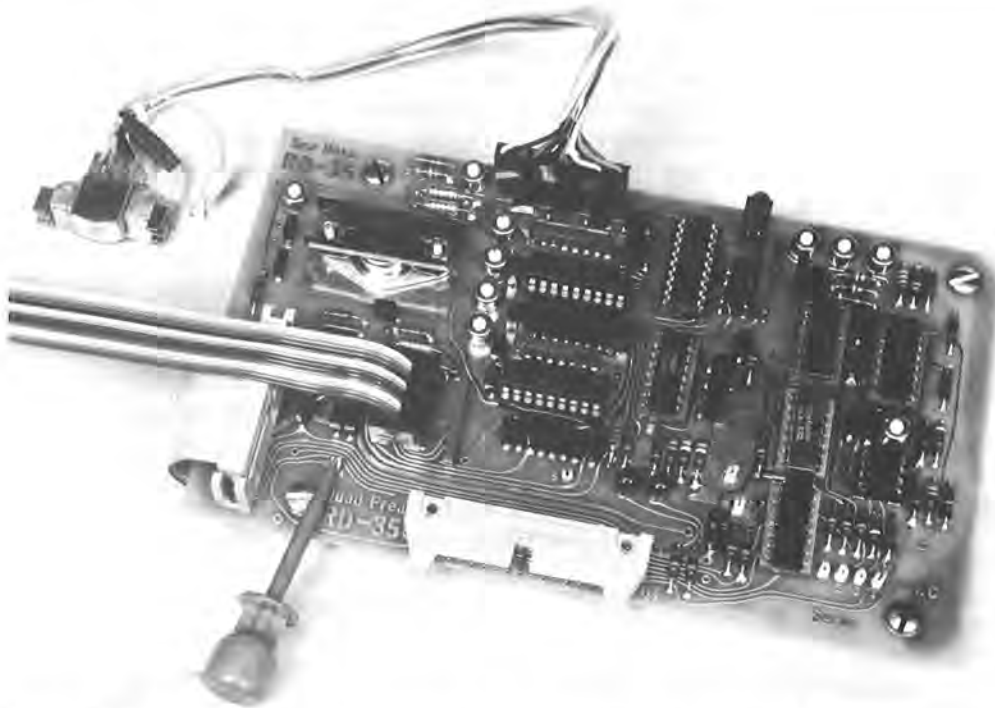


Figure 12.7. Several connection techniques are illustrated in this digital recorder printed-circuit card. The tape head connects via an in-line connector (which mates with a row of Wire-Wrap-type pins), and other signals are brought out with “mass-termination” ribbon connectors and a dual in-line ribbon connector. A test lead is shown clipped onto a “test point” terminal. This board also illustrates PC board heat sinking (upper left), a logic-state indicator (upper right), miniature single-turn trimmers, and single in-line (SIP) resistor networks.

your hands often to prevent deposition of oily film on the Mylar, and use alcohol to wipe any areas that might become oily. Use an Xacto knife with curved blade for tape and outline cutting, and learn not to cut through the Mylar. Press the tape down firmly after positioning; otherwise it will eventually curl up. Allow generous overlap where tape meets pad, etc. When laying out tape, don't hold it under tension; it will shrink and pull away from pads. Use precut bends and circles for the larger tape widths (0.062 inch or wider) when navigating tight turns. After the Mylar is completely taped, check it against the schematic by going over each connection on the diagram with a red pencil. When all seems OK, seal up

flaws on the Mylar with an indelible black felt-tip pen.

Precut PC graphics patterns are available from several manufacturers. Table 12.1 shows some recommended types. The Bishop Graphics catalog (5388 Sterling Center Drive, Westlake Village, CA 91359) includes extensive information on PC board layout and execution.

### 12.06 Stuffing PC boards

Your worries aren't over when you've got a finished board. You've got some decisions to make (e.g., whether or not to use IC sockets) and some important things to do (e.g., defluxing and lead trimming). Herewith, some thoughts on these subjects.

TABLE 12.1. SELECTED PC GRAPHIC PATTERNS

| Pattern <sup>a</sup>        | Bishop     | Datak   |
|-----------------------------|------------|---------|
| Small pads (0.150" OD)      | D203       | JD-145  |
| Medium pads (0.187" OD)     | D104       | JD-146  |
| Large pads (0.250" OD)      | D108       | JD-150  |
| Giant pads (0.300" OD)      | D293       | JD-343  |
| 0.150" thermal relief – pos | 5272       | JDS-532 |
| 0.150" thermal relief – neg | 5278       | –       |
| 0.187" thermal relief – pos | 5232       | –       |
| 0.187" thermal relief – neg | 5238       | –       |
| 16-pin DIP                  | 6109       | JD-64   |
| 16-pin DIP with in-betweens | 6946       | JD-179  |
| 20-pin DIP                  | 6999       | JD-575  |
| 20-pin DIP with in-betweens | –          | JD-585  |
| 28-pin DIP                  | 6904       | JDS-398 |
| 28-pin DIP with in-betweens | –          | JDS-591 |
| TO-5 transistor             | 6077       | –       |
| TO-18 transistor            | 6274       | JD-88   |
| TO-92 transistor            | –          | JD-91   |
| 0.100" connector pads       | 5004       | JD-145  |
| 0.100" edge-connector strip | 6714       | JD-123  |
| 0.156" edge-connector strip | 6722       | JD-121  |
| 0.031" black tape           | 201-031-11 | –       |
| 0.040" black tape           | 201-040-11 | –       |
| 0.050" black tape           | 201-050-11 | –       |
| 0.062" black tape           | 201-062-11 | –       |
| 0.100" black tape           | 201-100-11 | –       |
| 0.200" black tape           | 201-200-11 | –       |
| 0.062" universal corners    | CU601      | –       |
| 0.100" universal corners    | CU607      | –       |
| 0.200" universal corners    | CU609      | –       |

(a) for 2:1 artwork originals.

### Sockets

There is great temptation to use IC sockets everywhere, for ease of troubleshooting. However, if you're not careful, the sockets may well cause more trouble than they prevent. In general, sockets are a good idea at the prototyping phase, where IC substitution may be necessary to convince you that the trouble you're having is a

design error, not a bad component. They should also be used for expensive ICs (e.g., a D/A converter, microprocessor, or the like), ICs that you're likely to want to change from time to time (e.g., a program ROM), and ICs that have a good chance of being damaged sooner or later (e.g., chips that buffer input or output signals from outside the instrument).

The problem is that a poorly designed socket may prove unreliable over extended time periods. A nonsoldered joint must have a gas-tight seal, such as that created by a mechanical metal-to-metal wiping action, with the seal then being left undisturbed. PC edge connectors, for example, used to be somewhat unreliable; with time, manufacturers learned some good tricks: bifurcated contacts (two independently sprung contacts for each finger), gold plating on the socket and on the edge fingers, and good mechanical design to ensure firm contact pressure during wiping and afterward. Joints that aren't gas-tight can be expected to fail after some time, perhaps a year or so. This sometimes happens inadvertently, e.g., by inserting a component in a PC board and then forgetting to solder it. Such connections have the maddening property of working fine at first, then becoming intermittent months or years later, owing to the formation of corrosion. A different problem can arise when heavy ICs (24 pins or more) are held in sockets. They can work their way out of the sockets after repeated vibration or shock.

We have found that the pin-and-jack type of IC socket (popularized by the Augat 5xx-AG series), although expensive compared with many other socket types, gives good reliability.

### Soldering and defluxing

The usual procedure is to insert some components, turn the board over and bend the leads aside to hold the components

in place, then solder them using a thermostated soldering iron and fine solder. ICs can be inserted easily with an insertion tool (highly recommended), and it is best to use a lead bender on resistor leads, etc., in order to prevent slivers of solder being shaved off the leads during insertion. We recommend the adjustable “flip-over” board holders (marketed by OK Industries), which hold the components firmly in place with foam rubber during soldering. After soldering, the leads should be trimmed with a snipper.

Now comes a very important step: Solder flux should be removed from the board. If it isn’t, the board will look just terrible in a few years, when you’re not around to defend it! Rules for defluxing:

1. Do it.
2. Do it soon. The stuff gets much harder to remove with time.
3. Use a solvent such as Freon, alcohol, or some other organic solvent recommended for this purpose. Use a small brush to help dislodge stubborn globs of flux.

Commercial board manufacturers clean their boards with *vapor-phase degreasers*, in which hot vapor from a bath of boiling solvent condenses on the board (suspended above), dissolves the flux, and drips back into the bath. This method works well, especially because the distillation process continually bathes the board with clean (and hot) solvent. Because organic solvents aren’t exactly good for your health, people have experimented with water-based cleaning processes. One method uses a water-based “saponifier” to emulsify the rosin so it will wash away; alternatively, you can use a water-soluble solder flux (rather than the usual flux, which is made from insoluble tree rosin). As nice as they sound, these water-based cleaning methods turn out to be tricky to do correctly; if you’re not careful you can leave corrosive residue that ruins the board, in time. For the do-it-yourselfer we recommend organic

solvent cleaning; take care not to breathe the stuff or get it on your hands.

When defluxing boards, keep in mind that mere removal of the rosin residues is not enough – you have to get rid of the ionic “activators” also. Incomplete defluxing may do more harm than good, since it can release the activators from the area of the solder joints and spread them all over the board. Note also that solder flux comes in several grades of aggressiveness. We routinely use “RA” (rosin, activated), which is the most active flux for electronic soldering, because it produces good solder joints even when surface oxidation is present; RA flux residues should be thoroughly removed after soldering. The alternative, “RMA” flux (rosin, mildly activated), is less active and is often specified for government contracts or for applications where defluxing is not possible.

### 12.07 Some further thoughts on PC boards

The solderability of PC boards tends to decrease with time, owing to oxide formation, so it is best to stuff the components soon after the board is made. For the same reason, you should keep unstuffed boards in plastic bags, away from corrosive fumes. Good circuit boards should be made from 1/16 inch FR-4-type board material (sometimes referred to as “epoxy fiberglass”) clad with 2 ounce copper. Remember that a circuit on a PC board is basically sitting on a piece of glued-together stuff; the board can absorb moisture and develop electrical leakage. Another pathology of PC board material is “hook,” the variation of dielectric constant with frequency; the consequent variation of stray capacitance can make it impossible to build an amplifier with flat frequency response, for instance. Oscilloscope manufacturers are very aware of this bizarre effect.

PC runs with large currents passing through them have to be widened to prevent excessive heating and voltage drops. As a rough guide, here is a table of approximate conductor widths that give temperature rises of 10°C or 30°C for the currents listed, for 2 ounce copper PC boards. For other foil thicknesses, just scale the widths accordingly.

|           |        |        |        |
|-----------|--------|--------|--------|
|           | 0.5A   | 1.0A   | 2.0A   |
| 10°C rise | 0.004" | 0.008" | 0.020" |
| 30°C rise | 0.002" | 0.004" | 0.010" |
|           | 5.0A   | 10.0A  | 20.0A  |
| 10°C rise | 0.070" | 0.170" | 0.425" |
| 30°C rise | 0.030" | 0.080" | 0.200" |

PC runs with high voltage need correspondingly wider spacing – a good rule is 5 volts per mil (0.001"). It is also a good idea to avoid sharp corners and points, in favor of rounded contours.

### Tools

As a starting point, we lined up the most heavily used tools on our bench and came up with the following part numbers:

|                   |                                                                  |
|-------------------|------------------------------------------------------------------|
| Long-nose pliers  | Erem 11d, Utica 321-4 $\frac{1}{2}$ ,<br>C.K 3772H, Xcelite 72CG |
| Snippers          | Erem 90E, C.K 3786HF                                             |
| Soldering iron    | Weller WCTP-N, Ungar<br>"Ungarmatic"                             |
| Solder            | Ersin Multicore 22ga,<br>Sn63 alloy, RA flux                     |
| IC inserter       | Solder Removal 880                                               |
| Lead bender       | Production Devices PD801                                         |
| Solvent dispenser | Menda 613                                                        |
| Solder sucker     | Edsyn Soldapullt DS017                                           |

For the most effective desoldering of PC boards, it's worth investing in a controlled-vacuum desoldering station. These work well even on solder-filled plated-through holes, with less tendency to damage the delicate foil pads (when compared with the simple spring-powered plunger types). These desoldering stations do have a tendency to clog with solder, however. They are manufactured by several companies,

including Edsyn, OK, Pace, Ungar, and Weller. Lots of useful gimmicks for PC assembly are listed in the Contact East catalog (335 Willow Street South, N. Andover, MA 01845) and the Marshall Claude Michael catalog (9674 Telstar Avenue, El Monte, CA 91731).

### 12.08 Advanced techniques

Hand-drawn schematics, manually converted first to PC board layouts and then to hand-taped Mylars, are the traditional steps in PC board design technique, and they were used nearly universally through the mid-1970s. It still makes sense to do simple boards this way, particularly if you aren't trying to break new records in board component density. With a little kit of double-sized patterns, some opaque tape, a gridded table, and some clear Mylar, you're ready to go. And you don't have to buy expensive CAD software (and learn how to use it!), pay photoplottng charges, etc.

However, as soon as you strive for high-density boards, paved wall to wall with fifty or a hundred IC's, often requiring four or six layers laid out to 10-10 design rules, the honeymoon is over. Even with extraordinary effort and quadruple-sized Mylar, you can hardly achieve the precision interlayer alignment required. Furthermore, you'll need a month's vacation to recover from the concentration required to work out the routing. The first board you get back from the PC house is unlikely to be error-free, and any serious changes to the taped Mylars are sheer hell, often requiring massive areas of wiring to be pulled up and repositioned (which often introduces new errors). We aren't kidding.

### CAD/CAM

The solution is CAD/CAM (computer-aided design/computer-aided manufacture). These software packages are powerful, and with the advent of desktop



workstations with multimegabyte memory, dazzling graphics, and processing speeds of tens of MIPS (million instructions per second), you don't need special hardware to run them. Some popular PC CAD systems come from Valid Logic, Mentor Graphics, and Daisy Systems. They aren't cheap, but at least they're not getting any more expensive. Here's a quick tour of board design using CAD/CAM.

- *Schematic capture.* You begin by "capturing" your schematic directly into the graphics workstation. You can draw and edit diagrams, using standard electronic symbols from a "library." Just as with word processing, you can pull up old work, extract sections you want to reuse (e.g., a DRAM memory control circuit, an active filter, etc.). With a mouse (or joystick, trackball, or graphic tablet) you can move things around, with the wiring following you around (not always totally successfully!). You name signals, give the ICs part numbers, and so on. Good CAD systems assist you by looking up the ICs and assigning pin numbers. And good CAD systems are *hierarchical*: You can, for example, do a top-level diagram in which the major subcircuits are indicated by big boxes; each box opens up to reveal its sub-boxes, eventually displaying the lowest (gate-level) description. At any stage of the design you can get handsome laser-printed schematics. See Appendix E for further thoughts on schematic diagrams.

- *Checking and simulation.* The output from the schematic capture consists of a set of drawings (Fig. 12.8) and a *netlist*, which is simply a list of every signal, telling every "node" (component pin) it is connected to.

At this stage in the design you should, of course, spend plenty of time with the schematics, making sure they're what you want. A good CAD system can help you out, by flagging obvious errors, for example an output tied to ground, or to another

output. You can ask the system to simulate digital circuit operation, but you must provide it with test "vectors" that describe machine states, and you must describe the function of each chip in your circuit that is not already described in the library. Since a complex chip like a microprocessor requires lots of description, some CAD systems have a socket into which you can plug such a chip; it then uses the chip itself for hardware modeling.

At this point, too, you can ask the CAD system to disgorge voluminous documentation showing (a) for each named signal, all the pins ("nodes") it visits and (b) for each pin of each component, all other nodes to which it is connected. Good systems will even inform you if you have violated loading rules for digital outputs, etc.

- *Placement and routing.* The next step is component placement and routing. One of the great beauties of an integrated CAD/CAM system is that it automatically extracts the netlist from its representation of the schematic – if the schematic is correct, the final PC board will be too. Although some CAD systems claim to be able to juggle the component placements into a near-ideal configuration, it's usually best to do component placement manually. As with schematic capture, this is done with mouse, joystick, or graphic tablet. You begin by setting up the board outline (it's often the same as some previous board you have designed), within which you now position the parts. Since the component library has outline and pin information, you get to play with the characteristic IC and component shapes. Good CAD systems can flag design-rule errors as you make them, so you can see if the parts are too close together, etc.

Once the parts have been placed, the wiring must be routed. At this point it is conventional to display what's called the "rat's nest," a display of the board with all the connections shown as straight lines

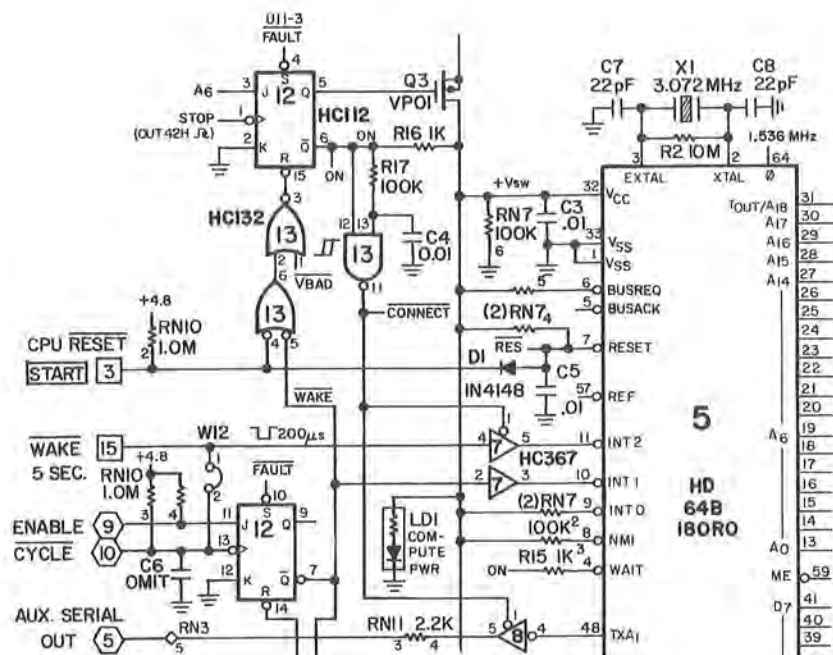
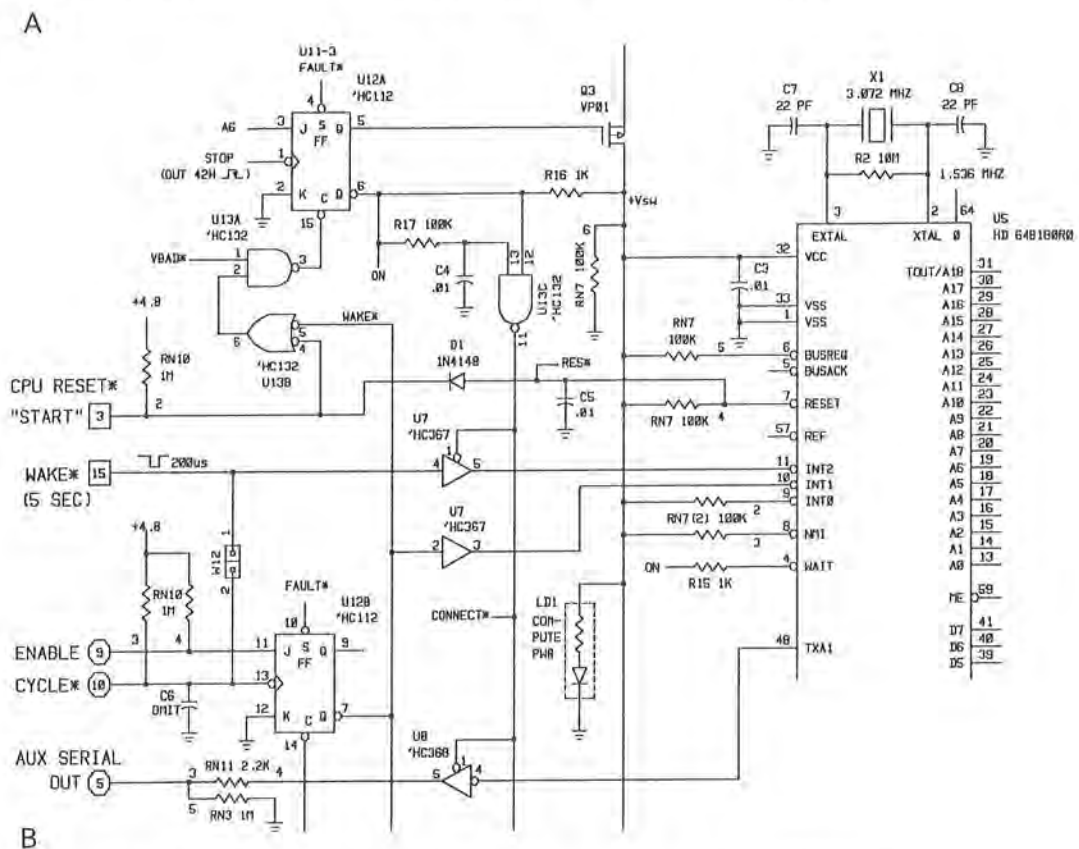


Figure 12.8. Portion of a complex schematic diagram from which the PC board in Figure 12.11 was produced.  
A. Manually inked on a drafting table.  
B. Machine-drawn, from Case Technologies schematic capture.

connecting their respective pins. It's a tangled mess. With creative use of color you can make some sense of it, for example by selectively turning on only power and ground connections.

Early CAD systems made you work hard at this point. You were forced to route every connection by hand, with just a little help from the computer. Current systems provide *autorouting*, which means what you hope it does: The software finds connection paths, taking care not to violate the design rules that specify not only line widths and clearances but also parameters such as the maximum number of "vias" (plated-through holes used only for a signal trace, not a component lead). The best CAD machines can do 100% autorouting, although the result often lacks the full elegance of hand-routed lines. For example, it may put a via daringly close to a component pad; although its placement meets all the rules, you can create a board that's more easily manufactured and soldered by shifting the via away from the pad. You usually spend a few hours cleaning up the machine's work in this way, pushing and pulling traces around a bit (Fig. 12.9).

In complex digital circuits you can often simplify the routing considerably by reassigning gates or sections within chips, or even swapping gates between IC packages. As an example, you might have a byte-wide data bus connected to an octal *D* register. Your initial assignment of bus bits to corresponding register bits might force each line to leapfrog the next, whereas by reversing the order of bits in the register the resulting routing requires no crossings. Good CAD systems can do this automatically (the library specification of each chip identifies duplicated functions), even updating ("back-annotating") the schematic via the modified netlist.

The final step is to check the proposed routing for any rule violations, and also for exact equivalence to the schematic. The

latter step is important because you may have made manual changes to the routing netlist that altered the circuit. At this point all CAD systems will produce paper pen-plots of the routing, showing any designated combination of components, wiring, labels, etc., in a dizzying array of colors.

- *Photoplot and drill tape.* If all has gone well, the last stage of the process consists of producing a machine-readable board description suitable for the PC house. For this you need two things: a tape that tells a photoplotter how to draw a precision plot of the "artwork" (the individual foil patterns for each layer, a pattern for the "solder mask," and a silkscreen pattern for printed legends) and a "drill tape," which tells each hole size and its precise position. The artwork tape is usually written in "Gerber format," named after a brand of photoplotter that exposes film by moving a large easel under a stationary projector in response to commands written on tape. (A newer form of photoplotter uses laser scanning, producing large plots in minutes rather than the hours required by the early Gerber machines.) Some PC board fabricators want you to supply the finished photoplots (Fig. 12.10), whereas others ask for the Gerber tape itself. A surprising number of board houses ignore the drill tape, preferring to figure out where the holes go by manually digitizing the photoplotted artwork (if you ask, they'll tell you that customer-supplied drill tapes sometimes have crazy commands that break their drill bits). Drill tapes, believe it or not, are not magnetic tapes; they are *paper* tapes!

*Board production.* In this world you don't manufacture your own PC boards; you go to a PC board house. They're all over the place, and they just love to make boards (Fig. 12.11), in exchange for money of course. Some specialize in

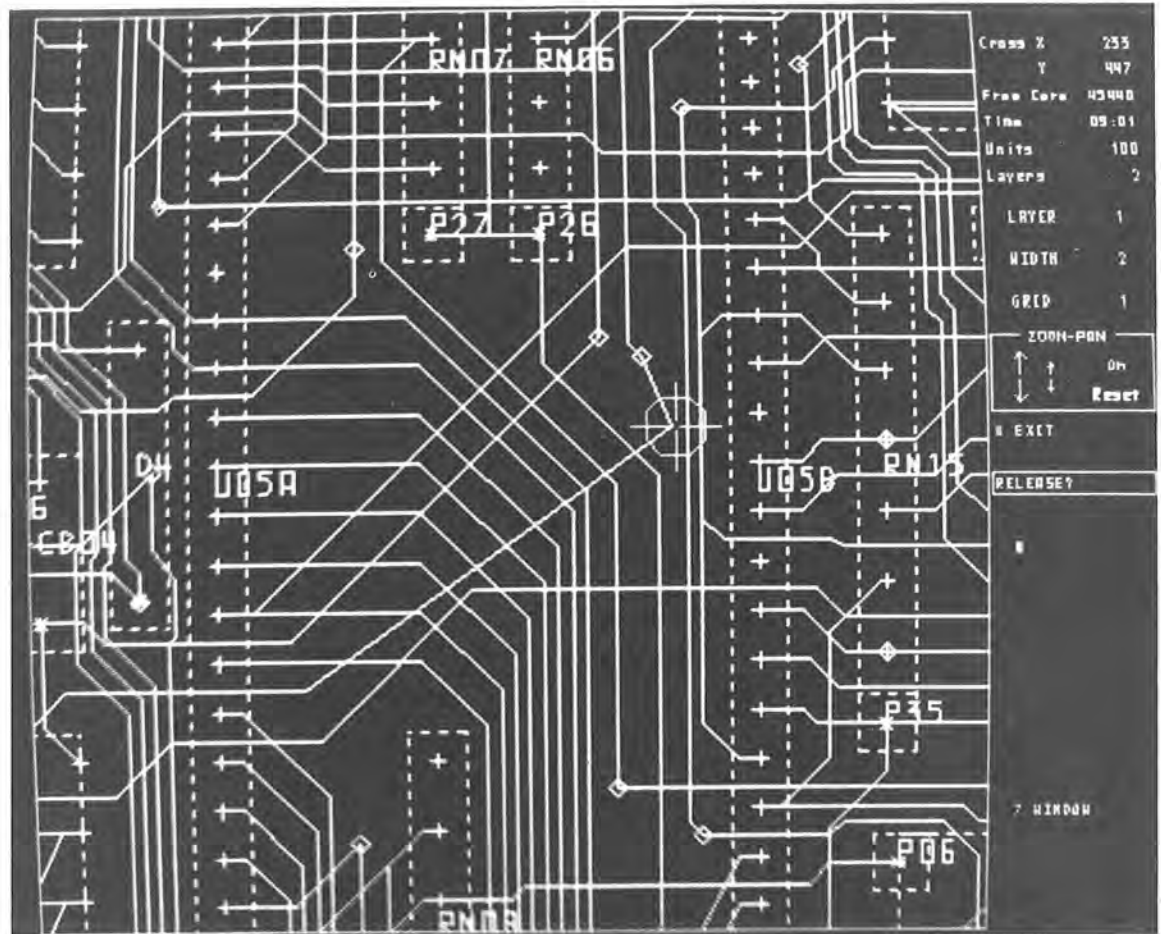


Figure 12.9. Photograph of PC board design system screen display (Racal-Redac), showing manual “cleanup” of trace routing. The cursor, steered via a “graphics tablet” stylus, shows the circuit trace that is being modified. This display is in color, with different colors and intensities designating the different layers, component outlines, legends, etc.

small-quantity, fast-turnaround prototyping, whereas others may do only production quantities. Your first task is to find out which companies do a good job. This is harder than it sounds. You can’t tell much on the telephone, because it is possible to *know* how to make good boards, but still make bad ones. Here are some tricks we’ve tried: (a) Look at the little logos on good boards you’ve seen (e.g., boards you find in instruments or computers sold by

major companies, who, like you, usually buy their boards from external board houses) – companies like IBM and Apple are extremely careful about their board suppliers, (b) Ask everyone you know whom they use to make their boards, (c) Go and visit the board house; while there, look for cleanliness, high morale, native intelligence, diligence, and pride in workmanship; you might also ask to see some boards, (d) Finally, if you know



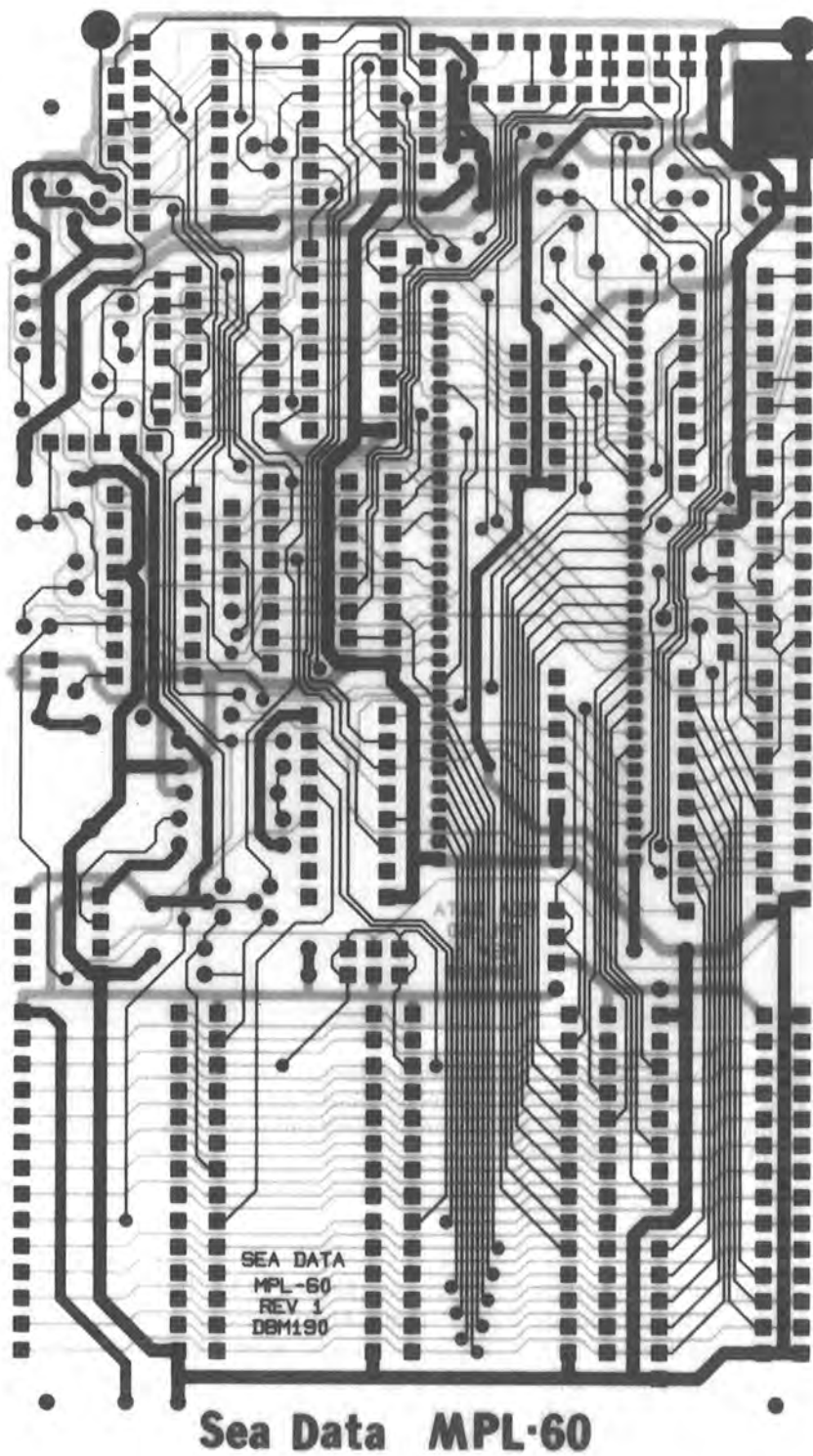


Figure 12.10. Photoplot for the double-sided board shown in Figure 12.11. Both layers are overlaid in this photograph, showing the view from the solder side.

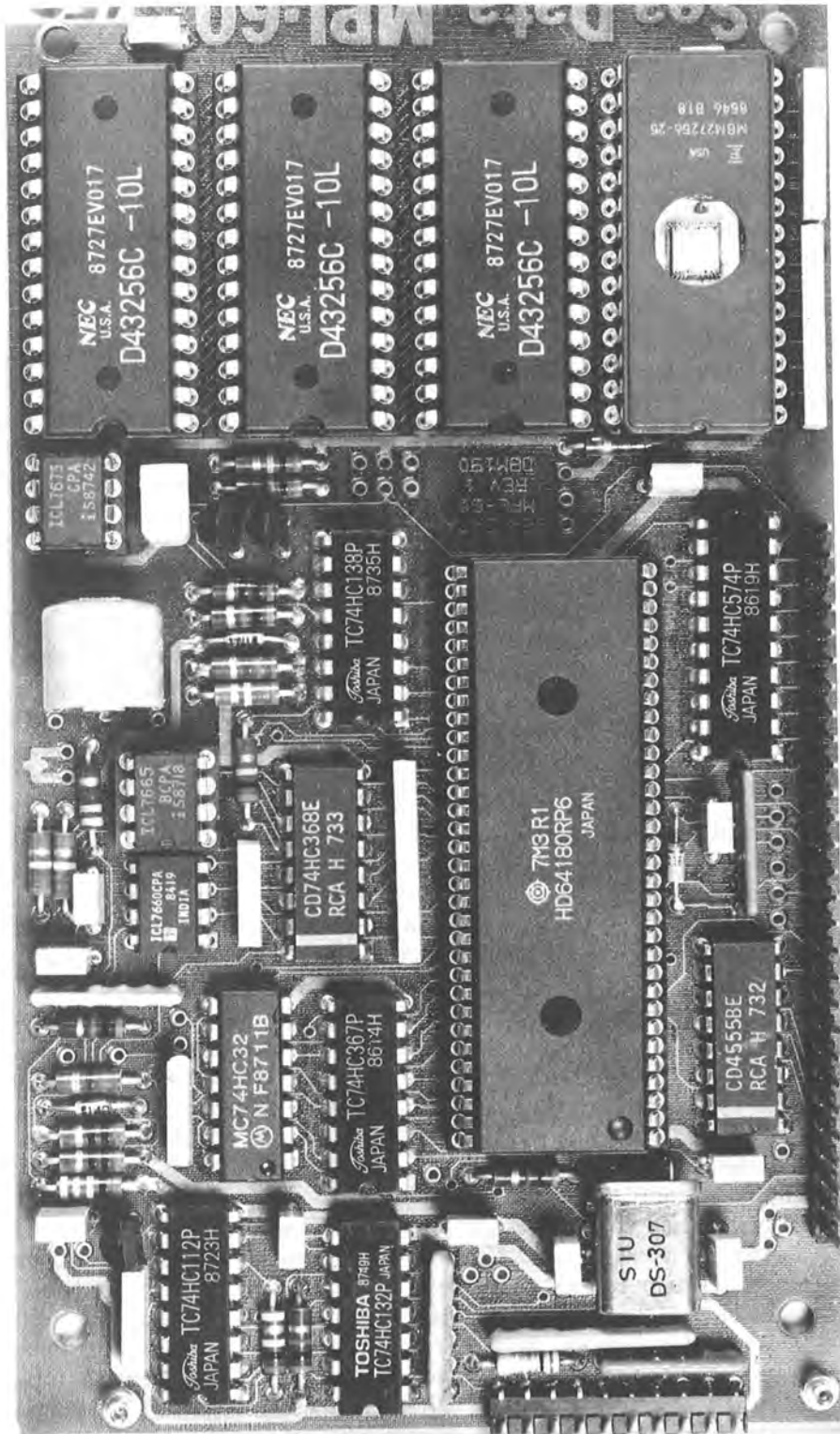


Figure 12.11. Finished microprocessor board, designed to double-sided 12-12 rules. Higher component density can be achieved only by going to tighter design rules, or, preferably, a multilayer board.

someone who works at a bare-board testing house, cajole him into telling you who makes good boards: Those guys really know, but they're not supposed to tell!

Next you have to get a quote for price and delivery. Be prepared to answer the following set of questions, which every board house will ask you:

1. board dimensions?
2. number of layers (sides)?
3. design rules (minimum conductor width and spacing)?
4. any gold-plated edge-connectors?
5. any fancy board cutouts or shapes?
6. solder mask? wet or dry? to cover the vias ("tented") or not?
7. board and copper thickness?
8. number of holes?
9. number of hole sizes?
10. silkscreen 2 sides, 1 side, or none?
11. any unusual options, e.g. "dry mask over bare copper"?

Having agreed on terms, you must now supply the following things:

1. actual-size layer positive transparencies (i.e., opaque = copper)
2. soldermask patterns(s)
3. silkscreen pattern(s)
4. drill drawing (hole pattern, coded by sizes)
5. mechanical drawing (precise outline, aligned to specific holes)
6. drill tape (a paper tape; often optional)
7. additional specifications and notes

The last item is important. You should strictly specify the minimum trace width (an over-etched board has thin traces), the minimum conductor spacing (under-etched board), the minimum annular ring (caused by misaligned holes or misaligned layers), the tolerances for finished hole diameters, minimum thickness of copper (and solder) plating, the maximum number of repairs permitted per board, and so on. The industry generally adheres to

some typical numbers (e.g., the IPC-600 standards), but it's best to have it clearly in writing, in case you get some boards you're unhappy with. Furthermore, your board may have unusual requirements; for example, if you've used unusually small pads for your vias, your minimum annular ring specification for vias will be less forgiving than usual.

*Testing.* For any new PC board design, *always get a few prototypes made first.* Stuff them with sockets, and test the dickens out of the circuit. You can expect several kinds of bugs, namely (a) your circuit had a design error, (b) the circuit was OK, but you made an error in schematic capture (which you should have caught earlier), (c) an error crept in during routing (which also should have been caught earlier), (d) the board has flaws, usually shorts or opens, due to blemishes, misalignments, or other problems during manufacture (see below), or, finally, (e) there are some bad components or solder joints.

Assuming circuit changes are needed during debugging, the usual procedure is to cut existing foil traces with a small knife, then solder in wires to make the new connections. This may be impossible in a multilayer board with buried signal layers, however. In that case you have to make your changes at the final destinations – the IC or component pins. A good way is to stack up two IC sockets, with a pin removed from the upper one.

During circuit testing, mark up your schematics with circuit fixes, for later entry into the schematic capture (or hand-drawn circuit). The second pass through the CAD system is usually easy and fast; but be *very* careful, because you probably will go directly to a production run, and any errors remaining in the production board will have to be fixed on each unit. If that becomes necessary, you use the same procedures of cutting and jumpering as above, taking care to secure the wiring

(a hot-wax gun seems to work well). These extra wire patches are sometimes called “roach wires”; the official term is “ECO,” for “engineering change order.”

There are plenty of things that can go wrong in board production (see the fine handbook by Coombs for gory details). With high-density boards (Fig. 12.12) even a minor blemish can create a short or open circuit somewhere. And plated-through holes have all sorts of pathologies of their own that can lead to open circuits. It’s an axiom in electronics that the cost to repair a flaw multiplies at each stage of production – a bad board connection is easily fixed at the board house, fixed with greater investment of time during circuit checkout, and repaired with considerable expense once shipped into the field. So it makes good sense to have the board house do a “bare-board” test on all your production boards. This they do on a curious little setup that jabs the board with a set of pins that align with the pads (they fabricate the “bed of nails” from your drill tape, of course). Oddly enough, none of the testing houses compares the bare board with your *netlist*; instead, they just test a bunch of boards, and assume that if a number of boards are identical, they are also correct. Although sharp-eyed readers will see a flaw in the logic, they can rest assured that the system usually works anyway.

- **CAD-based Wire-Wrap.** The netlist output from the schematic-capture phase contains all that is necessary to construct the circuit. Many CAD systems offer output formats compatible with automated Wire-Wrap fabrication. You begin by choosing component placements on a standard Wire-Wrap panel (see the Augat catalog for numerous specimens), after which the CAD system constructs a special form of netlist called a “From-To” list. Each entry tells the coordinates of the two pins connected by a wire, and the wrap level

(height above board) of each end’s wrap. If you send this tape to a wire-wrapping house such as DataCon, they will produce a machine-wrapped board. It typically takes a few weeks and costs about 6 cents per wire plus the price of the Wire-Wrap panel (about \$2 per IC position). As with PC boards, there can be flaws. Our experience is that you might find one or two problems per board, usually in the form of a broken wire that looks OK from the outside.

A more serious question about Wire-Wrap prototypes is “Why bother?” If your objective is to make PC boards, you’re going to have to do a PC board routing eventually, so why not do the work for the prototype? Furthermore, a wire-wrapped board behaves differently than a PC board, owing to lead inductance and differences in ground-pin connection length and ground-plane inductance. The effects are most noticeable if you have fast logic such as 74F, 74AS, or 74AC(T), or memory arrays with widely bused signals. In general, a good multilayer PC board performs far better than a comparable wire-wrapped board, so you do yourself a favor by building the prototype with a PC board.

Of course, if you are building only a handful of units, you may want to avoid the one-time (“nonrecurring”) expenses associated with PC board production and just build wrapped boards as the finished product. In that case the machine-wrapped boards are a rational choice. Don’t overlook the alternative, however, which is the following.

- **CAD-based Multiwire.** “Multiwire” is a trademark of the Kollmorgen Corporation, which manufactures a clever hybrid of custom-routed PC board. The basic board has pads in standard or custom groups to hold your ICs and to provide power and ground. The signal interconnections are made not with PC traces, however, but with fine insulated wire (#34 or #38) that





is machine-placed onto a sticky coating on the board. The wire ends are attached to the IC pads with a hole plating process.

Multiwire has good signal properties because of its complete groundplane, and it lets you make circuit changes in subsequent production runs relatively painlessly. Since the (insulated) wires can cross over each other, you can achieve very high component density, comparable with that of a 10-layer PC board. Kollmorgen claims that Multiwire is thus a cheaper solution than custom multilayer PC boards, particularly in small production runs and in situations where your circuit is evolving.

□ *CAD-base ASIC.* We finish our CAD narrative by pointing out that you would not dream of designing custom or semicustom (gate-array) ICs by hand (Fig. 12.13). These belong in the general class of “application-specific integrated circuits,” or ASICs. Once again, CAD comes to the rescue, with capture and simulation software that gives you some assurance that your chip has a chance of working. ASICs are rational in the context of substantial production runs, and we confidently assume that the situation will only improve with time. By the time of the next revision of this book it may have become



Figure 12.13. At the *chip*-design level, CAD/CAM isn't a luxury – it's a necessity. (Courtesy of Performance Semiconductor Corporation, Sunnyvale, California.)

commonplace to design custom ASICs, promptly delivered by your local silicon foundry, for just the sort of circuits for which you now design a PC board filled with standard ICs. We see no fundamental reason to prevent the ASIC solution from becoming the less expensive (and better performing) alternative.

## INSTRUMENT CONSTRUCTION

### 12.09 Housing circuit boards in an instrument

Circuit cards, whether printed circuits, Wire-Wrap panes, or breadboarding cards, have to be mounted in some sort of enclosure and connected to power supplies, panel controls and connectors, and other circuitry. In this section we will discuss some of the popular methods of putting instruments together so that circuits are neatly mounted and accessible for testing and repair. We will begin with methods for holding the circuit cards themselves, then discuss the business of cabinets, front and rear controls, power-supply mounting, etc.

#### *Circuit card mounting*

In simple instruments you may have only a single circuit card, whether printed circuit, Wire-Wrap, breadboard card, or whatever. In that case a simple solution is to drill holes near the corners and mount the card with screws (and standoff bushings) to a flat surface, component side up. Connections can then be made with a card-edge connector socket (if the card has plated fingers), with flat cable terminated with a connector to mate with a plug on the board, or with individual soldered connections to swaged terminals on the board. With edge or ribbon connectors, the card will support the connector adequately, so no extra connector supports have to be used. Whatever

the method of connection, it is wise to arrange the wiring such that the board can be tipped upward for access to the underside so that you can make modifications or repairs.

In a system with several circuit boards, the best way to arrange things is with some sort of card "cage," a rigid assembly with guides for individual cards to slide into and aligned holes along the rear so that you can mount edge connectors to mate with the cards. There's lots of flexibility in card width, spacing, and number of cards that can fit into a card cage. A very common size accommodates cards 4.50 inches wide with 44-pin dual-readout (22 pins each side) edge connections on 0.156 inch centers. There are plenty of other formats, often with tighter pin spacing (0.1 inch is common), multiple groups of edge connectors on a single card, or the more reliable "2-part" connectors (in which one mating portion solders to the board as a component); the most popular example of the latter is the so-called VME connector, which has 64 or 96 pins. Cards can be spaced as little as 0.5 inch apart, if necessary, although 0.6 inch is a more comfortable spacing; if space is no problem, 0.75 inch spacing will allow plenty of room, even with Wire-Wrap pins and bulky components. It is best to look at some catalogs to see what's available. You can get variations with plastic card guides or just dimples in the metal sides to align the cards, and there are various types of card ejectors (attached to the PC cards) to help remove a card.

Card cages are available with simple flange mounting to a flat surface parallel to the cards, as well as in various configurations that fit nicely into rack enclosures, etc. You can even get modular enclosures that include an integral card cage, with some additional room for power supplies, panel controls, etc.

Warning: Instruments built with the circuit spread over a number of plug-in

circuit boards achieve a nice modularity and ease of repair. But this method of construction can cause difficulty in circuits with low-level signals (less than a millivolt), and in circuits with high-frequency signals (above a few megahertz). The basic problem is that you often cannot provide a sufficiently stable and low-inductance grounding system to the set of boards by way of the connectors at the ends of the boards. Combining low-level analog circuitry with digital switching waveforms is particularly dangerous. This problem is worst with a hand-wired backplane, in which the ground distribution depends on a few wires running between the connectors. Typical symptoms are low-level 60Hz or 120Hz pickup, miscellaneous “fuzz” at the level of fractions of a millivolt, and coupling of radiofrequency signals between circuits that are supposed to be isolated. If the boards have a bare groundplane and are plugged into a metal card cage, the symptoms often change as you press on the cards, altering the uncontrolled ground contact.

We’ve suffered with such problems on more than one occasion, and we offer the following advice. First, it may be best to avoid interconnected small boards entirely, building all the critical circuitry on one larger board with integral groundplane. On that board you may need to use coaxial lines or twisted-pairs to make connections between separated portions of the circuit. Second, if you do use interconnected boards, you will achieve better ground distribution with a PC motherboard provided with wide ground traces, rather than a hand-wired backplane. In radiofrequency systems you sometimes see springy metallic fingers used along the card guides, to provide a continuous and stable ground connection. Third, the use of coaxial line or twisted-pairs, combined with differential inputs (or “pseudo-differential” ground-sensing inputs, see Fig. 7.70), is often the best way to handle microvolt

signals, which are otherwise prone to ground-loop noise and pickup. Finally, except for the ground-sensing method just described, we suggest as many redundant ground connections as possible – multiple connections to the chassis, doubled connector pins and wires, etc. – to reduce the inductance through which ground currents must flow. Don’t worry about “ground loops” in a digital or RF circuit; that’s a microvolt audio circuit issue. Refer to Section 7.24 for additional ground philosophy.

### **Backplane connections**

Card-edge connectors are available with lugs for solder connections, with Wire-Wrap pins, and with small pins for insertion into PC boards. In many cases it is best to wire up the intercard connections with point-to-point wiring between card-edge connector pins, using the edge connectors with solder lugs. A neat job requires some cabling of wire bundles, with the wires running in straight lines parallel to the card cage dimensions. In other cases it may be preferable to use Wire-Wrap connections on this backplane, especially if there are many connections between backplane pins with relatively few connections to other points in the instrument, and if there is no need for shielded-cable connections to the backplane.

The third possibility is to use a motherboard backplane, a PC board designed just to hold the card-edge sockets. Motherboards are popular in bused systems (they are nearly universal in computers) and should be considered in any case if the instrument is intended for production in significant quantities. With double-sided motherboards, you can have the advantage of a groundplane (lower inductance and coupling of signal lines), or you can use both sides for signals if the intercard wiring is complicated. With bused systems the backplane is usually simple, with lines



connecting corresponding bus pins on all cards. In computer backplanes you sometimes see a motherboard used with Wire-Wrap pins sticking through. This is very handy if you want the motherboard to do all the bus and power-supply connections, leaving the unused pins to be connected in a custom configuration by wire-wrapped connections. Figure 12.14 shows a simple PC motherboard.

### 12.10 Cabinets

Depending on the intended use, an electronic instrument might be housed in a benchtop cabinet (complete with rubber feet and hinged front “bail”), in a cabinet or panel designed for mounting in a standard “relay rack” 19 inches wide (either screwed directly to the rack flanges that run vertically up from the floor or mounted

on ball-bearing rack slides for simplified access), in a modular instrument case designed to plug into slots in a larger rack-mounted “bin,” “cage,” or “crate” (the latter usually provide dc power connections through standardized connectors at the rear), or perhaps in some other format such as a free-standing pedestal-mounted case.

There are many cabinet configurations available in both benchtop and rack-mounting formats. Among the most popular are cabinets 17 inches wide, available in various heights (always multiples of 1.75 inch) and depths, that accept optional rack-mounting flanges or slides (a rack 19 inches wide has about 17 1/2 inches of clearance between the flanges). That way you can convert an instrument from rack-mounted to benchtop format, or vice versa, by just changing a bit of cabinet hardware.

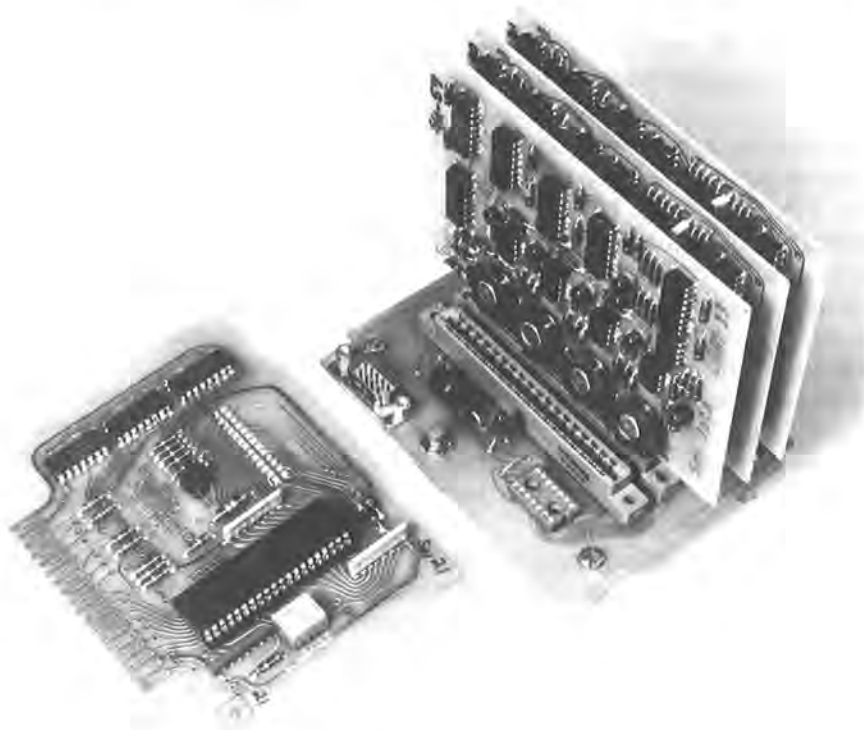


Figure 12.14. A “motherboard” provides a convenient method of interboard connection, reducing hand wiring and the possibilities for error considerably, while simultaneously providing superior electrical performance. In large systems, the motherboard and its connectors would probably be mounted rigidly at the rear of a card cage.

A point to check: Some of these convertible cabinets require removal of the outer case for rack mounting, whereas others let you keep the cabinet intact.

In the category of modular instrumentation, the NIM bin is popular in nuclear and atomic measurements, the CAMAC crate has established itself in some computer interfacing, and several manufacturers have defined modules and bins (e.g., the TM500 series from Tektronix and the EFP series of blank modules from Vector). Blank chassis are available in each of these formats, complete with rear connectors to mate with the dc power receptacle of the mother bin.

### 12.11 Construction hints

Rather than attempt to list the enormous variety of manufactured cabinets by name or style, we will simply offer some general comments on construction of instruments. These suggestions, together with the figures in this chapter, should help you choose electronics enclosures wisely and fill them up with circuitry in a sensible way.

In general, you use the front panel for indicators, meters, displays, keypads, etc., as well as controls and frequently used connectors. It is common to put seldom-used adjustments and connectors that don't require frequent access on the rear panel, along with large connectors, line cord, fuses, etc. (Fig. 12.15).

The traditional way to lay out a professional-looking front panel is by silkscreening the legends onto the painted or anodized aluminum surface. Although the result doesn't look bad, it tends to erode with frequent rough handling (although a clear overcoat can help considerably). If you look at a recent Fluke, Tektronix, or Hewlett-Packard instrument, you'll see the modern alternative – an adhesive Lexan panel of pleasing appearance and unbelievable toughness. These are made by silkscreening the legends on the *back* side of a matte-textured 0.010 inch Lexan film,

to which a tough adhesive is then applied. You can apply lettering and patterns in several colors, and you can specify colored plastic windows or cutouts. You can get these custom laminated panel coverings in quantity from many labeling services; you just provide them with the "artwork" (usually an actual-size positive or negative). You might prepare the artwork with dry-transfer rub-on lettering or, better yet, use some desktop publishing software and a laser printer.

Perhaps the most important thing to remember when laying out an instrument is the need for good accessibility to circuit cards and controls. It should be possible to replace any component in the instrument without great pain. This means neat cabling of wiring, so that modular units can be raised up without having to unsolder anything, and careful planning, so that circuit cards can be tested while operating in the instrument. For instance, a card cage might be mounted with the cards vertical; to get at them, you remove the top panel from the cabinet, then plug in extender cards to make the circuit cards accessible. If the cards are mounted horizontally, you might make the front panel removable, or hinged, to provide access. At all costs, fight the temptation to lay down the circuit in "layers," with circuitry nicely covering other circuitry. Figure 12.16 shows an example of neat and accessible front-panel cabling built in a cabinet with a removable front panel.

### 12.12 Cooling

Instruments that consume more than a few watts will usually require some sort of forced air cooling. As a rule of thumb, a small instrument running more than 10 watts, or a larger (rack-width) instrument consuming more than about 25 watts, will probably benefit from a blower. It is important to keep in mind that a box full of electronics may run at

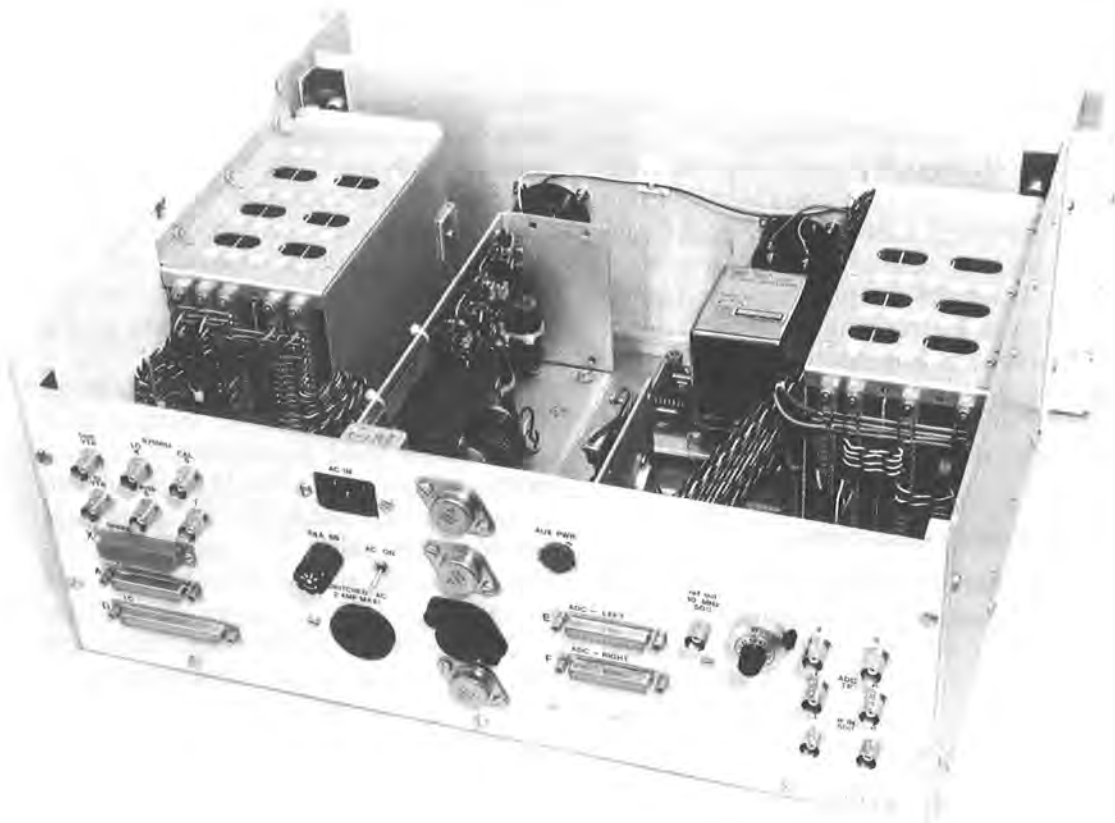


Figure 12.15. In this instrument we used the rear panel for heat-sinking power components, and for seldom-changed controls and connectors. The front panel detaches for access to the nine circuit boards housed in card cages.

a nice temperature when sitting on the bench with the top cover removed, but when installed in a rack with other heat-producing equipment (where the ambient temperature may reach 50°C), complete with its outer cover, it is likely to run very hot, leading to early failure of components and generally unsatisfactory operation.

Instruments running at moderately low power, say at the figures mentioned earlier, can often be cooled adequately with simple convective cooling. In such cases you might perforate the top and bottom covers, paying attention to the location of major heat-producing components (power resistors and transistors). It may be best to mount the high-power components on the

rear panel, using heat sinks with their fins aligned vertically (see Section 6.04). Circuit boards will also be better ventilated if mounted vertically, although heat dissipation in circuit cards is often negligible. If simple convective cooling doesn't keep things cool enough, you have to resort to a blower.

The simple "Muffin-type" venturi instrument blower, with flow rate of about 100 cubic feet per minute (CFM) assuming relatively unimpeded airflow, will adequately cool instruments running 100 watts or more. Here's the relevant formula:

$$\text{air temperature rise} (^{\circ}\text{C}) = \frac{1.6 \times P(\text{watts})}{\text{airflow}(\text{CFM})}$$

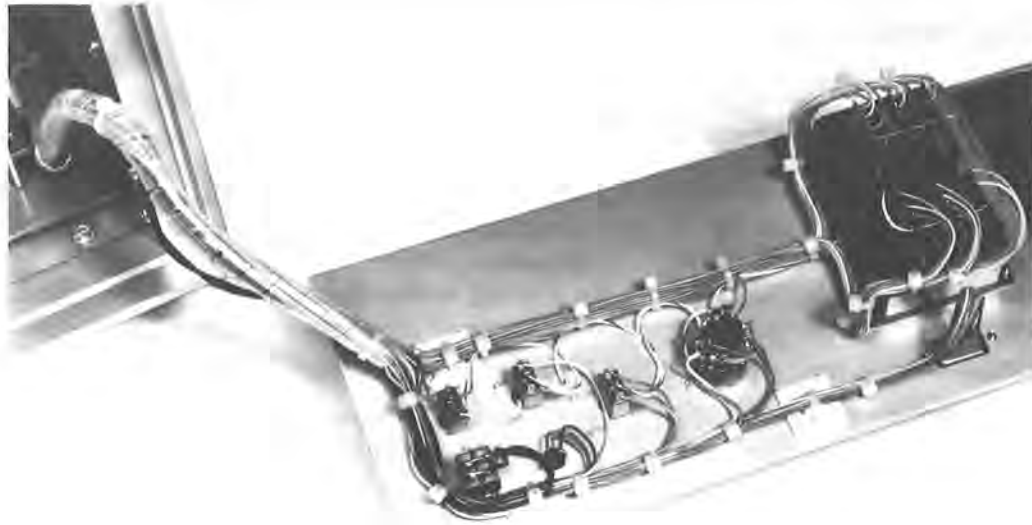


Figure 12.16. One way to ensure good accessibility to panel controls is to bring all wiring away at one end, so the panel can be hinged or otherwise detached from the instrument. In this example the panel slides into a slotted instrument case. Note the use of cable “ties” and self-adhesive supports to keep the wiring tidy.

If less airflow will suffice, a quieter version of the venturi fan is available from most manufacturers. Table 12.2 shows some part numbers. The airflow of these fans is greatly reduced when operating against high back-pressure. Figure 12.17 shows a graph.

All the blower manufacturers now sell brushless dc fans, in addition to the traditional line-voltage-operated ac fans. Running the fan from low-voltage dc (usually 12V or 24V) makes it easy to control the fan speed electronically, in response to the actual temperature inside the instrument. You can either do the temperature sensing and speed-control feedback yourself (perhaps using an outboard module like the “SmartFan” controllers made by Control Resources) or buy a fan with built-in “thermal speed control” (e.g., the Rotron “ThermaPro-V” series). Either way, using a variable-speed blower makes good sense, because under most conditions the fan noise will be far less than the

worst-case (full-speed) condition, which is needed only when operating an instrument at its maximum ambient temperature.

When laying out an instrument designed for forced air cooling, try to arrange things so that the air enters the box at one end, flows around the components, and exits at the far end. In an instrument with an interior horizontal chassis partition, for example, you might punch some inlet perforations at the bottom rear, perforate the internal chassis near the front of the instrument, and mount the exhaust fan at the top rear, thus forcing the airflow to pass through all parts of the instrument. Keep in mind that a circuit board will block airflow, and plan accordingly. If there is significant impedance to the flow of air (high back-pressure), a centrifugal blower will work better than the propeller type. The blades of the latter go into “stall” when the back-pressure exceeds about 0.3 inch of water, rendering the fan totally ineffective. Finally, in any



TABLE 12.2. VENTURI FANS

| Manuf   | Standard<br>4.7" square<br>105–120 CFM | Quiet<br>4.7" square<br>70 CFM | Very quiet<br>4.7" square<br>50 CFM | Mini<br>3.1" square<br>35 CFM |
|---------|----------------------------------------|--------------------------------|-------------------------------------|-------------------------------|
| Rotron  | MU2A1                                  | WR2H1                          | WR2A1                               | SU2A1                         |
| IMC     | 4715FS-12T-B50                         | 4715FS-12T-B20                 | 4715FS-12T-B00                      | 3115FS-12T-B30                |
| Pamotor | 4600X                                  | —                              | 4800X                               | 8500D                         |
| Torin   | A30108                                 | A30390                         | A30769                              | A30473                        |

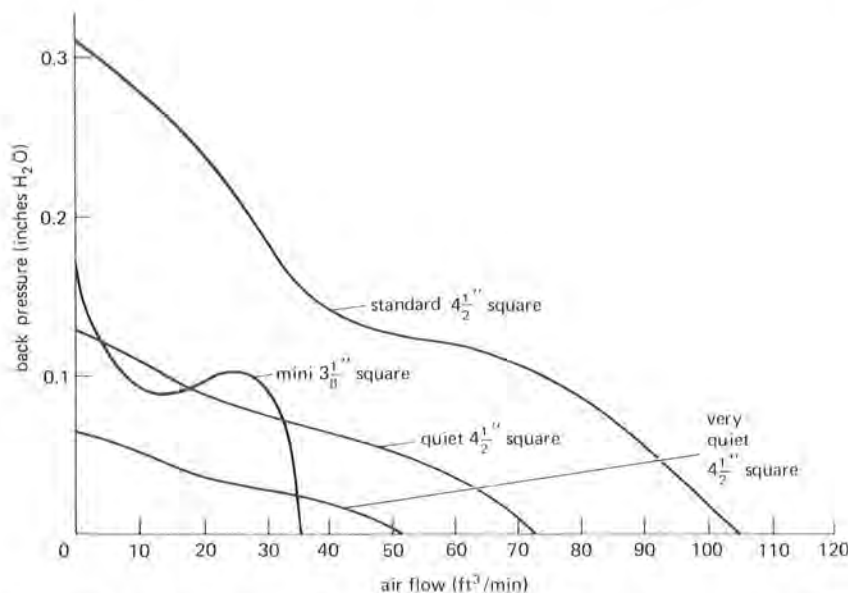


Figure 12.17. Airflow versus back-pressure for the venturi blowers listed in Table 12.2.

cooling situation it is a good idea to design conservatively; failure rates for electronic instruments rise dramatically when equipment is operated hot. Figure 12.18 illustrates good instrument design in regard to cooling and accessibility.

### 12.13 Some electrical hints

#### Unreliable components

The most unreliable components in any electronic system will be the following (worst first):

1. Connectors and cables
2. Switches
3. Potentiometers and trimmers

Keep them in mind as your brainchild proliferates in complexity.

#### RF line filters

As we mentioned earlier, it is a good idea to use RF filters on the ac power-line inputs. These are manufactured by a number of companies, including Corcom, Cornell-Dubilier, and Sprague. They are

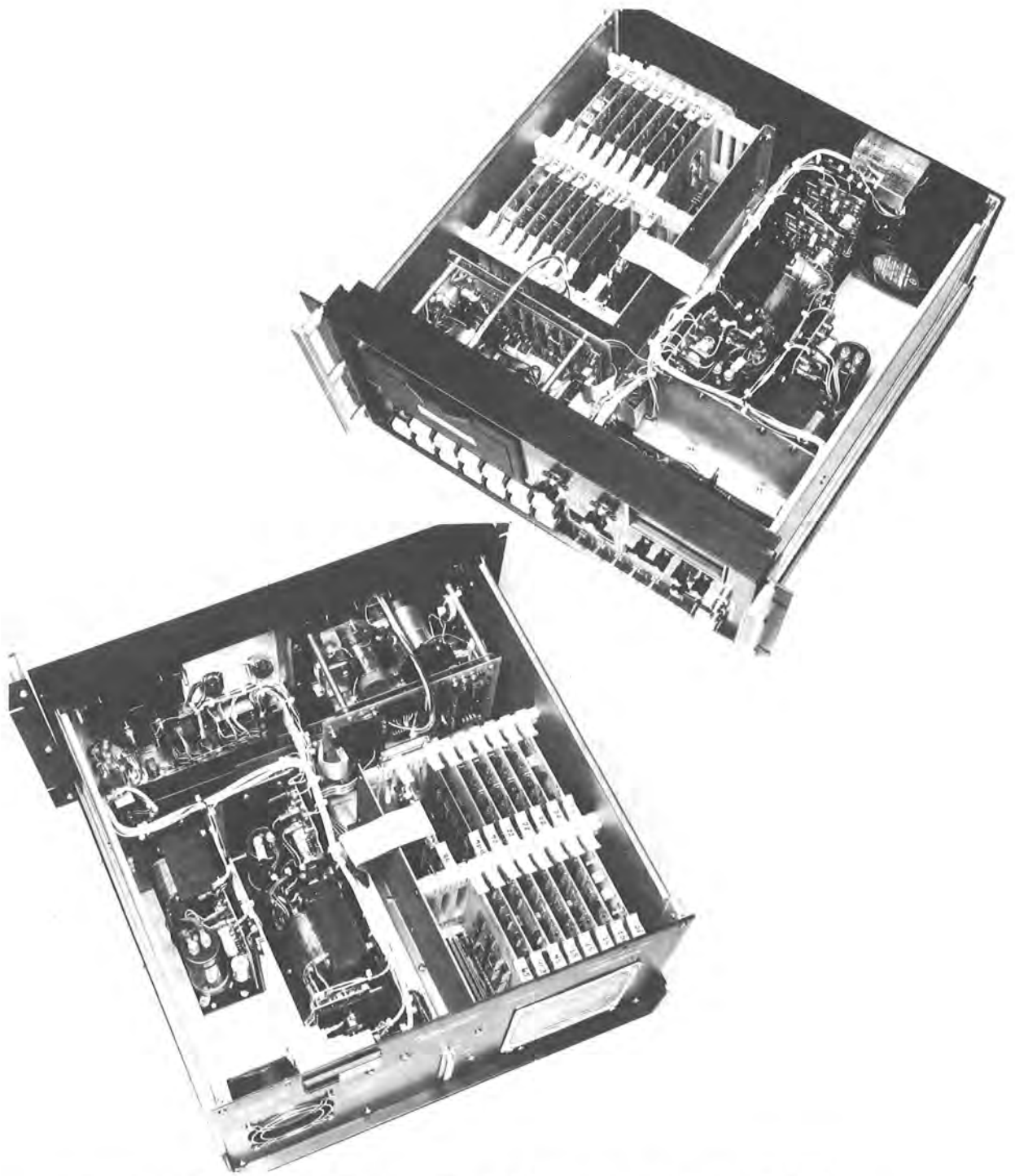


Figure 12.18. Views of a complete instrument (a digital cassette tape reader) illustrating several techniques of support and interconnection. Most of the electronics are housed in a card cage (with hand-wired backplane and mass-termination connections), and the electronics associated with the tape drive are on two boards near the drive motor (with mass-termination, in-line, and DIP plug connections). Adjustments and test points are accessible near the edges of each circuit card. Note the cooling path: Air is sucked in behind the card cage; it flows between the cards, then around the central partition and back over the power supplies before being blown out by the exhaust fan at the right rear.

available as simple modules with solder-lug terminals or in configurations that include an integral chassis-mounting ac line plug to mate with the standard IEC line cord. These filters provide excellent rejection of RF signals on the power line (as well as preventing their emission by the instrument itself), and they are also partially effective in reducing line transients. As an example, the Corcom 3R1 filter (rated at 3A, 115V) has 50dB rejection of RF at 200kHz and more than 70dB rejection for signals above 0.5MHz (see Section 6.11 and Table 6.3).

We favor the use of complete "power entry modules," rear-mounted units that combine 3-prong IEC power connector, RF line filter, fuse(s), line-voltage selector, and power switch. Look in the catalog pages of Corcom, Curtis, Delta, Power Dynamics, and Schaffner for representative units.

### **Transient suppressors**

Power-line transient suppressors are also a good idea in any instrument to prevent malfunction (or even damage) from the occasional 1kV to 5kV spikes that occur on everyone's ac power line. You just put them across the power-line terminals downstream of the fuse; they act like bidirectional zeners with enormous peak current capability. They come in packages that look like disc ceramic capacitors or power diodes; the inexpensive and small GE V130LA10A, for instance, costs about a dollar, begins conducting at 185 volts, and can handle peak currents of 4000 amps (see Section 6.11 and Table 6.2 for more details).

### **Fusing**

A line fuse is *mandatory* in every line-operated electronic instrument, without exception. As we indicated in Section 6.11, the wall socket is fused at a current designed to prevent fire hazard in the wall

wiring, typically 15 or 20 amps. That won't prevent a major disaster in a malfunctioning instrument, such as one with a power-supply capacitor failure, where the instrument may begin to draw 10 amps or so (more than 1kW heating in the power transformer). Important fact (learned the hard way by the authors): The lead from the power line goes to the innermost terminal of the fuse holder, so when you insert a new fuse, you can't get your fingers onto the "hot" terminal. Use a slow-blow fuse with rating 50%–100% greater than worst-case current drain of the instrument.

### **Cold-switching philosophy**

Whenever possible, it is a good idea to avoid running logic signals or analog waveforms to panel controls; this is to prevent cross-coupling and signal degradation that otherwise might occur. Instead, run dc control levels to the panel switches and pots, and use on-board circuitry to handle the actual signal switching, etc. This is especially important in noisy environments, or when dealing with high-speed or low-level signals, since the dc control signals can be thoroughly bypassed, whereas fast signals cannot. For example, use select gates (multiplexers) rather than routing the logic signals through a switch, and use a voltage-controlled oscillator rather than an *RC* oscillator for panel control of frequency. The few extra components you need to do things this way will buy you increased reliability and simplified assembly (no shielded cable, for example).

### **12.14 Where to get components**

Getting the parts you need to build some piece of electronic equipment can present some real difficulties. Most of the large electronics distributors have abandoned over-the-counter sales, making it nearly impossible for the small purchaser to go

down to the store and buy a few parts. Fortunately, the large (well-stocked) distributors will still take an order by telephone, with cash pickup at the "will call" counter. When playing this game, it is essential to know exactly what you want, by part number and manufacturer (for ICs, you may have to know the full part number, prefixes, suffixes, and all).

Many distributors are extremely hesitant to sell in small quantities, so you're often forced to buy at least five or ten of each item. Add to this the fact that a given distributor handles only a fraction of the brands you may need, and you're faced with a major chore. The consumer-oriented electronics stores (Radio Shack, etc.) will deal in small quantities and do have counter sales, but they tend to stock an extremely limited range of parts. The parts distribution system seems to be aimed at the industrial user, with his quantity ordering. Electronics manufacturers are treated well by the distributors, who pay them frequent visits, give them handsome catalogs and data books, and engage in competitive pricing.

Some special cautions are in order for buying ICs. Many kinds of ICs are manufactured without 100% testing; instead, a sample of each batch is tested, with the whole batch being rejected if the sample shows excessive failure rate. As a result, you can, on occasion, get a perfectly worthless chip straight from a reputable

manufacturer's production line. As a rough guide, you can expect up to 0.1% of new chips to be defective. That's not too serious, and in any case you can have your chips tested if a lower reject rate is necessary. Furthermore, all manufacturers test their LSI chips, and some manufacturers (AMD, for example) perform 100% testing of all their ICs.

A more serious problem arises when those rejected batches find their way into the hands of a "relabeler," a kind term for a junk peddler. Labeling machines are inexpensive; consequently, "counterfeit ICs" are all too common. Our experience has been that the large distributors (Arrow, Hamilton/Avnet, Newark, Schweber, and Wyle, to name a few) are reliable, at least for the brands for which they are authorized distributors. Surprisingly, most mail-order houses seem to ship good merchandise, often at very good prices, but there is an element of risk involved. Be suspicious of any IC without a date code. Because of the extra time and annoyance involved in finding bad ICs in a circuit, we generally recommend that you play it safe by using regular distributors for all IC buying, in spite of the generally higher prices. However, two mail-order houses that we have used successfully are Digi-Key (full-range catalog; Thief River Falls, MN) and Microprocessors Unlimited (computer chips; Beggs, OK).





# HIGH FREQUENCY AND HIGH-SPEED TECHNIQUES

## CHAPTER 13

### HIGH-FREQUENCY AMPLIFIERS

In this chapter we will discuss the important subject of high-frequency and radio-frequency techniques, as well as the digital equivalent, high-speed switching. High-frequency techniques find wide application in communications and broadcasting and in the domain of radiofrequency laboratory measurements (resonance, plasmas, particle accelerators, etc.), whereas high-speed switching techniques are essential for the fast digital instrumentation used in computers and other digital applications. High-frequency and high-speed techniques are extensions of our ordinary linear and digital techniques into the domain where the effects of interelectrode capacitance, wiring inductance, stored charge, and short wavelength begin to dominate circuit behavior. As a result, circuit techniques depart radically from those used at lower frequencies, with such bizarre incarnations as stripline and waveguide and devices like Gunn diodes, klystrons, and traveling-wave tubes. To give an idea of what is possible, there are now commercially available digital ICs (counters, etc.) that

operate at pulse rates of 3GHz and higher and linear circuit elements (amplifiers, etc.) that operate at frequencies in excess of 100GHz.

We will begin with a discussion of high-frequency transistor amplifiers, complete with simple transistor and FET models. After a few examples, we will move to the important subject of radiofrequency techniques, followed by a discussion of communications concepts and methods, including modulation and detection. Finally, we will look at high-speed switching techniques in some detail. Because of the specialized nature of these subjects, this chapter could be passed over in a first reading.

#### 13.01 Transistor amplifiers at high frequencies: first look

Amplifiers of the type we discussed earlier (e.g., common-emitter amplifiers with resistive collector load) show a rolloff of gain with increasing signal frequency, mostly owing to the effects of load capacitance and junction capacitance. Figure 13.1 shows the situation in its

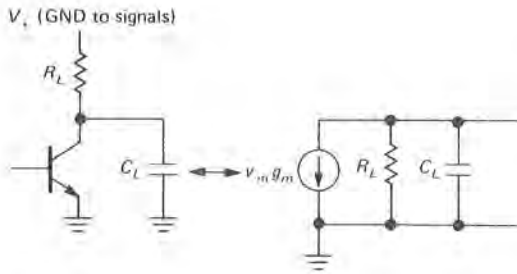


Figure 13.1

simplest form (we'll complicate things soon enough!).  $C_L$  represents the effective capacitance from collector to ground and forms a low-pass filter of time constant  $R_L C_L$  in combination with the amplifier's collector load resistance  $R_L$ . Remember that at signal frequencies  $V_+$  is the same as ground; hence the equivalent circuit shown.  $C_L$  includes collector-to-emitter and collector-to-base capacitances, as well as load capacitance. At frequencies approaching  $f \approx 1/R_L C_L$  the amplifier's gain begins dropping rapidly.

#### Reducing load capacitance effects

The simplest therapy consists of measures to reduce the product  $R_L C_L$ . For example:

1. Choose a transistor (or FET) with low interelectrode (junction and lead) capacitance; these are usually designated as RF or switching transistors.

2. Isolate the load with an emitter follower, thus reducing the capacitive load seen at the collector.

3. Reduce  $R_L$ . If you keep  $I_C$  constant, the gain drops, owing to reduced  $g_m R_L$ . Remember that for a transistor,  $g_m = 1/r_e$ , or  $I_C(mA)/25$  for an amplifier with bypassed emitter. To keep the gain constant with decreasing  $R_L$ , you have to raise the collector current by keeping  $V_+$  constant. Thus,

$$f_{\max} \approx 1/R_L C_L \propto I_C/C_L$$

which accounts for the rather high currents often used in high-frequency circuits.

#### 13.02 High-frequency amplifiers: the ac model

Load capacitance is not the only effect reducing amplifier gain at high frequencies. As we mentioned earlier (see the discussion of Miller effect in Chapter 2), the feedback capacitance ( $C_{cb}$ ) from output to input can dominate the high-frequency rolloff, especially if the input signal source impedance is not low. In order to determine where an amplifier will roll off, and what to do about it, it is necessary to introduce a relatively simple ac model of transistors and FETs. We will do that now, with a worked example of a high-frequency amplifier to illustrate how to use it.

#### ac model

The common-emitter (or source) models diagrammed in Figure 13.2 are just about

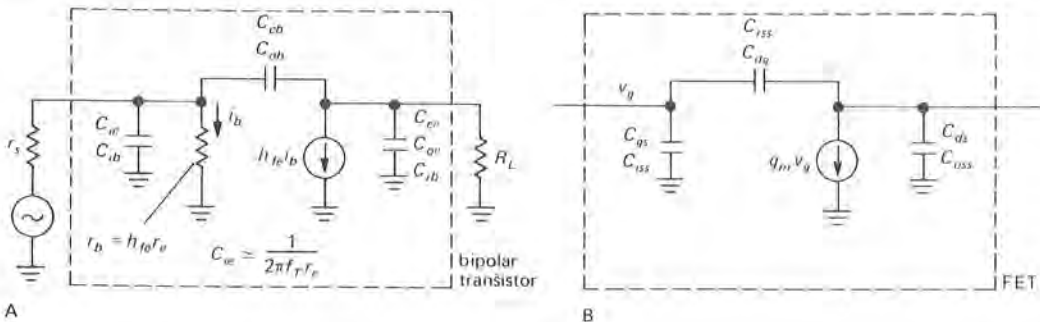
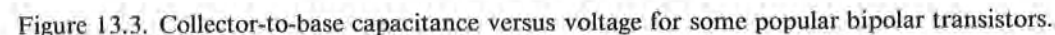


Figure 13.2. Bipolar transistor and FET high-frequency equivalent circuits.

The input capacitance  $C_{ie}$  is different, since you're dealing with a forward-biased junction. In this case the effective capacitance rises dramatically with increasing base current, since  $V$  is near  $V_d$ , and it would make little sense to specify a value for  $C_{ie}$  on a transistor data sheet. However, it turns out that the effective  $C_{ie}$  increases with increasing  $I_E$  (and therefore decreasing  $r_e$ ) in such a way that the  $RC$  product ( $r_b C_{ie}$ )





remains roughly constant. As a result, the transistor's gain at a particular frequency depends primarily on the ratio between current lost into  $C_{ie}$  and current that actually "drives the base" and is not strongly dependent on collector current. Therefore, instead of attempting to specify  $C_{ie}$ , the transistor manufacturer usually specifies  $f_T$ , the frequency at which the current gain ( $h_{fe}$ ) has dropped to unity. It is easy to show that  $f_T$  is given by

$$f_T = \frac{1}{2\pi C_{ie} r_e}$$

or, equivalently,

$$C_{ie} = \frac{1}{2\pi f_T r_e}$$

for particular values of  $C_{ie}$  and  $r_e$  at some collector current. Transistors intended for radiofrequency applications have  $f_T$  in the range of 500MHz to 10GHz whereas "gen-

eral purpose" transistors have  $f_T$ s in the range of 50MHz to 250MHz. Figure 13.4 shows the variation of  $f_T$  with collector current for typical transistors.

### □ 13.03 A high-frequency calculation example

Let's apply our simple model to the design of a high-frequency broadband amplifier. We will show the driving stage also, so the driving (source) impedance is known. As it will turn out, the amplifier will exhibit poor performance and severe loading of the driving stage. The sort of performance problems you will see are characteristic of real-life circuit design, and we will talk about ways to improve performance by changes in circuit configuration and operating points. Figure 13.5 shows the circuit fragment. This subcircuit is assumed

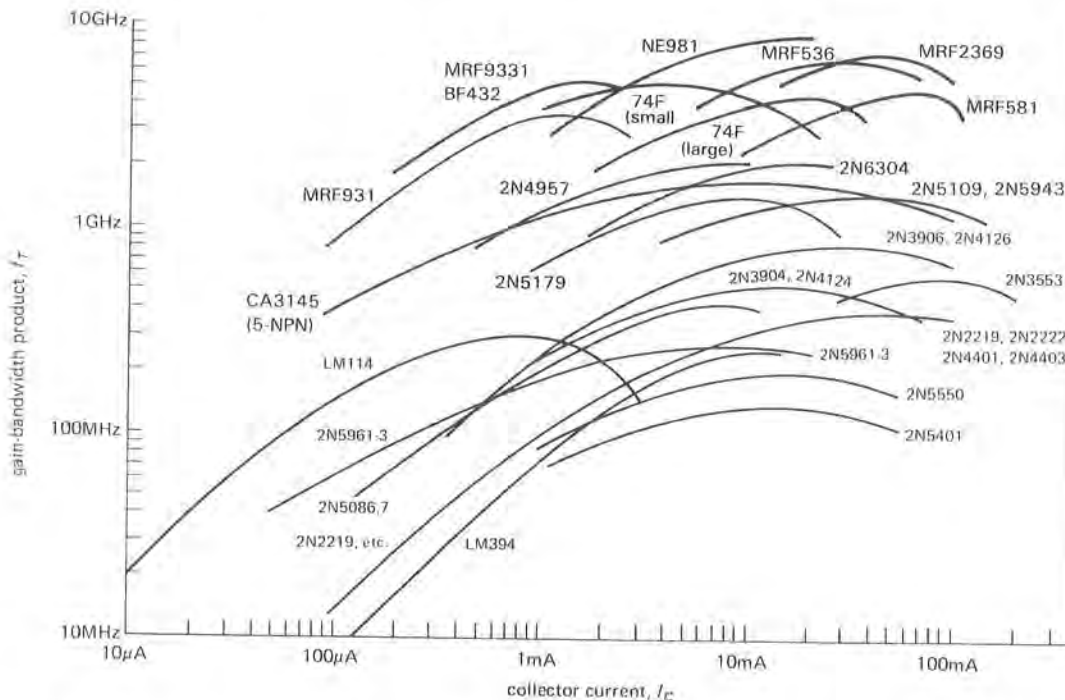


Figure 13.4. Gain-bandwidth product,  $f_T$ , versus collector current for some popular bipolar transistors.

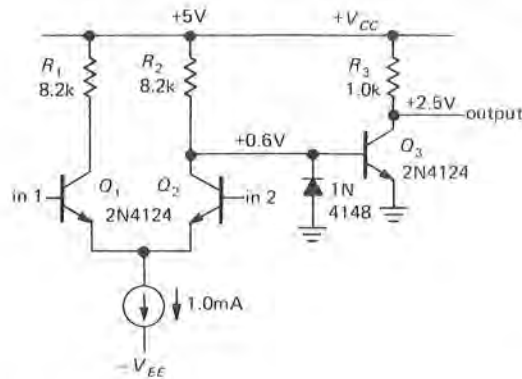


Figure 13.5

to lie within an overall amplifier circuit with feedback at dc to stabilize the quiescent point at  $\frac{1}{2} V_{CC}$ ; it would not be biased stably as shown. Since we are interested in high-frequency performance, we won't worry further about how the biasing is accomplished. Note that the differential stage has very little common-mode input range, extending to perhaps +0.25 volt or so and limited in the negative direction by the compliance of the emitter current source.

#### □ Analysis of high-frequency rolloff

The differential stage has calculable gain and output impedance, allowing us to analyze the output-stage rolloff in detail. Our analysis of the gain of the  $Q_3$  amplifier stage consists of the following:

1. Find the low frequency gain with zero source impedance. Then find the 3dB point, owing to output capacitance and feedback capacitance in combination with load resistance:

$$f_{-3\text{dB}} = \frac{1}{2\pi R_L (C_L + C_{cb})}$$

2. Find the input impedance, the combination of base input impedance ( $r_b$  and  $C_{ie}$ ) and effective feedback capacitance ( $G_V C_{cb}$ ).

3. Compute the 3dB point due to input loading of the source; compare with the

“output 3dB point” calculated in step 1, to see where the high-frequency bottleneck is.

4. Improve performance, if necessary, by alleviating the problem that dominates the high-frequency rolloff.

Note that the feedback capacitance  $C_{cb}$  appears in both the output and input circuit calculations, multiplied by the voltage gain (Miller effect).

Let's try our simple method of analysis on this circuit, modeled in Figure 13.6. The 2N4124 is parametrized by  $C_{cb} = 2.4\text{pF}$  at 2.5 volts,  $h_{fe} \approx 250$ , and  $f_T = 300\text{MHz}$ .

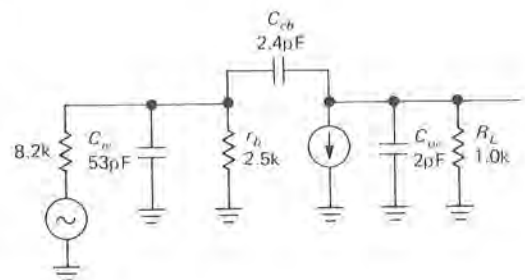


Figure 13.6

1. Assuming  $Q_3$  is driven by a voltage source, its low-frequency voltage gain is 100, since  $r_e = 10$  ohms at 2.5mA collector current. The 3dB point set by output capacitance is at roughly 40MHz (2.4pF in parallel with 2pF, driven by 1.0k). Note that in this simple calculation we have ignored the load capacitance and stray wiring capacitance.

2. The input resistance is roughly 2.5k ( $h_{fe} r_e$ ), paralleled by the Miller capacitance (240pF) and by  $C_{ie}$ ; the latter works out to about 53pF, using the formula given earlier.

3. The 3dB point due to input capacitances comes out roughly at 280kHz ( $R = 8.2\text{k}$  paralleled with 2.5k;  $C = 240\text{pF} + 53\text{pF}$ ), dominated by the Miller-effect capacitance  $C_{cb} G_V$  in combination

with the relatively high impedance at the base. Note that the low-frequency gain is actually less than 100, considering the input signal to be the *unloaded* output if the differential stage, owing to the loading of the previous stage by the low input resistance; when you include this effect, the gain at low frequencies is actually  $100 \times 2.5 / (2.5 + 8.2)$ , or about 23.

The combination of excessive previous stage loading and low 3dB rolloff frequency makes this a poor circuit, included here to illustrate the real-life problems of high-frequency amplifier design. In practice, you would improve performance by using greatly reduced collector impedances or by going to a different amplifier configuration intended to reduce or eliminate the effects of input capacitance ( $f_T$ ) and feedback capacitance ( $C_{cb}G_V$ , Miller effect).

#### 13.04 High-frequency amplifier configurations

As the preceding example illustrates, the Miller effect can dominate the high-frequency performance of an amplifier driven by moderately high source impedance. In that example, an  $f_T$  of 300MHz and an output time constant corresponding to 3dB point of 40MHz were swamped by an input time constant giving a 3dB point of 280kHz.

#### Three cures for Miller effect

Besides the brute-force approach of reducing collector resistances enormously, there are several interesting configurations that aim to reduce driving (source) impedance or reduce feedback capacitance or both. Figure 13.7 shows these configurations, drawn in their simplest forms, without regard to bias or power supplies (i.e., the signal-frequency circuit alone is drawn).

In the first circuit, an emitter follower reduces the driving impedance seen at the input of a common emitter amplifier. This greatly reduces the degradation of high-frequency performance caused by  $f_T$  and  $C_{cb}G_V$ . The second circuit is the popular cascode, in which a common-emitter stage drives a common-base stage, eliminating  $C_{cb}G_V$  Miller effect ( $Q_4$ 's emitter is pinned by the fixed base voltage; it just passes  $Q_3$ 's collector current through to  $R_L$ ). In the third circuit a follower drives a common-base stage, eliminating Miller effect and reducing the driving impedance at the same time; this circuit is the familiar differential amplifier, with unbalanced collector resistors and one input grounded.

#### More techniques

In addition to these circuit configurations, there are two other approaches to the input and feedback capacitance problem, namely

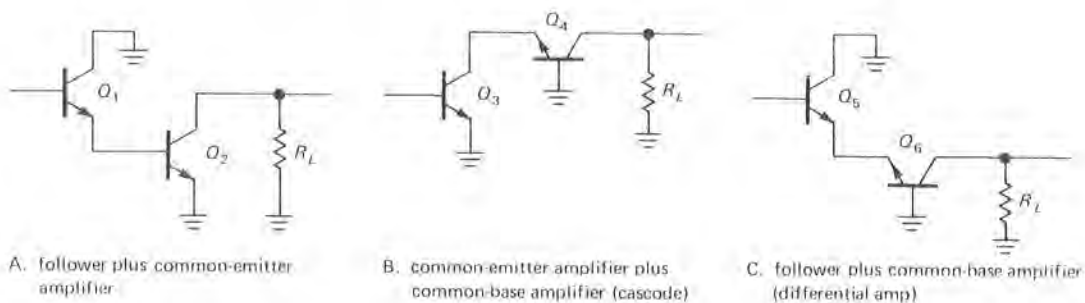


Figure 13.7. Simplified high-frequency amplifier configurations.

(a) the use of a simple grounded-base amplifier alone, if the driving impedance is low enough, and (b) the use of tuned circuits at the input and output of a common-emitter (or other) amplifier, to “tune away” the effects of interelectrode capacitance. Note that such a tuned amplifier does not have broadband response, but amplifies only a narrow range of frequencies (which may be an advantage, depending on the application). In addition, *neutralization* may be necessary. We will discuss narrow-band tuned amplifiers in a later section of the chapter. An in-between approach involves the use of “peaking” inductances of a few microhenrys in series with collector load resistances to cancel some of the effects of capacitance and hold up the gain at frequencies somewhat above the normal high-frequency rolloff (Fig. 13.8).

In order to be able to estimate the high-frequency performance of circuits involving followers and grounded-base stages, we will need simple ac transistor models for these configurations (Fig. 13.9). Note that in the emitter follower model the impedances depend on source and load impedances (reactance as well as resistance). We will apply these models in the next example.

#### □ 13.05 A wideband design example

As an example of an improved broadband amplifier design, consider the circuit in Figure 13.10, configured to eliminate almost entirely the rolloff caused by Miller effect. It uses emitter follower inputs (high input impedance) to a differential amplifier; the output is isolated by an emitter follower from the cascode-connected output section of the differential amplifier. The design is based on the use of a good high-frequency transistor such as the 2N5179, with an  $f_T$  of 1000MHz (specified as  $h_{fe} = 10$  at 100MHz) and a  $C_{cb}$  of 0.5pF at 2 volts. The approximate equivalent circuit, in terms of junction and stray

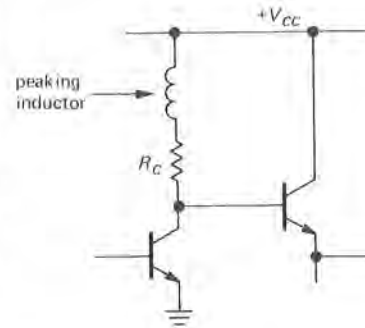


Figure 13.8

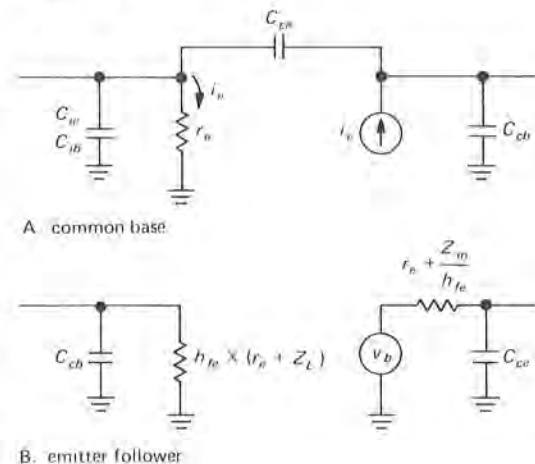


Figure 13.9

capacitances and their shunt resistances, is shown in Figure 13.11.

To determine the high-frequency rolloff point of this amplifier, you have to go through each stage, analyzing the various  $RC$ 's by substituting the appropriate equivalent circuits. There is usually one stage that sets the lowest limit, and with some intuition and guesswork you can often put your finger right on it. In this case the limiting performance is set by the finite driving impedance ( $300\Omega$ ) to  $Q_7$ 's base, in combination with the capacitance of  $Q_7$  and the load capacitance  $C_L$  as seen buffered at  $Q_7$ 's base (remember that  $h_{fe}$  drops approximately as  $1/f$ , so at very high frequencies the isolating effect of an emitter follower is seriously degraded).

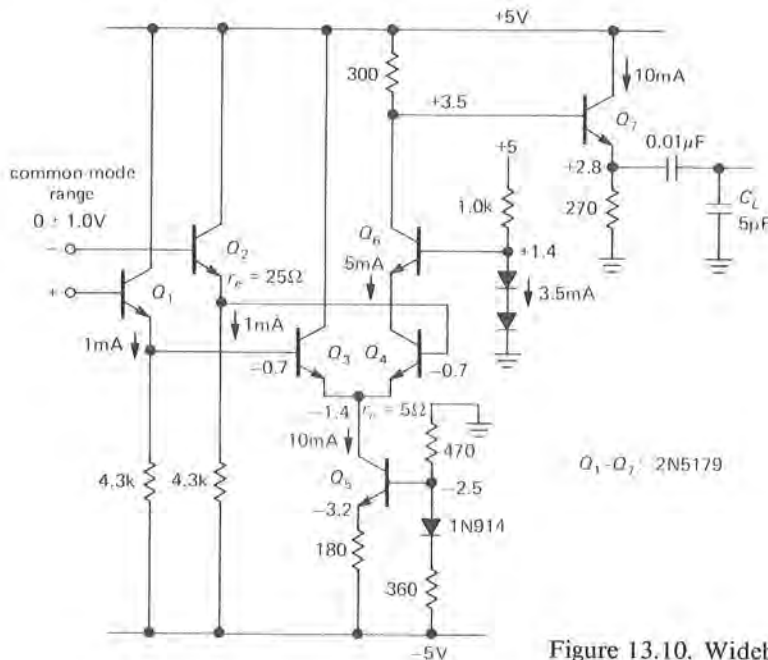


Figure 13.10. Wideband differential amplifier.

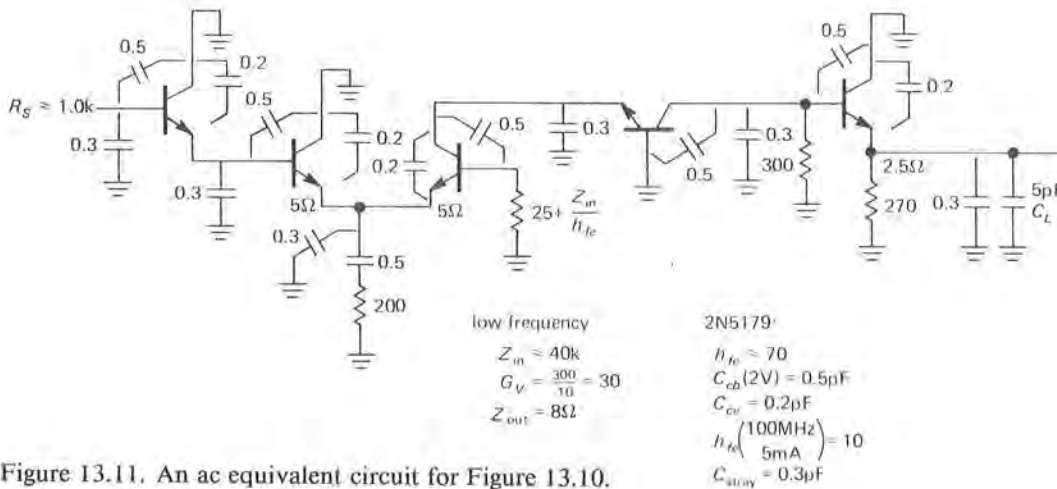


Figure 13.11. An ac equivalent circuit for Figure 13.10.

The simple method we used to figure the 3dB point goes something like this: Apply the emitter follower equivalent circuit to  $Q_7$  to get the impedance looking into the base, knowing the load capacitance, junction capacitances, and wiring capacitance (we used  $C_{cb} = 0.5pF$ ,  $C_{ce} = 0.2pF$ ,

and  $C_{stray} = 0.3pF$ ). Since the impedance looking into the base depends on  $h_{fe}$  you have to do the calculation as a function of frequency (assuming  $h_{fe} \approx 1/f$  at high frequency); we chose instead to do the calculation at a few high frequencies, guessing that the 3dB point would be in the



neighborhood of a few hundred megahertz. Figure 13.12 summarizes the process. At frequencies of 100MHz, 200MHz, and 400MHz we took the load impedance, multiplied it by the transistor beta (of  $Q_7$ , assuming  $h_{fe} \approx 1/f$ ), combined it with the other impedances already present at the base, and then calculated the resultant magnitude of impedance in order to get the relative output swing as a function of frequency. As can be seen, the output was down 3dB somewhere around 180MHz.

Now, using this as an estimate of the 3dB point, we went through the rest of the circuit, checking to see if other  $RC$ s gave significant attenuation at this particular frequency. As an example,  $Q_4$ 's collector circuit would be down 3dB at about 1000MHz, using the value of tran-

sistor beta at 180MHz ( $h_{fe} \approx 5$ ); in other words, the cascode portion of the circuit does not degrade overall performance.

In a similar manner, it is a relatively straightforward process to verify that no other portion of the circuit sets as low a 3dB point. When dealing with the input stage, you have to assume some value of driving (source) impedance. If you assume  $Z_s = 1000$  ohms (rather high for a video circuit like this), you find, finally, that the combination of source resistance in combination with input capacitance (1.0k, 0.8pF) contributes a 3dB point at about 200MHz. Thus, the overall circuit has good performance up to about 200MHz, for source impedances somewhat less than 1k, but this will be degraded for source impedances comparable to or exceeding 1k.

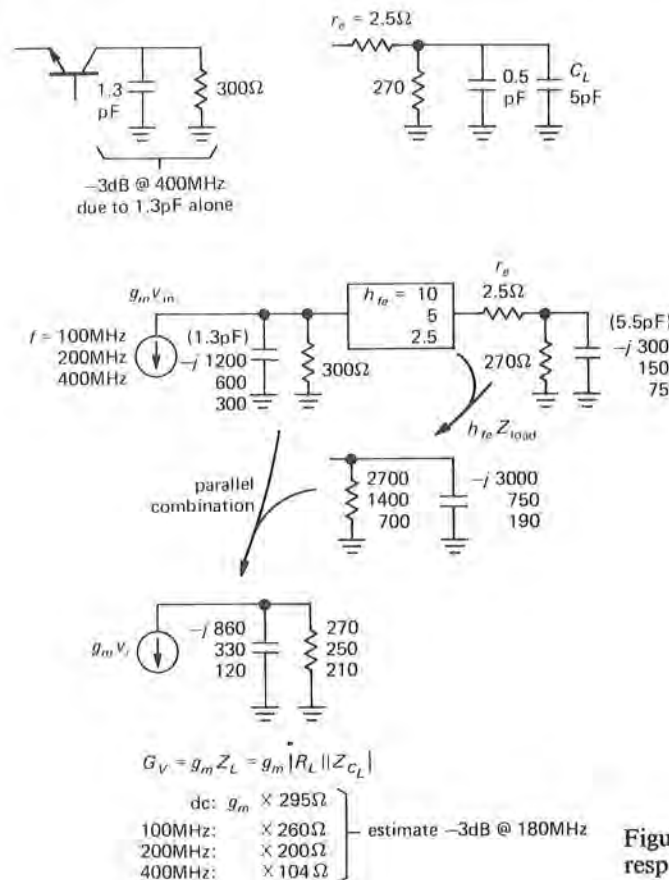


Figure 13.12. Calculating amplifier response rolloff.

This is a considerable improvement over the earlier circuit we analyzed. □ **13.07 The shunt-series pair**

□ **13.06 Some refinements to the ac model**

□ **Base spreading resistance**

It is worth noting that the models we have been using are somewhat simplistic, and they neglect some important effects, e.g., the finite resistance  $r'_b$  of the base contact. Transistors intended for high-frequency use often specify the parameter  $r'_b C_{cb}$ , the “collector-base time constant.” For the 2N5179 it is 3.5ps (typ), implying a base contact “spreading” resistance of about 7 ohms. When analyzing performance at extremely high frequencies, it is necessary to include such effects in the calculations; in this example it has no effect on the conclusions we reached earlier.

**Pole splitting**

Another simplification in the preceding treatment is the assumption that each  $RC$  rolloff acts independently of the others. It is intuitively easy to see that there must be some interaction, by the following argument: The Miller effect itself is a form of high-frequency negative feedback. Since it samples output *voltage*, it must therefore act to lower the output impedance of the transistor stage, particularly at higher frequencies, where its “loop gain” is high (of course, it also lowers the voltage gain, which is the whole problem). The resultant reduced impedance at the collector drives the  $R_L C_L$  rolloff up to a higher frequency, since the collector impedance parallels  $R_L$ . Thus, lowering the frequency of the Miller-effect rolloff (by raising  $G_V$  or  $C_{cb}$ ) raises the rolloff due to collector and load capacitance. This is known as “pole splitting.”

A popular circuit for broadband low-gain amplifiers is the shunt-series pair (Fig. 13.13). The idea is to make an amplifier of relatively low gain (perhaps 10dB)

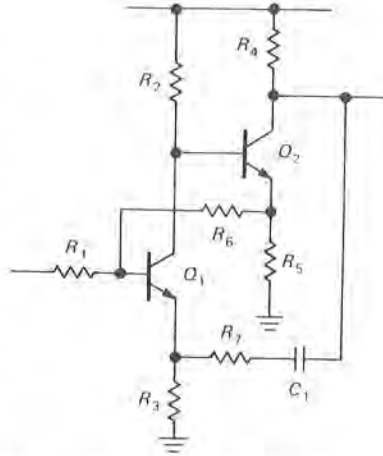


Figure 13.13. Shunt-series pair.

with flat response over a broad range of frequencies. That sounds like an application for negative feedback. However, negative feedback can be troublesome at radiofrequencies, owing to uncontrolled phase shifts in a high-loop-gain feedback path. The shunt-series pair overcomes this difficulty by having several feedback paths, each with relatively low loop gain.

In the preceding circuit, both  $Q_1$  and  $Q_2$  operate as low-gain voltage amplifiers, since their emitter resistors are not bypassed.  $R_6$  provides feedback around  $Q_1$  alone, since  $Q_2$  is used as a follower for that loop. Once  $Q_1$ 's overall voltage gain is set ( $R_6/R_1$ ),  $R_4$  is chosen to set  $Q_2$ 's open-loop gain ( $R_4/R_5$ ). Finally, the feedback to  $Q_1$ 's emitter is added to reduce the gain to its design value.

The shunt-series pair is a convenient amplifier building block because it is extremely stable and easy to design. Amplifiers with bandwidth to 300MHz or so

are easily constructed with this technique. You can get gains of 10dB to 20dB per amplifier, cascading several stages, if necessary, to obtain greater gain.

In Section 13.11 we will discuss the techniques used to construct tuned (narrow-band) amplifiers, as contrasted with the broadband design we have been talking about so far. Since signals of interest are often confined to a narrow band of frequencies in instruments that operate at radiofrequencies, tuned amplifiers are extremely useful.

#### □ 13.08 Modular amplifiers

From the foregoing discussion of RF amplifiers it might seem that any project at high frequencies would become a formidable design effort, with messy calculations and numerous trial designs. Luckily, there are complete packaged amplifier *modules* available from more than a dozen suppliers, in configurations to meet almost any need. In fact nearly every RF component can be obtained as a module, including oscillators, mixers, modulators, voltage-controlled attenuators, power combiners and dividers, circulators, hybrids, directional couplers, etc. We will describe some of these other circuit elements in Section 13.12.

In its most basic form, the prepackaged RF amplifier comes as a thin-film hybrid circuit with gain over a wide band, packaged in a 4-pin transistor or surface-mount package (Fig. 13.14). Two of the pins are input and output terminals, with convenient 50 ohm impedance levels, and the remaining pins are for ground and the dc supply. There are dozens of different amplifiers available, some optimized for low noise and others for high power or large dynamic range. Individual amplifiers may be designed for operation over a very wide frequency range, or for a particular band of frequencies used in communications. As an example, the UTO-514 from Avantek

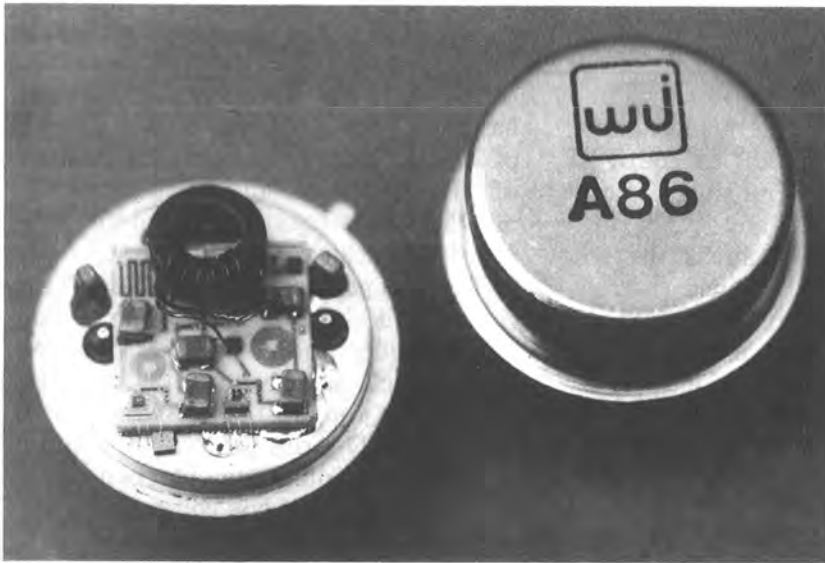
has 15dB of gain over the frequency range of 30MHz to 200MHz, with a noise figure of 2dB (maximum) and a gain flatness of  $\pm 0.75$ dB. It is packaged in a 4-pin TO-8 transistor package.

The high performance Avantek UTO series and Watkins-Johnson A-series of modular amplifiers include almost a hundred models each, with bandwidths to 2GHz. We have found the less expensive Avantek GPD-series (or Watkins-Johnson EA-series) modules to be particularly affordable and handy. For example, the GPD-201 spans 5–200MHz with 30dB gain (min) and 3.5dB NF (typ), and costs \$29.

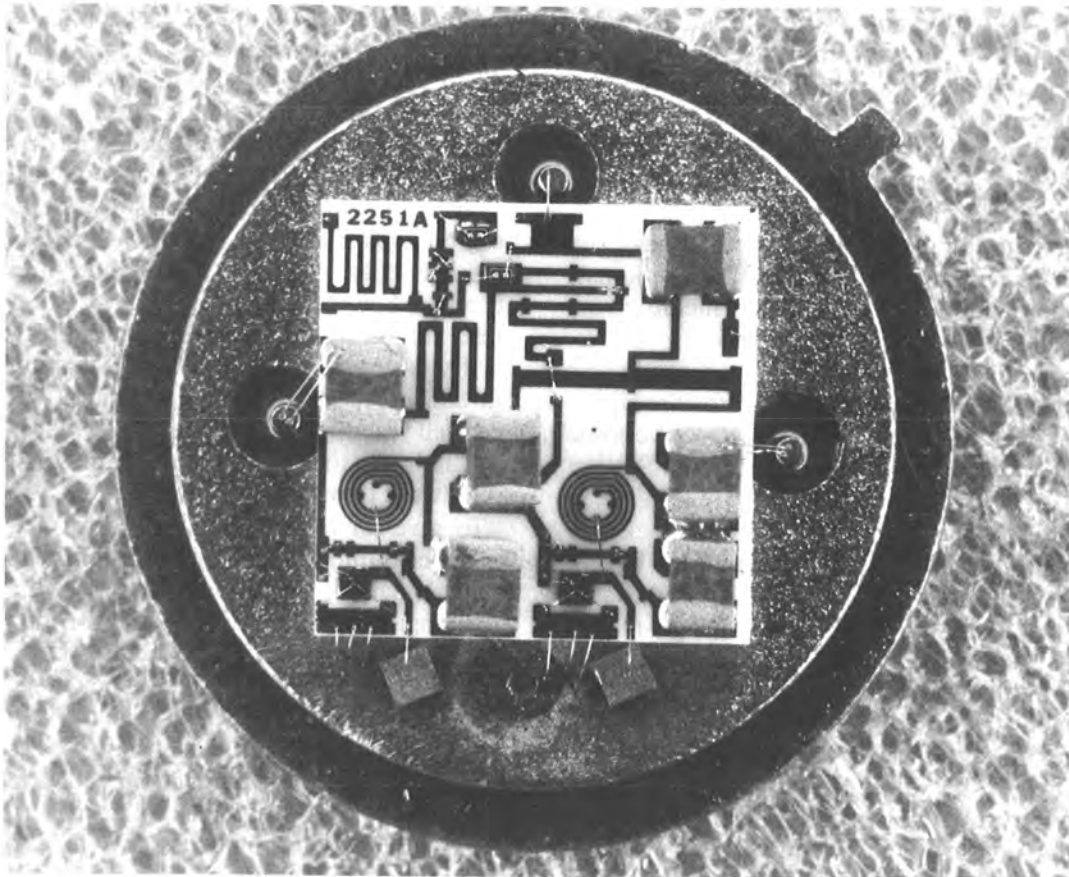
These hybrid amplifiers can be used singly or in cascade, usually as part of a stripline (Section 13.21). To make life even easier, the amplifier manufacturers have thoughtfully provided completed amplifier building blocks as prepackaged modules. These beasts typically occupy a small metal box, perhaps  $2 \times 2 \times 1$  inch, with SMA-type RF coax connectors for input and output. You can choose from a list of these standard “connected” amplifiers, or, if you prefer, you can specify a custom cascade of your choosing. Avantek will even sell you the little boxes and PC boards (that hold up to four modular amplifiers), if you want to make your own (Fig. 13.15).

To give an idea of what you can get, we have thumbed through the impressive Avantek catalog and come up with the following: The AMG-1020 is a nice low-noise amplifier with 34dB gain and a 2.7dB noise figure over the frequency range 50MHz to 1000MHz. For even wider bandwidth you might choose the UTC20-211 which spans 10MHz to 2000MHz with a 5dB noise figure and 26dB of gain. Wideband amplifiers are available to 18GHz and beyond, using GaAsFET (and HEMT) technology.

Amplifiers for use over a narrow band of frequencies can be optimized for low-noise performance; extremely good



A



B

Figure 13.14. A. 10-200MHz cascable amplifier.  
B. Detail of hybrid construction technique on ceramic substrate, showing chip capacitors, thin-film inductors and resistors, transistors, and wire bonding. (Courtesy of Watkins-Johnson Company.)

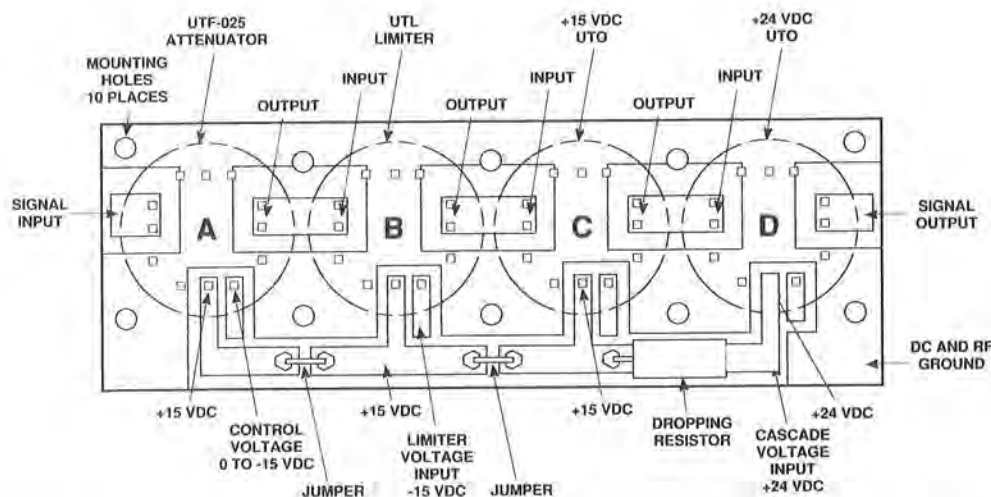


Figure 13.15. "Microstrip" (stripline) board for modular RF components. (Courtesy of Avantek, Inc.)

amplifiers are available for the communications bands. For example, for your own backyard satellite-downlink receiver, Avantek's AM-4285 has 50dB of gain ( $\pm 0.5$ dB) in the 3.7–4.2GHz band, with an excellent 1.5dB noise figure ( $T_N = 120^\circ\text{K}$ ). In the 7.25–7.75GHz band, their AM-7724 delivers a gain of 35dB ( $\pm 0.25$ dB) with an astounding 1.8dB noise figure.

With custom assembly and tuning it is possible to go considerably beyond the performance of these excellent commercial amplifiers. For example, radioastronomers routinely achieve noise figures of 0.7dB with custom L-band (1–2GHz) GaAs FET amplifiers, and 0.15dB noise figure when operated at liquid-nitrogen temperature (the corresponding noise temperatures are 50°K and 10°K; see Section 7.12). Recent designs employing high-electron-mobility FETs (HEMTs) have achieved a phenomenal 8°K noise temperature at 8.5GHz. The outstanding practitioner of this fine art is Sandy Weinreb at the National Radio Astronomy Observa-

tory (Charlottesville, VA). And a small company known as Berkshire Technologies (Oakland, CA) will build you custom amplifiers with this sort of stunning performance; their cooled amplifiers are currently achieving 5°K noise temperature at L-band, and 15°K at 8.5GHz.

To measure noise figure (noise temperature) in microwave amplifiers, you invariably use the hot-load/cold-load method. Look back at Section 7.19 to remind yourself how that works.

There is plenty of commercial competition in these amplifier modules, as well as other RF modular components. For complete amplifier modules, some of the larger suppliers are Aerotech/TRW, Avantek, Aydin Vector, Hewlett-Packard, Narda, Scientific Communications, and Watkins-Johnson. In practice, when designing an RF system you might well choose to thumb through catalogs of available (and custom) modules in order to assemble a system (Fig. 13.16). Screw them all down to a plate, connect them together with coaxial cable, and off you go!





Figure 13.16. Radiofrequency modules are available with connectors, with through-board PC pins, and with surface-mount pads. (Courtesy of Avantek, Inc.)

#### **Wide-bandwidth op-amps**

If you are like most people, you probably think of op-amps as relatively low frequency amplifiers, not good for anything much above 100kHz, or perhaps 1MHz. That's certainly true for garden variety op-amps, with their  $f_T$  values typically in

the range of 1MHz to 5MHz (see Table 4.1). However, as Table 7.3 shows, there is a class of accurate op-amps with gain bandwidth products up to 100MHz or so. In fact, if you can tolerate input offsets of 10mV or so, you can get op-amps with  $f_T$  values up to 1GHz. Unlike "video amplifier" ICs (which are single-ended,

TABLE 13.1. RF TRANSISTORS

| Type    | Case  | $P_{diss}$<br>$T_C=25^\circ\text{C}$<br>(W) | $V_{CE0}$<br>(V) | $V_{CB0}^a$<br>(V) | $I_C$<br>max<br>(A) | $h_{FE}$<br>@<br>typ | $I_C$<br>(mA) | $C_{cb}$<br>@10V<br>(pF) | $f_T$<br>(MHz) | $I_C$<br>(mA) | Power<br>gain<br>(dB) | Power<br>@ $f$<br>(W) | Output<br>power<br>@ $f$<br>(W) | Output<br>@ $f$<br>(MHz) | Comments                         |
|---------|-------|---------------------------------------------|------------------|--------------------|---------------------|----------------------|---------------|--------------------------|----------------|---------------|-----------------------|-----------------------|---------------------------------|--------------------------|----------------------------------|
| 2N3375  | TO-60 | 12                                          | 40               | 65                 | 1.5                 | $10^m$               | 250           | 12                       | 550            | 120           | 6                     | 175                   | 10                              | 100                      | low cost, easy to mount          |
| 2N3553  | TO-39 | 7                                           | 40               | 65                 | 1.0                 | $10^m$               | 250           | 7                        | 600            | 100           | 15                    | 175                   | 7                               | 100                      | low cost, popular                |
| 2N3866  | TO-39 | 5                                           | 30               | 55                 | 0.4                 | 50                   | 50            | 3                        | 900            | 40            | 15                    | 400                   | $1^m$                           | 400                      | oscillator                       |
| 2N4427  | TO-39 | 2                                           | 20               | 40                 | 0.4                 | 50                   | 100           | 4                        | 800            | 50            | 15                    | 175                   | $1^m$                           | 175                      |                                  |
| 2N5016  | TO-60 | 30                                          | 30               | 65                 | 4.5                 | $10^m$               | 500           | 25                       | 600            | 500           | 8                     | 400                   | 30                              | 100                      |                                  |
| 2N5109  | TO-39 | 2.5                                         | 20               | 40                 | 0.4                 | 60                   | 50            | 3                        | 1500           | 70            | 15                    | 200                   | —                               | —                        | low noise, popular               |
| 2N5179  | TO-72 | 12                                          | 20               | 20                 | 0.05                | 70                   | 3             | 0.7                      | 1500           | 10            | 20                    | 200                   | 0.02                            | 500                      | low noise, $r'_b C_c=7\text{ps}$ |
| 2N5994  | strip | 35                                          | 30               | 65                 | 5                   | —                    | —             | 70                       | —              | —             | 10                    | 100                   | 35                              | 175                      | VHF power                        |
| 2N6267  | strip | 20                                          | 50               | 50                 | 1.5                 | —                    | —             | 13                       | —              | —             | 10                    | 2000                  | 10                              | 2000                     | microwave power                  |
| 2N6603  | strip | 0.5                                         | 15               | 25                 | 0.03                | 80                   | 15            | 0.5                      | 5500           | 10            | 16                    | 1000                  | —                               | —                        | microwave small signal           |
| 2N6679  | strip | 0.9                                         | 20               | 30                 | 0.07                | 100                  | 15            | 0.3                      | —              | —             | 11                    | 4000                  | 0.07                            | 4000                     |                                  |
| MRF931  | strip | 0.05                                        | 5                | 10                 | 0.005               | 70                   | 0.25          | $0.4^b$                  | 3000           | 1             | 12                    | 1000                  | —                               | —                        | battery-powered, telemetry       |
| NE981   | strip | 0.3                                         | 6                | 10                 | 0.03                | 100                  | 10            | 0.4                      | 7000           | 20            | 14                    | 1000                  | 0.01                            | 7000                     | NEC                              |
| MRF571  | strip | 1                                           | 10               | 20                 | 0.07                | 100                  | 30            | 0.7                      | 8000           | 50            | 15                    | 1000                  | 0.1                             | 1000                     | low noise, low cost              |
| MRF941  | strip | 0.4                                         | 10               | 20                 | 0.05                | 100                  | 5             | 0.2                      | 8000           | 15            | 13                    | 2000                  | —                               | —                        | low noise                        |
| MRF951  | strip | 1                                           | 10               | 20                 | 0.1                 | 100                  | 5             | 0.3                      | 7500           | 30            | 13                    | 2000                  | —                               | —                        | low noise                        |
| MRF9331 | strip | 0.05                                        | 8                | 15                 | 0.001               | 80                   | 0.5           | $0.2^b$                  | 5000           | 1             | 20                    | 1000                  | —                               | —                        | micropower                       |
| AT41485 | strip | 0.5                                         | 12               | 20                 | 0.06                | 150                  | 10            | 0.2                      | 8000           | 25            | 12                    | 2000                  | 0.1                             | 2000                     | Avantek, low noise               |
| AT42085 | strip | 0.5                                         | 12               | 20                 | 0.08                | 150                  | 35            | 0.3                      | 8000           | 35            | 10                    | 4000                  | 0.1                             | 4000                     | Avantek, low noise               |
| AT64020 | strip | 3                                           | 20               | 40                 | 0.2                 | 50                   | 100           | —                        | —              | —             | 10                    | 2000                  | 0.5                             | 4000                     | Avantek                          |

(a) because the base is reverse-biased when the collector tuned circuit goes high,  $V_{CB0}$  is often the relevant breakdown voltage.(b) at  $V_{CB}=1\text{V}$ . (m) minimum.

TABLE 13.2. WIDE-BAND OP-AMPS

| Type    | Mfg <sup>a</sup> | Supply                   |                                | Bandwidth    |            |                                        | Ext comp | DIP pkg avail? | Input cap (pF) | Slew rate typ (V/ $\mu$ s) | Settling time to 1% (ns) | Small sig rise time (ns) | Output curr (mA)      | V <sub>out</sub> (peak)                 |                              | Class A output? | Curr lim? | Comments |                                         |
|---------|------------------|--------------------------|--------------------------------|--------------|------------|----------------------------------------|----------|----------------|----------------|----------------------------|--------------------------|--------------------------|-----------------------|-----------------------------------------|------------------------------|-----------------|-----------|----------|-----------------------------------------|
|         |                  | V <sub>os</sub> max (mV) | I <sub>in</sub> max ( $\mu$ A) | -3dB         |            | V <sub>50<math>\Omega</math></sub> (V) |          |                |                |                            |                          |                          |                       | V <sub>150<math>\Omega</math></sub> (V) |                              |                 |           |          |                                         |
|         |                  |                          |                                | G=1 (MHz)    | G=10 (MHz) |                                        |          |                |                |                            |                          |                          |                       |                                         | large sig <sup>d</sup> (MHz) |                 |           |          |                                         |
| AH0010  | OE               | 20                       | 20 <sup>1</sup>                | $\pm 18$     | 5          | —                                      | —        | —              | —              | 1500                       | 100 <sup>k</sup>         | —                        | 100                   | 5                                       | 10                           | 3               | —         | —        | buffer                                  |
| CLC110  | CL               | 8                        | 50                             | $\pm 7$      | 15         | —                                      | —        | —              | 1.6            | 800                        | 5                        | 0.4                      | 70                    | 3                                       | 4                            | 2               | —         | —        | buffer                                  |
| CLC400  | CL               | 6                        | 25                             | $\pm 7$      | 15         | —                                      | —        | —              | 0.5            | 700                        | 10                       | 1.6                      | 70                    | 3                                       | 4                            | 2               | —         | —        | 0.1% in 15ns                            |
| SL541   | PL               | 5                        | 25                             | $\pm 12, -6$ | 16         | —                                      | —        | —              | —              | 175                        | 50                       | —                        | 6.5                   | —                                       | —                            | —               | —         | —        | —                                       |
| VA707   | VT               | 6                        | 1.1                            | $\pm 6$      | 7          | —                                      | —        | —              | 3              | 105                        | 150 <sup>k</sup>         | 9                        | $\pm 50$              | $\pm 2.7$                               | $\pm 4$                      | 5               | —         | —        | decomp, G <sub>V</sub> >12 <sup>p</sup> |
| MSK737  | KE               | 11                       | 0.25                           | $\pm 18$     | 37         | —                                      | —        | —              | —              | 2800                       | 35                       | —                        | $\pm 120$             | $\pm 6$                                 | $\pm 12$                     | 10              | —         | —        | —                                       |
| MSK738  | KE               | 0.1                      | 0.06                           | $\pm 18$     | 37         | —                                      | —        | —              | 3              | 3200                       | 25                       | 2.5                      | $\pm 120$             | $\pm 6$                                 | $\pm 12$                     | 10              | —         | —        | precision, low I <sub>b</sub>           |
| MSK739  | KE               | 0.03 <sup>t</sup>        | 75pA <sup>1</sup>              | $\pm 18$     | 25         | —                                      | —        | —              | —              | 5500                       | 8                        | —                        | $\pm 120$             | $\pm 6$                                 | $\pm 12$                     | 10              | —         | —        | precision, FET                          |
| AD844A  | AD               | 0.4                      | 0.3                            | $\pm 18$     | 7          | 0.3                                    | —        | —              | 2              | 2000                       | 70 <sup>k</sup>          | —                        | $\pm 50$              | $\pm 2.5$                               | $\pm 7.5$                    | 5               | —         | —        | —                                       |
| AD846A  | AD               | 0.2                      | 0.5                            | $\pm 18$     | 5          | 15                                     | —        | —              | 2              | 450                        | 50                       | —                        | 50                    | 5                                       | 10                           | 3               | —         | —        | 0.01% in 80ns                           |
| 1467    | TP               | 0.1                      | 25                             | $\pm 16$     | 35         | —                                      | —        | —              | 2              | 300                        | 10                       | 40                       | $\pm 10$              | —                                       | —                            | —               | —         | —        | precision                               |
| EL2003  | EL               | 40                       | 25                             | $\pm 15$     | 10         | —                                      | —        | —              | —              | 1000                       | —                        | —                        | 6                     | 12                                      | 2                            | —               | —         | —        | buffer                                  |
| EL2020C | EL               | 10                       | 15                             | $\pm 18$     | 9          | 40                                     | —        | —              | —              | 500                        | 50                       | 5                        | $\pm 30$              | $\pm 1.5$                               | $\pm 5$                      | 3               | —         | —        | good data sheet                         |
| EL2022  | EL               | 2.5                      | 20                             | $\pm 20$     | 18         | 15                                     | —        | —              | 1.3            | 1900                       | 22 <sup>k</sup>          | 2.1                      | $\pm 100$             | $\pm 5$                                 | $\pm 11$                     | 5               | E         | —        | good data sheet                         |
| SL2541  | PL               | 10 <sup>1</sup>          | 20                             | $\pm 12, -5$ | 25         | —                                      | —        | —              | 3.5            | 1400                       | 30                       | 1.6                      | 15                    | —                                       | —                            | —               | —         | —        | —                                       |
| CA3450  | RC               | 15                       | 0.35                           | $\pm 7$      | 30         | —                                      | —        | —              | —              | 400                        | 35                       | —                        | 75                    | $\pm 4$                                 | —                            | 6               | —         | —        | —                                       |
| AD5539S | AD               | 3                        | 13                             | $\pm 8$      | 14         | —                                      | —        | —              | —              | 600                        | 12                       | —                        | 15                    | —                                       | $\pm 2.3$                    | 1               | —         | —        | —                                       |
| LM6364  | NS               | 6                        | 5                              | $\pm 18$     | 5          | —                                      | —        | —              | 3              | 300                        | 100 <sup>k</sup>         | —                        | $\pm 30$              | $\pm 1.5$                               | $\pm 5$                      | 3               | —         | —        | decomp, G <sub>V</sub> >5 <sup>r</sup>  |
| AD9610  | AD               | 1                        | 50                             | $\pm 18$     | 21         | 15                                     | —        | —              | 2              | 3500                       | 18 <sup>k</sup>          | 3.5                      | $\pm 50$              | $\pm 2.5$                               | $\pm 8$                      | 5               | —         | —        | —                                       |
| AD9611  | AD               | 3                        | 5                              | $\pm 6$      | 74         | 5                                      | —        | —              | 3              | 1900                       | 13                       | 1.4                      | $\pm 50$              | $\pm 2.5$                               | $\pm 3$                      | 5               | —         | —        | —                                       |
| 9826    | OE               | 20                       | 50                             | $\pm 15$     | 15         | —                                      | —        | —              | 2              | 1000                       | 10                       | —                        | 100                   | $\pm 2.5$                               | $\pm 7.5$                    | —               | —         | —        | hybrid                                  |
| SL9999  | PL               | 15                       | 18                             | $\pm 12, -5$ | 35         | —                                      | —        | —              | —              | 1300                       | 24                       | —                        | $\pm 50$ <sup>1</sup> | —                                       | —                            | —               | —         | —        | "A/D driver"                            |

(a) see footnote to Table 4.1. (b) input current ( $\mu$ A) if current-feedback type. (c) nV/ $\sqrt{\text{Hz}}$  at 10kHz. (d) full-swing output. (e) number of 150 $\Omega$  (double-terminated video) loads with video levels. (f) E - external. (g) at 1kHz. (h) G<sub>V</sub> = +2, 50 $\Omega$  input. (i)  $\pm 3\text{V}$  out. (j) typical, over full temp. (k) 0.1%. (l) programmable. (m) min/max. (n)  $\mu\text{V}$  rms, 10Hz to 100kHz. (o) above 5MHz. (p) VA706 for G = 1; VA708 for G > 3. (q) 0.5MHz. (r) 6361 for G=1, 6365 for G>25. (s) typical.

(a) see footnote to Table 4.1. (b) input current ( $\mu$ A) if current-feedback type. (c) nV/ $\sqrt{\text{Hz}}$  at 10kHz. (d) full-swing output. (e) number of 150 $\Omega$  (double-terminated video) loads with video levels. (f) E - external. (g) at 1kHz. (h) G<sub>v</sub> = +2, 50 $\Omega$  input. (i)  $\pm 3$ V out. (j) typical, over full temp. (k) 0.1%. (l) programmable. (m) min/max. (n)  $\mu$ V rms, 10Hz to 100kHz. (o) above 5MHz. (p) VA706 for G = 1; VA708 for G > 3. (q) 0.5MHz. (r) 6361 for G = 1; 6365 for G > 25. (s) typical.

and operate at fixed gain), these are true op-amps (you use external feedback to determine the configuration and gain), and they are usable as closed-loop amplifiers to 100MHz or more. Many of these wide bandwidth op-amps use the vertical *pnp* process to get good performance. In a deviation from normal op-amp practice, you will often find unsymmetrical input impedances, with *current feedback*. Take a look at Table 13.2 for a sampling of these fast beasts.

## RADIOFREQUENCY CIRCUIT ELEMENTS

### 13.09 Transmission lines

Before proceeding to the subject of communications circuits, it is necessary to deal briefly with the interesting subject of transmission lines. You have met these earlier in connection with digital signal communications in Chapter 9, where we introduced the ideas of characteristic impedance and line termination. Transmission lines play a central role in radiofrequency circuits, where they are used to pipe signals around from one place to another within a circuit, and often to an antenna system. Transmission lines provide one of the most important exceptions to the general principles (see Chapter 1) that a signal source ideally should have a source impedance small compared with the impedance of the load being driven and that the load should present an input impedance large compared with the source impedance driving it. The equivalent rule for transmission lines is that the load (and possibly the source) should present an impedance equal to the characteristic impedance of the line. The line is then "matched."

Transmission lines for signals of moderate frequency (up to 1000MHz, say) come in two major types: parallel conductors and coaxial line. The former is typified by the inexpensive molded 300 ohm "twin

lead" used to bring the signal from a television antenna to the receiver, and the latter is widely used in short lengths with BNC fittings to carry signals between instruments. (Fig. 13.17).

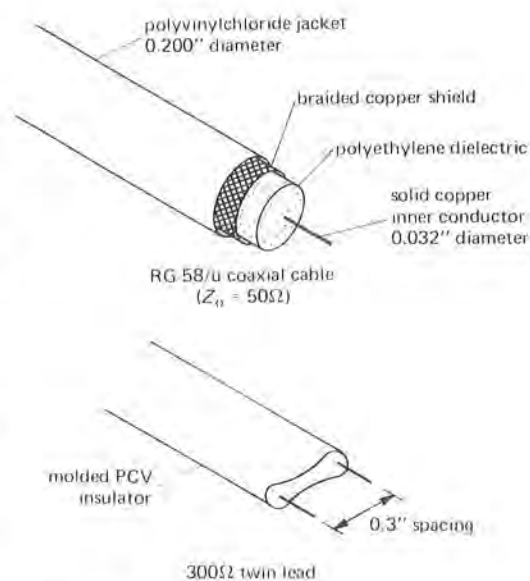


Figure 13.17

In the domain of ultra-high-frequency circuitry there are "stripline" techniques that involve parallel-conductor transmission lines as part of the actual circuit, and at the higher "microwave" frequencies (upwards of 2GHz, say) conventional lumped circuit elements and transmission lines are replaced by cavity and waveguide techniques, respectively. Except at these extremes of frequency, the familiar coaxial cable is probably the best choice for most radiofrequency applications. Compared with parallel conductor line, a properly matched coax line has the advantage of being totally shielded, i.e., there is no radiation or pickup of external signals.

#### Characteristic impedance and matching

A transmission line, whatever its form, has a "characteristic impedance"  $Z_0$ , meaning

that a wave moving along the line has a ratio of voltage to current equal to  $Z_0$ . For a lossless line,  $Z_0$  is resistive and equal to the square root of  $L/C$ , where  $L$  is the inductance per unit length and  $C$  is the capacitance per unit length. Typical coaxial lines have impedance in the range of 50 to 100 ohms, whereas parallel-conductor lines have impedances in the range of 300 to 1000 ohms.

When used with high-frequency (or short-rise-time) signals, it is important to "match" the load to the characteristic impedance of the line. The important facts are the following: (a) A transmission line terminated with a load equal to its characteristic impedance (resistance) will transfer an applied pulse to the termination without reflection. In that case all the power in the signal is transferred to the load. (b) The impedance looking into such a terminated line, at any frequency, is equal to its characteristic impedance (Fig. 13.18).

This is surprising at first, since at low frequencies you tend to think of a length of coax as a small capacitive load, generally a pretty high (capacitive) impedance. Also, at low frequencies (wavelength  $\gg$  length of cable) there is no need to match the line's impedance, provided you can handle the capacitance (typically 30pF per foot). If the cable is terminated with a resistor, on the other hand, it magically becomes a pure resistance at all frequencies.

### Mismatched transmission lines

A mismatched transmission line has some interesting, and occasionally useful, properties. A line terminated in a short circuit produces a reflected wave of opposite polarity, with the delay time of the reflected wave determined by the electrical length of the line (the speed of wave propagation in coax lines is about two-thirds the speed of light, because of the solid dielectric spacing material). You can see the reason for this, since the short circuit enforces a point

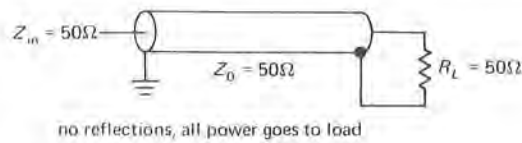


Figure 13.18

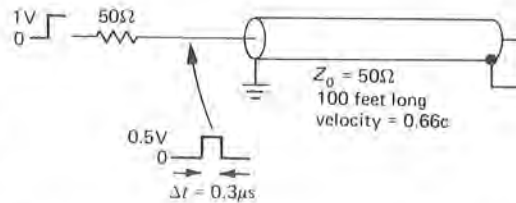


Figure 13.19. Pulse generation with shorted transmission line (inverted reflection).

of zero voltage at the end; the cable produces this obligatory condition by creating a wave of opposite phase at the short. In similar manner, an open-circuited cable (boundary condition of zero *current* at the end) produces a noninverted reflection of amplitude equal to the applied signal.

This property of a shorted cable is sometimes exploited to generate a short pulse from a step waveform. The step input is applied to the cable input through a resistance equal to  $Z_0$ , with the other end of the cable shorted. The waveform at the input is a pulse of width equal to the round-trip travel time, since the reflected step cancels the input (Fig. 13.19).

Cables terminated with a resistance  $R$  unequal to  $Z_0$  also produce reflections, although of lesser amplitude. The reflected wave is inverted if  $R < Z_0$  and uninverted if  $R > Z_0$ . The ratio of reflected wave amplitude to incident wave amplitude is given by

$$A_r/A_i = (R - Z_0)/(R + Z_0)$$

### Transmission lines in the frequency domain

Looked at in the frequency domain, a transmission line matched at the far end looks like a load of impedance  $Z_0$ , i.e., a



pure resistance if line losses are neglected. That makes sense, because it just swallows any wave you apply, all the power going into the matching resistor. This is true independent of cable length or wavelength. It is when you deal with mismatched lines that things begin to get interesting in the frequency domain. Since, for a given line length, the reflected wave arrives back at the input with a phase (relative to the applied signal) that depends on applied frequency, the impedance seen looking into the input depends on the mismatch and on the electrical length of the transmission line, in wavelengths.

As an example, a line that is an odd number of quarter wavelengths long terminated in an impedance  $Z_{\text{load}}$  at the far end presents an input impedance  $Z_{\text{in}} = Z_0^2/Z_{\text{load}}$ . If the load is resistive, the input will look resistive. On the other hand, a line that is an integral number of half wavelengths long presents an input impedance equal to its terminating impedance (Fig. 13.20).

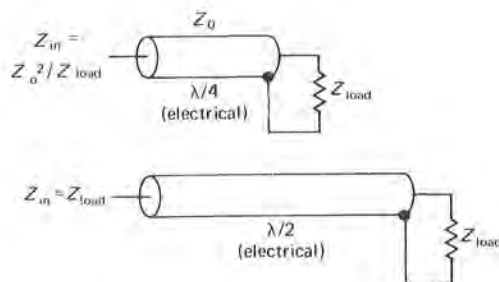


Figure 13.20

The presence of reflected signals on a transmission line is not necessarily bad. For operation at a single frequency, a mismatched line can be driven (through a line tuner) in such a way as to match its resultant input impedance, often with only negligibly greater line losses (due to higher voltages and currents for the same forward power) than with a matched load. But

a mismatched line has different properties at different frequencies (the famous “Smith chart” can be used to determine transmission-line impedances and “standing-wave ratio,” or SWR, a measure of the amplitude of reflected waves), making it undesirable for broadband or multifrequency use. In general, strive to terminate a transmission line in its characteristic impedance, at least at the receiving end.

### □ 13.10 Stubs, baluns, and transformers

There are some interesting applications of transmission lines that exploit the properties of mismatched sections or generally use sections of line in an unconventional way. The simplest is the quarter-wave matching section, which exploits the relationship  $Z_{\text{in}} = Z_0^2/Z_{\text{load}}$ . This can be rearranged to read  $Z_0 = (Z_{\text{in}}Z_{\text{load}})^{1/2}$ . In other words, a quarter-wave section can be used to match any two impedances by choosing the characteristic impedance of the matching section appropriately.

In a similar manner, a short length of transmission line (a “stub”) can be used to “tune” a mismatched load by simply putting the stub across or in series with the mismatched line, choosing the stub length and termination (open or shorted) and its position along the mismatched line correctly. In this sort of application the stub is really functioning as a circuit element, not a transmission line. At very short wavelengths the use of sections of transmission line as circuit elements is common (Fig. 13.21).

Sections of transmission line (or a transformer made with several interconnected windings) can be used to construct a “balun,” a device for matching an unbalanced line (coax) to a balanced load (e.g., an antenna). There are simple configurations for making fixed-impedance transformations at the same time (1:1 and 4:1 are common). Perhaps the nicest circuit

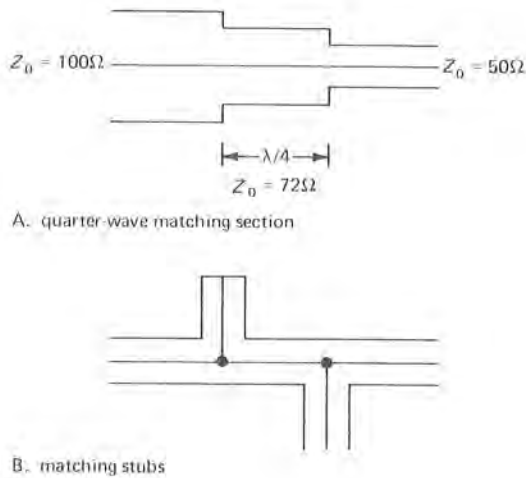


Figure 13.21

element made from transmission line is the broadband transmission-line transformer. These gadgets consist simply of a few turns of miniature coax or twisted-pair wound on a ferrite core, suitably interconnected. They avoid the high-frequency limitations of conventional transformers (caused by the resonant combination of “parasitic” winding capacitance and inductance) because the coils are arranged so that the winding capacitance and inductance form a transmission line, free of resonances. They can provide various impedance transformations with excellent broadband performance (e.g., less than 1dB loss from 0.1MHz to 500MHz), a property not shared by transformers constructed from simple coupled inductors. Transmission-line transformers are available from the Vari-L Co. and Mini-Circuits, among others, as packaged modules. Figure 13.22 shows a few examples of baluns and a transmission-line transformer.

### 13.11 Tuned amplifiers

In radiofrequency circuits intended for communications, or for other applications where the operating frequency is confined

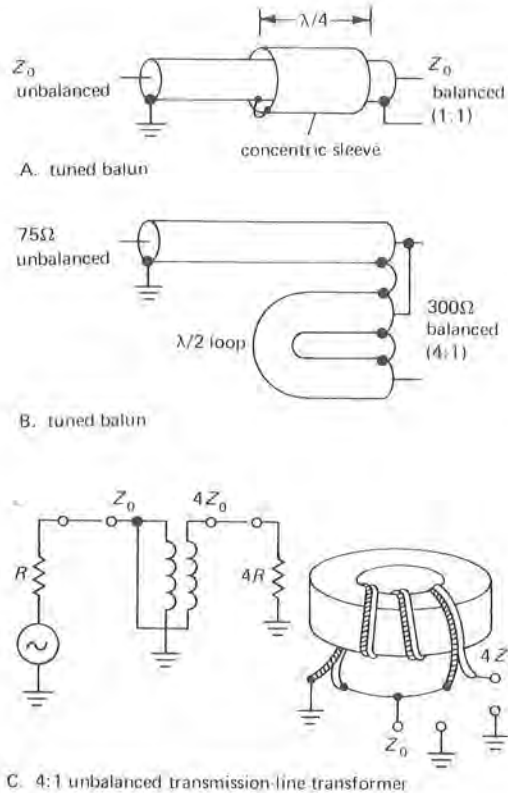


Figure 13.22. Transmission-line transformers.

to a narrow range, it is common to use tuned LC circuits as collector or drain loads. This has several advantages: (a) higher single-stage gain, since the load presents a high impedance at the signal frequency ( $G_V = g_m Z_{load}$ ) while allowing arbitrary quiescent current; (b) elimination of the undesirable loading effects of capacitance, since the LC circuit “tunes out” any capacitance by making it part of the tuned circuit capacitance; (c) simplified interstage coupling, since an LC circuit can be tapped or transformer-coupled (or even configured as a resonant matching network, as in the popular pi network) to achieve any desired impedance transformation; (d) elimination of out-of-band signals and noise owing to the frequency selectivity of the tuned circuits.

### Examples of tuned RF circuits

You will see tuned RF amplifiers in their natural element when we discuss communications circuits shortly. At this point we would simply like to illustrate the use of tuned circuits in oscillators and amplifiers with a few examples. Figure 13.23 illustrates the classic tuned amplifier. A

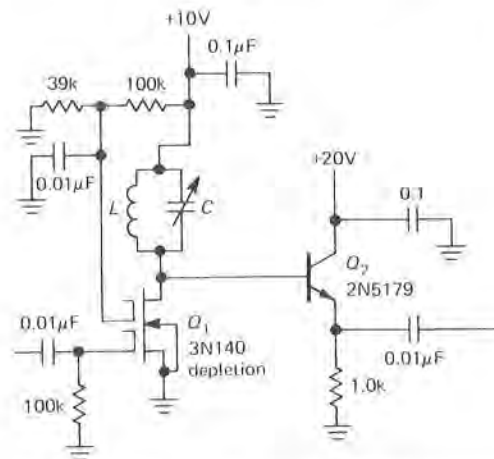


Figure 13.23. Dual-gate MOSFET (cascode) tuned amplifier.

dual-gate depletion-mode FET is used to eliminate the problems of Miller effect, since the input is untuned. By operating the lower gate at dc ground, the stage runs at  $I_{DSS}$ . The parallel tuned LC sets the center frequency of amplification, with output buffering via  $Q_2$ . Since the drain sits at +10 volts, the output follower requires a higher collector voltage. This sort of circuit has quite high voltage gain at resonance, limited by the LC circuit  $Q$  and loading by the follower.

In the circuit shown in Figure 13.24 a carefully constructed and tunable LC circuit is used to set the frequency of an oscillator. This is known as a VFO (variable-frequency oscillator); it is used as the tunable element in some transmitters and receivers, as well as variable-frequency RF signal sources. In this circuit

a JFET provides the necessary power gain, with positive feedback from the source coupled into a “link” on  $L_1$ . The link has fewer turns than the inductor, providing voltage gain and therefore oscillation. By adding a varactor diode, which acts as a voltage-variable capacitor (see Fig. 5.44), you can make such an oscillator voltage-tunable. Note the use of a feed-through capacitor and decoupling RF chokes on the power-supply leads; this is nearly universal practice in radiofrequency circuits.

The circuit in Figure 13.25 is a 200MHz common-emitter transistor amplifier stage. This circuit illustrates neutralization, the technique of canceling out capacitively coupled signal from output to input by adding a current of the opposite phase.  $C_N$  is the neutralizing capacitor, driven from the bottom of the collector “tank” circuit where the phase is opposite to that at the collector. This circuit matches the output impedance to the line by tapping down on the “tank” (the collector LC circuit), a simple but inflexible method.

This last circuit (Fig. 13.26) is a 25kW RF amplifier, using a zero-bias grounded-grid triode. Vacuum tubes are still used in high-power frequency amplifiers, because no solid-state device can match their performance (for example, the 8973 power triode delivers 1.5MW at 50MHz!). The grounded-grid configuration requires no neutralization. The output circuit is the popular pi network, driven by blocking capacitor  $C_8$ .  $C_9$ ,  $L_4$ , and  $C_{10}$  form the actual network, with their values determined by the desired resonant frequency, impedance transformation, and loaded  $Q$  ( $Q$ , or quality factor, is a measure of the sharpness of resonance, see Section 1.22). The RF choke at the output prevents dc voltage appearing there, and the plate RF choke is used to apply plate voltage while allowing signal swing at the operating frequency.

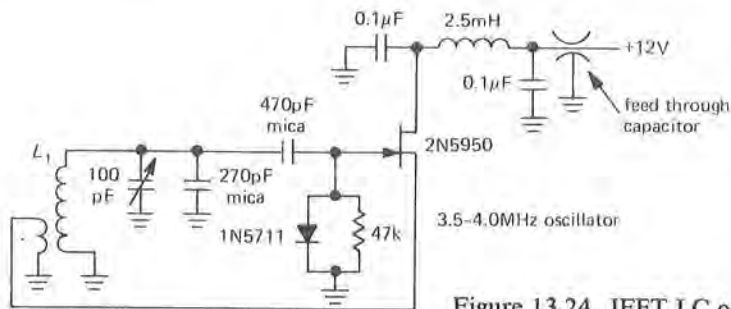


Figure 13.24. JFET LC oscillator.

### 13.12 Radiofrequency circuit elements

In radiofrequency circuits you meet several kinds of specialized circuit modules that have no equivalents in low-frequency circuitry. Before going on to consider RF communications circuits, we will mention some of these, since they enjoy widespread use in the generation and detection of radiofrequency signals.

#### Oscillators

If great stability is not important, a simple  $LC$  oscillator of the type just illustrated will generate a radiofrequency signal with adjustability over an octave or more by varying either  $C$  or  $L$  (the latter is sometimes called a "permeability-tuned oscillator," or PTO). With careful design and attention to detail during construction, such a VFO (variable-frequency

oscillator) can have drifts of less than a few parts per million measured over hours and will be quite satisfactory for use in receivers and noncritical transmitting applications.  $LC$  oscillators can be operated at frequencies ranging from audiofrequencies to hundreds of megahertz.

Just as with the amplifier modules we mentioned in Section 13.08, prepackaged oscillator modules with excellent performance are readily available. Tunable oscillator modules use varactors (voltage-variable capacitors) to adjust the operating frequency of an  $LC$  oscillator in response to an externally applied voltage. A fancier version of tunable oscillator for gigahertz frequencies uses a yttrium/iron/garnet (YIG) sphere as a magnetically tunable resonant cavity; YIG-tuned oscillators provide high spectral purity and tuning linearity. A recent technique for making inexpensive oscillators of good stability

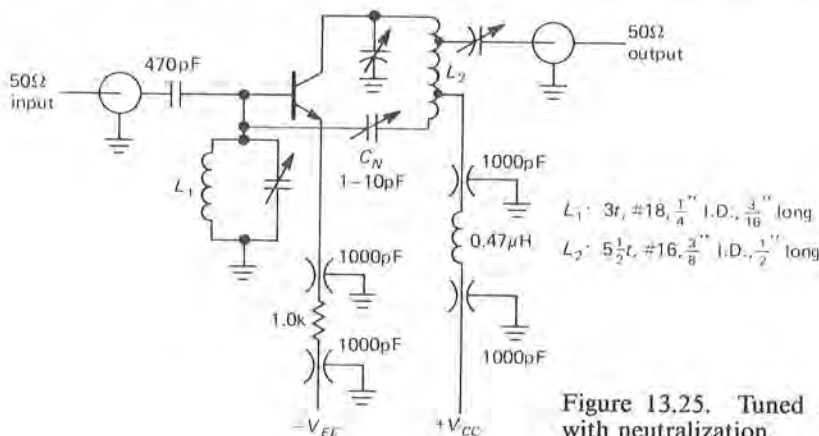


Figure 13.25. Tuned 200MHz RF amplifier with neutralization.

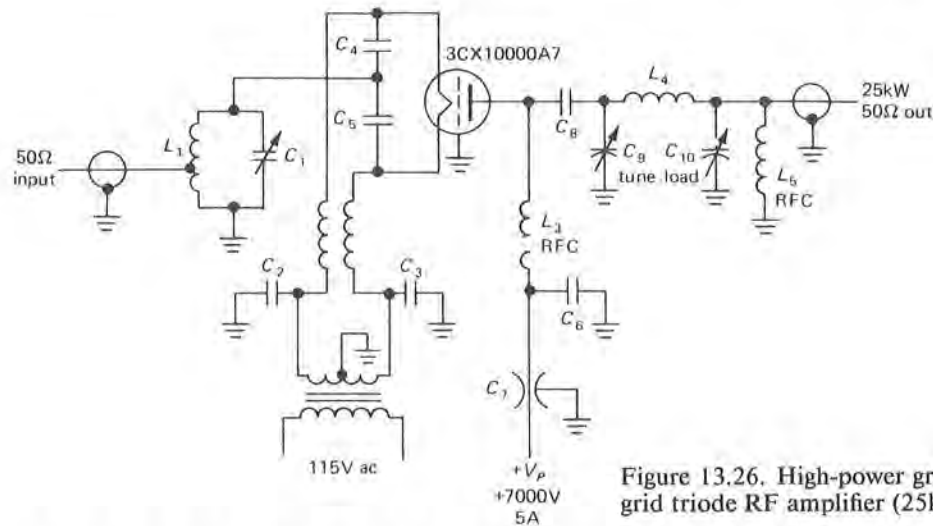


Figure 13.26. High-power grounded-grid triode RF amplifier (25kW output).

in the gigahertz range uses a dielectric “pill” resonator as the feedback element for a GaAs FET (or bipolar) oscillator. Oscillators using this “dielectrically stabilized” technique are simple and stable and have low noise.

For high stability, the best kind of oscillators use quartz crystals to set the operating frequency. With off-the-shelf garden-variety crystals, you can expect overall stabilities of a few parts per million, with tempco of order 1 ppm/degree or better. A temperature-compensated crystal oscillator (TCXO), which uses capacitors of controlled tempco to offset the crystal’s frequency variation, can deliver frequency stability of 1 ppm over a temperature range of 0°C to 50°C or better. For the utmost in performance, oscillators with the crystal maintained in a constant-temperature “oven” are available, with stabilities of a few parts per billion over time and temperature. Even the so-called atomic oscillators (rubidium, cesium) actually use a high-stability quartz oscillator as the basic oscillating element, with frequency adjusted as necessary to agree with a particular atomic transition frequency.

Crystal oscillators are commercially available in frequencies ranging from about 10kHz to about 100MHz in all of the

variations just mentioned. There are even little DIP and transistor can (TO-5) oscillators, with logic outputs. Only a slight electrical adjustment of frequency is possible, so the frequency must be specified when the oscillator or crystal is ordered.

To get both adjustability and high stability, a frequency synthesizer is the best choice. It uses tricks to generate any desired frequency from a single source of stable frequency, typically a 10MHz crystal oscillator. A synthesizer driven from a rubidium standard (stability of a few parts in  $10^{12}$ ) makes a nice signal source.

### Mixers/modulators

A circuit that forms the product of two analog waveforms is used in a variety of radiofrequency applications and is called, variously, a modulator, mixer, synchronous detector, or phase detector. The simplest form of modulation, as you will see shortly, is amplitude modulation (AM), in which the high-frequency *carrier* signal is varied in amplitude according to a slowly varying *modulating* signal. A multiplier obviously performs the right function. Such a circuit can also be used as a variable gain control, thinking of one of the inputs as a dc



voltage. There are convenient ICs to do this job, e.g., the MC1495 and MC1496.

A mixer is a circuit that accepts two signal inputs and forms an output signal at the sum and difference frequencies. From the trigonometric relationship

$$\begin{aligned}\cos \omega_1 t \cos \omega_2 t \\ = \frac{1}{2} \cos(\omega_1 + \omega_2)t + \frac{1}{2} \cos(\omega_1 - \omega_2)t\end{aligned}$$

it should be clear that a “four-quadrant multiplier,” i.e., one that performs the product of two input signals of any polarity, is in fact a mixer. If you input two signals of frequency  $f_1$  and  $f_2$ , you will get out signals at  $f_1 + f_2$  and  $f_1 - f_2$ . A signal at frequency  $f_0$  mixed with a band of signals near zero frequency (band-limited to a maximum frequency of  $f_{\max}$ ) will produce a symmetrical band of frequencies around  $f_0$ , extending from  $f_0 - f_{\max}$  to  $f_0 + f_{\max}$  (the spectrum of amplitude modulation, see Section 13.15).

It is not necessary to form an accurate analog product in order to mix two signals. In fact, any nonlinear combination of the two signals will produce sum and difference frequencies. Take, for instance, a “square-law” nonlinearity applied to the sum of two signals:

$$\begin{aligned}(\cos \omega_1 t + \cos \omega_2 t)^2 \\ = 1 + \frac{1}{2} \cos 2\omega_1 t + \frac{1}{2} \cos 2\omega_2 t \\ + \cos(\omega_1 + \omega_2)t + \cos(\omega_1 - \omega_2)t\end{aligned}$$

This is the sort of nonlinearity you would get (roughly) by applying two small signals to a forward-biased diode. Note that you get harmonics of the individual signals, as well as the sum and difference frequencies. The term “balanced mixer” is used to describe a circuit in which only the sum and difference signals, not the input signals and their harmonics, are passed through to the output. The four-quadrant multiplier is a balanced mixer, whereas the nonlinear diode is not.

Among the methods used to make mixers are the following: (a) simple nonlinear transistor or diode circuits, often using Schottky diodes; (b) dual-gate FETs, with one signal applied to each gate; (c) multiplier chips like the MC1495, MC1496, SL640, or AD630; (d) balanced mixers constructed from transformers and arrays of diodes, generally available as packaged “double-balanced mixers.” The latter are typified by the popular M1 series of double-balanced mixers from Watkins-Johnson spanning the frequency range to 4000MHz with 20dB to 50dB of signal isolation, or the inexpensive SBL-1 mixer (1–500MHz) from Mini-Circuits Lab. Mixers are widely used in the generation of radio-frequency signals at arbitrary frequencies; they let you shift a signal up or down in frequency without changing its spectrum. You will see how it all works shortly.

The equations above show that the simple quadratic-law mixer produces outputs of equal amplitudes at both sum and difference frequencies. In communications applications (e.g., the “superheterodyne” receiver), where mixers are often used to shift frequency bands, it is sometimes desirable to suppress one of those mixer products. We’ll see in Section 13.16 how to make such an *image-reject* mixer.

### Frequency multipliers

A nonlinear circuit often is used to generate a signal at a multiple of the input signal’s frequency. This is particularly handy if a signal of high stability is required at a very high frequency, above the range of good oscillators. One of the most common methods is to bias an amplifier stage for highly nonlinear operation, then use an LC output circuit tuned to some multiple of the input signal; this can be done with bipolar transistors, FETs, or even tunnel diodes. A multiplier like the 1496 can be used as an efficient doubler at low radiofrequencies by connecting the input

signal to both inputs, thus forming the square of the input waveform. The square of a sine wave contains frequencies at the second harmonic only. Prepackaged frequency doublers that use balanced mixers are commercially available; they are very broadband (for example, the Watkins-Johnson FD25 accepts an input frequency from 5MHz to 2400MHz), with good suppression of both the input frequency (the "fundamental") and unwanted harmonics (typically 30dB). Exotic devices such as SNAP diodes and varactors are also used as multipliers. A frequency-multiplier circuit should include a tuned output circuit or should be followed by tuned amplifiers, since, in general, many harmonics of the input signal are generated in the nonlinear process.

#### Attenuators, hybrids, circulators

There are some fascinating passive devices that are used to control the amplitude and direction of radiofrequency signals passing between circuit modules. All of these are broadband transmission-line (or waveguide) components, meant to be inserted in a line of fixed impedance, usually 50 ohms. They are all widely available as modules.

The simplest is the attenuator, a device to reduce the amplitude of a signal. They come with a big knob and accurately calibrated steps of attenuation, or as voltage-controlled attenuators. The latter are simply balanced mixers with the control current serving as one of the multiplying inputs. Fixed attenuators (Fig. 13.27) are convenient for reducing signal levels between components, as you are assembling the modules of a 50 ohm radiofrequency system; they also reduce any impedance mismatch that may be present.

A hybrid (also known as a "rat race," magic T, 3dB coupler, or iso-T) is a clever transmission-line configuration with four ports. A signal fed into any port emerges

from the two closest ports, with specific phase shifts (usually  $0^\circ$  or  $180^\circ$ ). A hybrid that has one port terminated in its characteristic impedance is a 3-port "power splitter/combiner." Splitter/combiners can be cascaded to make multiport splitter/combiners. A close cousin of the hybrids is the directional coupler, a 3-port device that couples a small fraction of a passing wave out to a third port. Ideally there is no output at the third port for a wave passing through in the opposite direction.

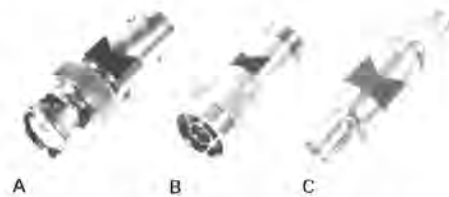


Figure 13.27. Fixed attenuators. (Courtesy of Merrimac Industries, Inc.)

- A. BNC.
- B. Type N.
- C. SMA.

The most magical devices in this general category are the circulators and isolators. They employ exotic ferrite materials and magnetic fields to achieve the impossible: a device that will transmit waves in only one direction. The isolator has two ports and allows transmission in one direction only. Circulators have three or more ports, and they transmit an incoming signal at any port to the next port in succession.

#### Filters

As you will see, frequency selectivity is often needed in the design of radiofrequency circuits. The simple tuned  $LC$  amplifier provides a good measure of selectivity, with the peakiness of the response adjustable via the  $Q$  factor of the  $LC$  circuit. The latter depends on losses in the inductor and capacitor, as well as loading by the associated circuitry.  $Q$  values as

high as several hundred can be easily obtained. At very high frequencies, lumped *LC* circuits are replaced by stripline techniques, and at microwave frequencies you use cavity resonators, but the basic idea remains the same. Tuned circuits can also be used to *reject* a particular frequency, if desired.

For applications where it is necessary to have a filter that passes a very narrow band of frequencies relatively unattenuated, with a sharp dropoff outside the limits of the band, a superior bandpass filter can be made from a set of piezoelectric (ceramic or quartz-crystal) or mechanical resonators. There are commercially available 8-pole and 16-pole crystal-lattice filters with center frequencies in the range of 1MHz to 50MHz and bandwidths ranging from as little as a few hundred hertz to several kilohertz. These filters are extremely important in setting receiver selectivity and in the generation of certain kinds of modulated signals. Recently, surface acoustic wave (SAW) filters have become popular and inexpensive; these, too, can have level passband characteristics with extremely steep skirts. This desirable characteristic is usually expressed as a “shape factor,” e.g., the ratio of  $-3\text{dB}$  bandwidth to  $-40\text{dB}$  bandwidth, with values as small as 1.1. In a typical application SAW filters are used in television receivers and cable systems to limit the received passband.

Of course, in situations where such narrow passbands are not needed, filters can be designed with multiple resonant *LC* sections. Appendix H shows some low-pass and high-pass *LC* filter examples.

### Detectors

The bottom line in the extraction of information from a modulated radiofrequency signal involves *detection*, the process of stripping the modulating signal from the “carrier.” There are several methods, de-

pending on the form of modulation (AM, FM, SSB, etc.), and we will discuss this important topic next, along with communications concepts.

### 13.13 Measuring amplitude or power

As we will soon see, detection of an AM signal is simply the generation of a voltage proportional to the instantaneous amplitude of the RF signal carrying the modulation. In many other applications as well (radioastronomy, laboratory RF measurements, signal generator “leveling,” filter design, surveillance, etc.) it’s important to be able to measure the amplitude or power of radiofrequency signals. Before moving on to communication, therefore, let’s take a moment to look at these circuits and methods.

#### Signal rectification

In Section 1.30 we showed how to use a simple diode to derive an output voltage proportional to the amplitude of a signal. Since the diode’s “knee” is somewhat soft, and offset by about 0.6 volt, we showed how to compensate the diode drop with a second diode. In Section 4.18 we showed how to circumvent the diode’s nonlinearity and offset by putting the diode in the feedback path of an op-amp, thus forming a precision rectifier (or absolute-value) circuit.

Each of these circuits has problems. The simple diode detector has the advantage of working over an enormous range of frequencies (up to gigahertz, with the right kind of diodes), but it is nonlinear at low levels. The use of Schottky (“hot-carrier”) diodes helps somewhat, because of their lower forward drop. You can improve the situation considerably by pre-amplifying the signal before rectification (this is used, for example, in the UTD-1000 amplifier/diode “level detector” from Avantek); but that limits dynamic range,

owing to amplifier saturation (the UTD-1000 has a 30dB range and operates from 10 to 1000MHz). The active rectifier, by contrast, is highly linear; but it works well only at relatively low frequencies, in common with most op-amp circuits. You can use fast op-amps (Section 13.08) to improve the situation, but you will still be limited to frequencies below 10MHz or so.

### Synchronous (homodyne) detection

An interesting method that combines dynamic range, accuracy, and speed is *synchronous detection*, also called “homodyne detection.” In this method (Fig. 13.28) the output is rectified by being inverted during alternate half cycles. This obviously requires a clean signal at the same frequency as the signal being detected, either supplied externally or regenerated internally with a phase-locked loop (Section 9.27). Synchronous detection works well up to frequencies of a few megahertz, at least; the big disadvantage is the need for a coherent reference signal. We’ll see this same circuit in Section 15.15, in the guise of a *phase detector* (in which form it also made a brief appearance in Section 9.27).

### Current-source drive

Another solution to the problem of rectifier diode nonlinearity is to drive the rectifier circuit with a *current*, rather than a voltage; the output is then loaded resistively, to produce a proportional output voltage

(Fig. 13.29). Figure 13.30 shows a good implementation with voltage-controlled transistor current source; the performance is shown in Figure 13.31. You can understand this circuit in another way: With no input signal, the amplifier’s output is decoupled from the rectifier network, producing very high voltage gain (with its current-sink load); thus, only a very small input signal is needed to turn on the diodes. At that point the voltage gain drops to  $G_V = R_L / (R_E + r_e)$  (in this case,  $G_V \approx 3$ ), preventing saturation. With a wide-band amplifier and fast diodes, this circuit design will operate to 100MHz or more.

### Post-detection diode compensation

From Hewlett-Packard (HP Journal, 10/80) comes the circuit shown in Figure 13.32, which uses matched Schottky diodes in a clever arrangement in which each diode sees the same signal. Since the op-amps operate on the detected (low-frequency) signal, the bandwidth is limited only by the diode circuit. The designers of this circuit deserve a pat on the back (and three “attaboys”).

### Amplitude-tracking detector

Figure 13.33 shows another clever idea: Eliminate diode nonlinearities and offsets by using a locally generated signal, detected in a symmetrical circuit, to cancel the unknown current. Feedback adjusts

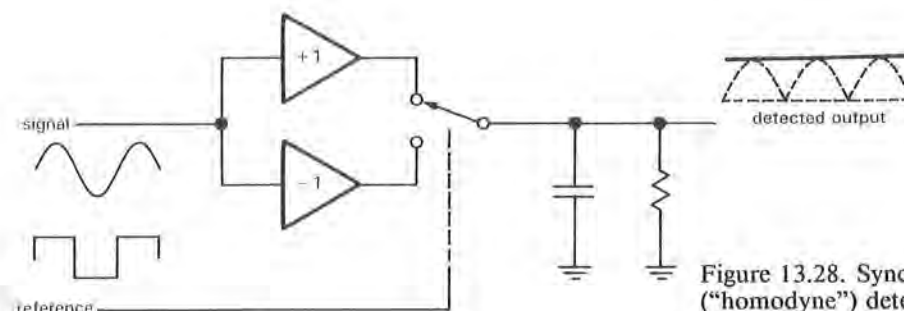


Figure 13.28. Synchronous (“homodyne”) detector.



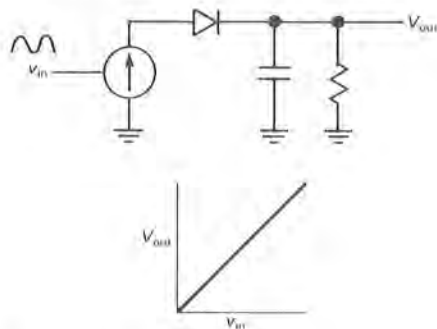


Figure 13.29

the amplitude of the local low-frequency signal until the rectified outputs balance. The frequency of the nulling signal is low enough so that its amplitude can be accurately measured with an op-amp precision rectifier. A good implementation of this circuit will operate linearly down to a few millivolts, and up to a gigahertz.

#### Power detection

The foregoing methods all measure the *amplitude* of a high-frequency signal. There

are many occasions on which you really want to know *power*. Of course, for a sine wave, the two are simply related by  $P = V_{\text{rms}}^2 / R_{\text{load}}$ , so you can simply convert a measured amplitude to a calculated power. For nonsinusoidal waveforms, however, a true power measurement can be made only by averaging the square of the actual voltage waveform. In the language of radiofrequency measurements, you need a “square-law detector.”

There are numerous methods available. For signals of low to moderate frequency, you do well with a “function module,” for example the AD637 monolithic rms-to-dc converter from Analog Devices. These devices use the exponential diode characteristic within feedback loops to form the square of the input signal, which is then low-pass filtered and fed to an analogous square-root circuit. You get excellent linearity and dynamic range, and good bandwidth, from these circuits. For example, AD637 has 8MHz bandwidth at full level, 0.02% nonlinearity, and 60dB dynamic range; it even has a logarithmic (dB) output.

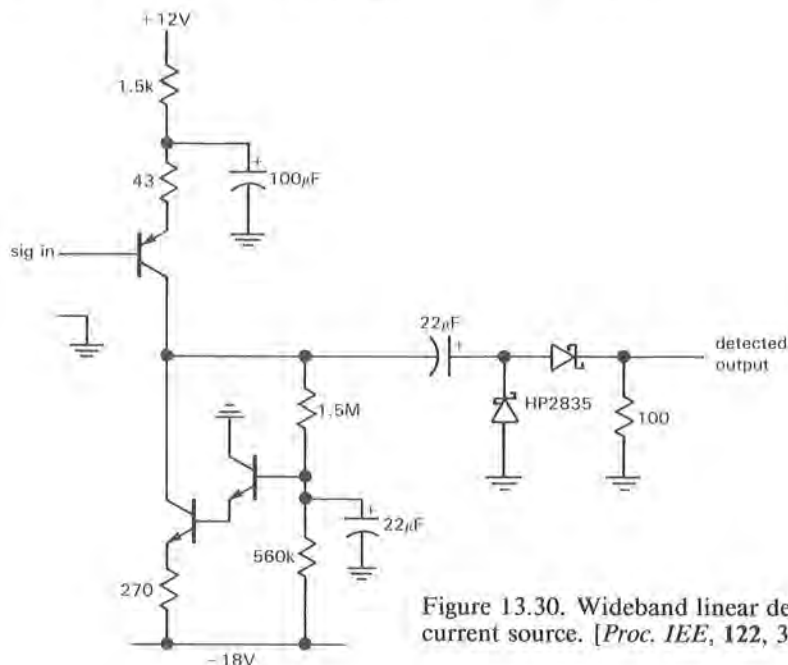


Figure 13.30. Wideband linear detector using current source. [*Proc. IEE*, 122, 3, 249 (1975).]



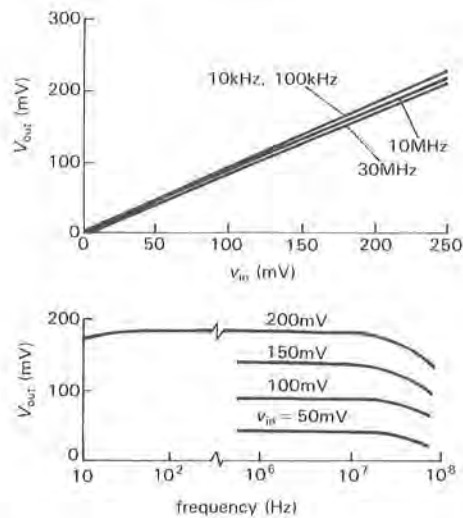


Figure 13.31. Performance curves for wide-band detector.

For frequencies above a few megahertz these square/square-root methods of rms conversion fail, owing to inadequate op-amp loop bandwidth. However, there are other methods you can use. Figure 13.34 shows a simple square-law detector circuit that uses a *back diode*, which is nothing but a tunnel diode (Section 1.06) used in the nontunneling direction (where it has zero forward voltage drop). We got this circuit from the radioastronomers at the Haystack Observatory and were amazed

at its extraordinary power linearity (Fig. 13.35).

An interesting square-law technique of considerable ancestry is the *bolometric* method, in which the input signal (perhaps amplified) is used to power a resistive heater, whose temperature is then measured. Since the heater power is accurately proportional to  $V^2$ , this method is intrinsically square-law. An example of a bolometric module is the LT1088 from Linear Technology. It uses a matched pair of resistive heaters coupled to a matched pair of temperature-measuring diodes. The input signal is applied to one heater, and feedback is applied to the reference heater to maintain the diodes at the same temperature. The reference heater's drive voltage is the output (Fig. 13.36).

The bolometric technique is intrinsically wideband, and accurately square-law. It has limited dynamic range, however, because it's hard to measure microscopic amounts of heat, and it's easy to blow it out with macroscopic amounts of heat! The exemplary LT1088, for example, works from dc to 300MHz, but has only 25dB of dynamic range. With careful bolometer design it is possible to extend the bandwidth to very high frequencies, and to wider dynamic range. The 432–438 series of bolometric power meters

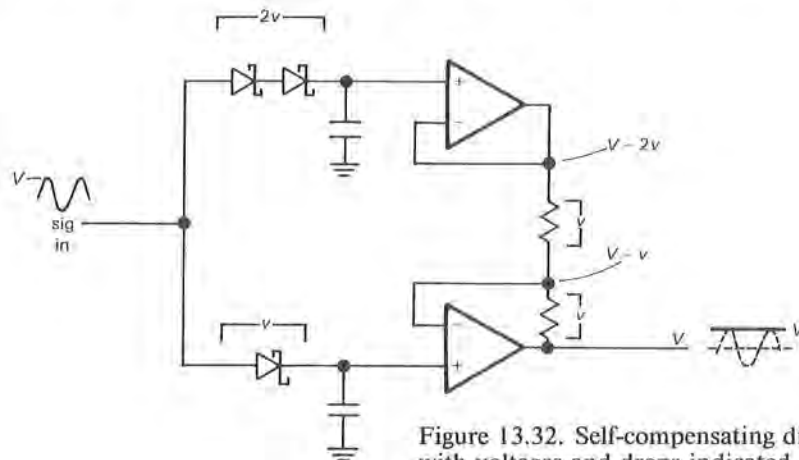


Figure 13.32. Self-compensating diode detector, with voltages and drops indicated.

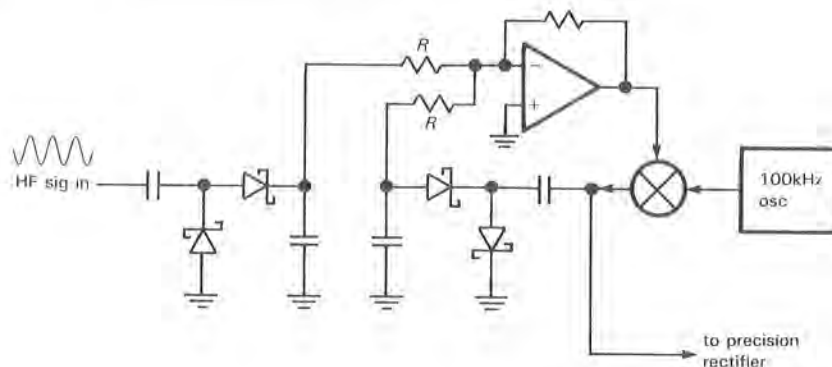


Figure 13.33.  
Amplitude-tracking  
detector.

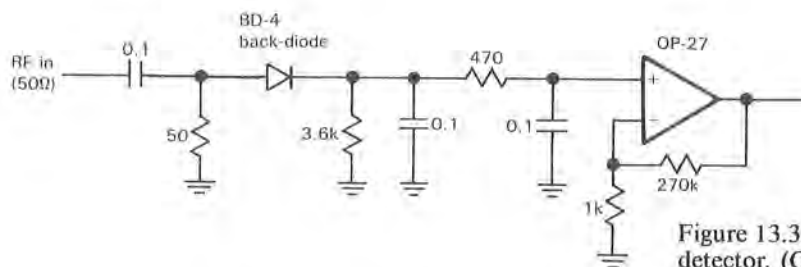


Figure 13.34. Back-diode square-law  
detector. (Courtesy of Alan Rogers,  
Haystack Observatory.)

from Hewlett-Packard span the frequency range from 100kHz to 50GHz, using a set of interchangeable bolometric power sensors. These cover the range +44 dBm (25W) to -70dBm (100pW), a total span of 114dB (factor of  $2.5 \times 10^{11}$  in power), although any one bolometer can accommodate at most a 50dB range.

#### RADIOFREQUENCY COMMUNICATIONS: AM

Since radiofrequency techniques find their greatest application in communications, it is important to understand how signals can be modulated and demodulated, i.e., how radiofrequencies are used to carry information from one place to another. Besides, how would you feel if, after taking a course in electronics, someone asked you how a radio works and you didn't know?

#### 13.14 Some communications concepts

In communications theory we speak of a communications "channel," a means of conveying information from A to B. For example, the channel might consist of a cable or an optical-fiber link. In radio-frequency communications the channel is the electromagnetic frequency spectrum, which, roughly speaking, extends from very low frequencies (VLF) of a few kilohertz, through the "short waves" of a few megahertz to a few tens of megahertz, the very high frequencies (VHF) and ultrahigh frequencies (UHF) extending up to several hundred megahertz, and the microwave region beginning at about 1GHz.

A signal, consisting of speech, say, is sent on a radiofrequency channel by having it modulate a radiofrequency "carrier." It is important to understand why this is done at all, rather than transmitting the speech directly. There are basically two

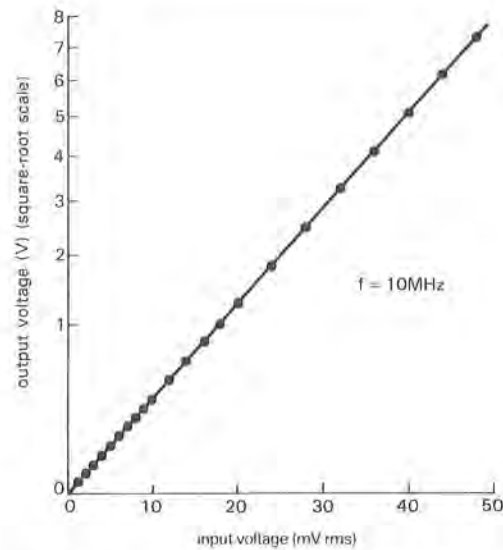


Figure 13.35. Performance of square-law detector.

reasons. First, if the information were transmitted at its natural band of frequencies with radio waves (in this case, in the VLF portion of the spectrum), any two signals would overlap and jam each other; i.e., by encoding the information onto carriers in separate portions of the spectrum,

it is possible to “frequency-multiplex” the signal, and thereby maintain many channels simultaneously. Second, some wavelengths are more conveniently generated and propagated than others. For instance, in the region from 5MHz to 30MHz, signals can travel around the world by multiple reflections from ionosphere, and at microwave frequencies antennas of modest size can form narrow beams. Consequently, the HF (short-wave) region is used for over-the-horizon communication, whereas microwaves are used for line-of-sight repeaters and radar.

There are several ways to modulate a carrier. Roughly speaking, all methods have in common the property that the modulated signal occupies a bandwidth at least comparable to the bandwidth of the modulating signal, i.e., the bandwidth of the information being sent. Thus, a high-fidelity audio transmission will occupy 20–40kHz of spectrum, regardless of the carrier frequency. A perfect unmodulated carrier has zero bandwidth and conveys no information. A transmission of low information content, e.g., telegraphy, occupies a

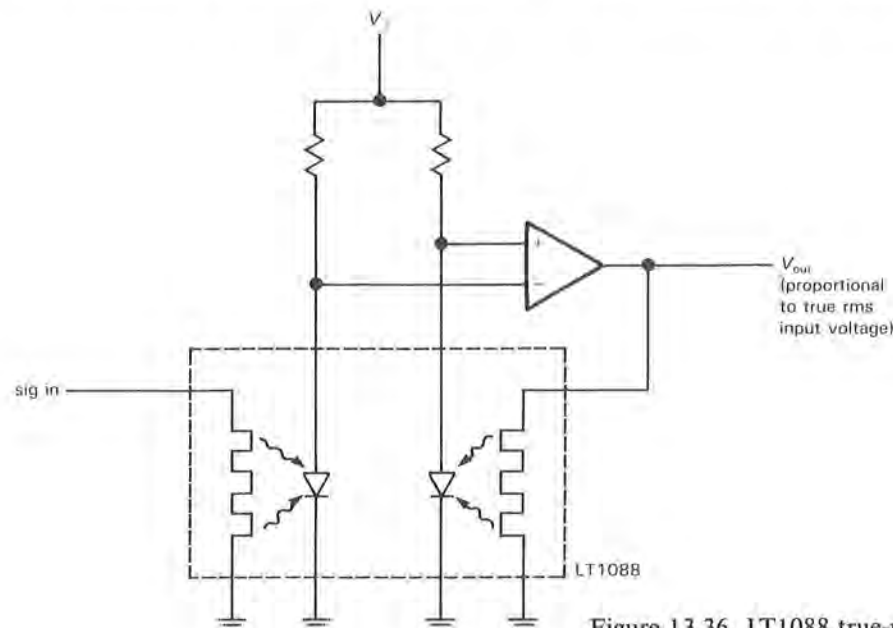


Figure 13.36. LT1088 true-rms detector.

relatively narrow slice of spectrum (perhaps 50–100Hz), whereas something like a television picture requires several megahertz. For completeness it should be pointed out that more information can be sent on a channel of given bandwidth if there is sufficiently high signal/noise ratio (SNR). Such “frequency compression” takes advantage of the fact that “channel capacity” equals bandwidth times  $\log_2(1 + \text{SNR})$ .

### 13.15 Amplitude modulation

Let’s begin with the simplest form of modulation (AM), taking a look at its frequency spectrum and methods of detection. Imagine a simple carrier,  $\cos \omega_c t$ , varied in amplitude by a modulating signal of much lower frequency,  $\cos \omega_m t$ , in the following way:

$$\text{signal} = (1 + m \cos \omega_m t) \cos \omega_c t$$

with  $m$ , the “modulation index,” less than or equal to 1. Expanding the product, you get

$$\begin{aligned} \text{signal} = & \cos \omega_c t + \frac{1}{2}m \cos(\omega_c + \omega_m)t \\ & + \frac{1}{2}m \cos(\omega_c - \omega_m)t \end{aligned}$$

i.e., the modulated carrier has power at frequency  $\omega_c$  and at frequencies on either side  $\omega_m$  away. Figure 13.37 shows the signal and its spectrum. In this case the modulation ( $m$ ) is 50%, and the two “sidebands” each contain  $\frac{1}{16}$  of the power contained in the carrier.

If the modulating signal is some complex waveform  $[f(t)]$ , like speech, the amplitude-modulated waveform is given by

$$\text{signal} = [A + f(t)] \cos \omega_c t$$

with the constant  $A$  large enough so that  $A + f(t)$  is never negative. The resulting spectrum simply appears as symmetrical sidebands around the carrier (Fig. 13.38).

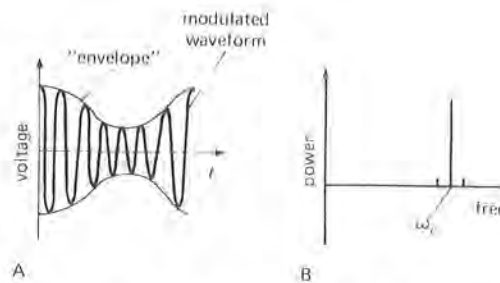


Figure 13.37. Amplitude modulation.

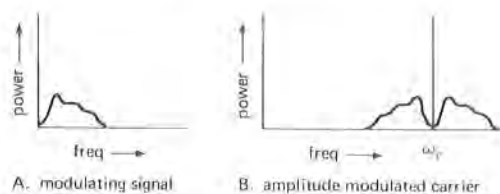


Figure 13.38. AM spectrum for a band of modulating frequencies (speech).

### AM generation and detection

It is easy to generate amplitude-modulated RF. Any technique that lets you control the signal amplitude with a voltage in a linear manner will do. Common methods involve varying the RF amplifier supply voltage (if the modulation is done at the output stage) or using a multiplier chip such as the 1496. When the modulation is done at a low-level stage, all following stages of amplification must be linear. Note that in AM the modulating waveform must be biased up so that it never assumes negative values. Look at the graphs in Figure 13.39.

The simplest receiver of AM consists of several stages of tuned RF amplification, followed by a diode detector (Fig. 13.40). The amplifier stages provide selectivity against signals nearby in frequency, and they amplify the input signals (which may be at the microvolt level) for the detector. The latter simply rectifies the RF waveform, then recovers the smooth “envelope” with low-pass filtering. The

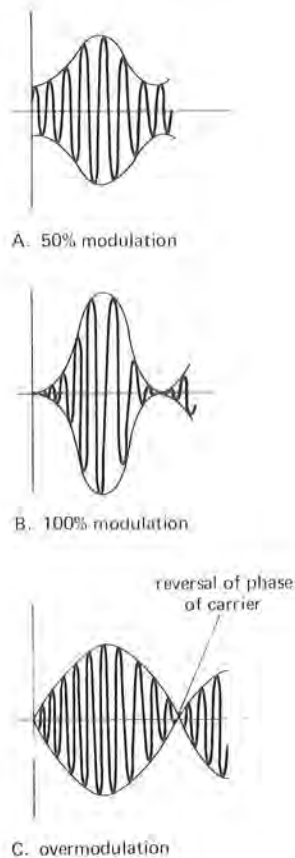


Figure 13.39

low-pass filter should reject RF while passing the audiofrequencies unattenuated. This simple scheme leaves much to be desired, as you will see. It is really just a glorified crystal set.

### 13.16 Superheterodyne receiver

A receiver consisting of a set of tuned RF amplifiers is undesirable for several

reasons. First of all, the individual amplifiers must be tuned to the same frequency, requiring either great coordination by someone with a lot of hands or extremely good tracking of a set of simultaneously tuned  $LC$  circuits. Second, since the overall frequency selectivity is determined by the combined responses of the individual amplifiers, the shape of the passband will depend on the accuracy with which the individual amplifiers are tuned; the individual amplifiers cannot have as sharp a response as would be desirable, since tuning would then be practically impossible. And since the signal being received can be at any frequency within the tuning range of the amplifiers, it isn't possible to take advantage of crystal-lattice filters to generate a flat passband with steep falloff on either side (steep "skirts"), a very desirable passband characteristic.

A nice solution to these problems is the superheterodyne ("superhet") receiver shown in Figure 13.41. The incoming signal is amplified with a single stage of tuned RF amplification, then mixed with an adjustable local oscillator (LO) to produce a signal at a fixed intermediate frequency (IF), in this case 455kHz. From then on the receiver consists of a set of fixed-tuned IF amplifiers, including selective elements such as crystal or mechanical filters, finally terminating in a detector and audio amplifier. Changing the LO frequency tunes the receiver, since a different input frequency then gets mixed to the IF passband frequency. The input RF amplifier must be gang-tuned with the LO, but the alignment is not critical. Its purposes are (a) to

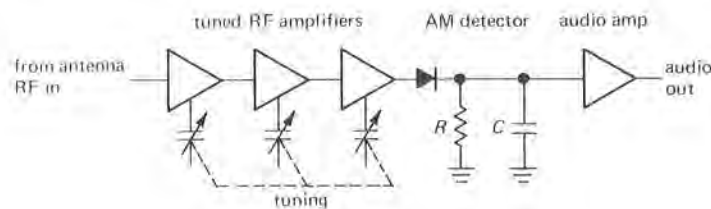


Figure 13.40



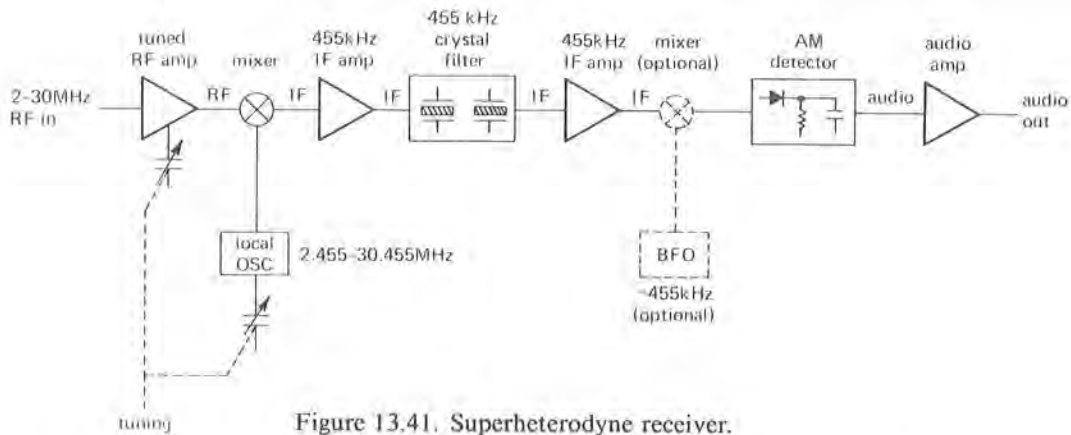


Figure 13.41. Superhetrodyne receiver.

improve the sensitivity with a stage of low amplification prior to mixing and (b) to reject signals at the “image” frequency, in this case input signals at a frequency of 455kHz *above* the LO (remember that a mixer generates sum and difference frequencies). In other words, the superheterodyne receiver uses a mixer and local oscillator to shift a signal at the (variable) input frequency over to a fixed intermediate frequency where most of the gain and selectivity are concentrated.

### Superhet potpourri

There are some additional features often added to a superheterodyne receiver. In this example a beat frequency oscillator (BFO) is shown; it is used in the detection of some signals with modulation other than AM (telegraphy, suppressed carrier telephony, frequency-shift keying, etc.). It can even be used for AM detection in what is known as a “homodyne” or “synchronous” detector. Receivers often have more than one mixing stage (they’re called “multiple-conversion” receivers). By using a high first IF, image rejection is improved (the image is twice the IF frequency away from the actual received signal). A lower second IF makes it easier to use sharp-cutoff crystal filters, and a third IF is sometimes generated to allow the use of audio-type notch filters, low-frequency ceramic

or mechanical filters, and “product detectors.”

Recently, the use of direct up-conversion (an IF higher than the input signal frequency) in a front-end balanced mixer, with crystal filters at the  $\sim 40$  MHz IF, followed by detection with no further mixing, has become popular. Such a single-conversion scheme offers better performance in the presence of strong interfering signals, and it has become practical with the availability of good VHF crystal-lattice filters and low-distortion wide-range balanced mixers with good noise performance.

### Image-reject mixers

The superhet receiver requires a tuned RF amplifier in order to reject the image band, which is separated by twice the IF frequency from the desired in-band RF signal frequency. The RF amplifier must be selective enough to reject the image band (i.e., its response to image-frequency signals must be much less than its response to in-band signals), and it must be tuned in order to keep its bandpass a constant (IF) frequency away from the LO, as the latter is adjusted to tune the receiver.

There is another way to suppress response at the image frequency, without using a tuned RF amplifier. Look at Figure

13.42, which shows an *image-reject* mixer: You begin with a pair of mixers, driven with quadrature LOs (“quadrature” means “differing in phase by 90°”); then combine the IF output signals, once again introducing a 90° phase shift in one path. The pair of 90° phase shifts adds for one sideband, and subtracts for the other, causing cancellation of the image band. Reversing the sign of the final 90° phase shift interchanges image band and signal band. In practice you usually use “4-port quadrature hybrids” to do the phase shifting, resistively terminating the unused output in each case. If you assemble an image-reject mixer from standard broadband components, you can expect something like 20dB suppression of the image sideband, with operation over one or two octaves of frequency. It’s sometimes essential to be able to move around rapidly in frequency (called “frequency agility”) without having to tune a tracking RF amplifier; in that case image-reject mixers are just what you want.

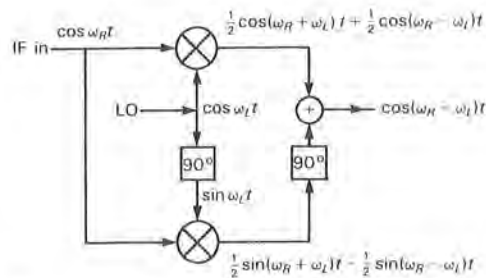


Figure 13.42. Image-reject mixer.

An interesting subtlety: As we remarked in Section 13.12, a mixer can be thought of as a modulator, and vice versa. The language you use to describe it depends on whether you are using the device to translate a low-frequency “baseband” of information up to high frequencies (in which case you call it a “modulator”) or using it to translate a modulated RF band down

to baseband [or perhaps an intermediate (IF) band along the way], where you demodulate it to extract the original modulating signal (in which case you call it a “mixer”). When you turn things around this way, what we called the image band becomes the other sideband. Our two methods of image rejection (RF filter, image-reject mixer) become the two classic methods of single-sideband modulation, namely the “filter” method and the “phasing” method. This may make more sense after you read the next section (but don’t worry if it doesn’t; we just couldn’t resist trying to explain this unifying idea).

## ADVANCED MODULATION METHODS

### □ 13.17 Single sideband

From a glance at the spectrum of an AM signal it is obvious that things can be improved. Most of the power (67%, to be exact, at 100% modulation) is in the carrier, conveying no information. AM is at most 33% efficient, and that only when the modulation index is 100%. Since voice waveforms generally have a large ratio of peak amplitude to average amplitude, the modulation index of an AM signal carrying speech is generally considerably less than 100% (although speech-waveform “compression” can be used to get more power into the sidebands). Furthermore, the symmetrical sidebands, by conveying the identical information, cause the signal to occupy twice the bandwidth actually necessary.

With a bit of trickery it is possible to eliminate the carrier [a balanced mixer does the job; note that  $\cos A \cos B = \frac{1}{2} \cos(A + B) + \frac{1}{2} \cos(A - B)$ ], creating what is known as “double-side-band-suppressed carrier,” or DSBSC. (This is just what you will get if the audio signal multiplies the carrier directly, without first being biased so that the audio

waveform is always positive, as in normal AM.) Then, either by using sharp crystal filters or by using a method known as “phasing,” one of the remaining sidebands can be eliminated. The “single-sideband” (SSB) signal that remains forms a highly efficient mode of voice communication and is widely used by radio amateurs and commercial users for long-range high-frequency telephony channels. When you’re not talking, there’s nothing being transmitted. To receive SSB, you need a BFO and product detector, as shown in the last block diagram, to reinsert the missing carrier.

#### □ Modulation spectra

Figure 13.43 shows representative spectra of voice-modulated AM, DSBSC, and SSB. When transmitting SSB, either sideband can be used. Note that SSB consists simply of the audio spectrum translated upward in frequency by  $f_c$ . When SSB is being received, the BFO and mixer combine to translate the spectrum down to audiofrequencies again. If the BFO is slightly mistuned, all audiofrequencies will be offset by the amount of mistuning. This dictates good stability for the LO and BFO in a receiver used for single-sideband.

Note that a mixer (modulator) can always be thought of as a frequency translator, especially when combined with suitable filters to eliminate the undesired outputs: When used as a modulator, a low-frequency band of frequencies is shifted up by the carrier frequency, to form a band

centered around  $f_c$ . When used as a mixer, a band of frequencies around  $f_c$  is shifted down to audiofrequencies (“baseband”), or to a band centered around the IF frequency, by the action of a high-frequency LO.

#### □ 13.18 Frequency modulation

Instead of modulating the amplitude of a carrier, as in AM, DSBSC, and SSB, it is possible to send information by modulating the frequency or phase of the carrier:

$$\text{signal} = \cos[\omega_c t + k \int f(t) dt]$$

frequency modulation (FM)

$$\text{signal} = \cos[\omega_c t + k f(t)]$$

phase modulation (PM)

FM and PM are closely related and are sometimes referred to as “angle modulation.” FM is familiar as the mode used in the 88–108MHz VHF broadcast band, and AM is used in the 0.54–1.6MHz broadcast band. Anyone who has tuned an FM receiver has probably noticed the “quieting” of background noise characteristic of FM reception. It is this property (the steep rise of recovered SNR with increasing SNR of the channel) that makes wideband FM preferable to AM for high-quality transmission.

Some facts about FM: When the frequency deviation  $k f(t)/2\pi$  is large compared with the modulating frequency [highest frequency present in  $f(t)$ ], you have “wideband FM” as used in FM broadcasting. The modulation index,  $m_f$ , equals

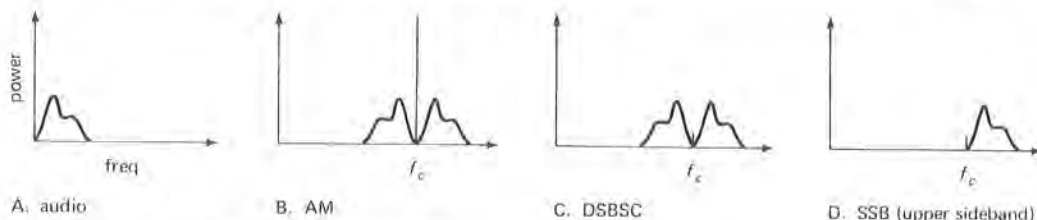


Figure 13.43. Suppressed-carrier spectra.

the ratio of frequency deviation to modulating frequency. Wideband FM is advantageous because under the right conditions the received SNR increases 6dB per doubling of FM deviation. The price you pay is increased channel bandwidth, since a wideband FM signal occupies approximately  $2f_{\text{dev}}$  of bandwidth, where  $f_{\text{dev}}$  is the peak deviation of the carrier. FM broadcasting in the 88–108MHz band uses a peak deviation  $f_{\text{dev}}$  of 75kHz, i.e., each station uses about 150kHz of the band. This explains why wideband FM is not used in the AM band (0.54–1.6MHz): There would be room for only six stations in any broadcasting area.

### FM spectrum

A carrier that is frequency-modulated by a sine wave has a spectrum similar to that shown in Figure 13.44. There are

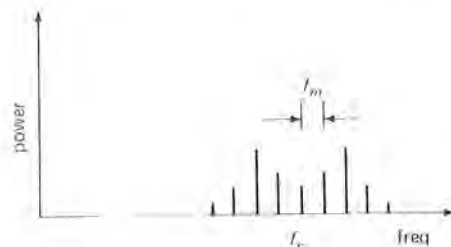


Figure 13.44. Wideband FM spectrum.

numerous sidebands spaced at multiples of the modulating frequency from the carrier, with amplitudes given by Bessel functions. The number of significant sidebands is roughly equal to the modulation index. For narrowband FM (modulation index  $<1$ ), there is only one component on either side of the carrier. Superficially this looks the same as AM, but when the phase of the sidebands is taken into account, you have a waveform of constant amplitude and varying frequency (FM), rather than a waveform of varying amplitude and constant frequency (AM). With wideband FM

the carrier amplitude may be very small, with correspondingly high efficiency, i.e., most of the transmitted power goes into the information-carrying sidebands.

### □ Generation and detection

FM is easily produced by varying an element of a tuned circuit oscillator; a varactor (a diode used as a voltage-variable capacitor, Section 5.18) is ideal. Another technique involves integrating the modulating signal, then using the result to do phase modulation. In either case it is often best to modulate at low deviation, then use frequency multiplication to increase the modulation index. This works because the *rate* of frequency deviation is not changed by frequency multiplication, whereas the deviation is multiplied along with the carrier.

To detect FM, an ordinary superhetrodyne receiver is used, with two differences. First, the final stage of IF amplification includes a “limiter,” a stage run at constant (saturated) amplitude. Second, the subsequent detector (called a discriminator) has to convert frequency deviation into amplitude. There are several popular methods of detection:

1. A “slope detector,” which is nothing more than a parallel  $LC$  circuit tuned off to one side of the IF frequency; as a result, it has a rising curve of response versus frequency across the IF bandwidth, thereby converting FM to AM. A standard envelope detector converts the AM to audio. There are improved versions of the slope detector involving a balanced pair of  $LC$  circuits tuned symmetrically to either side of the IF center frequency.
2. The Foster-Seely detector, or its variant, the “ratio detector,” using a single tuned circuit in a fiendishly clever diode arrangement to give a linear curve of amplitude output versus frequency over the IF bandpass. These discriminators are superior to the simple slope detector (Fig. 13.45)



3. A “phase-locked loop” (PLL). This is a device that varies the frequency of a voltage-controlled oscillator to match an input frequency, as we discussed in Section 9.31. If the input is the IF signal, the control voltage generated by the PLL is linear in frequency, i.e., it is the audio output.

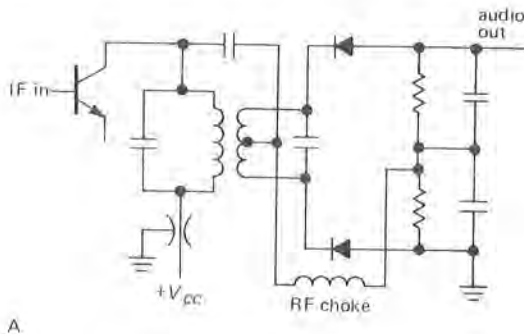
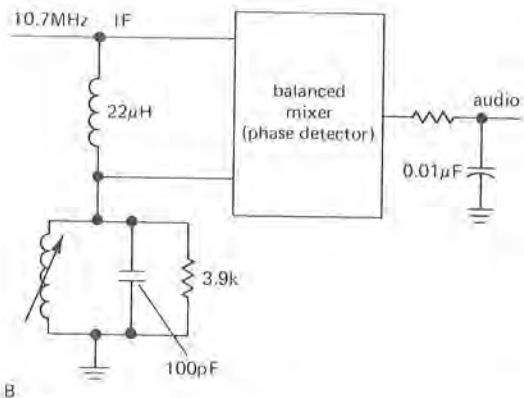


Figure 13.45. FM discriminators.  
A. Foster-Seeley.  
B. Balanced quadrature detector.

4. An averaging circuit, in which the IF signal is converted to a train of identical pulses at the same frequency. Averaging this pulse train generates an output proportional to IF frequency, i.e., the audio output plus some dc.

5. A “balanced quadrature detector,” which is a combination of a phase detector (see Sections 9.27 and 9.31) and a phase-shifting network. The IF signal is passed through a network that produces a

shift varying linearly with frequency across the IF passband (an  $LC$  circuit would do nicely). The resultant signal and the original signal are compared in a phase detector, giving an output that varies with relative phase. That output is the desired audio signal (Fig. 13.45).

It is often pointed out that FM provides essentially noise-free reception if the channel has sufficient SNR, as compared with AM, where the rejection of interference improves only gradually with increasing signal power. This makes sense when you remember that FM signals pass through a stage of amplitude limiting before detection. As a result, the system is relatively insensitive to interfering signals and noise, which appear as amplitude variations added to the transmitted signal.

#### □ 13.19 Frequency-shift keying

Transmission of digital signals (radioteletype, RTTY) is usually done by shifting a continuous-running carrier in frequency between two closely spaced frequencies according to the 1's and 0's being transmitted; 850Hz of shift is a typical value. The use of frequency-shift keying (FSK), rather than on/off modulation, is extremely effective in the presence of large signal fading from changing propagation conditions. To demodulate FSK, you simply use a differential amplifier looking at the outputs from a pair of filters set at the two detected audio frequencies. You can think of FSK as digital FM. Narrow-shift FSK has been used to circumvent selective fading between the two signal frequencies. However, the shift cannot be reduced below the information bandwidth of the keyed signal itself, roughly the “baud” rate (number of bit cells per second), or about 100Hz for ordinary radioteletype.

#### □ 13.20 Pulse-modulation schemes

There are several methods whereby analog signals can be transmitted as pulses.



The basic fact that makes digital transmission of analog signals possible is expressed in the Shannon sampling theorem, which states that a band-limited waveform is fully described by sampling its amplitude at a rate equal to twice the highest frequency present. Thus, a method that conveys the amplitude of a waveform, by digital methods or whatever, at instants of time separated by  $1/2f_{\max}$  can be used instead of a continuous modulation scheme. Several methods are shown in Figure 13.46.

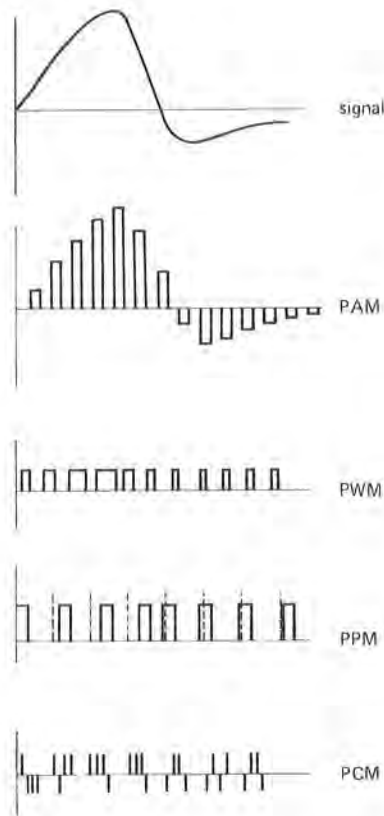


Figure 13.46. Pulse-modulation schemes.

In pulse-amplitude modulation (PAM), a train of pulses of amplitude equal to the signal is transmitted at regular intervals. This scheme is useful for *time multiplexing* of several signals on one information

channel, since the time between samples can be used to transmit the samples of another signal (with an increase of bandwidth, of course). In pulse-width modulation (PWM), the width of fixed-amplitude pulses is proportional to the instantaneous signal amplitude. PWM is easy to decode, using simple averaging. In pulse-position modulation (PPM), pulses of fixed width and amplitude are delayed or advanced relative to a set of fixed times, according to the amplitude of the signal.

#### □ Pulse-code modulation

Finally, in pulse-code modulation (PCM) the instantaneous amplitude of the signal is converted to a binary number and transmitted as a serial string of bits. In the illustration, a 4-bit offset binary code corresponding to 16 levels of quantization has been used. PCM excels when error-free transmission is required over noisy channels. As long as 1's and 0's can be identified unambiguously, the correct digital code, and hence a replica of the original signal, can be recovered. PCM is particularly useful in repeater application, e.g., transcontinental telephone channels, where the signal must pass through many stations and be amplified along the way. With any of the linear modulation schemes (AM, FM, SSB) noise accumulated in transit cannot be removed, but with PCM the digital code can be correctly regenerated at each station. Thus the signal starts anew at each station.

There are variations of PCM (known as coded PCM) in which techniques other than simple serial binary sequences are used to encode the quantized samples; for instance, a burst of one of 16 tones could be used in the preceding example. PCM is routinely used for telemetry of images from space vehicles, owing to its error-free properties. It is also used for "compact-disc" digital audio, in which each stereo channel is sampled and converted to a

16-bit number 44,100 times per second. In any PCM application the bit rate must be chosen low enough to ensure a low probability of error in bit recognition. In general, this limits transmission on a given channel to speeds much below what could be used with direct analog modulation techniques.

## □ RADIOFREQUENCY CIRCUIT TRICKS

In this chapter we are attempting to highlight some of the motivation and techniques of circuitry operated at radiofrequencies. In such limited space it is not possible to consider circuit design and construction in as much detail as we have generally attempted in the other chapters, nor would that even be desirable in a book intended as a broad introduction to electronics. In keeping with this philosophy, we would like to give some idea of the techniques that are ordinarily used in RF circuits. These are generally aimed at reducing stray inductance and capacitance and coping with circuitry whose dimensions are often comparable with a wavelength. There will be no attempt to weave these together into a coherent methodology; just think of them as a bag of tricks.

### □ 13.21 Special construction techniques

RF “chokes” (small inductors, in the range of microhenrys to millihenrys) are used extensively as signal-blocking elements. Power-supply voltages will usually be brought into a shielded enclosure with shielded “feedthrough capacitors” (bypass to ground combined with a mechanical feedthrough terminal), with an RF choke in series. A variation is to use ferrite beads on leads of transistors, FETs, etc. These are used because of the tendency of RF circuits toward “parasitic” oscillations, encouraged by unintentional tuned circuits at UHF formed by the wiring itself. Stringing

a few beads on a base or collector lead here and there adds a lossy series inductance to prevent the oscillation (if you’re lucky, that is!).

Inductors play a major role in RF design, and you see plenty of open coils and “slug-tuned” inductors and transformers (such as the little metal IF transformer cans you see everywhere in receiver circuits). Small-value air-variable capacitors are equally popular.

As suggested earlier, RF circuits are constructed in shielded enclosures, often with internal grounded partitions between sections of the circuit to prevent coupling. It is common to build circuits on double-sided PC board, with one side used as a ground-plane. Alternatively, a circuit may be constructed immediately adjacent to a shield or other grounded surface. Grounds can’t be wishy-washy at RF; you’ve got to solder a shield along its whole length, and you have to use a lot of screws to mount a partition or cover.

When building circuits at higher radio frequencies, it is absolutely essential to keep component leads as short as possible. That means snipping off leads right at the resistor or capacitor and soldering them with no visible lead showing (the components get plenty hot, but they seem to survive). At VHF and UHF you often use ceramic capacitor “chips,” soldered to PC strips, etc., without leads at all. Watch out when using conventional capacitors, because their internal series inductance can lead to self-resonance effects, sometimes at frequencies as low as a megahertz. The use of wide straps or metal ribbon, rather than ordinary wire, reduces inductance and is a favorite at UHF. At these frequencies you get into stripline and microstrip techniques, where every lead is itself a transmission line, complete with impedance matching. In fact, strips of sheet metal can be used as parts of tuned circuits; here’s a specification for an inductor in a 440MHz circuit (ARRL handbook):

" $L_1 - L_3$ , incl.  $-2\frac{5}{8} \times \frac{1}{4}$ -inch strip of brass, soldered to the enclosure on one end and to the capacitor at the other. Input and output taps are 1/2-inch up from the ground end." Of course, at microwave frequencies all such techniques give way to waveguide and cavity circuits, complete with exotica such as circulators and "magic T's" (Fig. 13.47).

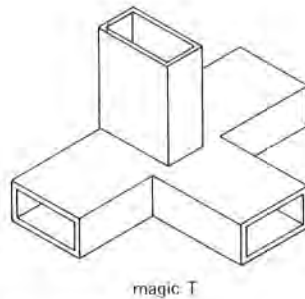


Figure 13.47

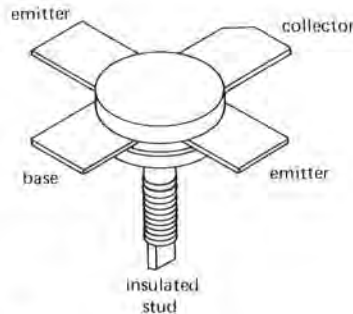


Figure 13.48

A facet of RF design that surprises beginners is the use of test instruments combined with "cut-and-try" techniques. You see widespread use of sweep generators (RF signal sources that sweep repetitively through a range of frequency), grid dip meters (for measuring resonances), SWR bridges, and spectrum analyzers, with plenty of circuit experimentation. At these frequencies you just can't predict everything; it takes some trial and error (and

lots of experience) to make things work well.

#### □ 13.22 Exotic RF amplifiers and devices

Familiar devices such as bipolar transistors and FETs are used at radiofrequencies, although often in somewhat different incarnations. Transistors intended for use at VHF and above come in strange-looking packages, with flat strips radiating out from the center for connection to a stripline or PC board (Fig. 13.48). There are also devices and circuits with no low-frequency analog, such as the following.

**Parametric amplifiers.** These devices amplify by varying a *parameter* of a tuned circuit. An analogy is a pendulum formed by hanging a weight on a length of string. Imagine that the motion represents output signal. You can build up the swing by gently shoving the weight at the resonant frequency; this is analogous to an ordinary amplifier, with a transistor or other active device providing the "shove." But there's another completely different way to get the thing swinging, namely by pulling up and down on the string (varying its length, a parameter of the system) at *twice* the natural resonant frequency. Try it (Fig. 13.49). The pendulum is closely analogous



Figure 13.49. Pendulum analogy to the parametric amplifier.

to the Adler parametric amplifier. In a paramp you can vary the capacitance of a tuned circuit with a varactor (voltage-variable capacitor) by driving it with a

“pump” signal. Paramps are used for low-noise amplification.

*Masers.* “Maser” is an acronym for microwaver amplification by stimulated emission of radiation. These things are basically atomic or molecular amplifiers, tricky to make and use but delivering the lowest noise of any amplifier.

*GaAs FETs.* The latest word in simple microwave amplifiers. Performance is comparable to that of paramps, without the fuss and bother. Present-day commercial GaAs FETs will deliver 28dB of gain at 10GHz with a 2dB noise figure. The latest in low-noise GaAs FETs are the so-called HEMTs – high-electron-mobility transistors – which can deliver astonishing noise figures [e.g., 0.12dB( $T_N = 8^\circ\text{K}$ ) at 8.5GHz in an experimental chilled amplifier].

*Klystrons and traveling-wave tubes.* Vacuum-tube amplifiers used at microwave frequencies, klystrons and TWTs take advantage of transit-time effects within the tube. A variation known as a reflex klystron works as an oscillator by bouncing its electron beam into its guts. There are klystrons available that can continuously deliver 0.5MW RF output at 2000MHz.

*Magnetrons.* The heart of radar and microwave ovens: a high-power oscillator tube, full of little resonant cavities, and operated in a large magnetic field to make the electrons spiral around inside.

*Gunn diodes, IMPATT diodes, PIN diodes.* These exotic devices are used extensively at UHF and microwave frequencies. Gunn diodes are used as low-power oscillators in the 5–100GHz range, delivering output powers of 100mW or so. IMPATT diodes are analogous to klystrons, with capabilities of a few watts at a few gigahertz. PIN diodes behave as voltage-variable resistances and are used to switch microwave

signals on and off by becoming a short circuit across a waveguide. PIN diodes can be used as lumped attenuators or switching elements at lower radiofrequencies. A leader in this field is Unitrode Corp. (Lexington, MA), whose “PIN diode designers’ handbook and catalogue” is indispensable.

*Varactors, SNAP diodes.* Varactors are reverse-biased diodes used as variable capacitances for tuning purposes, or in paramps. Because of their nonlinearity they are also used for harmonic generation, i.e., as frequency multipliers. SNAP diodes are also popular for harmonic generation, since they exhibit sub-picosecond rise times.

*Schottky diodes, back-diodes.* You have seen Schottky diodes earlier as high-speed diodes with low forward drop. They’re often used as mixers, as are back-diodes, variations of tunnel diodes. Look at Section 13.13 for a nice back-diode application (square-law detector).

## HIGH-SPEED SWITCHING

The same effects that limit linear amplifier performance at high frequencies (the combination of junction capacitance, feedback capacitance, with its Miller effect, and stray capacitance in combination with finite source and load resistance) impose speed limitations on high-speed digital circuits. Many of these problems don’t affect the designer directly, since they’ve been handled well in the design of digital ICs themselves. The average circuit designer would have a difficult time even coming close to the performance of TTL circuits, for example, using discrete transistor design.

Nevertheless, there are plenty of occasions when you’ve got to know how to design fast switching circuits. For example, in driving some external high-voltage or high-current load (or a load of opposite



polarity) from a logic output, it is quite easy to lose a factor of 100 in switching speed through careless design. Furthermore, there are situations in which no packaged digital logic is used at all, and you're on your own all the way.

In this section we will begin with a simple transistor model useful for switching-circuit calculations. We will apply it to a few example circuits to show how it goes (and how important the choice of transistor can be). We will conclude by illustrating transistor switching design with a complete high-speed circuit (a photomultiplier preamp/discriminator).

### 13.23 Transistor model and equations

Figure 13.50 shows a saturated transistor switch, connected as an inverter, driven

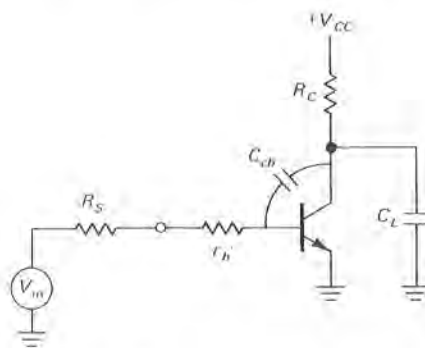


Figure 13.50

from a source of pulses with extremely fast rise and fall times.  $R_s$  represents the source impedance,  $r'_b$  is the relatively small intrinsic transistor base "spreading" resistance (of the order of  $5\Omega$ ),  $C_{cb}$  is the all-important feedback capacitance, and  $R_C$  is the load resistance, paralleled by load capacitance  $C_L$ . The effects of finite load resistance can be included by letting  $R_C$  represent the Thévenin equivalent of the combined resistances, with  $V_{CC}$  suitably

modified. The collector-to-emitter capacitance has been absorbed into  $C_L$ , and  $C_{be}$  has been ignored, since  $C_{cb}$  always dominates owing to Miller effect.

Figure 13.51 shows a typical output waveform from this circuit when driven

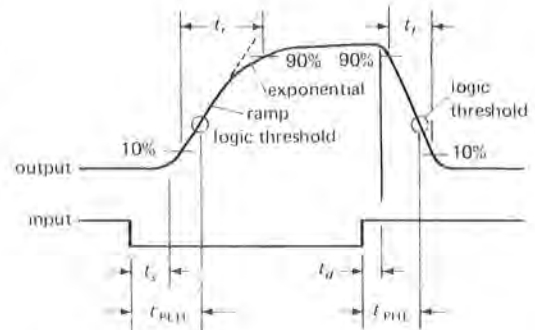


Figure 13.51. Waveform parameters of a transistor switch.

by a clean negative-going input pulse. The rise time,  $t_r$ , is usually defined as the time to go from 10% to 90% of the final value, with the corresponding definition of fall time,  $t_f$ . Note especially the relatively long storage time,  $t_s$ , that is required for the transistor to come out of saturation, compared with the correspondingly shorter delay time,  $t_d$ , to bring the transistor into conduction. These definitions are all conventionally taken between the 10% and 90% points. More useful for digital logic purposes are the propagation times,  $t_{PLH}$  and  $t_{PHL}$ , defined as the time from the input transition until the output passes through the logic threshold (rising or falling, respectively). Other symbols are in common use, e.g.,  $t_{pd1}$  or  $t_{pr}$  is often used for what we've called  $t_{PLH}$ .

Let's use the circuit model to estimate rise and fall times for a given circuit. In the process you will even come to understand why the rising portion of the output waveform sometimes ends with an exponential.



### □ Estimation of rise time

After the input signal has dropped to its LOW state and  $t_s$  has elapsed (more on that later), the collector begins to rise. Two effects limit the rate of rise: (a)  $R_C$  in combination with  $C_{cb}$  and  $C_L$  sets a time constant, generating an exponential rise toward  $V_{CC}$ , but (b) if that rate of collector rise is great enough, the resulting current through  $C_{cb}$  generates forward base bias across the source impedance ( $R_s + r'_b$ ), and it can turn on the base, with the resultant effect of slowing the collector rise through negative feedback. What you have in the latter case is an integrator, and the collector waveform is a ramp. In general (depending on circuit values and transistor parameters), the collector waveform may begin as a ramp and change over to an exponential, as shown previously.

The scope photo in Figure 13.52 illustrates these effects. Instead of an *npn* transistor, however, we've used an *n*-channel enhancement-mode MOSFET. It behaves similarly, but the larger gate turn-on voltage makes the waveforms particularly clear. In addition, MOSFETs do not have storage or delay time effects, and they have no dc input current, which keeps things simple. Figure 13.53 shows the circuit,

with our intentionally exaggerated signal source impedance. Note how feedback capacitance pins the gate voltage at the turn-on threshold while the drain switches. Note also that the rising drain waveform changes over to an exponential, when  $R_D$  is large.

A simple way to estimate circuit behavior is the following:

1. Compute the "integrator-limited" rate of rise of collector voltage, according to

$$\frac{dV_c}{dt} = \frac{V_{BE} - V_{in}(LOW)}{C_{cb}(R_s + r'_b)}$$

2. Find the collector voltage  $V_X$  at which the output waveform changes from a ramp to an exponential, according to

$$V_X = V_{CC} - \left( \frac{V_{BE} - V_{in}(LOW)}{R_s + r'_b} + C_L \frac{dV_c}{dt} \right) R_c$$

This allows you to determine the collector waveform and rise time, as we will illustrate with examples presently. If  $V_X$  comes out negative, that means the entire collector rise is exponential: The capacitive load dominates, and the base is never turned on via current through the feedback capacitor. The term  $r'_b$  is usually negligible.

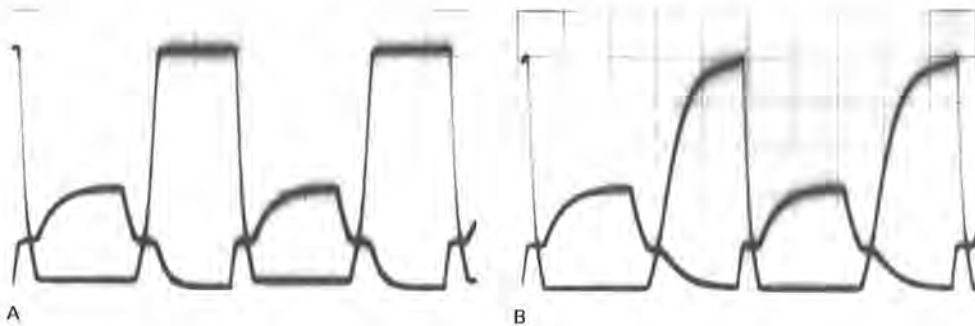


Figure 13.52. Gate and drain voltage switching waveform showing dynamic Miller clamping with (exaggerated) 100k source impedance. Vertical: 2V/div; signal frequency: 6kHz.

A. With 10k drain resistor.  
B. With 200k drain resistor.

**EXERCISE 13.1**

Derive the two preceding formulas. Hint: For the second formula, equate the feedback current needed to bring the base into conduction with the available collector pull-up current less the current needed to derive the (capacitive) load.

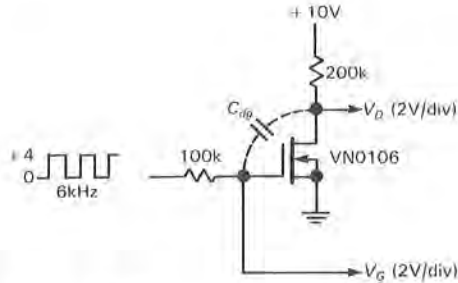


Figure 13.53

□ **Estimation of fall time**

Following the short delay time after the input has gone HIGH, the collector begins to drop, on its way to saturation. With a little bookkeeping of currents, it is easy enough to see that the collector current is given by

$$I_C = \left( \frac{V_{in}(\text{HIGH}) - V_{BE}}{R_S + r'_b} + C_{cb} \frac{dV_c}{dt} \right) h_{fe}$$

$$= \frac{V_{cc} - V_c}{R_c} - (C_L + C_{cb}) \frac{dV_c}{dt}$$

where the first line is the net base current multiplied by  $h_{fe}$  and the second line is the available collector current through  $R_C$ , less the current needed to drive the capacitance seen at the collector. Remember that  $dV_C/dt$  is negative. Rearranging, we get

$$-\frac{dV_c}{dt} = \frac{1}{C_L + (h_{fe} + 1)C_{cb}} \times$$

$$\left( \frac{V_{in}(\text{HIGH}) - V_{BE}}{R_S + r'_b} h_{fe} - \frac{V_{cc} - V_c}{R_c} \right)$$

where the first term in parentheses is recognizable as  $h_{fe}i_{drive}$  and the second is

$i_{pull-up}$ . You are now licensed to try some circuits; you will be able to see what sort of rise times and fall times can be expected and which capacitance dominates. First, however, a word on delay and storage times.

□ **Delay and storage times**

In general, delay times are very short. The main effect is the time constant involved in moving the base capacitance up to  $V_{BE}$ , a time constant of order

$$T \sim (R_s + r_b)(C_{cb} + C_{be})$$

When working at extremely high speeds, transistor transit-time effects may also become important.

Storage times are another matter. A transistor in saturation has charge stored in the base region, and even after the base drive signal has gone close to ground (or even a bit negative), it requires a relatively long time for the extra injected minority carriers from the emitter to be swept from the base region by the collector current. Transistors differ widely in storage time; it can be shortened by using less base overdrive during saturation and by reverse-biasing the base to reverse the base current when switching the transistor OFF. This equation for storage time,  $t_s$ , makes these points:

$$t_s = K \ln \frac{I_B(\text{ON}) - I_B(\text{OFF})}{\frac{I_C}{h_{FE}} - I_B(\text{OFF})}$$

where  $I_B(\text{OFF})$  is negative for “discharge” reverse base currents. The constant  $K$  includes a “minority-carrier lifetime” term, which is greatly reduced by gold doping. However, such doping reduces  $h_{FE}$  and increases the leakage current. This explains the good speed performance of TTL, along with its low breakdown voltage (7V).

Storage times can be as long as several hundred nanoseconds, and they are typically an order of magnitude longer than delay times. The popular general-purpose

2N3904, for instance, has a specified maximum delay time of 35ns and a storage time of 200ns under standardized test conditions, which include driving the base negative by two diode drops.

Since storage times can turn out to be a severe limitation on the performance of high-speed switching circuits, there are several measures that can be taken to circumvent the problem. One solution is to avoid saturation altogether. A Schottky clamping diode (a “Baker clamp”) from base to collector will accomplish this by robbing current from the base when the collector is nearing saturation. It prevents transistor saturation, since its forward voltage drop is less than that of the collector-base junction. The Schottky families of TTL logic use this trick. A small “speedup” capacitor (25–100pF) across the base driving resistor is often a good idea in addition, since it can reduce storage time by providing a pulse of current to remove base charge at turn-off, and in addition it increases base drive current during turn-on transitions. Figure 13.54 illustrates these methods.

### 13.24 Analog modeling tools

It should be evident that designing both high frequency amplifiers and high speed switching circuits with discrete components is a complex subject, particularly when the full effects of parasitic capacitances and inductances are properly accounted for. Our simplified models provide good circuit intuition, but they may often be inadequate if you are trying to squeeze maximum performance from an amplifier in the gigahertz range of frequencies, for example. The traditional approach consisted of a combination of more complex modeling (plenty of calculation!) and lots of breadboarding.

A pleasant development is the maturation of computer-aided analog modeling,

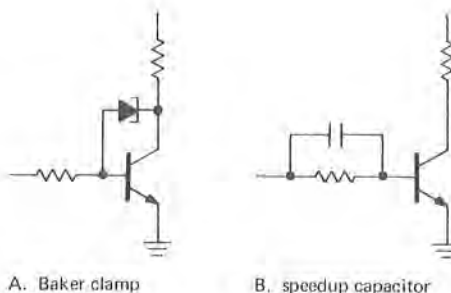


Figure 13.54

in particular a program called “SPICE” and its several commercial extensions. SPICE (originated by L. W. Nagel) models the operation of your trial circuit (using a library of sophisticated component models), predicting the gain, distortion, noise, frequency response, etc. You can ask it to show you the waveforms (voltage *and* current) at any point in the circuit – a computerized oscilloscope! Thus you can fiddle with your hypothetical circuit, trying speedup capacitors, exploring component and circuit trade-offs, and so on. In fact, good simulators even let you see the effects of component tolerance, via either conventional worst-case analysis or a more sophisticated “Monte Carlo” statistical treatment. Some modeling programs also perform a “sensitivity” analysis, telling you which components affect performance the most.

Computer-aided modeling programs are inexpensive and available for desktop computers as well as mainframes (e.g., IsSpice from Intusoft, or PSpice from Microsim). Most accept “netlists” as generated by schematic capture programs (Section 12.08). Modeling programs are not restricted to high frequency design. They can be used to model circuits with op-amps, digital circuits, and the innards of ICs themselves. Although SPICE dominates the current scene, there are at least a dozen competing modeling programs, optimized for specific applications.

## □ SOME SWITCHING-SPEED EXAMPLES

In this section we will analyze the performance of a few simple circuits, based on the methods just discussed.

### □ 13.25 High-voltage driver

Let's begin with the circuit in Figure 13.55. It is a simple inverting stage intended for driving a piezoelectric crystal with 100 volt pulses, generated originally with TTL logic. The TTL output, and therefore the base driving signal, is roughly as indicated. In these calculations we will ignore  $\tau'_b$ , which is small compared with the source impedance.

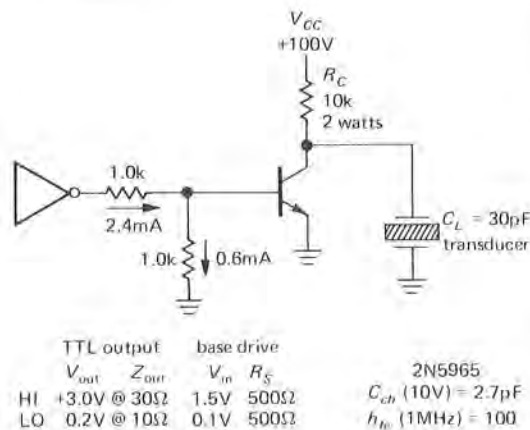


Figure 13.55

#### □ Rise time

We begin by calculating the integrator-limited collector rise:

$$\frac{dV_c}{dt} = \frac{V_{BE} - V_{in(LOW)}}{C_{cb}R_S} \approx 450V/\mu s$$

from which the estimated rise time will be

$$t_r = \frac{0.8V_{cc}}{dV_c/dt} \approx 180ns$$

Now we find the collector voltage at which the rise changes from an integrator-limited ramp to an exponential:

$$V_X = V_{CC}$$

$$-R_C \left( \frac{V_{BE} - V_{in(LOW)}}{R_S} + C_L \frac{dV_C}{dt} \right)$$

$$\approx -50V$$

This means that the collector rising waveform is exponential the whole way, with the feedback current ( $C_{cb}dV_C/dt$ ) insufficient to pull the base up into conduction, given the source impedance. The collector time constant is  $R_C(C_L + C_{cb})$ , or  $0.33\mu s$ , with a rise time (10% to 90%) of 2.2 time constants, or  $0.73\mu s$ . It is clear that the combination of collector resistor and load capacitance dominates the rise.

#### Fall time

To analyze the fall time, we use the formula derived earlier to find

$$\frac{dV_c}{dt} = \frac{1}{C_L + (h_{fe} + 1)C_{cb}} \times \left\{ h_{fe} \left( \frac{V_{in(HIGH)} - V_{BE}}{R_S} \right) - \frac{V_{CC} - V_C}{R_C} \right\}$$

$$\approx 530V/\mu s$$

$$t_f = \frac{0.8V_{CC}}{dV_C/dt} \approx 0.15\mu s$$

The last term depends on  $V_C$ , but is negligible compared with the first term in parentheses. If it weren't, you would have to evaluate it at several values of collector voltage to get a good picture of the falling waveform. At this point it should be noted that the calculated fall time corresponds to a frequency of about 3MHz, and therefore the value of  $h_{fe} = 100$  we used is realistic ( $f_T = 300$ ). If a calculated rise time or fall time corresponds to a frequency much higher than originally assumed, it is generally necessary to go back and recompute the transition time, using a new  $h_{fe}$  based on a better estimate of the transition time. This iteration process will usually give a satisfactory answer on the second pass.

### □ Switching waveform

For this circuit, then, the collector waveform is as shown in Figure 13.56. The rise is dominated by the time constant of the load capacitance and collector resistor, whereas the fall is dominated by the feedback capacitance in combination with the source impedance. To put it another way, the collector voltage falls at a rate such that the current through the feedback capacitance is almost sufficient to cancel the base drive current and bring the base out of conduction. Note that we have assumed throughout that the TTL output waveform is much faster than the output of our circuit. With typical rise and fall times of about 5ns, that is a good approximation.

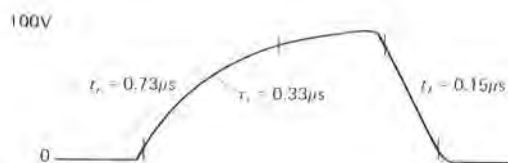


Figure 13.56

### □ 13.26 Open-collector bus driver

Suppose we want to drive an open-collector TTL bus from the output of an NMOS circuit. We can do it by interposing an *n*pn inverting stage, as in Figure 13.57. The base resistors are necessarily large because of the low output sourcing capability of NMOS operating from +5 volts (see Section 9.09). We have chosen two popular transistors in order to illustrate the effect that parameters like  $C_{cb}$  can have.

For rise time, we calculate as before, and we find the following integrator-limited rise times:

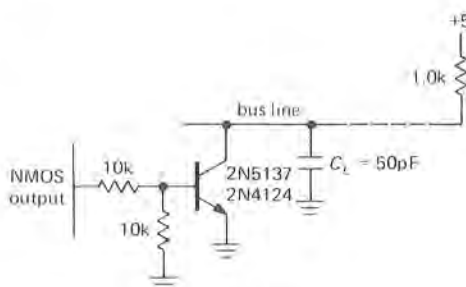
|           | 2N5137      | 2N4124      |
|-----------|-------------|-------------|
| $dV_C/dt$ | $8.5/\mu s$ | $76V/\mu s$ |
| $t_r$     | 470ns       | 53ns        |

The crossover to exponential is calculated as:

|            | 2N5137 | 2N4124 |
|------------|--------|--------|
| $V_X$      | 4.4V   | 1.1V   |
| time const | 66ns   | 52ns   |

The falling waveform comes out:

|           | 2N5137       | 2N4124       |
|-----------|--------------|--------------|
| $dV_C/dt$ | $-11V/\mu s$ | $-78V/\mu s$ |
| $t_f$     | 360ns        | 51ns         |



|           | NMOS output | base drive |
|-----------|-------------|------------|
| $V_{out}$ | $V_{in}$    | $R_S$      |
| HI        | +3.5 @ 1kΩ  | 1.7V 5.5k  |
| LO        | 0.0V @ 200Ω | 0.0V 5.1k  |

|                 | 2N5137 | 2N4124 |
|-----------------|--------|--------|
| $C_{cb}$ (10V)  | 16pF   | 1.8pF  |
| $h_{fe}$ (1MHz) | 100    | 100    |

Figure 13.57

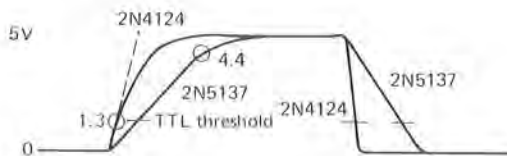


Figure 13.58

### □ Choice of transistor

The situation is as shown in Figure 13.58. The inferior performance of the 2N5137 is due entirely to the effects of feedback capacitance, aggravated in this example by the relatively high value of source impedance. The transition times of the 2N4124 are probably a bit optimistic,



since they correspond to a frequency of about 10MHz at which  $h_{fe}$  is somewhat lower than assumed.

It is interesting to look at the time required to reach the TTL threshold voltage of about 1.3 volts, the relevant parameter in a system in which TTL gates are driven by the bus signals. Ignoring storage and delay times, the times to reach the TTL threshold voltage are the following:

|                   | 2N5137      |             | 2N4124      |             |
|-------------------|-------------|-------------|-------------|-------------|
|                   | <i>calc</i> | <i>meas</i> | <i>calc</i> | <i>meas</i> |
| rise( $t_{PLH}$ ) | 150ns       | 130ns       | 17ns        | 30ns        |
| fall( $t_{PHL}$ ) | 340ns       | 360ns       | 47ns        | 52ns        |

The rise and fall times we measured are in reasonable agreement with the predictions of our somewhat simplistic model, except perhaps for the rise time of the 2N4124 circuit. There are a few possible explanations why the rise time we predicted is too small in that case: The calculation used the value for  $h_{fe}$  at 10MHz, whereas a 17ns rise time implies somewhat higher frequencies, and hence a lower value of  $h_{fe}$ . Also, by actual measurement the particular transistor has  $C_{cb} = 2.2\text{pF}$  at 10 volts and  $C_{cb} = 3\text{pF}$  at 2 volts. Curiously, the 2N5137 we used had a much lower  $C_{cb}$  (about 5pF) than specified on the data sheet, and so we added a small external capacitor to bring it up to specifications. This probably represents a change in the manufacturing process since the original data sheet was published.

#### EXERCISE 13.2

Verify the results just calculated for  $dV_C/dt$  (rise and fall) and  $V_X$ .

#### □ Pull-up to +3 volts

Note that the times to reach TTL threshold from a HIGH state are much longer than the times from a LOW state, even though the output slew rates (in the case of the 2N4124 circuits) are almost the same. That's because the TTL threshold voltage

is not symmetrically positioned between +5 and ground, forcing the collector to slew through a larger voltage on the way down. For this reason TTL buses are often terminated to a source of +3 volts (a series pair of diodes tied to +5 is one trick sometimes used), or each line of the bus can be terminated with a voltage divider, as in Figure 13.59.

#### EXERCISE 13.3

Calculate the rise and fall times and the propagation delays for a 2N4124 driving the preceding bus with  $C_L = 100\text{pF}$ . Show your work.

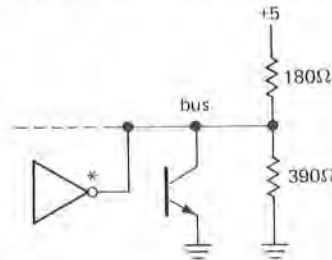


Figure 13.59

#### □ 13.27 Example: photomultiplier preamp

As we will discuss in Chapter 15, a device called a photomultiplier tube (PMT) is an extremely useful light detector, combining high sensitivity with high speed. Photomultipliers are also useful in applications where the quantity being directly measured isn't light, e.g., high-energy-particle detectors in which a scintillator crystal generates light flashes in response to particle bombardment. To take advantage of a photomultiplier's properties, it is necessary to use a charge-sensitive high-speed discriminator, a circuit that generates an output pulse when an input pulse of charge exceeds some threshold corresponding to the detection of a photon of light.

Figure 13.60 shows a circuit of a high-speed photomultiplier preamp and discriminator that illustrates the high-frequency and switching techniques discussed in this chapter. The output of the

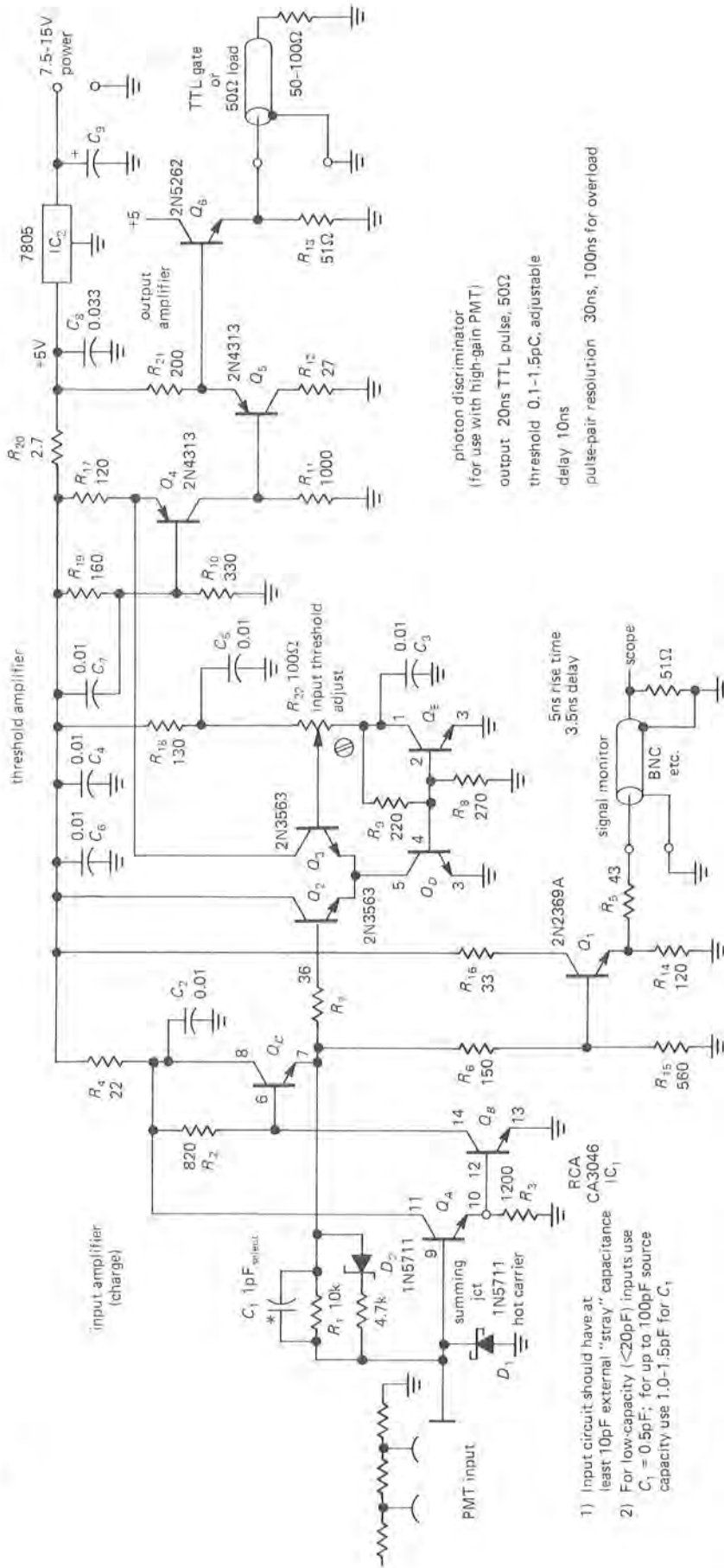


Figure 13.60. High-speed charge-sensitive amplifier for photon counting with photomultiplier tubes.

photomultiplier tube consists of negative pulses of charge (electrons are negative), each pulse having a width of perhaps 10–20ns. The larger pulses correspond to detected photons (quanta of light), but there are also lots of smaller pulses that arise from noise within the photomultiplier tube itself and that should be rejected by the discriminator.

#### □ Circuit description

The circuit begins with an inverting input amplifier ( $Q_A - Q_C$ ) with current and charge feedback via  $C_1$  and  $R_1$ . The input follower presents a low driving impedance to  $Q_B$  (which provides the voltage gain) to reduce the effects of feedback capacitance ( $C_{cb}$ ). The follower at the output of the gain block,  $Q_C$ , provides a low output impedance while allowing  $Q_B$  to have a reasonable amount of gain. The signal at this point is a small positive pulse corresponding to the negative-charge input from the PMT; dc feedback stabilizes  $Q_C$ 's output at about  $2V_{BE}$ .  $Q_1$  is biased as a class A emitter follower, giving a low-impedance “monitor” output of the amplified photomultiplier pulses before discrimination.

Differential amplifier  $Q_2$  and  $Q_3$  form the discriminator; the threshold is set by  $R_{22}$  referenced to a voltage (set by  $Q_E$ , operating as an “adjustable diode”) that tracks the input amplifier's  $2V_{BE}$  quiescent point. This diode-drop tracking occurs because the transistors  $Q_A - Q_E$  are in a monolithic transistor array (CA3046) and are all at the same temperature.  $Q_4$  forms an inverted cascode with  $Q_3$ , both for high speed and level shifting. Two stages of follower, arranged with opposite-polarity transistors to cancel  $V_{BE}$  offsets, complete the circuit.

There are several interesting points in this circuit. Transistor quiescent currents are set rather high (the differential pair  $Q_2Q_3$  has 11mA emitter current,  $Q_5$  idles

at 20mA, and the output transistor has to source 120mA to drive a 50Ω load) in order to get a good high-speed performance. Note that the cascode base ( $Q_4$ ) is bypassed to  $V_+$ , not ground, since its input signal is referenced to  $V_+$  via  $R_{17}$ . The comparator's emitter current source is a current mirror, convenient since  $Q_E$  is already used for the threshold reference.  $D_1$  and  $D_2$  are used to improve overload performance. Although it complicates the circuit, clamping diode  $D_1$  can be returned to  $Q_E$ 's collector (bypassed to ground) to put a tighter limit on negative (overload) swings at the input.

#### □ Performance

Figure 13.61 shows a graph of output pulse shape and timing versus input pulse size (measured as quantity of charge). The output pulses are stretched by large input overloads, but the overall performance is quite good, measured by usual photomultiplier preamp standards.

## SELF-EXPLANATORY CIRCUITS

### 13.28 Circuit ideas

Figure 13.62 shows a few wideband circuit ideas.

## ADDITIONAL EXERCISES

(1) In this problem you are to work out in detail the high-frequency behavior of the circuit in Figure 13.10, summarized briefly in Section 13.05. (a) Begin by repeating the calculation of driver/output-stage rolloff diagrammed in Figure 13.12. Be careful as you combine complex impedances. Write to one of the authors if you find an error! (b) Now check to see that the high-frequency rolloffs of the

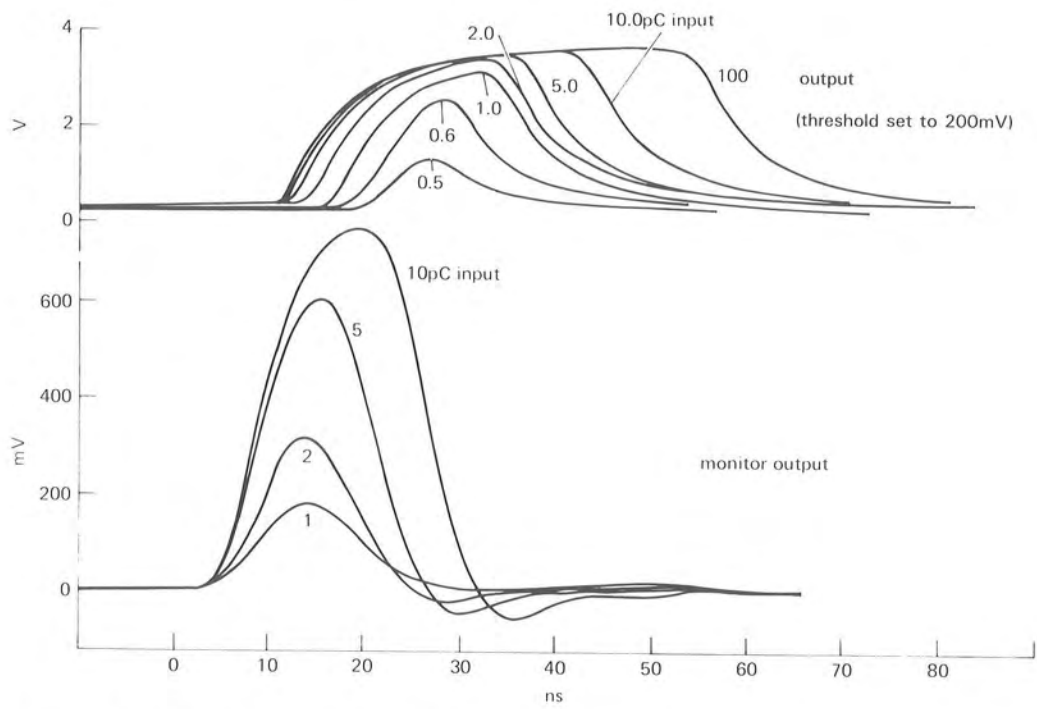


Figure 13.61. Pulse response of the photon-counting amplifier.



Circuit ideas

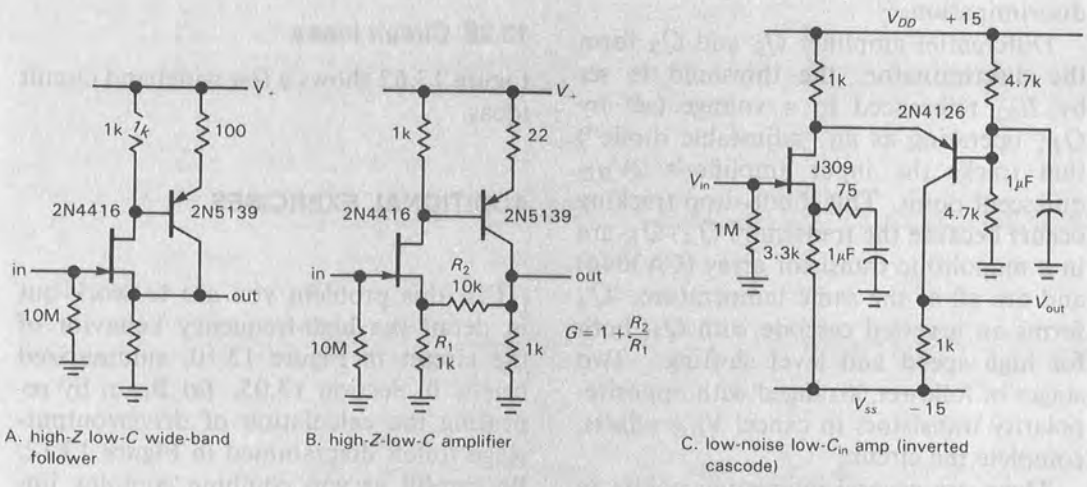


Figure 13.62

previous stages are significantly higher in frequency than the  $\sim 180\text{MHz}$  3dB frequency of the output stage and its driver. In particular, check the following rolloffs:  $Q_1$ 's output (emitter) impedance driving a capacitive load (see Fig. 13.11);  $Q_2$ 's output driving a slightly different capacitive load (because  $Q_4$ 's collector is not grounded);  $Q_3$  and  $Q_4$ 's emitters driving a capacitive load;  $Q_4$ 's collector driving a capacitive load.

(2) What is the impedance looking into a length of coaxial cable that is (a) open-circuited at the far end, and a quarter wave long, electrically, at the frequency of interest, (b) short-circuited at the far end, and a quarter wave long, electrically, at the frequency of interest, (c) same as (a), but a half wave long, (d) same as (b), but a half wave long? The result in (d) is the basis of the so-called choke joint used in waveguides.

(3) Work out in detail the rise time and fall time of the high-voltage switching circuit in Figure 13.55, as summarized in Section 13.25. Use  $V_{BE} = 0.7\text{V}$ .

(4) "The Rise and Fall of a Bus Driver": Calculate the rise and fall times for the TTL bus driver circuit of Figure 13.57, as summarized in Section 13.25. Use  $V_{BE} = 0.7\text{V}$ .

(5) Design a video amplifier with a gain of +5 and a rolloff of 20MHz or more. The input impedance should be 75 ohms, and the output should be able to drive a 75 ohm load with 1 volt pp output capability. A nice way to achieve the noninverting gain is to use a common-base input stage with an emitter follower output, as suggested in Figure 13.63. If you like the circuit, finish the design by choosing operating currents, resistor values, and biasing components. You can, of course, use something like a differential amplifier/cascode/follower combination, if you prefer. Note that the gain must be noninverting, or the image will be reversed.

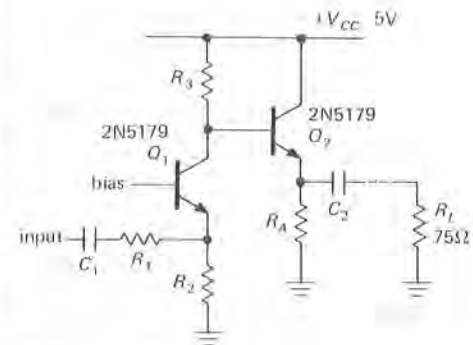


Figure 13.63





# LOW-POWER DESIGN

CHAPTER

14

## INTRODUCTION

Lightweight hand-held instruments, data loggers that make measurements at the ocean bottom, digital modems that power themselves from the “holding current” in the telephone line – these are just a few of the applications that invite (or demand) low-power electronic design techniques. Within such instruments you often find examples of all the areas of design already treated in this book – regulated supplies, linear circuits (both discrete and op-amp), digital circuits (almost invariably CMOS) and associated conversion techniques, and, increasingly, microprocessor circuits of considerable complexity. Although we have occasionally discussed power consumption and speed/power trade-offs (when comparing logic families, for example), the design of micropower electronic instruments involves special techniques and cautions throughout and requires a chapter all its own.

We will begin by considering a range of applications where low-power consumption is desirable or essential.

They’re not all of the exotic transmitter-attached-to-wild-moose variety, and in fact there are plenty of places where ac line power is plentiful, but where battery power is more convenient or performs better. Next we will review the power sources that make sense for low-power electronics, beginning with the ubiquitous “primary” (nonrechargeable) cells (alkaline, mercury, silver, lithium) and the closely related “secondary” (rechargeable) cells (nickel-cadmium, lead-acid, “gel” cells). We’ve hounded the manufacturers of batteries mercilessly for their latest technical literature, and so we’ve got really useful battery comparison data – energy content, discharge characteristics, effects of temperature, discharge rate, and storage conditions, etc. We’ll help you pick the right battery for your application.

Batteries aren’t the only power sources to consider, so we’ll continue by talking about those little black plastic “wall-plug-in” modules you get with the consumer electronics gizmos you can buy. Wall plug-ins are remarkably inexpensive and are available as simple transformers,

or unregulated dc supplies (transformer-rectifier-capacitor), or as complete regulated dc supplies; they also come in dual- or triple-supply varieties. Solar cells are useful in some unusual applications, so we'll discuss them too. Finally, we'll mention the use of signal currents, for example the dc current you find on the telephone line, or the ac relay energizing current in a thermostat or doorbell, to activate a micropower instrument.

We'll then talk about low-power design at the same level of detail that we've used in the rest of the book, mirroring a similar progression of subjects: regulators and references, then linear circuitry (discrete and op-amp), digital circuits and conversion techniques, and finally microprocessors and memories. There are, in addition, techniques that have no counterpart in ordinary design, for example "power switching," in which a normal circuit is rendered micropower by applying dc for extremely short intervals; for example, a microprocessor data logger might be powered for 20ms every 60 seconds. We'll talk about some of those tricks, and the curious pitfalls that await the uninitiated.

Finally, we'll talk about packaging – including the small plastic instrument cases that are widely available, complete with back door for access to the batteries. Micropower instruments are generally much easier to package than conventional instruments are, since they're usually lightweight, they don't produce heat, and they don't need the usual paraphernalia of power cords, line filters, and fuses. Micropower design is different and fun, and it presents new challenges for the jaded electronic designer. Read on to learn more about this exciting (but often neglected) subject!

We hate to say it, but this chapter could be passed over in a first reading of the book.

## 14.01 Low-power applications

We've collected together most of the reasons that might motivate you to do low-power circuit design. They're listed here, in no particular order.

### *Portability*

You can't carry it around with you if it has a power cord running off to a wall outlet. "It" could be a commercial product such as a calculator, wristwatch, hearing aid, "walkabout" tape recorder or receiver, paging radio, or digital multimeter. Or it might be a custom portable instrument, for example a small transmitter used to study herd migration and physiology. Because batteries have a finite energy content, you've got to keep power consumption low in order to have reasonable battery life with acceptable battery weight. A multimeter that runs 1000 hours on a single 9 volt battery will outsell a competing unit that requires four D cells and gives only 100 hours of service. A portable transmitter for animal migration studies is useless if it runs only two days on a fresh set of batteries. Thus, low-power circuitry is at a premium in instruments designed for portability. In the special case of extremely small instruments (e.g., a wristwatch), the tiny energy content of the self-contained batteries dictates micropower design, with total current drains of just a few microamps.

### *Isolation*

Instruments that are powered from ac line current are not suitable for some kinds of "floating" measurements at high potential. For example, you might want to measure microamp charged-particle-beam currents at the +100kV terminal of a particle accelerator. It may well be that you can't make the measurement by lifting the low-voltage end off ground (as in Fig. 4.79), because

the power transformer of the high-voltage supply causes significant 60Hz currents to be capacitively coupled to the high-voltage supply through its transformer (or perhaps because of corona discharge and other high-voltage leakage effects, which add spurious current as measured at the ground return). If you try to build an ac-powered instrument to measure the current, say by using a differential amplifier connected across a precision resistor in the high-voltage lead, the power supply of your instrument will have to use a special power transformer rated at 100kV insulation breakdown so that the measuring circuit (op-amps, readouts) can float at 100kV. Since such a transformer is almost impossible to find, this is a good place for a circuit powered by a battery (or possibly by a solar cell, illuminated by light beam across the potential gap), intrinsically isolated from the power line and from earth ground.

In the foregoing example, an ac-powered circuit would have another problem, namely that it would impress some 60Hz ripple onto the circuit it was attempting to measure, due to capacitive coupling and leakage of 60Hz currents through its power transformer. So the power transformer would have to be of special design to ensure low inter-winding capacitance and low leakage currents. This problem of coupled 60Hz ripple can also crop up in conventional circuits that deal with signals at very low levels, for example weak audio signals. Although such problems can usually be solved by careful design, the isolation provided by a battery-powered preamp can be a real advantage in these situations.

#### ***Little power available***

Dialers, modems, remote data-acquisition systems that send through the telephone lines – these are examples of instrument designs that can be powered by the holding

current of the telephone line itself (it's about 50V dc open-circuit, driven from an impedance of about 600Ω; you must load it so that the dc voltage is below 6V, in order for the telephone company to think you're "off hook," and therefore to maintain the connection). Likewise, "smart" thermostats for heating systems often use NiCd rechargeable cells for their dc power, charged during intervals when the relay is not activated by the low-current ac then available (usually a 24V ac transformer in series with a relay coil of a few hundred ohms resistance).

This same trick of powering your circuit with ac signaling currents could also be used with doorbell circuits, and any other application where low-voltage ac relays are used. Another example of extracting power from a signaling current is the use of "industrial sensor current loops," in which a dc current in the range 4mA to 20mA (or, sometimes, 10mA to 50mA) is used to send analog sensor measurements over a two-wire system. Modules using this standard typically permit a voltage drop of 5 to 10 volts; hence the opportunity to power remote instrumentation from the signal current itself.

For these applications you have available a power supply delivering currents of the order of a few milliamps across a few volts, which is enough to power relatively complex low-power circuits. It is certainly attractive to attempt signal-current-driven low-power design, given the more cumbersome alternative of separate ac power sources for the same instrument.

A final example of a power source that limits you to a few milliwatts is the use of solar cells to power instruments (and/or charge their batteries). There are wristwatches and inexpensive pocket calculators built this way, and they have the advantage of (a) staying sealed and (b) remaining inexpensive, respectively.

**No power available**

Battery operation really becomes essential when there's nothing else available. Examples include physical oceanography, where you may wish to deploy a set of sensors on the ocean bottom for six months, quietly logging ocean currents, sediments, salinity, temperature, and pressure, as well as environmental studies, where remote measurements of pollutants at inaccessible sites are required. In these applications you usually want extended operation on a set of batteries, sometimes up to a year or more; hence the need for careful micropower design.

There are other situations where ac power is available, but not convenient. Household examples include smoke detectors and wall clocks.

**Minimizing heat management**

Digital circuitry constructed with ECL or Schottky technologies can easily consume 10 watts or more per board, and a system of several such boards requires forced air cooling. On the other hand, the newer high-speed CMOS logic families (with names such as 74ACxx and 74ACTxx) deliver performance comparable to that of their advanced Schottky cousins, with negligible static power consumption and greatly reduced dynamic power consumption (Figs. 8.18 and 9.2). That means smaller power supplies, closed dirt-free enclosures (no fan), and better long-term reliability.

The same considerations apply to linear design, making low power consumption a desirable objective to keep in mind in almost any application, even when plenty of power is available.

**Uninterruptibility**

Momentary power interruptions often cause microprocessor-based instruments to re-initialize themselves, computers to crash, etc. A nice solution is the use of

an uninterruptible power supply (UPS), usually in the form of a battery-powered dc-to-ac inverter with 115 volt 60Hz output, able to switch on automatically within a few milliseconds of a power interruption. Uninterruptible power supplies are available with power ratings of many kilowatts. The big ones are expensive and bulky; however, there are compact units, powered by a small bank of lead-acid gel cells (see Section 14.02), for systems that use less than a kilowatt of ac power. For truly low-power systems, a small UPS inverter or direct dc battery backup (as in Fig. 1.83) is convenient and is a good reason to practice low-power design.

**POWER SOURCES****14.02 Battery types**

The Duracell "Comprehensive Battery Guide" lists 133 off-the-shelf batteries, with descriptions like zinc-carbon, alkaline manganese, lithium, mercury, silver, zinc-air, and nickel-cadmium. There are even subclasses, for example Li/FeS<sub>2</sub>, Li/MnO<sub>2</sub>, Li/SO<sub>2</sub>, Li/SOCl<sub>2</sub>, and "lithium solid state." And from other manufacturers you can get sealed lead-acid and gel-type batteries. For the truly exotic application you might even want to consider fuel cells or radioactive thermal generators. What are all these batteries? How do you choose what's best for your portable widget?

The foregoing list divides into so-called *primary* and *secondary* batteries. Primary batteries are designed for a single discharge cycle only, i.e., they're nonrechargeable. Secondary cells (NiCd, lead-acid, and gel-type in the foregoing list), by comparison, are designed to be recharged, typically from 200 to 1000 times. Among primary batteries, you usually make your choice of chemistry based on trade-offs among price, energy density, shelf life, constancy of voltage during discharge, peak current capability, temperature range, and availability.



Once you've picked the right battery chemistry, you figure out which battery (or series combination of batteries) has enough energy content for the job.

Fortunately, it's pretty easy to eliminate most of the batteries in the catalogs, if you follow our first suggestion: *Avoid hard-to-get batteries*. Besides being hard to find, they're usually not fresh. So it's usually better to stick with the varieties available at the drugstore, or perhaps photography store, even if it results in somewhat less than optimum design. We particularly recommend the use of commonly available batteries in the design of any consumer electronic device; as consumers ourselves, we shun those inexpensive marvels that use exotic and expensive batteries. (Remember those early smoke detectors that used an 11.2V mercury battery?)

#### □ Primary batteries

Now for details. Table 14.1 compares the characteristics of the various primary cells, and Table 14.2 and Figure 14.1 give actual numbers for the most popular cells.

The old-fashioned "dry cell" with a cat on the outside is a LeClanche cell. Inside it's as primitive as you might guess, with a carbon rod stuck down into a cathode mixture of manganese dioxide, carbon, and ammonium and zinc chloride electrolytes. There's a cylindrical separator made of flour-and-starch paste, then a zinc anode outer can. The top is sealed on with wax and asphalt seals, designed to vent the innards if too much pressure builds up. These cells are the cheapest you can buy, but you don't get too much for your money. In particular, the voltage drops and the impedance rises steadily as the battery is used; furthermore, the battery's capacity drops drastically if used at high currents.

The "heavy-duty" dry cells are similar, but with a higher proportion of zinc chloride and correspondingly different mechanical design to accommodate greater

gassing. Although their total energy content is only slightly greater than that of LeClanche cells, these cells are considerably better in delivering most of their rated capacity even when operated at high currents. For example, a LeClanche D cell delivers 4.2 amp-hours (Ah) into a 150 ohm load, 1.2Ah into 15 ohms, and 0.15Ah into 1.5 ohms; the equivalent zinc chloride cell delivers 5.6, 5.4, and 1.4 amp-hours, respectively. The zinc chloride cell also shows less dropoff of capacity at low temperatures.

The alkaline manganese cell, generally sold as simply "alkaline," is better still in high-current-discharge and low-temperature operation. It is inside-out, compared with zinc-carbon, having the powdered-zinc negative anode and potassium hydroxide electrolyte in the middle, surrounded by a manganese-dioxide-and-carbon outer positive cathode. For comparison with the numbers above, an alkaline D cell delivers 10 amp-hours into 150 ohms, 8Ah into 15 ohms, and 4Ah into 1.5 ohms. Because of its particular chemistry, an alkaline battery maintains a low and slowly increasing internal resistance as it discharges, compared with the rapidly rising internal resistance of both types of zinc-carbon cells. It also works better at low temperatures. Alkaline batteries have a longer shelf life than LeClanche or zinc chloride. As Figure 14.1 suggests, the cell voltage-versus-discharge curve for all three types of batteries lets you easily estimate the condition of the battery. Figure 14.2 shows comparative performance for the three kinds of "dry cells."

Mercury, silver oxide, lithium – these are the real premium cells, with greatly superior performance compared with alkaline and zinc-carbon. The mercury cell uses an amalgamated zinc anode, mercuric-oxide-plus-carbon cathode, and sodium or potassium hydroxide electrolyte. It excels in constancy of open-circuit

TABLE 14.1. PRIMARY BATTERIES

| Type                                                   | Advantages                                                                                                                                                                                                                                       | Disadvantages                                                                                                                                                                       |
|--------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Zinc-carbon (LeClanche)<br>(standard "dry cell")       | least expensive<br>widely available                                                                                                                                                                                                              | lowest energy density (1–2Wh/in <sup>3</sup> )<br>sloping discharge curve<br>poor high-current performance<br>impedance increases as discharged<br>poor low-temperature performance |
| Zinc-carbon (zinc chloride)<br>("heavy duty" dry cell) | less expensive than alkaline<br>better than LeClanche at high curr<br>and low temp                                                                                                                                                               | low energy density (2–2.5Wh/in <sup>3</sup> )<br>sloping discharge curve                                                                                                            |
| Alkaline manganese<br>("alkaline" dry cell)            | moderate cost<br>better than zinc chloride at high curr<br>and low temp<br>maintains low impedance as discharged<br>moderate energy density (3.5Wh/in <sup>3</sup> )<br>widely available                                                         | sloping discharge curve                                                                                                                                                             |
| Mercury                                                | high energy density (7Wh/in <sup>3</sup> )<br>flat discharge curve<br>good at high temperatures<br>good shelf life<br>low and constant impedance<br>open-circuit voltage 1.35V±1%                                                                | expensive<br>poor at low temp (<0°C)                                                                                                                                                |
| Silver oxide                                           | high energy density (6Wh/in <sup>3</sup> )<br>flat discharge curve<br>good at high & low temp (to -20°C)<br>excellent shelf life                                                                                                                 | expensive                                                                                                                                                                           |
| Lithium oxyhalide                                      | high energy density (8Wh/in <sup>3</sup> )<br>highest energy density per unit weight<br>flat discharge curve<br>excellent at high & low temp (to -55°C)<br>extraordinary shelf life (5–10 yrs @70°C)<br>light weight<br>high cell voltage (3.0V) | expensive                                                                                                                                                                           |
| Lithium solid-state                                    | high energy density (5–8Wh/in <sup>3</sup> )<br>excellent at high & low temp<br>(-40°C to 120°C)<br>unbelievable shelf life (>20y @ 70°C)<br>light weight                                                                                        | expensive<br>low current drain only                                                                                                                                                 |

TABLE 14.2. BATTERY CHARACTERISTICS

| Type          | R <sub>int</sub><br>(Ω) | V <sub>oc</sub><br>(V) | Capacity <sup>a</sup><br>continuous, to 1V/cell |        |                    |        | Size<br>(in) | Weight<br>(gm) | Connect <sup>b</sup> | Comments                                         |
|---------------|-------------------------|------------------------|-------------------------------------------------|--------|--------------------|--------|--------------|----------------|----------------------|--------------------------------------------------|
|               |                         |                        | (mAh)                                           | @ (mA) | (mAh)              | @ (mA) |              |                |                      |                                                  |
| 9V "1604"     |                         |                        |                                                 |        |                    |        |              |                |                      |                                                  |
| Le Clanche    | 35                      | 9                      | 300                                             | 1      | 160                | 10     | 0.65x1x1.9   | 35             | S                    | 280mAh@100mA<br><br>Kodak Li-MnO <sub>2</sub>    |
| Heavy Duty    | 35                      | 9                      | 400                                             | 1      | 180                | 10     | "            | 40             | S                    |                                                  |
| Alkaline      | 2                       | 9                      | 500                                             | 1      | 470                | 10     | "            | 55             | S                    |                                                  |
| Lithium       | 18                      | 9                      | 1000                                            | 25     | 950                | 80     | "            | 38             | S                    |                                                  |
| 1.5V Alkaline |                         |                        |                                                 |        |                    |        |              |                |                      |                                                  |
| D             | 0.1                     | 1.5                    | 10000                                           | 10     | 8000               | 100    | 1.3Dx2.2L    | 125            | B                    | 4000mAh @ 1A                                     |
| C             | 0.2                     | 1.5                    | 4500                                            | 10     | 3200               | 100    | 1.0Dx1.8L    | 64             | B                    |                                                  |
| AA            | 0.4                     | 1.5                    | 1400                                            | 10     | 1000               | 100    | 0.55Dx1.9L   | 22             | B                    |                                                  |
| AAA           | 0.6                     | 1.5                    | 600                                             | 10     | 400                | 100    | 0.4Dx1.7L    | 12             | B                    |                                                  |
| Mercury       |                         |                        |                                                 |        |                    |        |              |                |                      |                                                  |
| 625           | —                       | 1.35                   | 250 <sup>c</sup>                                | 1      | 250 <sup>c</sup>   | 10     | 0.62Dx0.24L  | 4              | B                    |                                                  |
| 675           | 10                      | 1.35                   | 190 <sup>c</sup>                                | 0.2    | —                  | —      | 0.64Dx0.21L  | 2.6            | B                    |                                                  |
| 431           | —                       | 11.2                   | 1000 <sup>c</sup>                               | 25     | —                  | —      | 1.0Dx2.9L    | 115            | S                    |                                                  |
| Silver        |                         |                        |                                                 |        |                    |        |              |                |                      |                                                  |
| 76            | 10                      | 1.55                   | 180                                             | 1      | —                  | —      | 0.46Dx0.21L  | 2.2            | B                    |                                                  |
| Li-Oxyhalide  |                         |                        |                                                 |        |                    |        |              |                |                      |                                                  |
| D             | —                       | 3.9                    | 14000 <sup>d</sup>                              | 175    | 10500 <sup>d</sup> | 350    | 1.3Dx2.3L    | 113            | B,T                  | SOCl <sub>2</sub> /BrCl                          |
| D             | —                       | 3.95                   | 14000 <sup>d</sup>                              | 175    | 12000 <sup>d</sup> | 1000   | "            | 110            | B,T                  | SO <sub>2</sub> Cl <sub>2</sub> /Cl <sub>2</sub> |
| D             | —                       | 3.5                    | 9500 <sup>d</sup>                               | 175    | 8500 <sup>d</sup>  | 1000   | "            | 120            | B,T                  | SOCl <sub>2</sub>                                |
| Li solid      | —                       | 4.0                    | 350 <sup>d</sup>                                | 1μA    | 175 <sup>d</sup>   | 0.1    | 1.2Dx0.23L   | 16             | T                    | high impedance                                   |
| Ni-Cd         |                         |                        |                                                 |        |                    |        |              |                |                      |                                                  |
| D             | 0.009                   | 1.3                    | 4000 <sup>c</sup>                               | 800    | 3500 <sup>c</sup>  | 4000   | 1.3Dx2.3L    | 130            | B                    | Safe/Powerasonic                                 |
| 9V            | 0.84                    | 8.1                    | 100 <sup>c</sup>                                | 10     |                    |        | 0.65x1x1.9   | 35             | S                    |                                                  |
| Pb-acid       |                         |                        |                                                 |        |                    |        |              |                |                      |                                                  |
| D             | 0.006                   | 2.0                    | 2500 <sup>e</sup>                               | 25     | 2000 <sup>e</sup>  | 1000   | 1.3Dx2.6L    | 180            | T                    |                                                  |

(a) see Fig. 14.1 for discharge curves. (b) B - button; S - snap; T - solder tabs. (c) to 0.9V/cell. (d) to 2.5V/cell. (e) to 1.75V/cell

voltage (1.35V, stable to 1%) as well as constancy of voltage during discharge (a "flat discharge curve"); see Figure 14.1. It performs well at temperatures up to 60°C, but performance is seriously degraded below -10°C.

The silver oxide cell is similar to the mercury cell, but with the mercuric oxide replaced by silver oxide. It, too, has a very flat discharge curve, but with higher open-circuit voltage (1.6V) and improved performance at low temperatures (to -20°C).

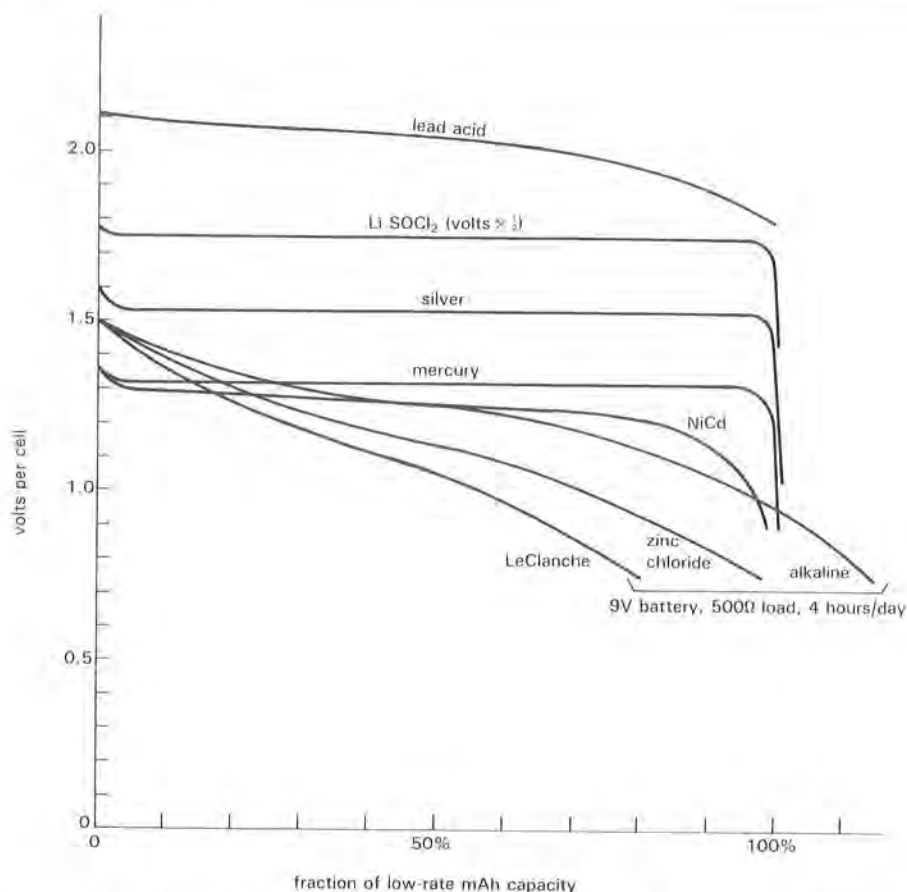
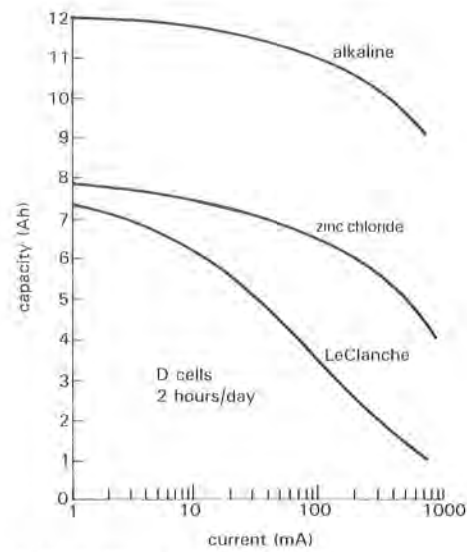


Figure 14.1. Discharge curves for primary batteries. (This and subsequent figures in this chapter are adapted from battery literature by Arco Solar, Duracell, Electrochem Industries, Eveready, Gates, Kodak, PowerSonic, Solavolt, and Yuasa.)

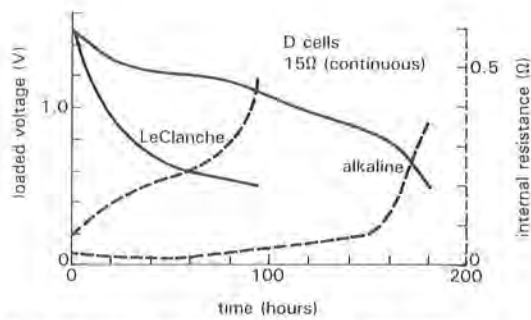
Lithium cells are the newest of the commercially available primary cells, with several different chemistries available. They have the highest energy density per unit weight. They are the best performers at very high and low temperatures, and they have extraordinary shelf life at all temperatures. For example, a D-size lithium thionyl chloride ( $\text{Li/SOCl}_2$ ) cell delivers more than three times the energy (10Ah at a terminal voltage of 3.5V) of an alkaline D cell, with comparable size and weight. Lithium batteries will operate down to  $-50^\circ\text{C}$  and up to  $70^\circ\text{C}$  (see Fig. 14.3), with 50% of their room-temperature service at  $-40^\circ\text{C}$ , a temperature at which other primary batteries cease to function

at all. Lithium batteries have shelf lives of 5 to 20 years at room-temperature and can be stored for 1–2 years at  $70^\circ\text{C}$ , a temperature that makes other batteries wither. They have a flat discharge curve. Their long shelf life and 3–3.5 volts make lithium batteries ideal for on-board CMOS memory backup.

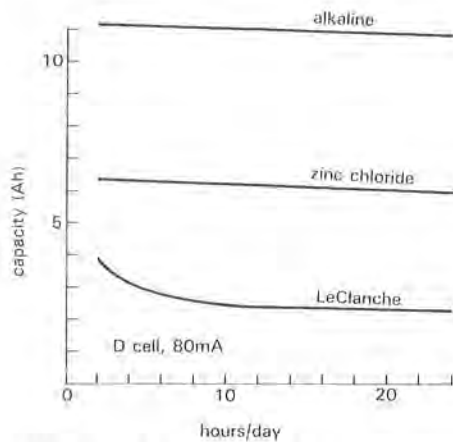
Each lithium chemistry system has its own peculiarities. For example, lithium thionyl chloride batteries have a tendency to develop an electrode passivation that raises their internal resistance enormously; it can be “burned off” by momentary operation at high current. Lithium sulfur dioxide has been implicated in some battery explosions:



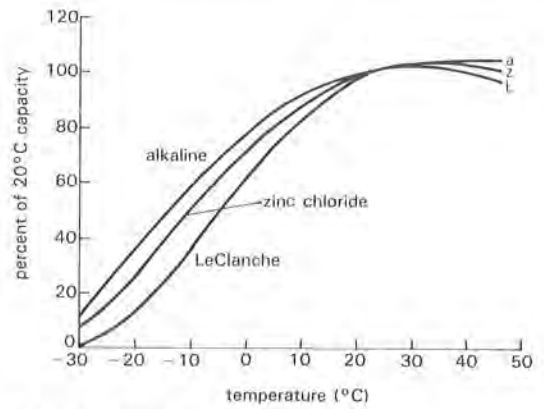
A. D-cell capacity versus load current



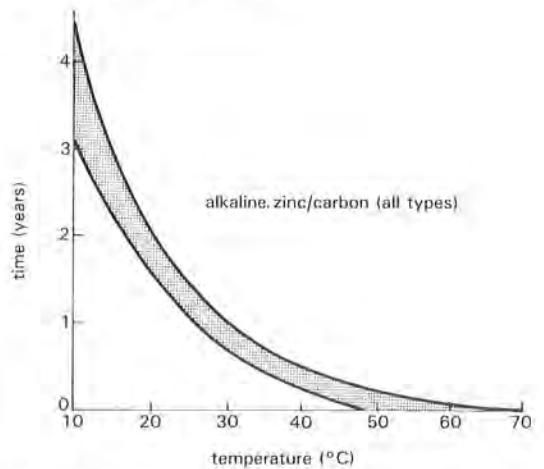
B. D-cell voltage and internal resistance during discharge



C. D-cell capacity versus duty cycle



D. capacity versus temperature



E. shelf life (10% loss of capacity) versus temperature

Figure 14.2. Zinc “dry-cell” performance comparison.

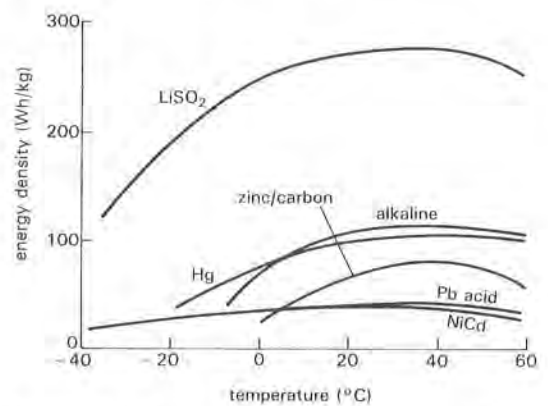


Figure 14.3. Energy density versus temperature for primary cells.



**Warning:** There have been incidents in which lithium batteries have exploded, in some cases causing severe personal injury. Having warned you about this, we will not be responsible for any calamities you might experience at the hands of a lithium battery.

#### □ Secondary batteries

For use in electronic equipment, your choices are (a) nickel-cadmium ("nicad") or (b) sealed lead-acid. Both have lower energy content than primary cells (Table 14.2), but they *are* rechargeable. Nicad cells provide 1.2 volts, are generally available in the 100mAh–5Ah range, and work down to  $-20^{\circ}\text{C}$  (and up to  $+45^{\circ}\text{C}$ ); lead-acid batteries provide 2 volts per cell, are generally built to provide 1 to 20 amp-hours, and will work down to  $-65^{\circ}\text{C}$  (and up to  $+65^{\circ}\text{C}$ ). Both types have relatively flat discharge curves. Lead-acid batteries have low self-discharge rates and are claimed to retain two-thirds of their charge after a year's storage at room temperature (though our experience leads us to be skeptical); nicad batteries have relatively poor charge retention, typically losing half their charge in 4 months (which we do believe!) (see Fig. 14.4). A nicad D cell provides 5Ah (at 1.2V), whereas a lead-acid D cell provides 2.5Ah (at 2V); the comparable alkaline cell provides 10Ah at 1.5 volts.

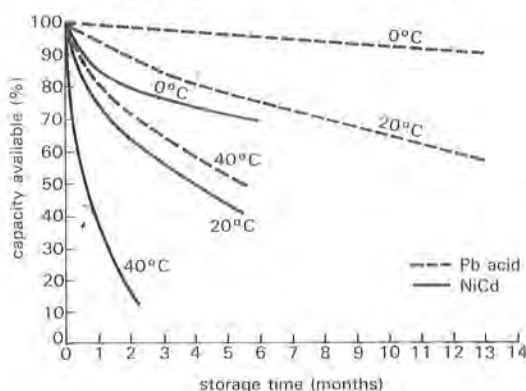


Figure 14.4. Secondary battery charge retention.

Both nicad and sealed lead-acid batteries claim to be good for 250–1000 charge/discharge cycles (more if they are only partially discharged each time; less if completely discharged, or charged/discharged rapidly). Nicads have an overall life expectancy of 2–4 years if held at a constant trickle charge current (see below); the comparable life for sealed lead-acid batteries held at constant "float" voltage is claimed to be 5–10 years.

It's worth pointing out that these rechargeable batteries really are sealed; they won't drip mysterious and terrible chemicals. In particular, although the name "lead-acid" conjures up images of husky car batteries with corroded terminals and leaking acid, the sealed types really are clean batteries: You can run them in any position, they don't drip or ooze, and they're generally well-behaved. In our experience you can design them into real electronic instruments without fear of your circuit boards dissolving into a white crusty plague, or the bottom of your expensive enclosure becoming awash with foul-smelling bilge juices.

Secondary batteries die young if they aren't charged correctly. The procedures are different for nicad and lead-acid. It's conventional to designate charging rates in terms of the ampere-hour capacity of the battery; for example, charging at "C/10" means applying a charging current equal to one-tenth of the ampere-hour capacity of the fully charged battery. For the nicad D cell above, that would be 500mA.

□ **Nickel-cadmium.** Nicad cells are designed to be charged at constant current and to withstand continuous charging at C/10. Because of inefficiencies in the charge/discharge cycle, you have to charge at this rate for 14 hours to guarantee a full charge; you can think of this as charging the battery 140%.

Although it's OK to overcharge nicad batteries indefinitely at  $C/10$ , it's better to switch over to a "trickle" charge, typically at  $C/30$  to  $C/50$ . However, nicads are funny, with a "memory" effect, so that a trickle rate may fail to revive a fully discharged battery; a minimum of  $C/20$  is recommended.

There are applications where you can't wait a whole day for nicads to recharge. Nicad literature gives you permission to charge normal cells at a "high rate" of  $C/3$  to  $C/10$ , if you don't do it too much. Up to three days at  $C/3$  is about the limit. There may be some venting of gases under these conditions, in contrast to "normal" charging at  $C/10$  in which internally evolved oxygen gets recombined within the cell. There are special "fast-charge" nicad cells designed to be charged at  $C/1$  to  $C/3$  in a special charger that senses the fully charged condition by monitoring cell temperature (they have internal chemistry that makes them heat rapidly once they are charged). Unlike the situation with lead-acid batteries, you can't reliably determine when a nicad is fully charged by monitoring terminal voltage, because it changes with repetitive cycling, temperature, and rate. Nicads should not be charged by a constant voltage, nor held "floated" at a fixed voltage.

You can buy handy little nicad chargers from several companies, including the battery manufacturers themselves. They typically let you charge all the popular sizes (D, C, AA, and 9V).

Nicads have a pathology all their own. If you're like us, you probably take it personally when your rechargeable calculator dies during a tax audit. As the graphs show (Fig. 14.5), nicads have "memory" effects, so that the first discharge after a long period of charging may be poor. They are intolerant of reverse polarity; thus, the first cell to discharge suffers horribly if a series string of them is fully discharged.

Likewise, nicads shouldn't be connected in parallel. You'll find people promoting various snake remedies, such as periodic "deep discharge," or shock therapy in the form of a substantial electrolytic capacitor discharged across a moribund nicad. Although we're skeptical of the latter, periodic deep discharge is important for nicad health.

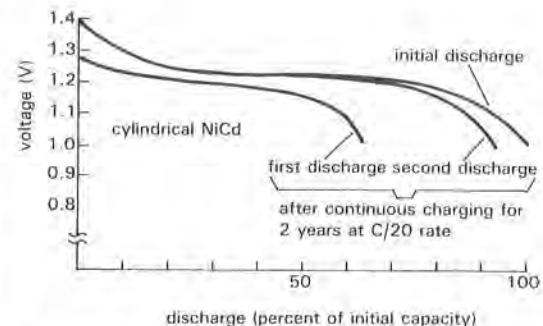


Figure 14.5. Nicad battery restored to good health by "deep discharge."

- **Lead-acid.** These versatile batteries can be charged by applying a current-limited constant voltage, a constant current, or something in between. With current-limited constant-voltage charging, you apply a fixed voltage (typically between 2.3V and 2.6V per cell); the battery initially draws a high current (up to  $2C$ ), but tapers down as it charges, eventually leveling off to a trickle current that maintains the battery in a fully charged state. A higher applied voltage gives you a faster charge, at the expense of greater required charger current and reduced overall battery life. A simple implementation is to use a 3-terminal regulator like the 317 to supply a current-limited fixed voltage. The battery's charge can be held indefinitely by maintaining a fixed "float" voltage between 2.3 and 2.4 volts per cell (corresponding to a trickle current of  $C/1000$  to  $C/500$ ).

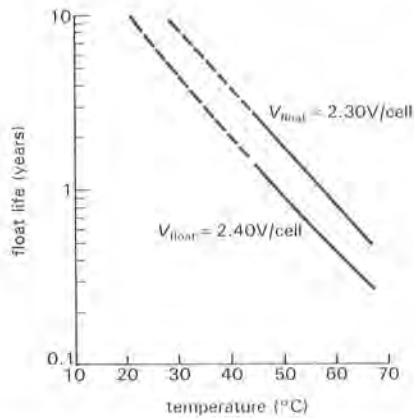


Figure 14.6. Excessive float voltage reduces life of lead-acid batteries.

Figure 14.6 shows the trade-offs. These charging and floating voltages are mildly temperature-dependent and should be adjusted by  $-4\text{mV}/^\circ\text{C}$  for operation at extremes of temperature.

With constant-current charging (which is not often used) you apply a fixed current, typically  $C/5$  to  $C/20$ ; the battery voltage rises gradually as the battery charges, then increases dramatically as full charge is reached. At this point (indicated by a terminal voltage of  $2.5\text{V}/\text{cell}$ ) you must reduce the current, typically to a fixed  $C/500$  rate, which will maintain full charge indefinitely. Sealed lead-acid batteries will give 8–10 years of service while being charged at a  $C/500$  rate.

A nice lead-acid charging method is the so-called two-step technique (Fig. 14.7): After a preliminary “trickle” charge, you begin with a high-current “bulk-charge” phase, applying a fixed high current  $I_{\text{max}}$  until the battery reaches the “overcharge voltage,”  $V_{\text{OC}}$ . You then hold the voltage constant at  $V_{\text{OC}}$ , monitoring the (dropping) current until it reaches the “overcharge transition current,”  $I_{\text{OCT}}$ . You then hold a constant “float voltage,”  $V_F$ , which is less than  $V_{\text{OC}}$ , across the battery. For a 12 volt 2.5Ah lead-acid battery, typical values are  $I_{\text{max}} = 0.5$  amp,  $V_{\text{OC}} = 14.8$  volts,  $I_{\text{OCT}} = 0.05$  amp, and  $V_F = 14.0$  volts.

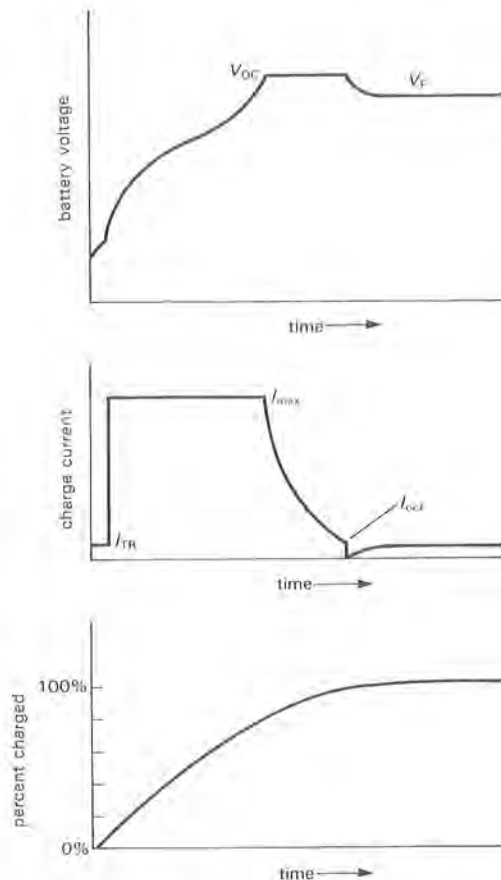


Figure 14.7. Recommended lead-acid battery charging cycle (“two step”).

Although this all sounds rather complicated, it results in rapid recharge of the battery without damage. Unitrode makes a nice IC, the UC3906, that has just about everything you need to do the job. It even includes an internal voltage reference that tracks the temperature characteristics of lead-acid cells and requires only an external *pnp* pass transistor and four parameter-setting resistors.

#### Battery availability and recommendations

As we said at the outset, it’s really a good idea to design your instrument to use a popular and readily available battery.

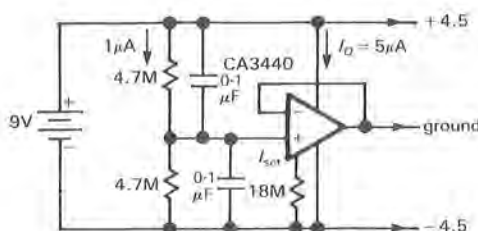


Figure 14.8. Creating a split supply from a single battery.

Tops on the list are the 9 volt “transistor” batteries, known generically as NEDA 1604 (1604, LeClanche; 1604D, heavy duty; 1604A, alkaline; 1604M, mercury; 1604LC, lithium; 1604NC, nicad). You can buy 9 volt alkaline batteries in every corner grocery store (or open-air market) in the world. Op-amps work well on  $\pm 9$  volt supplies; you can even use  $\pm 4.5$  volts if you use a resistive divider and a follower to generate a midpoint “ground” (Fig. 14.8; discussed further in Section 14.08). There are nice little plastic instrument cases, complete with 9 volt battery compartment underneath, available from many manufacturers at very attractive prices. We recommend using alkaline, rather than zinc-carbon, because of the improved electrical characteristics discussed earlier. Kodak’s new “Ultralife” lithium 9 volt battery looks like a real winner, with 1000mAh capacity, long shelf life (80% retention after 10 years), and flat discharge curve (Fig. 14.9); they wisely used 3 cells, not 2, so its terminal voltage

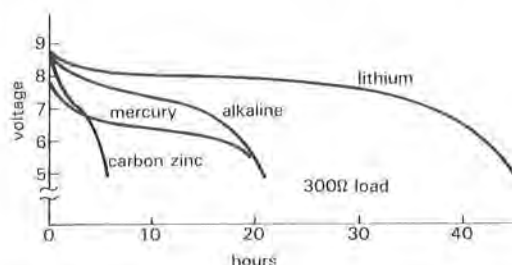


Figure 14.9. 9 volt battery discharge curves; “lithium” is the 3-cell Kodak “Ultralife.”

is close to 9 volts, the same as alkaline. Our preliminary measurements on some early samples showed a rather high internal impedance, however.

The familiar alkaline AA, C, and D cells have more energy capacity (and correspondingly lower internal impedance) than the 500mAh of the 1604A (3, 9, and 20 times as much, respectively), and they’re just as easy to get. But they are somewhat less convenient because of the problem of holding and connecting reliably to a group of series-connected cells. Everyone has noticed that if you shake a dim flashlight, it usually gets brighter. The problem is compounded by the tendency of some types (alkaline, mercury) to grow white deposits on the terminals (this is officially known as “salting”).

Nicads are also available (though not in every drugstore) in the standard battery sizes (AA, C, D, and 9V), for applications where it makes sense to use secondary batteries. But you get only about 25%–50% the energy capacity, and reduced battery voltage (1.2V versus the alkaline’s 1.5V per cell).

Lithium batteries are available in the same standard battery sizes, although they provide 3 volts or more per cell. Most manufacturers also provide them with solder tabs for more reliable connections; this makes good sense, considering their extended shelf life. Lithium cells are also available in flat “button” shapes with solder tabs, for use as CMOS memory backup, or to power calendar clock chips. Kodak’s 9 volt lithium battery has a nice twist, namely gold-plated snap tabs, for more reliable connections. See our earlier warning on lithium-battery detonation.

Sticking to batteries you can get in any small town, in most photo stores you’ll find a selection of mercury, silver, and lithium cells. They’re meant to go into cameras (and calculators and watches) and are generally of the “button” variety. For

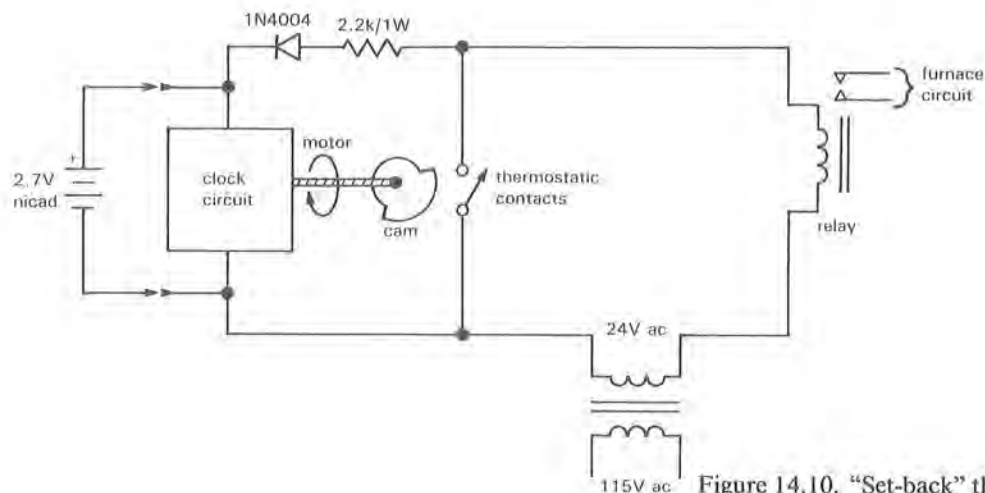


Figure 14.10. "Set-back" thermostat.

example, there's the popular 625 mercury button, hardly larger than a coat button and good for 250mAh. The smaller 76 silver oxide cell (and energetically equivalent type 675 mercury cell) has an interesting twist, namely an offspring lithium 3 volt cell (NEDA 5008L) of the same diameter and twice the height, intended to substitute for a pair of the 1.5 volt cells. At this voltage you can run CMOS logic directly, as well as low-voltage op-amps like the LM10 and the ICL7610 series and TI's versatile "LinCMOS" series of op-amps (the TLC251-254 series) and comparators (TLC372/4, TLC339/393, and TLC3702/4).

If your application requires the rechargeability and high peak currents of sealed lead-acid batteries, or some exotic form of primary battery, you'll generally have to deal with the battery manufacturers or their distributors. Look for names like Gates, Powersonic, and Yuasa for lead-acid. Duracell and Eveready dominate the primary-cell market. All of these companies have helpful and extensive data books on batteries and battery lore.

In the next few sections we will consider alternative power sources – wall-plug-in modules, solar cells, and signal currents – for low-power equipment. It's worth remembering that each of these power sources can be used to charge secondary

batteries. For example, the popular "set-back" thermostats that turn down the heat at night use the high-impedance 24 volt ac relay signaling current to charge nicads and keep the clock running during periods of relay-ON (Fig. 14.10).

Table 14.3 summarizes our advice on the relative merits of various primary-cell battery types.

TABLE 14.3. PRIMARY-BATTERY ATTRIBUTES

|                      | 9V alkaline | 1.5V alkaline | Mercury | Silver | Lithium |
|----------------------|-------------|---------------|---------|--------|---------|
| <i>Properties</i>    |             |               |         |        |         |
| Inexpensive          | •           | •             | –       | –      | –       |
| Available            | •           | •             | –       | –      | –       |
| Wide temp range      | –           | –             | –       | –      | •       |
| Stable voltage       | –           | –             | •       | •      | •       |
| Reliable contacts    | •           | –             | –       | –      | •       |
| Good at high current | •           | •             | –       | –      | •       |
| Long shelf life      | –           | –             | –       | –      | •       |
| Miniature            | –           | –             | •       | •      | •       |
| <i>Applications</i>  |             |               |         |        |         |
| Linear circuits      | •           | –             | –       | –      | –       |
| Low voltage CMOS     | –           | •             | •       | •      | •       |
| 4000-series CMOS     | •           | –             | –       | –      | •       |
| CMOS backup          | –           | –             | –       | –      | •       |



### 14.03 Wall-plug-in units

Calculators, modems, tape recorders, telephone dialers, small measurement instruments – more and more low-power devices come with those familiar square black wall-plug-in power units (Fig. 14.11). Although they're usually labeled to match the instrument they power, you can easily get them in a variety of ratings, both in large quantities and small (distributors like Radio Shack and Digi-Key each stock a few types). The best news is the incredibly low price: A 9 volt 500mA (unregulated dc) wall plug-in costs about \$2.50 in quantity.

Wall plug-ins are a good way to power small instruments that need more power than you can get from batteries, or that need to keep their rechargeable batteries charged. They're cheaper than internally mounted discrete or modular power supplies, and by using them you save space and keep heat (and high voltage) outside the instrument. Furthermore, they usually satisfy UL and CSA safety requirements, which is important if you want to market an instrument without the lengthy UL approval process.

Wall plug-ins come in three flavors: plain step-down transformers, filtered but unregulated dc supplies, and complete regulated supplies (both linear and switchers). All come in a variety of voltages and currents, and the regulated supplies even come in useful combinations like +5 volts at 1 amp and  $\pm 15$  volts at 250mA. They have the usual features of IC regulators, namely current limiting and thermal shutdown, as well as optional overvoltage crowbar. You can get them with three-prong (grounding) wall plugs and with various output connectors; many of the larger units are also available as free-standing desk units, with an ac power cord. One word of caution: There is no standardization of connector type and voltage ratings. In fact, there is not even standardization of polarity! So it is effortless to blow out an instrument by plugging the wrong wall unit into it. Beware!

An extensive line of high-quality wall power units is manufactured by Ault (Minneapolis, MN). For inexpensive imported units, look at the catalogs of Conдор (Sunnyvale, CA) or Multi Products



Figure 14.11. Wall-plug-in power supplies. (Courtesy of Ault, Inc.)

International (Cedar Grove, NJ). Check the EEM (see Bibliography) for addresses and additional manufacturers.

#### □ 14.04 Solar cells

A combination of lead-acid or nicad battery plus silicon solar cells forms a good power source for a moderate-power instrument that is to be deployed at a remote site for extended periods. For example, you might want to tether a buoy that makes ocean measurements and transmits them periodically. If the average power consumption is 1 watt, primary batteries become prohibitively bulky (you'd have to use 500 alkaline D cells to last 1 year). Full sunlight delivers about 1kW per square meter after traversing the atmosphere; after accounting for the inefficiencies of solar cells (they're about 10% efficient when operating into the correct load) and the daylight and weather cycle at middle northern latitudes (where you average 100W/m<sup>2</sup> in winter, 250W/m<sup>2</sup> in summer), you can average about 25 watts (in July) or 10 watts (in January) per square meter from good quality solar cells, which cost about \$800 in 1986. In peak sunlight, such a solar module delivers 100 watts to a matched load.

With a bank of secondary cells for energy storage (lead-acid is better than nicad, because of its long life and wide operating temperature range), you can withdraw nearly this average power *continuously*; lead-acid cells are typically 70%–80% efficient, so, all factors (including weather) considered, you can withdraw something like 8 watts per square meter (winter) to 20 watts per square meter (summer), averaged over 24 hours.

For low-power instruments that only need to operate in bright light, you can omit the battery. Solar-powered CMOS calculators with liquid crystal displays are a boon to battery haters everywhere.

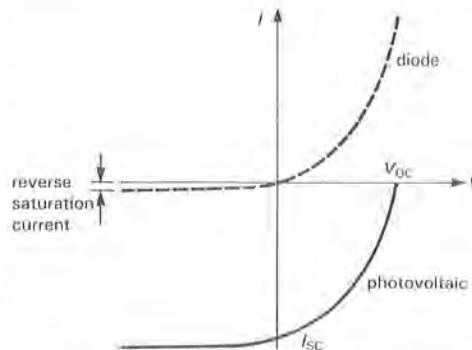


Figure 14.12. Solar-cell output voltage versus load current is simply a displaced diode  $V$ - $I$  curve.

#### □ VI characteristics

Silicon solar cells have a simple and very useful volt-ampere characteristic. It turns out that the open-circuit voltage is almost independent of light level and averages about 0.5 volt per cell; the  $V$ - $I$  curve is simply a displaced diode curve (Fig. 14.12). Typical solar panels consist of 36 series-connected cells, for an open-circuit voltage of about 18 volts. The terminal voltage stays nearly constant as load current is increased up to a maximum current, at which point the solar module becomes roughly a constant current source for further decreases in load impedance. The maximum current scales linearly with light level, giving a set of characteristic curves as shown in Figure 14.13. Solar cells work best when

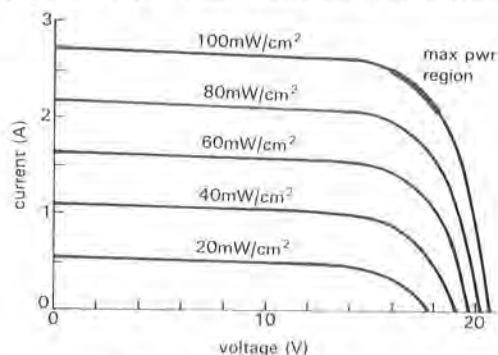


Figure 14.13. Solar-cell output versus irradiance (Solavolt MSVM4011).

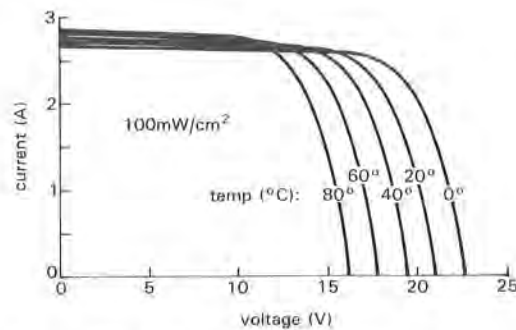


Figure 14.14. Solar-cell output versus temperature (Solavolt MSVM4011).

cold, since the open-circuit voltage drops with increasing temperature (Fig. 14.14).

For a given light level, maximum power is delivered when the operating point has maximum product  $VI$ ; in other words, the point on the  $V$ - $I$  curve that touches a family of hyperbolas (constant product  $VI$ ) plotted on the same axes. Roughly speaking, that's at the knee of the  $V$ - $I$  curve. Since the load impedance that corresponds to the knee changes rapidly with light level, you can't expect to maintain an optimum load (which would be a load impedance increasing inversely with light level, or, put another way, a load that draws a current proportional to light level, at roughly constant voltage). However, for low-power applications it isn't essential that the load extract maximum power – all that matters is that the load be powered under normal lighting conditions. That's the case for those solar-powered calculators, whose CMOS circuits draw so little current that there is plenty of reserve power except under very low light levels. Because of the wide voltage range of 74C/4000B “high-voltage” CMOS (3V to 18V), and the fact that solar cells have an open-circuit voltage that is relatively independent of light level, you don't need to use any voltage regulators; just power the CMOS directly from the module, with bypass capacitors, of course. A typical small module like the Solarex SX-2 provides 290mA at 8.5 volts

in sunlight and has an open-circuit voltage of 11 volts; you could use it, unregulated, for high-voltage CMOS, or, with a regulator, for any +5 volt logic family.

For any application that uses secondary cells for energy storage, it's worth noting the rather good match of solar-cell  $VI$  characteristics to the charging requirements of lead-acid cells. A solar module provides roughly constant charging current into a discharged battery, changing over to a constant-voltage “float” as the battery voltage rises at the end of charging. The temperature coefficient of open-circuit voltage ( $-0.5\%/^{\circ}\text{C}$ ) is a fair match to the recommended float-voltage tempco for lead-acid batteries ( $-0.18\%/^{\circ}\text{C}$ ). So some suppliers make solar modules that are intended to charge lead-acid batteries directly, for example the Arco M65 (2.9A @ 14.5V). The more usual way to match solar modules to the charge/float characteristics of lead-acid batteries is with a series or shunt regulator circuit designed for the job. Many solar modules are designed to work this way, with 20 volts open-circuit voltage and matching regulator module for charging 12 volt batteries. The regulators switch over from charging to temperature-compensated floating, with automatic load disconnection if the battery voltage drops too low. These systems are available for multiples of 12 volt systems (24V, 36V, 48V, etc.), and you can get accessories such as 60Hz inverters (to make ac), or dc-operated refrigerators, attic ventilators, etc.

Some of the bigger names in solar modules and systems are Arco Solar (Chatsworth, CA), Mobil Solar (Waltham, MA), Solarex (Rockville, MD), and Solavolt (Phoenix, AZ).

#### 14.05 Signal currents

Don't forget about the possibility of using signaling currents to power a micropower instrument. Four of the more

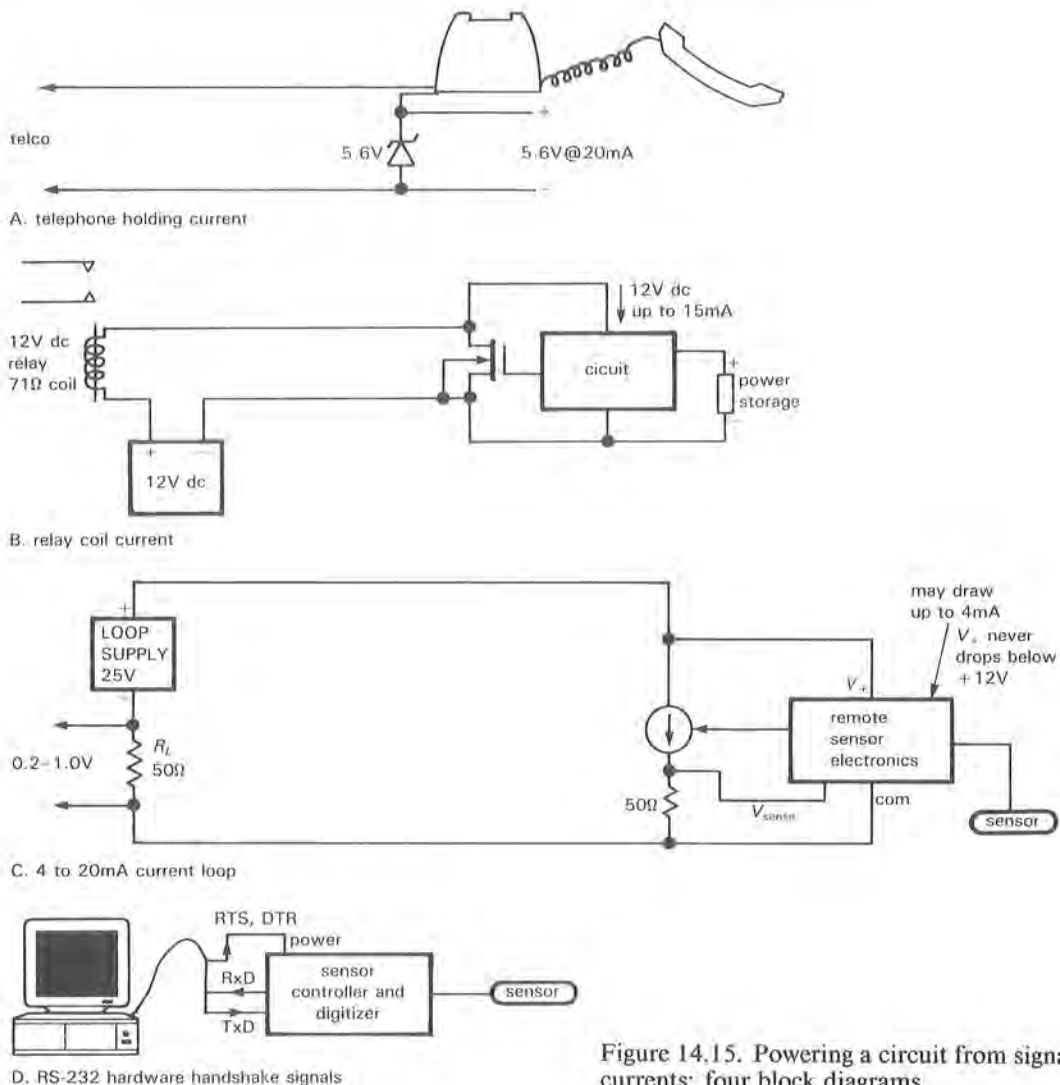


Figure 14.15. Powering a circuit from signal currents: four block diagrams.

common opportunities (Fig. 14.15) are (a) the dc holding current flowing through a telephone circuit that is “off hook,” (b) the ac or dc voltage available from a relay circuit when it is not energized, (c) the 4–20mA dc current used for industrial-sensor current-loop signaling, and (d) the serial-port RS-232C bipolarity “handshake” signals (RTS, DSR, etc.). In the first two schemes, your source of power is available only part of the time – power disappears when the phone is hung up (“on hook”) or when the relay circuit

is energized by your instrument. If you need power continuously, you’ll have to use rechargeable batteries, charged during periods of available power; for very low current loads, another possibility is a high-capacitance (up to 5 farads) “double-layer” capacitor, the same type used for CMOS memory retention.

Each of these power sources has rather strict limits on voltage compliance or maximum current. Here are their characteristics, and some hints on the parasitic use of these power sources.

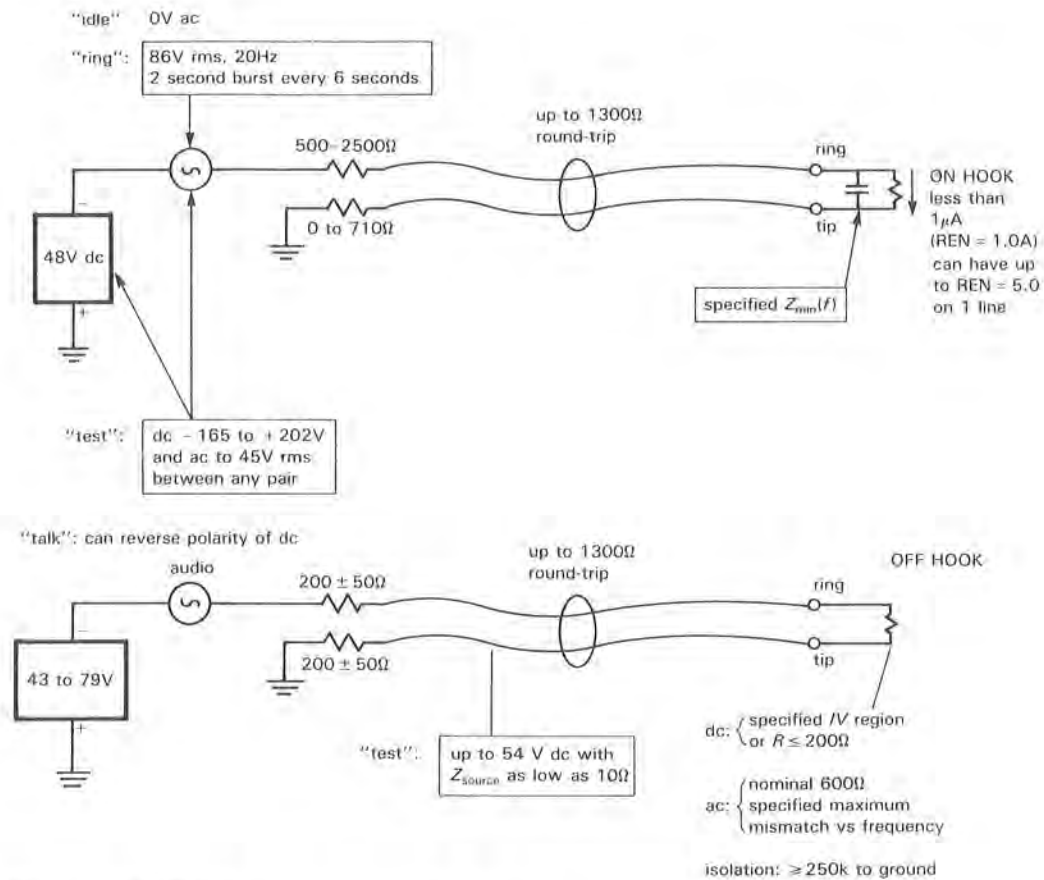


Figure 14.16. Telephone operating and test states.

#### □ Telephone-line power

There are several different states that the phone line can be in, depending upon what your phone is doing, and what the phone company is doing to you. The central office (or nearby equivalent) applies various dc (and ac) voltages to the two-wire phone loop (labeled "tip" and "ring") during these various stages of call progress (Fig. 14.16). In the *idle* state, the telephone company central office applies  $-48(\pm 6)$  volts dc in series with 500 to 2500 ohms to the "ring" line, and terminates the "tip" line to ground with 0 to 710 ohms. In addition, there is typically up to 1300 ohms of

external line resistance between the central office and you (the "subscriber"). When you go off-hook, the central office goes into *dialing* mode, applying a dial tone and a dc level of  $-43$  to  $-79$  volts in series with 200 ohms ( $\pm 50\Omega$ ) on "ring," and terminating "tip" with the same impedance to ground. The same dc voltage and source impedances are present in the *talking* state (after the connection is made), although the telephone company may, at its discretion, reverse the polarity of the dc voltage applied to "ring." Of course, in the talking state you also have audio signals superposed on the dc, which is the whole purpose of the telephone!



There are two other states. During *ring*, the phone company applies 86 ( $\pm 2$ ) volts rms, at 20Hz, on top of the usual  $-48$  volt dc bias. As with the dc, the ac ringing signal is applied to the “ring” lead. The official ringing specification is 2 second bursts at 6 second intervals. During *test* mode, the phone company applies various ac and dc test signals to make sure that the network is working properly. They can apply dc voltages in the range  $-165$  to  $+202$  volts, and ac voltages up to 45 volts rms, between any pair of conductors (ring, tip, ground) in the on-hook state, and up to 54 volts dc with source impedance as low as 10 ohms in the off-hook state. The phone company also specifies a range of lightning-induced “high-voltage surges” that you may find on your phone lines. They’re typically a few thousand volts, capable of sourcing a few hundred amps; equipment connected to the phone line is supposed to have transient suppressors so that it will survive such pulses. In addition, the phone company specifies “very high voltage surges” that may occur from a nearby lightning strike. These may reach 10kV and 1000 amps, and the idea is that even if your equipment gets fried, nobody should get hurt. So the specification says that the equipment shouldn’t shoot out pieces of stuff, catch fire, or electrocute anyone.

Permissible loads are specified in terms of “ringer equivalence number” (REN). Typical phones have an REN of 1.0A, which corresponds to (a) an on-hook dc resistance of  $50\text{M}\Omega$ , and on-hook ac impedance that stays above a specified curve of impedance versus frequency (satisfied by keeping  $|Z| > 125\text{k}$  from 4Hz to 3.2kHz, though it can be much lower over certain frequencies and voltages), (b) an off-hook dc characteristic that stays within the acceptable region of Figure 14.17 (or that measures  $200\Omega$  or less), and (c) an off-hook impedance that approximates 600 ohms from 200Hz to 3.2kHz (this is

actually specified in terms of reflection when driven by a  $600\Omega$  signal source: at least 3.5dB from 200Hz to 3.2kHz and 7dB from 500Hz to 2.5kHz). Loads connected to the phone lines must be dc-isolated from ground ( $50\text{M}\Omega$  on-hook, 250k off-hook). A total REN up to 5.0A is permissible, i.e., a load impedance as low as 1/5 the above values. The telephone company requires you to notify them of your total REN loading.

From the foregoing data it is obvious that the subscriber is not supposed to draw current in the on-hook state, and the  $50\text{M}\Omega$  minimum dc bridging resistance (for an REN of 1.0A) is really a leakage specification –  $50\text{M}\Omega$  corresponds to  $1\mu\text{A}$ . Nevertheless, with careful design and component selection you can maintain CMOS circuitry (digital or analog) in a quiescent state with a few microamps, and have it “wake up” when the line goes off-hook. Use a small tantalum electrolytic (or “double-layer” memory-retention capacitor, available in tiny packages up to 5 farads!) to keep things going during dialing or other transients. In the off-hook talking state, you are guaranteed 6 volts dc (7.8V after a few seconds) at a minimum of 26mA (see Fig. 14.17), which is enough

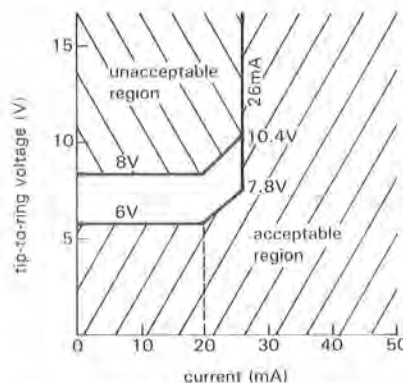


Figure 14.17. Allowable load conditions for “off-hook” telephone. The unshaded region is allowable beginning 1 second after going off-hook. (Adapted from Bell System Tech. Ref., Pub. 47001.)

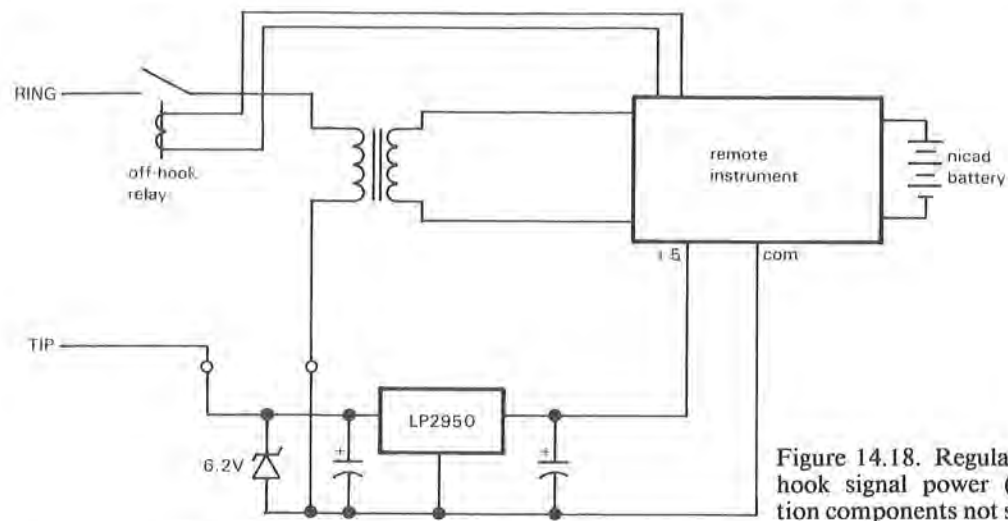


Figure 14.18. Regulated off-hook signal power (protection components not shown).

to operate lots of micropower circuitry; it's really comparable to the kind of power you get from a 9 volt battery. With a micropower low-dropout regulator like the LP2950 (80 $\mu$ A quiescent, 0.4V dropout at 100mA load) you can supply regulated 5 volts to digital circuitry, as in Figure 14.18. If you can be sure of having regular periods of off-hook operation, you can use the extra current then available to keep rechargeable batteries charged. For example, if you have an hour of off-hook operation per day, you can draw nearly a milliamp continuously.

**Warning:** Before designing any device for direct connection to the telephone system, be sure to get the relevant specifications. You must be in compliance with FCC regulations, which include test and approval procedures. Do not assume that the specifications quoted in this chapter are correct.

#### □ Relay circuits

Those popular "set-back" thermostats that turn down the heat at night, and turn it back up a half hour before you wake up, use the small ac power that can be extracted from the relay circuit without making the relay close. Typical mechanical

control relays draw 100mA or more at the coil rating of 24 volts ac and can be relied on to stay open at 10% of the normal coil current. So you can have 10mA or so, at nearly 24 volts ac, to power your instrument. Be sure to include rechargeable cells (or perhaps a memory-retention capacitor, if that's all that's needed), because the power source goes away when you close the contacts to energize the relay. Figure 14.10 shows the idea.

#### Industrial current loops

In industrial environments there is a standard for current-loop signaling, in which a remote sensor (a thermocouple, say; see Section 15.01) sends back its measurements by converting them to an analog current, which then flows through a loop. The dc bias for the loop is usually provided at the receiving end (Fig. 14.15C). There are two standards, namely full-scale ranges of 4 to 20mA and 10 to 50mA. The 4–20mA standard is more popular and usually uses a dc bias of 24 volts dc (though sometimes higher). For simplicity it is often desirable to use the signaling current to power the electronics at the remote end.

For this purpose you can use part of the loop bias for power. Commercially available current-loop modules generally specify that the data recipient has to provide a maximum load resistance  $R$  and minimum dc bias  $V_s$  such that  $(V_s - 12\text{V})/R_s$  is equal to the full-scale current. In other words, the remote module can drop up to 12 volts while still applying full-scale loop current. Of course, the module has to keep running when sending a loop current corresponding to minimum output. So the bottom line is that you always have available at least 12 volts at 4mA to power your equipment; you may have more, but don't count on it. That's plenty for even rather complex circuits, if you practice careful micropower design.

#### **RS-232 serial-port signals**

The RS-232C/D standard specifies bipolarity data and control signals of substantial drive capability (see Section 10.19); you can use one of the control signals (or even a data signal!) to run a low-power circuit. Officially an output must be able to assert  $\pm 5$  volt to  $\pm 15$  volt levels into a 3k to 7k load resistance. The RS-232 drivers in common use typically have an output impedance of a few hundred ohms, and current limit at 5 to 15mA. To become a parasite on this power source, you have to arrange your software to keep a known control line in a known (and stable) state. You can even use a *pair* of control lines, if available, to get split supply voltages ( $\pm 5\text{V}$ , min). Remember that control signals (RTS, DTR, etc.) are asserted HIGH, which is the opposite of the data signals.

Since there's usually plenty of commercial ac power available around a computer, you aren't really doing anything miraculous by sucking the life forces from the 25-pin *D* connector. However, for a simple serial-port hang-on circuit it is an elegant source of power. You can get commercial network interfaces and modems that work this way.

## **POWER SWITCHING AND MICROPPOWER REGULATORS**

### **14.06 Power switching**

You can tame your usual microprocessors, regulators, and other power-hungry components into a micropower application, if the design permits the circuit to be turned off (or put into a low-current standby state) most of the time, and only occasionally run at full current. For example, an oceanographic data logger might make a 10 second salvo of observations (temperature, pressure, salinity, ocean currents) once each hour for a 6 month period. Only the real-time clock need run continuously, with the analog signal conditioning circuitry, microprocessors, and data-recording media shut off except during actual data logging.

Even if you take pains to use micropower design techniques, you may still be forced to use some high-current devices, for example, if you need to use high-speed transducers or high-current actuators. You may need to use some specialized LSI digital circuits, op-amps, filters, or other circuits that are simply not available in low-power versions. In all these cases it is necessary to switch off power to the high-current portions of the circuit except when they must operate.

Such "power switching" can be the simplest form of micropower design, since ordinary design techniques with ordinary components can be used throughout. You've got to make sure the circuit "wakes up" gracefully (a linear circuit should be designed to avoid embarrassing momentary states, for example driving its outputs into saturation; a fully shut down microprocessor circuit would usually do a complete "cold boot"). Likewise, the circuit should be designed to shut down in an orderly manner.

There are several ways to do the power switching (Fig. 14.19):

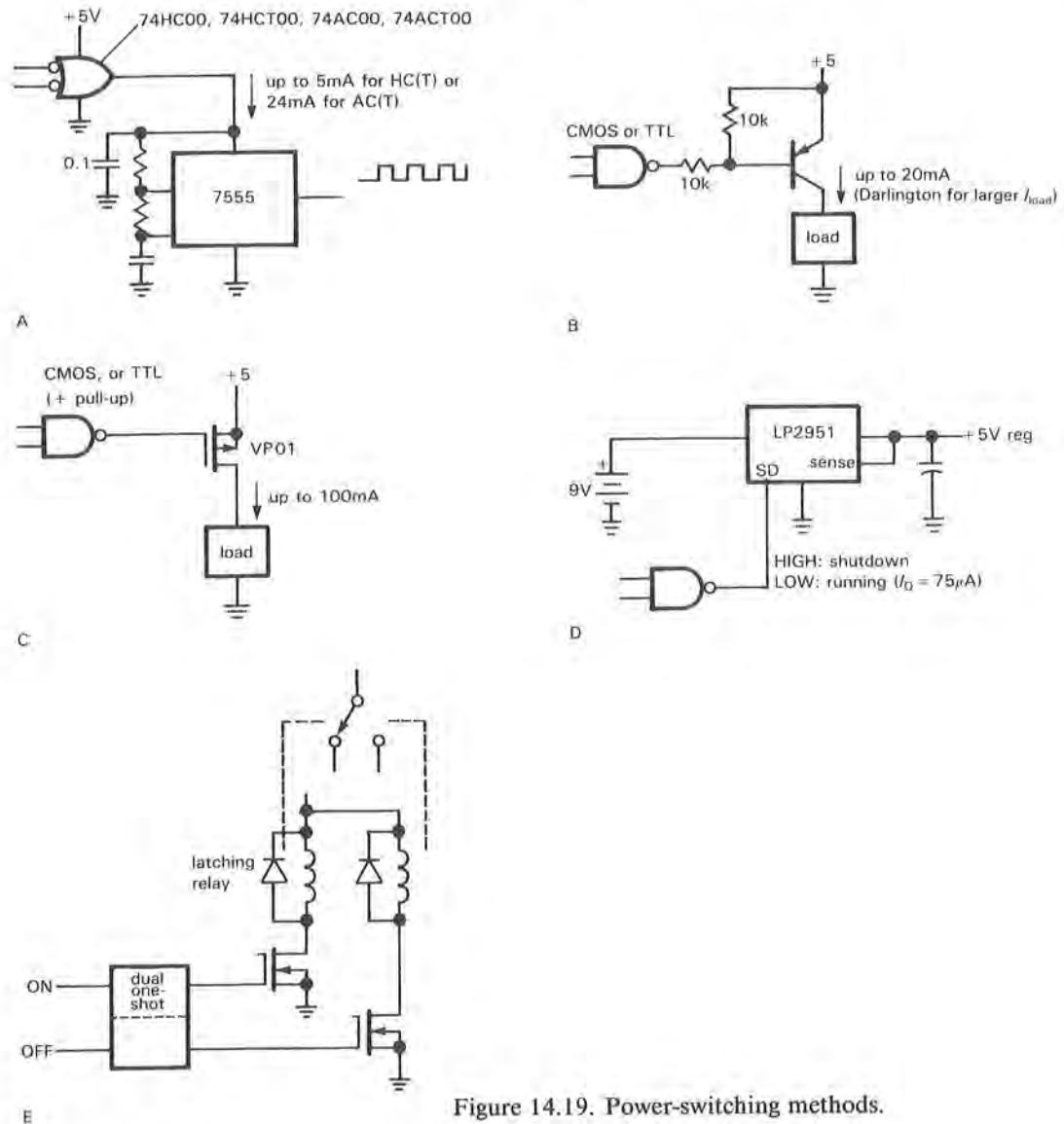


Figure 14.19. Power-switching methods.

1. If the switched components run at less than 5mA or so, you can power them directly from a CMOS logic output. The HC/HCT families can supply 5mA with only 0.5 volt drop below the positive rail; for higher current, several outputs can be used in parallel. The AC/ACT CMOS families are good for 24mA.

2. Use a power transistor, operating as a saturated switch (not a follower) to

minimize forward drop (thus *pnp* for a positive supply). The necessary base drive, chosen conservatively large to guarantee saturation, is a disadvantage, though it will probably be smaller than the current used by the switched circuit.

3. Use a power MOSFET. As with bipolar transistors, use as a switch, not a follower (thus *p*-channel for a positive supply). MOSFETs are easy to drive and have no gate current in either state.

4. Many of the low-power regulators include a “shutdown” input, with very low quiescent current in the standby mode (see Section 14.07). You can do power switching by commanding such a regulator into the active state.

5. Use a mechanical relay, perhaps a *latching* relay. There’s a good variety now available in DIPs and tiny metal cans, and they offer zero voltage drop, high overload capability, and the ability to switch bipolarity (or even ac) voltages. In addition, latching relays require no holding current. Be sure to use a diode to protect the relay driver from inductive spikes (Fig. 1.95).

### Current limiting

It is essential to limit the inrush current in a power-switched circuit, for two reasons: The high peak currents that would result from switching a battery (bypassed with a capacitor) into a load (similarly bypassed) could destroy the switch; this is true even for a small mechanical relay, which is most likely to fail by having its contacts fuse shut. Furthermore, the momentary collapse of the battery voltage during a high current-switching transient can cause volatile memory and other circuitry being held in a standby state to lose information (Fig. 14.20).

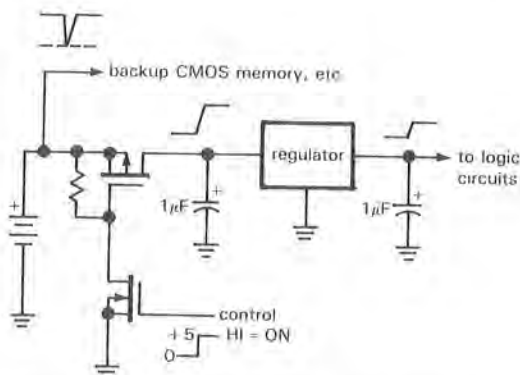


Figure 14.20. Inrush current can cause transient loss of battery voltage.

Several approaches are shown in Figure 14.21. As long as the switch can handle the transient, you can decouple the negative-going dip from the maintaining

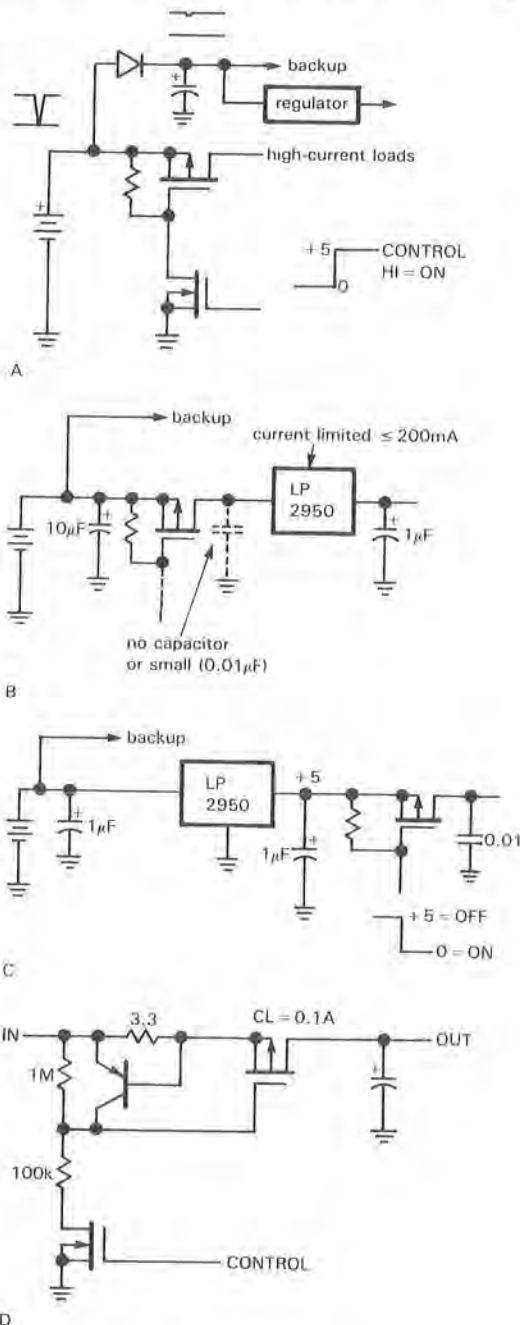


Figure 14.21. Four cures for inrush-current transients.



regulators with a diode, as in option 1. Alternatively, do the switching upstream of a current-limited regulator (keep its input bypass capacitor small), as in option 2, or put the switch after the regulator (option 3). The latter method isn't as good, because of degraded supply stiffness due to the switch's  $R_{ON}$ . Another method is to use upstream current-limited switching (option 4), in this case with 150mA current limit, to prevent collapse of  $V_{batt}$ .

#### 14.07 Micropower regulators

Until recently it had been difficult to find voltage regulator ICs with microamp quiescent currents capable of substantial output currents. The choice was (a) the Intersil 7663/4 or (b) build your own! Fortunately the situation is improving. Here is the current selection:

**ICL7663/4; MAX663/4/6** (Intersil; second-sourced by Maxim and others). These are multiterminal positive and negative regulators, with 1.5–16 volts operating range and maximum quiescent current of 10 $\mu$ A. The bad news is that they are slow (due to a "starved" servo amplifier; use lots of bypass capacitance) and only good for a few milliamps of load current (they're stiffer at higher input voltage, being CMOS devices); for example, with +9 volts input the output impedance is typically 70 ohms.

**LP2950/1** (National). These are positive regulators, available as a 3-terminal +5 volt regulator (2950) and an 8-terminal adjustable regulator (2951). Quiescent current is 80 $\mu$ A at zero load current, rising to 8mA at 100mA load current. These regulators use *pnp* pass transistors for low dropout voltage (80mV max at 100 $\mu$ A, 450mV max at 100mA) and are designed so that the quiescent current does not soar when the input voltage dips below dropout (a common disease of bipolar transconductance regulators). This last feature is particularly useful for battery-powered instruments that can continue to function with a low battery. The 2951

includes a shutdown input and dropout-detector output.

**LT1020** (Linear Technology). This is a multiterminal positive regulator with 40 $\mu$ A quiescent current, 2.5–35 volt output range, and 125mA maximum current. The *pnp* pass transistor gives low dropout voltage (20mV typ at 100 $\mu$ A, 500mV typ at 125mA). There is a shutdown input and dropout-detector output.

**TL580C** (Texas Instruments). This is a dual positive micropower *switching* regulator with 2.5–24 volt output range and 140 $\mu$ A quiescent current. Like all switching regulators, you get high efficiency (up to 80%) over a range of battery voltages and the flexibility to have output voltages greater than the unregulated input voltage.

**MAX630 series** (Maxim). These are micropower switching regulators, in a nice variety of options. The MAX630 is an adjustable (2V to 18V) positive step-up regulator (i.e.,  $V_{out} > V_{in}$ ), while the MAX634 is an inverting switcher (i.e., positive input, negative output). The MAX631–3 are fixed-voltage (5, 12, 15V) positive step-up switchers, with MAX635–7 being the inverting equivalents. The MAX638 is an adjustable positive step-down ( $V_{out} < V_{in}$ ) switcher. All are capable of output currents of a few hundred milliamps, with quiescent currents around 100 $\mu$ A and efficiencies around 80%.

**MAX644 series** (Maxim). These micropower switching regulators are designed to generate a +5 volt output when powered from a one- or two-cell battery. This clever design uses a two-part switching up-converter: One section runs continually, providing a low-current (<0.5mA) +5 volt output; it also supplies +12 volts dc needed to switch the MOSFET used for the high-current (up to 50mA) +5 volt output. In standby mode (low-current output only) the quiescent current is 80 $\mu$ A. The MAX644 is designed for 1.5 volt nominal input and operates down to input voltages of 0.9 volt.

TABLE 14.4. LOW-POWER REGULATORS

| Type     | Mfg <sup>a</sup> | Pins | I <sub>Q</sub><br>typ<br>( $\mu$ A) | Pol | V <sub>out</sub><br>(V) | V <sub>in</sub><br>(V) | Dropout          |                 | Shutdown | Tempco<br>typ<br>(ppm/°C) | Comments                             |
|----------|------------------|------|-------------------------------------|-----|-------------------------|------------------------|------------------|-----------------|----------|---------------------------|--------------------------------------|
|          |                  |      |                                     |     |                         |                        | V<br>(V)         | I<br>(mA)       |          |                           |                                      |
| ICL7663  | IL+              | 8    | 4                                   | +   | 1.5–16                  | 1.5–16                 | 0.8              | 20 <sup>b</sup> | •        | 200                       | MAX663, 7663S improved               |
| MAX664   | MA+              | 8    | 6                                   | –   | 1.3–16                  | 2–16.5                 | 0.2              | 20 <sup>b</sup> | •        | 100                       | also ICL7664                         |
| MAX666   | MA               | 8    | 6                                   | +   | 1.3–16                  | 2–16                   | 0.9              | 40              | •        | 100                       | MAX663 + dropout detec               |
| LT1020   | LT               | 14   | 40 <sup>c</sup>                     | +   | 0–35                    | 5–36                   | 0.5              | 125             | –        | 1%                        | dropout detector                     |
| LP2950   | NS               | 3    | 75 <sup>c</sup>                     | +   | 5                       | 5–30                   | 0.45             | 100             | –        | 20                        | dropout I <sub>Q</sub> = 110 $\mu$ A |
| LP2951   | NS               | 8    | 75 <sup>c</sup>                     | +   | 1.2–29                  | 2–30                   | 0.45             | 100             | •        | 20                        |                                      |
| MAX630   | MA               | 8    | 70                                  | +   | V <sub>in</sub> to 18   | 2–16.5                 | –                | 375             | •        | –                         | switching stepup                     |
| MAX635-7 | MA               | 8    | 80                                  | ±   | –5, –12, –15            | +2 to 16.5             | –                | 375             | –        | –                         | switching inverter                   |
| MAX634   | MA               | 8    | 100                                 | ±   | to –20                  | 2–16.5                 | –                | 375             | •        | –                         | switching inverter                   |
| MAX631-3 | MA               | 8    | 135                                 | +   | 5, 12, 15               | 1.5–V <sub>out</sub>   | –                | 325             | –        | –                         | switching stepup                     |
| MAX638   | MA               | 8    | 135                                 | +   | < V <sub>in</sub>       | 2–16.5                 | –                | 375             | •        | –                         | switching stepdown                   |
| TL580C   | TI               | 8    | 140                                 | +   | 2.5–24                  | 2.4–30                 | –                | 100             | –        | –                         |                                      |
| LM10     | NS+              | 8    | 300                                 | +   | 1–40                    | 1.1–40                 | 0.4              | 20              | –        | 30                        |                                      |
| LM2931   | NS+              | 5    | 400                                 | +   | 1.2–25                  | to 26V                 | 0.2              | 150             | •        | –                         | TO-220                               |
| LM2931-5 | NS+              | 3    | 400                                 | +   | 5                       | 5.2–26                 | 0.2              | 150             | –        | –                         | TO-92; 2931CT is adjustable          |
| TL750L05 | TI               | 3    | 1000                                | +   | 5                       | 5.2–26                 | 0.6 <sup>d</sup> | 150             | –        | –                         | TO-92; TL751 has shutdown            |
| LM317L   | NS+              | 3    | 2500 <sup>e</sup>                   | +   | 1.2–37                  | to 40V                 | 2                | 100             | –        | 0.7%                      | TO-92                                |
| LM337L   | NS+              | 3    | 2200 <sup>e</sup>                   | –   | 1.2–37                  | to –40V                | 2                | 100             | –        | 0.7%                      | TO-92                                |
| 78Lxx    | FS+              | 3    | 3000                                | +   | 5, 12, 15               | to 30V                 | 2                | 100             | –        | –                         | TO-92                                |
| 79Lxx    | FS+              | 3    | 2000                                | –   | –5, –12, –15            | to –35V                | 2                | 100             | –        | –                         | TO-92; LM320L also                   |
| LM330    | NS+              | 3    | 3000 <sup>c</sup>                   | +   | 5                       | 5.3–26                 | 0.6              | 150             | –        | –                         | TO-220                               |
| LM2930   | NS+              | 3    | 4000                                | +   | 5                       | 5.3–26                 | 0.6              | 150             | –        | –                         | TO-220; LM2935 also                  |

(a) see footnote to Table 4.1. (b) for V<sub>in</sub> = 9V. (c) no load. (d) over full temp range. (e) I<sub>L</sub> (min).

In addition, there are several “low-power” regulators (78L05, LM330, LM317L, LM2930/1), characterized by quiescent currents of a few milliamps. These are useful for instruments with some external power source, for example solar cells or telephone holding current. Also, don’t overlook the possibility of using a micropower voltage *reference*, rather than a regulator, if its voltage happens to be what you want. For example, the REF-43 from PMI is a 3-terminal 2.5 volt reference with 250 $\mu$ A maximum quiescent and excellent characteristics.

Look at Table 14.4 (which also includes the regulators above) for characteristics of most available micropower regulators.

### Negative supplies

With the exception of the ICL7664/MAX664, all of the linear micropower regulators are positive polarity only (though the LT1020 can be used to make a dual supply). If you need negative supply voltages, there are (besides the feeble 7664) several possibilities, namely (a) a “flying-capacitor” voltage converter chip like the 7662, (b) a discrete realization of a flying-capacitor voltage converter, using complementary power MOS transistors, (c) a voltage converter using a CMOS oscillator chip like the 7555 (that’s a CMOS 555) or the output of any CMOS logic gate that is driven by a square wave, (d) a switching supply, with inductive energy storage, or (e) the use of a single positive supply, with an op-amp-generated ground reference part way between ground and the positive rail. Let’s take them in turn:

1. The 7662 (and its predecessor 7660) is a CMOS IC introduced by Intersil and widely second-sourced (see Section 6.22). It has an oscillator and CMOS switches (Fig. 6.58), and with a few external capacitors you can use it to generate either  $-V_{\text{supply}}$  or  $+2V_{\text{supply}}$ , when powered by a positive voltage  $V_{\text{supply}}$ . Like most CMOS devices, it has a limited supply voltage

range; for the 7662,  $V_{\text{supply}}$  can range only from 4.5 to 20 volts (1.5V to 10V for the 7660). The output is not regulated, and it drops significantly for load currents greater than a few milliamps. In spite of these drawbacks, it can be very useful in special circumstances, for example to power an RS-232C driver chip on a board that otherwise runs on a single +5 volt supply. The MAX680 and LT1026 are flying-capacitor dual supplies that generate  $\pm 10$  volts (up to 10mA) from +5 volts (Fig. 6.60). There are also combination voltage converter and RS-232 driver/receivers available as single ICs, the LT1080 and MAX230–239 series. If your application requires RS-232 ports, you may be able to use the dual supply voltages generated by one of these RS-232 driver ICs to power your analog electronics.

2. To generate a larger negative voltage, you can use discrete MOS transistors in a flying-capacitor circuit (Fig. 14.22). The particular example shown idles at a few microamps and generates up to 30mA.

3. Figure 14.23 shows a simpler method, again somewhat limited in voltage range, using the CMOS 7555 timer chip. You can power the 7555 from a positive supply in the range 2 to 18 volts, thus generating up to  $-15$  volts or so. With a voltage multiplier (see Section 1.28) you can, of course, generate higher voltages, with correspondingly poorer regulation. If you have some CMOS logic in your circuit, you can use the output of a CMOS gate instead of the 7555. However, if you’re using a high-performance CMOS family such as HC/HCT or AC/ACT, then you are limited to 5 volt logic swings, whereas the older 4000 or 74C series permit 15 volt swings, albeit at lower current.

4. As we explained in Chapter 5, with inductive energy storage you can make switching supplies for which the output voltage is higher than the input, or much lower, or even negative, all with efficiencies of 75% or so, independent of input voltage. This is obviously useful in

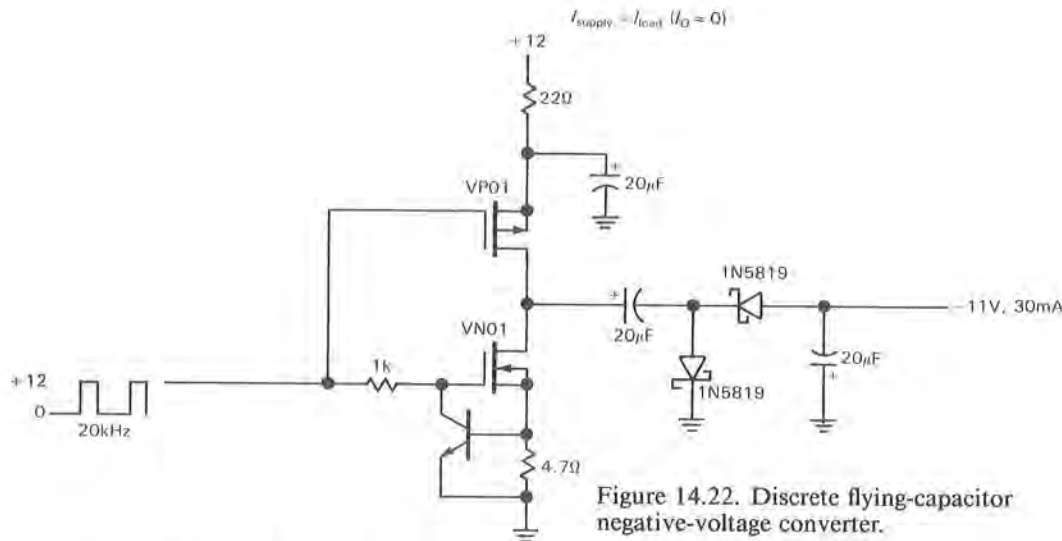


Figure 14.22. Discrete flying-capacitor negative-voltage converter.

micropower design, where the unregulated dc may be supplied by batteries whose voltage drops off with use. Switching supplies for micropower applications can be designed to maintain high efficiency even when unloaded (unlike ordinary high-current switchers), by using circuitry that shuts the oscillator off until the output drops, at which point it supplies a single charging pulse, then goes to sleep again. Figure 14.24 shows a +5 volt supply constructed with the low-power MAX631.

5. You may not need a separate negative supply, even if you are using op-amps with bipolarity output swings, etc. For example, you might generate a +4.5 volt ground reference (using a resistive divider and micropower op-amp follower) for an op-amp circuit running from a single 9 volt battery. Let's look at this method in some more detail.

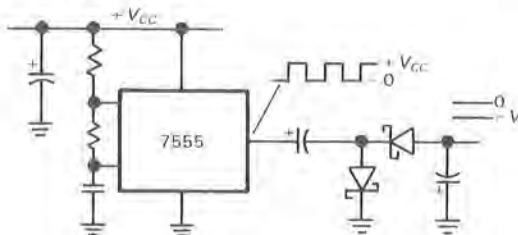


Figure 14.23. Negative-voltage generator from a positive square wave.

#### 14.08 Ground reference

Most of our op-amp circuits in Chapter 3 used symmetrical power supplies, usually  $\pm 15$  volts, because of the flexibility of dealing with signals near ground. As we mentioned in Section 4.22, however, it is possible to use only a single supply, by generating a reference voltage that substitutes for the ground potential of the usual bipolarity op-amp power supplies. When your power supply is a battery, there's an added incentive to keep things simple, preferably by using a single 9 volt battery.

The easiest way to generate an analog "common" is to split the battery voltage with a resistive voltage divider, then use a micropower op-amp follower to generate the low-impedance common. To the outside world that common voltage is "ground," with both ends of the battery floating; see Figure 14.8.

In the example circuit, we've chosen a 3440 CMOS programmable op-amp, biased to run at  $5\mu\text{A}$  quiescent current. The divider's unusually large resistors keep its contribution to the current drain small, with capacitive bypassing to keep the impedance low at the midpoint, which otherwise would be susceptible to hum and pickup of other signal

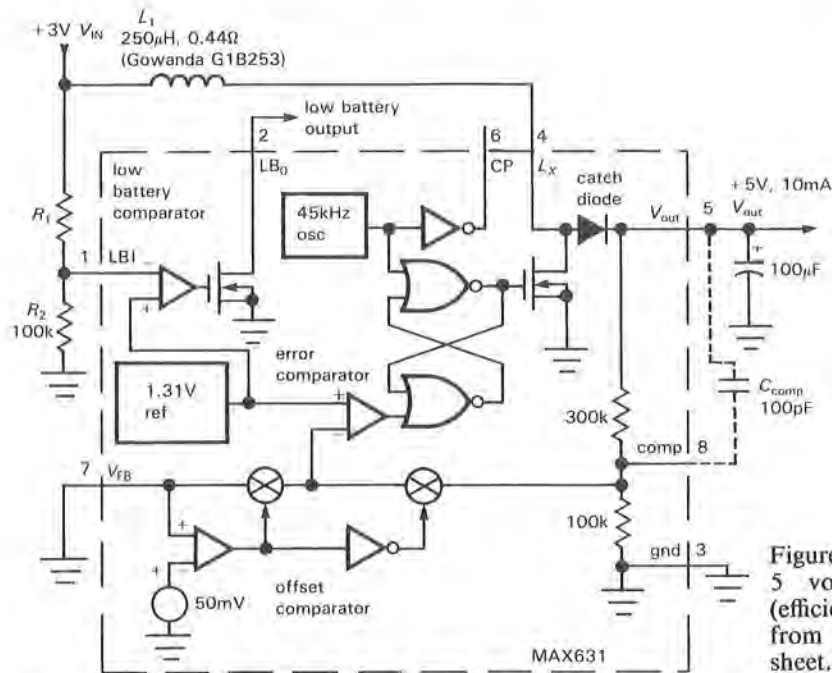


Figure 14.24. Low-power 5 volt switching regulator (efficiency = 74%). (Adapted from Maxim MAX631 data sheet.)

frequency interference. The 3440 is a good choice in this application, because it can sink or source substantial currents (up to a few milliamps) even when biased at 1μA; this property is not shared by all programmable op-amps, many of which have poor sourcing capability when operated at micropower levels. For example, the LM346 operating at 5μA can source only about 0.1mA, although it can sink 20mA (look ahead to Fig. 14.32).

Note that the reference voltage doesn't have to be half of the battery voltage; it may be best to split the supply unsymmetrically, to allow maximum signal swing. (We'll have an example in Section 14.12.) In some instances it may be preferable to put it at a fixed voltage from one rail, perhaps determined by a precision micropower voltage reference. That rail is then a regulated supply with respect to the common reference.

#### Output impedance

There are some situations in which you don't even need to use an op-amp to

generate the ground reference. For example, if the reference voltage goes only to op-amp inputs (which would have been connected to ground in the usual split-supply configuration), then a high-impedance resistive divider, bypassed to maintain a low impedance at signal frequencies, will usually suffice.

In the more usual case, however, the ground reference must present a low impedance, both at dc and at signal frequencies. For example, some ICs may use it as their negative rail; it might be the common point for low-pass filters, biasing networks, loads, etc. Look at almost any normal split-supply circuit and you'll find dc and signal currents flowing into and out of ground. As in the example above, be sure the op-amp you choose to generate the ground reference has the source and sink capability the circuit needs. Micropower op-amps tend to have rather high open-loop output impedances (Fig. 7.16), so at high frequencies (where there isn't much loop gain) the impedance of the ground reference may rise to several thousand ohms.



The obvious cure is to bypass the ground reference (Fig. 14.25A), but this is likely to cause ringing or even oscillation because of the lagging phase shift of the bypass capacitor in combination with the op-amp's relatively large output impedance, all of which is inside the feedback loop. Figure 14.25B shows one cure – a decoupling resistor of a few hundred ohms, which, however, raises the impedance at dc since it is outside the feedback loop. With two more parts, Figure 14.25C does the trick, maintaining dc feedback (via  $R_2$ ) and stability at the same time.

Whatever method you choose, make

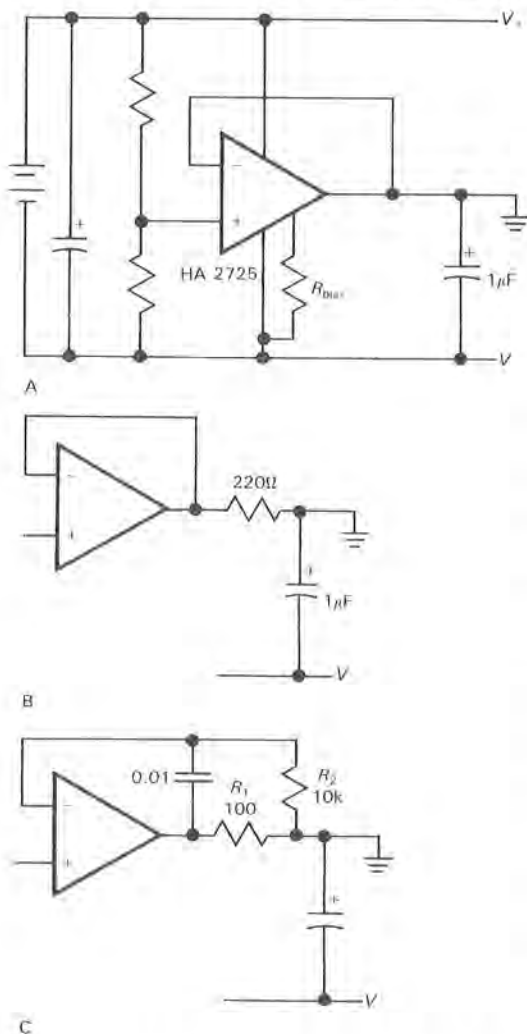


Figure 14.25. Bypassed split-supply generators.

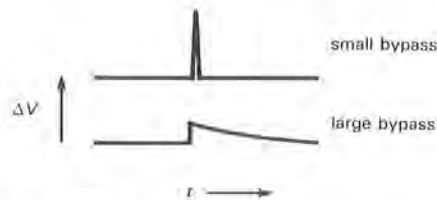


Figure 14.26

sure you test it under various steady-state and transient load conditions. A good way to test for transient behavior is to watch the voltage waveforms while applying a low-frequency square-wave load. There are some op-amps (e.g., the HA2725 and MC3476) that can drive directly into a large capacitive load without stability problems; apparently in these cases the external capacitor reflects back onto the internal compensation capacitor, pushing down the dominant pole in a “brute-force” compensation. In many cases, however, you’re more likely to wind up with a pair of nearby lagging phase shifts, which spells trouble.

Note that the choice of bypass capacitor value may involve some subtlety: For a load-induced spike of fixed charge injection into the ground reference node (i.e., a fixed amp-second product), a larger bypass capacitor will keep transient ground noise smaller, but will have a longer recovery time than a small capacitor (Fig. 14.26). For a high-gain low-speed circuit that may be worse, perhaps producing slow exponential recovery instead of harmless little spikes at the output.

When designing ground reference circuits, don’t overlook the reference voltage outputs that are sometimes provided on other ICs. For example, the LM322 timer provides a stable 3.15 volt output. Other chips that have external access to internal voltage references are A/D converters, V/F converters (e.g., the 331, with its 1.89V reference), and chips like the LM10, which has a 200mV reference and amplifier, in addition to an uncommitted op-amp. Figure 14.27 shows several buffered reference schemes.

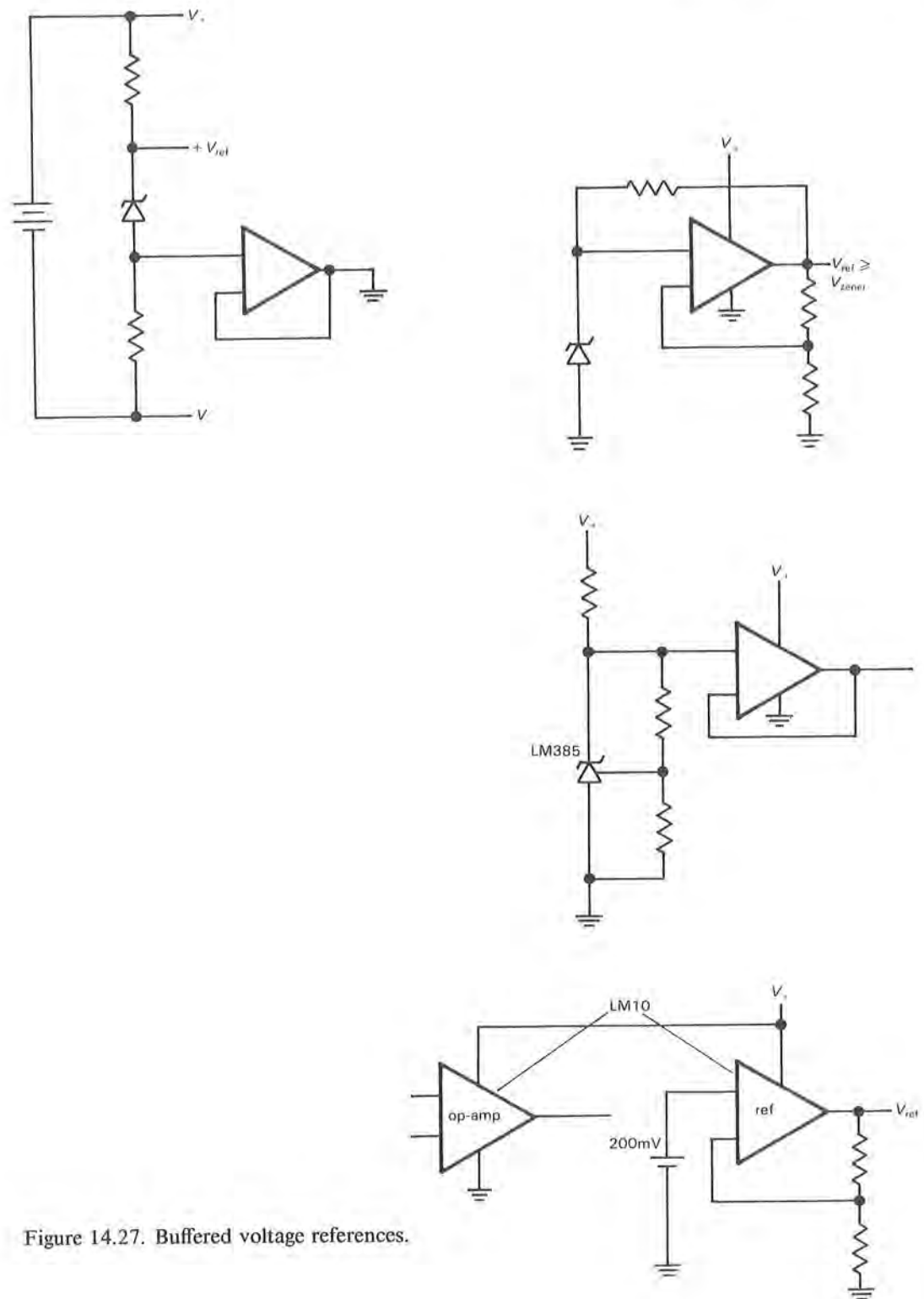


Figure 14.27. Buffered voltage references.

**14.09 Micropower voltage references and temperature sensors**

Most zener and bandgap references are relatively power-hungry and are not suitable for micropower circuits. As Table 6.7 demonstrates, most 3-terminal references run at about a milliamp, and most 2-terminal zener-like references are specified at similar operating currents.

Fortunately, there are several voltage references intended for micropower applications. The LM385 series includes a programmable 2-terminal bandgap reference (LM385, 1.24V–5.30V) and two fixed voltage references (LM385-1.2, 1.235V and LM385-2.5, 2.50V). The fixed-voltage models are specified to operate at currents down to 10 $\mu$ A, with dynamic impedances of 1 ohm at 40 $\mu$ A and 100 $\mu$ A, respectively. The minimum current of the programmable version goes from 10 $\mu$ A to 50 $\mu$ A, depending on voltage. All versions are available with tempcos down to 30ppm/ $^{\circ}$ C. The ICL7663/4 regulators (Section 14.07) can be used as 3-terminal references, with typical quiescent current of 4 $\mu$ A and dynamic output impedance of 2 ohms. The ICL8069 is a 2-terminal bandgap reference that operates down to 50 $\mu$ A (where the dynamic impedance is 1 $\Omega$ ), with tempco down to 50ppm/ $^{\circ}$ C. The AD589 has the same characteristics, with improved tempco (down to 10ppm/ $^{\circ}$ C). The LT1004 from Linear Technology is like the LM385-1.2, while their LT1034 is a dual 2-terminal reference (1.2V and 7.0V) with minimum operating currents of 20 $\mu$ A and tempco of 20ppm/ $^{\circ}$ C for its 1.2 volt reference; the 7 volt reference should be operated at 100 $\mu$ A (min), but is quieter than the bandgap references.

For better tempco at not-quite-micropower currents, there's the LM368 3-terminal reference, available in 5, 6.2, and 10 volt versions (0.05% accuracy). It draws 300 $\mu$ A, has low output impedance over frequency, and is available with tempcos

down to 10ppm/ $^{\circ}$ C. Even better is the REF-43, a 3-terminal 2.5 volt positive reference with 0.05% accuracy and 3ppm/ $^{\circ}$ C tempco (max). It has low  $Z_{out}$  (0.1 $\Omega$ ), excellent regulation (2ppm/ $V_{in}$ , max), output current to 10mA, and quiescent current of 250 $\mu$ A, max.

Table 14.5 lists currently available micropower references.

Finally, there are micropower ICs that convert temperature to current or to voltage. The AD590 and AD592 are 2-terminal current sources that run on 4 volts to 30 volts, and give a current of 1 $\mu$ A/ $^{\circ}$ K (e.g., 298.2 $\mu$ A at 0 $^{\circ}$ C). The LM334 is similar, but with a programming pin to set the conversion factor; the operating range is 1 $\mu$ A to 10mA. The LM34 (Fahrenheit) and LM35 (Centigrade) are 3-terminal temperature sensors with voltage output (thus 0V at 0 $^{\circ}$ F or 0 $^{\circ}$ C, and 10mV/ $^{\circ}$ F or  $^{\circ}$ C, respectively) and quiescent current of 100 $\mu$ A. The LM335 is a 2-terminal IC zener with a breakdown voltage of 10mV/K (e.g., 2.982V at 0 $^{\circ}$ C), operable down to 400 $\mu$ A. See Section 15.1 for additional information.

**LINEAR MICROPOWER DESIGN TECHNIQUES**

Thus far we have treated power sources, power-switching techniques, regulators, and references for the design of micropower instruments. Now, following the progression of topics in the rest of the book, we turn to the design of linear and digital circuits themselves. We will begin with a discrete linear circuit example (a high-gain micropower audio amplifier), then proceed to micropower op-amp design techniques. That will be followed by sections on digital and microprocessor design, and finally some comments on packaging techniques for low-power instruments.

TABLE 14.5. MICROPOWER VOLTAGE REFERENCES

| Type       | Mfg <sup>a</sup> | Term | B <sup>+</sup> gap/Zener | Trim | Voltage (V) | Acc'y (%) | Tempco max (ppm/°C) | Supply voltage min (V) | Supply curr <sup>b</sup> (μA) | Output curr max (mA) | Noise typ 10Hz-10kHz (μV,rms) | Drift typ (ppm/1kh) | Z <sub>out</sub> typ (Ω) | I <sub>ref</sub> @ (μA) | Comments                   |
|------------|------------------|------|--------------------------|------|-------------|-----------|---------------------|------------------------|-------------------------------|----------------------|-------------------------------|---------------------|--------------------------|-------------------------|----------------------------|
| TSC04A     | TS               | 2    | B                        | 2    | 1.26        | 2         | 50                  | -                      | 15                            | -                    | -                             | -                   | 0.3                      | 1000                    |                            |
| TSC05A     | TS               | 2    | B                        | 2    | 2.5         | 3         | 50                  | -                      | 20                            | -                    | -                             | -                   | 0.3                      | 1000                    |                            |
| REF25      | FE               | 2    | B                        | 2    | 2.5         | -         | 55 <sup>f</sup>     | -                      | 60                            | -                    | -                             | -                   | 1.5                      | -                       |                            |
| REF43F     | PM               | 3    | B                        | 8    | 2.5         | 0.05      | 10                  | 4.5                    | 350                           | 10                   | 5 <sup>f</sup>                | -                   | 0.03                     | -                       | 2ppm/V <sub>in</sub> , max |
| REF50      | FE               | 2    | B                        | 2    | 5.0         | -         | 55 <sup>f</sup>     | -                      | 60                            | -                    | -                             | -                   | 1.5                      | -                       |                            |
| LM385      | NS               | 2    | B                        | 3    | 1.2-5.3     | -         | 30                  | -                      | 9                             | -                    | 50 <sup>c</sup>               | 20                  | 0.4                      | 100 <sup>e</sup>        |                            |
| LM385-1.2  | NS               | 2    | B                        | 2    | 1.2         | 1         | 30                  | -                      | 10                            | -                    | 60                            | 20                  | 1                        | 40                      | LT1004-1.2                 |
| LM385-2.5  | NS               | 2    | B                        | 2    | 2.5         | 1         | 30                  | -                      | 20                            | -                    | 120                           | 20                  | 1                        | 100                     | LT1004-2.5                 |
| AD589M     | AD               | 2    | B                        | 2    | 1.24        | 2         | 10                  | -                      | 50                            | -                    | 5                             | -                   | 0.6                      | 500                     | MP5010                     |
| LT1034     | LT               | 2    | B                        | 2    | 1.22        | 1         | 20                  | -                      | 20                            | -                    | 4 <sup>d</sup>                | 20                  | 0.3                      | 100                     | dual ref                   |
| "          | LT               | 2    | Z                        | 2    | 7           | 3         | 40 <sup>f</sup>     | -                      | 100                           | -                    | -                             | 20                  | -                        | -                       |                            |
| LP2950ACZ  | NS               | 3    | B                        | 3    | 5.0         | 0.5       | 100                 | 5.4                    | 120                           | 100                  | 430 <sup>e</sup>              | -                   | 0.02                     | 100                     | regulator                  |
| ICL7663/4S | IL+              | 3    | B                        | 8    | 1.5-10      | -         | 100 <sup>f</sup>    | 1.5                    | 10                            | 40                   | -                             | -                   | 2                        | -                       | MAX663/4                   |
| ICL8069    | IL               | 2    | B                        | 2    | 1.23        | 2         | 50                  | -                      | 50                            | -                    | 5                             | 1                   | 1                        | 50                      |                            |
| TSC9491A   | TS               | 2    | B                        | 2    | 1.22        | 2         | 50                  | -                      | 50                            | -                    | -                             | -                   | 2                        | -                       |                            |

(a) see footnote to Table 4.1. (b) min oper curr (for 2-term refs); max quiescent curr (for 3-term refs). (c) at V<sub>ref</sub>. (d) 0.1-10Hz, pp.

(e) 10Hz-100kHz, 1μF bypass. (f) 0.1Hz-1kHz, max. (g) typical.

### 14.10 Problems of micropower linear design

In general, low-power linear design means low collector (or drain) currents and correspondingly large collector (drain) resistors. As a result, capacitance effects dominate, in the form of both Miller effect and normal  $RC$  rolloffs. You often resort to techniques that are ordinarily used only in radiofrequency design, for example the use of the cascode (Section 2.19), emitter followers, and the shunt-series pair (Section 13.07). RF transistors (with  $f_T$  of 1GHz or more) may be good choices even at low frequencies because of their extremely low feedback capacitance  $C_{ob}$ ; for example, the MRF931 has  $C_{ob} = 0.35\text{pF}$  at  $V_{CE} = 1$  volt and is specified for use down to 1 volt and 0.1mA ( $f_T = 3\text{GHz}$  at  $I_C = 1\text{mA}$  and  $V_{CE} = 1\text{V}$ ). In spite of measures like these, it still pays to choose low operating frequencies whenever possible, for example the clock rate for microprocessor or other CMOS digital systems.

Other undesirable effects of low-power operation are increased noise pickup (because of the relatively high signal source impedances), reduced drive capability (low operating currents, high impedances), and relatively high transistor noise voltage  $e_n$

(from Johnson noise in the relatively high  $r_e$ ; see Section 7.13). This latter problem also afflicts micropower voltage references; be sure to check the noise specifications. Even with emitter followers, output impedances can be unreasonably large ( $r_e = 25\text{k}\Omega$  at  $I_C = 1\mu\text{A}$ ).

In general, operation at low voltage is desirable, because of the correspondingly reduced collector resistors for the same operating current. In addition, at the same collector current the *power* is reduced proportional to supply voltage.

### 14.11 Discrete linear design example

Let us imagine we need a low-noise audio amplifier with lots of gain (at least 80dB) and low quiescent drain, to operate in some remote battery-powered application. Since the signal levels may vary over a wide range, it would be nice to include some provision for switching the gain over a range of, say, 60dB. To maintain shelf life from an alkaline 9 volt battery (500mAh) we should draw no more than  $20\mu\text{A}$  total (corresponding to a 3 year life); since other circuits probably run from the same battery, we will budget  $10\mu\text{A}$  for the amplifier.

The first thing to note is that a micropower op-amp cannot provide the performance we need. The exemplary "nanopower" CA3440, running at  $10\mu\text{A}$ , has a dc gain of 80dB (min) and a gain-bandwidth product of 300kHz; i.e., at 20kHz its gain is only 15 (24dB). We will discuss micropower op-amp design, and its limitations, in the next section. For now, all we need to know is that the objectives of op-amp design (dc coupling, accuracy, unity-gain compensation) are quite different from what we need for this example, and that we can do better with discrete design.

Let's begin by trying a stage of "series-feedback pair" shown in Section 4.27. Figure 14.28 shows our first try, where we have run a pair of superbeta low-noise

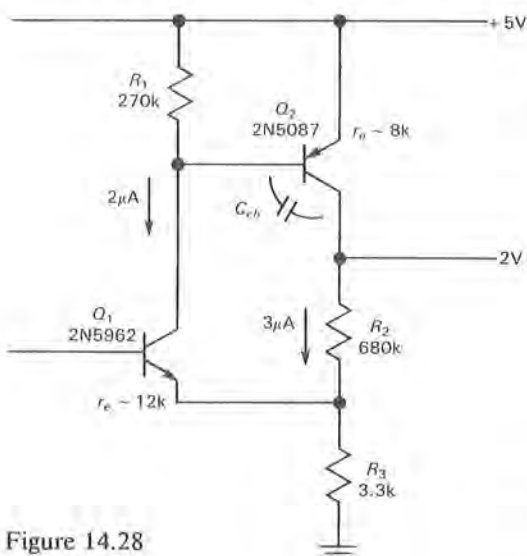


Figure 14.28



transistors at  $5\mu\text{A}$  combined collector currents, with a design gain ( $R_2/R_3$ ) of 200 (46dB). The bias arrangement is not shown.  $Q_1$ 's quiescent current is set by a  $V_{BE}$  drop across  $R_1$ , and  $R_2$  sets  $Q_2$ 's collector voltage, once its quiescent current is chosen. The intrinsic emitter resistances  $r_e$  are quite high, 12k and 8k, respectively, owing to the low collector currents.

Two such stages, with their 90dB gain and  $10\mu\text{A}$  quiescent current, would appear to do the job, perhaps with an emitter follower at the output. As we mentioned earlier, however, the effects of capacitance can be devastating in low-current high-impedance circuits. Let's see what capacitance does to this circuit's performance.

To evaluate Miller effect we need to know how the voltage gain is distributed between the two transistors.  $Q_2$  has  $r_e = 8\text{k}$ , so its voltage gain is about 85; the first stage, with feedback operating at the emitter, makes up the rest, with a gain of about 2.4. The high second-stage gain suggests that Miller effect there may dominate the amplifier's overall rolloff. In fact, that is the case: The 2N5087 has  $C_{cb} = 6.5\text{pF}$  at  $V_{cb} = 2$  volts, which acts like 550pF input capacitance to ground at the base. The reactance of this Miller capacitance equals  $R_1$  at 1kHz, which would cause 3dB of rolloff at that frequency in the absence of overall negative feedback. With feedback, the rolloff is higher, roughly 4.4kHz, but still far too low for an audio amplifier, which ideally should be flat to 20kHz.

The solution in this case is to notice that  $C_{cb}$  is the killer (via Miller) and use instead a transistor with very low  $C_{cb}$ . The *pnp* 2N4957 is a suitable choice, a low-noise high-gain UHF amplifier with  $C_{cb} = 0.6\text{pF}$  at 2 volts. Figure 14.29 shows the finished circuit, with a cascade of two series-feedback pairs and an output emitter follower. The second pair has a larger emitter resistor for proper biasing with dc coupling from the first stage. The gain switching is done with a 4066B

CMOS switch array, which has low noise and good inter-switch isolation. Power-supply decoupling is a good idea in a high-gain amplifier circuit like this. With all switches closed, this amplifier has 90dB of gain (switchable down to 30dB with various combinations of switch closures), 27kHz of bandwidth, and input noise voltage of  $12\text{nV}/\sqrt{\text{Hz}}$ ; with 50k source impedance its noise figure is 1.1dB. It's worth noting by comparison that the CMOS CA3440 that we considered initially has 20dB more noise ( $e_n = 110\text{nV}/\sqrt{\text{Hz}}$ ); even the excellent bipolar OP-90 from PMI has  $e_n = 40\text{nV}/\sqrt{\text{Hz}}$ , and that with  $20\mu\text{A}$  of supply current and a gain-bandwidth product of 25kHz (allowing a meager gain of only 2dB at 20kHz). For this application, discrete design techniques are mandatory.

It may seem paradoxical to be using UHF transistors for an audio application, but our example demonstrates that it makes good sense. There are some real gems in the RF data books. For example, the MRF9331 has  $C_{cb} = 0.25\text{pF}$  at 0 volts (!), and an  $f_T$  of 5GHz at 1mA. These devices are specified at low voltages and currents and are meant for battery-operated communications. For example, in actual measurements the MRF9331, operating at  $V_{CE} = 1.5$  volts, had  $h_{FE} = 30$  at 10nA, and 60 at  $1\mu\text{A}$ .

## 14.12 Micropower operational amplifiers

As with ordinary linear design techniques, it is tempting to abandon discrete low-power transistor design in favor of micropower op-amps, providing they can deliver the performance you need. Steady progress in bipolar linear ICs, along with recent improvements in CMOS IC fabrication, has led to a reasonable selection of micropower op-amps. There are some serious trade-offs and design constraints involved in the use of micropower op-amps, however. Let's look at some of these problems.

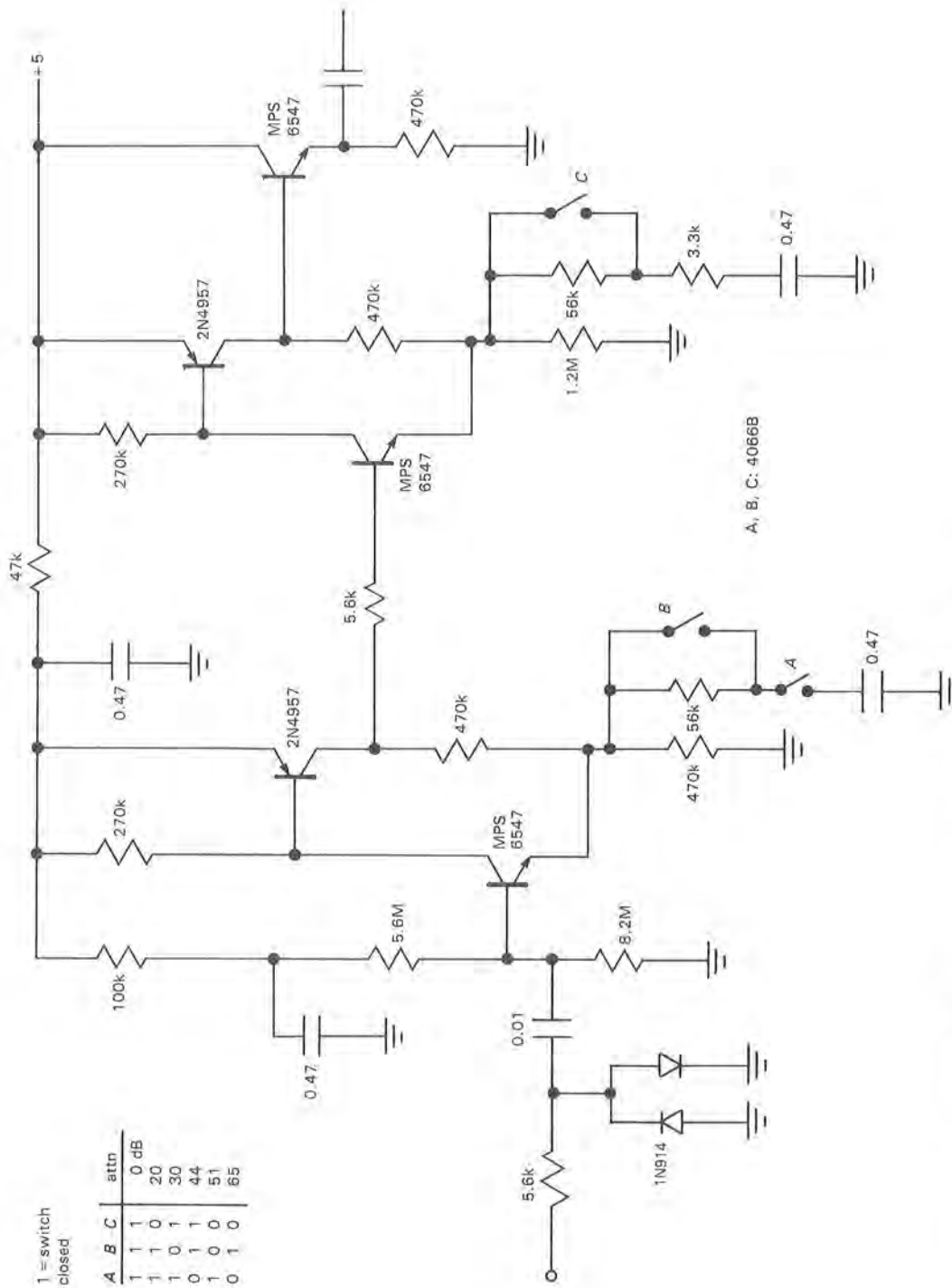


Figure 14.29. Micropower high-gain audio amplifier. The FET switches permit attenuation from maximum gain, as shown.

All other things being equal, reducing the operating current of an op-amp causes corresponding reductions in the unity-gain frequency  $f_T$  and the slew rate, and increases in the output impedance  $Z_{out}$ , crossover distortion, and input noise voltage  $e_n$ . In most cases it also tends to reduce the available output drive current  $I_{out}$ . In addition to these undesirable characteristics, the circuit compromises needed to design micropower op-amps can lead to additional pathology, for example the possibility of low-frequency oscillations (motorboating), latchup, or insufficient range of input offset trimming circuitry.

In many cases micropower design means battery-powered design, often with a single (and perhaps unregulated) dc supply voltage. An op-amp operating from a single low-voltage battery will have only limited output swing; in order to maintain good dynamic range and accuracy with these small signals, it is necessary to have smaller  $V_{os}$  compared with a conventional circuit using  $\pm 15$  volt supplies. For example, an op-amp that can swing only to within 2 volts of the power-supply rails is limited to 3 volts pp maximum swing when operating from a single 9 volt battery (which has dropped to 7V at the end of its life), compared with 26 volts pp when operating from  $\pm 15$  volt supplies. In this example you would need to keep op-amp input offsets nearly 10 times smaller in your battery-powered design to maintain equivalent accuracy.

#### **Programmable op-amps and micropower op-amps**

There are two categories of operational amplifiers that are suitable for micropower design: *programmable* op-amps, and *low-power* op-amps.

*Programmable* op-amps (Section 4.13) are like ordinary op-amps, but with an additional "programming" pin that lets you set the operating currents within the

IC. Programmable op-amps typically use current mirrors in various ratios to set the operating currents of their internal stages, so that the total quiescent operating current is some multiple of the programming current,  $I_{set}$ . The usual way to set  $I_Q$  is to hook a resistor from the programming pin to one of the supplies (usually  $V_-$ ), since the programming pin usually drives a current mirror directly. Quiescent currents down to a microamp or less are possible, though with correspondingly degraded dynamic performance (e.g., at  $I_Q = 1\mu A$ , a 4250 programmable op-amp has  $SR = 0.005V/\mu s$  and  $f_T = 0.01MHz$ ). Some popular programmable op-amps are the LM346 and the 4250 (both bipolar) and the CMOS CA3440.

*Low-power* op-amps are simply op-amps internally designed to operate at low quiescent current, without a programming pin. Examples are the precision OP-20 ( $45\mu A$ ) and OP-90 ( $12\mu A$  max) and TI's "LinCMOS" TLC27L2 ( $20\mu A$ ). A variation on this theme provides *pin-selectable* operating current, according to whether the programming pin is tied to  $V_+$  or  $V_-$  or is left open. Both the TLC271 and the ICL7612 work this way, with selectable operating currents of  $10\mu A$ ,  $100\mu A$ , and  $1mA$ .

#### **Op-amp design example: stuck-node tracer**

Let's start with a simple example, to see how micropower op-amp design goes. A tricky troubleshooting problem is a so-called stuck node, in which there is a short somewhere on a circuit board. It may be an actual short circuit in the wiring itself, or it may be that the output of some device (for example a digital three-state driver) is held in a fixed state. It's hard to find, because anywhere you look on that line you measure zero volts to ground.

A technique that does work, however, is to use a sensitive voltmeter to measure

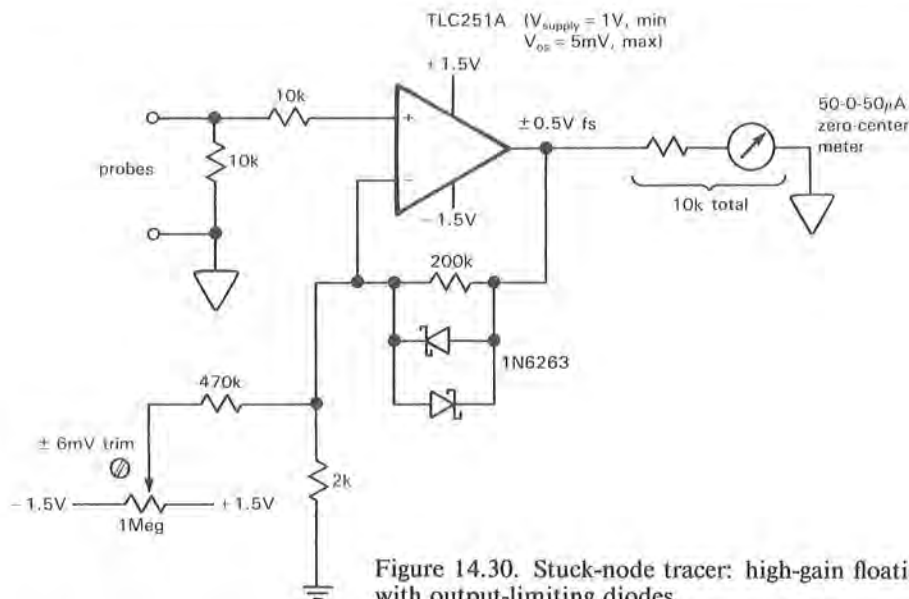


Figure 14.30. Stuck-node tracer: high-gain floating dc amplifier with output-limiting diodes.

voltage drops *along* the stuck trace. A typical signal trace on a printed-circuit board might be 0.012 inch wide and 0.0013 inch thick (1 ounce per square foot), which has a resistance along the trace of  $44\text{m}\Omega$  per inch. So if there's a device holding the line to ground somewhere, and you inject a diagnostic current of  $10\text{mA}$  dc somewhere else, there will be a voltage drop of  $440\mu\text{V}$  per inch in the direction of the stuck node.

Let's design a stuck-node tracer. It should be battery-powered so that it can float anywhere on the powered circuit under test. It should be sensitive enough to indicate a drop of as little as  $\pm 100\mu\text{V}$  on its zero-center meter, with larger meter deflections for larger drops. Ideally it should have a nonlinear scale, so that even for voltage drops of tens of millivolts the meter will not go off scale. With micropower design it should be possible to omit the on/off switch, since 9 volt batteries or AA-size cells give nearly their full shelf life (500mAh and 1400mAh, respectively) at drain currents less than  $20\mu\text{A}$ .

With a floating supply provided by

batteries, the simplest circuit is a high gain noninverting amplifier driving a zero-center meter (Fig. 14.30). Since the input and output are both intrinsically bipolarity, it's probably best to use a pair of AA cells, running the op-amp from  $\pm 1.5$  volt unregulated supplies. The back-to-back Schottky diodes reduce the gain at large output swings and prevent pegging; Figure 14.31 plots the resulting response versus  $V_{\text{in}}$ . Note the input-protection

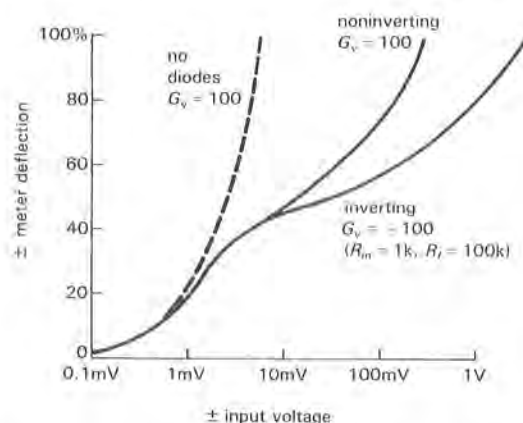


Figure 14.31. Stuck-node tracer achieves large dynamic range through nonlinear feedback.

resistor, in case voltages greater than  $\pm 1.5$  volts are applied across the inputs. The 10k resistor across the input keeps the output zero when you're not connected to the circuit under test.

The major difficulty in this design is in achieving an input offset less than  $100\mu\text{V}$  while maintaining micropower current drain, all with supply voltages of just  $\pm 1.5$  volts. The TLC251A is specified to operate down to 1 volt total supply voltage, and its CMOS output stage gives rail-to-rail swings. It has selectable operating currents of  $10\mu\text{A}$ ,  $150\mu\text{A}$ , or  $1\text{mA}$ ; naturally we choose  $10\mu\text{A}$  (by tying pin 8 to  $V_+$ ). This choice degrades slew rate and bandwidth, which we don't care about, but actually improves input offset drift ( $0.7\mu\text{V}/^\circ\text{C}$ ). The untrimmed input offset is  $5\text{mV}$ , which obviously must be trimmed. However, the fine print on the spec sheet says that "the amount of nulling range varies with the bias selection . . . . In low bias or when the TLC251 is used below 4V, total nulling may not be possible on all units."

If the conventional offset trimming circuit won't work, design your own! Here we've resorted to the circuit shown. It is guaranteed to work, since it can introduce more than  $5\text{mV}$  at the inverting input. And it adds only  $3\mu\text{A}$  to the supply current. But it is a compromise, because the trim depends on battery voltage, which isn't regulated. The trimming current is proportional to battery voltage, so in the worst case (a full  $5\text{mV}$  of initial input error) the offset voltage will drift  $50\mu\text{V}$  per percent of battery voltage change.

Until recently there was no clean solution to this problem. However, the OP-90 from PMI happens to provide a perfect solution here. It is a micropower op-amp with  $12\mu\text{A}$  of supply current and operation from supply voltages down to  $\pm 0.8$  volt. In the best grade (OP-90E), it has  $V_{os} = 150\mu\text{V}$  max, untrimmed.

Although it is bipolar, it swings clear down to the negative rail, and within a diode drop of the positive rail, which is good enough here. For this application it would probably make more sense to buy the cheap grade (OP-90G,  $0.5\text{mV}$ ) and trim the offset externally. One advantage of using a fixed-bias micropower op-amp, rather than a programmable op-amp, is that the offset trim is guaranteed to work.

### Miscellaneous micropower op-amps

The first programmable op-amp (in fact, the first low-power op-amp) was the bipolar 4250, introduced in 1967 by Union Carbide, who subsequently sold their line of linear products to Solitron. In 1970 the 4250 cost \$42.50. It became popular immediately (it still is) and is widely second-sourced. The 4250 is usable down to a microamp or so and will operate with 2 volts total supply voltage. It is cheap and delivers respectable performance; consider it an all-around "jellybean."

The 4250 does have one design peculiarity that can lead to problems when operated at low current. It has an amusing bias circuit that provides additional current to the output-stage drivers if the output load current to ground is large (compared with  $I_{\text{set}}$  multiplied by a pair of  $h_{FE}$ 's). This is supposed to help drive stiff loads, but the scheme can backfire if you overdo it, with the drivers robbing supply current from the rest of the op-amp. The op-amp then shuts off, relaxing via the compensation capacitor, then starts up again, etc., leading to a low-frequency oscillation at a few hundred hertz.

This problem was fixed in the quad bipolar LM346, which doesn't "motorboat," but which consequently has poor output-current sourcing capability at low supply current (see Fig. 14.32). The



TABLE 14.6. MICROPOWER OP-AMPS

| Type     | Mfg <sup>a</sup> | Supply<br>current<br>per amp<br>(V <sub>S</sub> =±5V)<br>typ<br>(μA) | #/pkg<br>1 2 4 | FET <sup>b</sup> | CM to rails? <sup>c</sup> | Total<br>supply<br>voltage |            | V <sub>os</sub><br>typ<br>(mV) | I <sub>b</sub><br>typ<br>(nA) | e <sub>n</sub> @10Hz<br>typ<br>(nV/√Hz) | Slew<br>rate<br>typ<br>(V/μs) |
|----------|------------------|----------------------------------------------------------------------|----------------|------------------|---------------------------|----------------------------|------------|--------------------------------|-------------------------------|-----------------------------------------|-------------------------------|
|          |                  |                                                                      |                |                  |                           | min<br>(V)                 | max<br>(V) |                                |                               |                                         |                               |
| CA3420   | RC               | 400                                                                  | * - -          | M N              |                           | 2                          | 20         | 2                              | 0.05pA                        | 350                                     | 0.5                           |
| AD821B   | AD               | 400 <sup>m</sup>                                                     | * * *          | J N              |                           | 4                          | 36         | 0.1                            | 0.01                          | 90                                      | 3                             |
| OP-97E   | PM               | 380                                                                  | * - -          | - -              |                           | 4.5                        | 40         | 0.01                           | 0.03                          | 17                                      | 0.2                           |
| LT1012C  | LT+              | 380                                                                  | * - -          | - -              |                           | 4                          | 40         | 0.01                           | 0.03                          | 17                                      | 0.2                           |
| LT1013C  | LT               | 370                                                                  | * * *          | - N              |                           | 4                          | 40         | 0.06                           | 15                            | 24                                      | 0.4                           |
| 358/324  | NS+              | 350                                                                  | - * *          | - N              |                           | 3                          | 32         | 2                              | 45                            | -                                       | 0.5                           |
| TSC911A  | TS               | 350                                                                  | * * *          | M N              |                           | 4                          | 16         | 0.005                          | 0.07 <sup>m</sup>             | 11 <sup>e</sup>                         | 2.5                           |
| TSC918   | TS               | 300                                                                  | * - -          | M N              |                           | 4.5                        | 16         | 0.005 <sup>m</sup>             | 0.1 <sup>m</sup>              | 4 <sup>e</sup>                          | 0.2                           |
| ALD1701  | AL               | 300                                                                  | * - -          | M B              |                           | 2                          | 12         | 2                              | 0.001                         | -                                       | 1                             |
| LM10     | NS+              | 270                                                                  | * - -          | - N              |                           | 1.1                        | 40         | 0.3                            | 10                            | 50                                      | 0.1                           |
| 312      | NS+              | 240                                                                  | * - -          | - -              |                           | 4                          | 40         | 2                              | 1.5                           | 45                                      | 0.1                           |
| MC34181  | MO               | 210                                                                  | * * *          | J -              |                           | 3                          | 36         | 0.5                            | 0.003                         | 65                                      | 10 <sup>i</sup>               |
| OP-80E   | PM               | 200 <sup>m</sup>                                                     | * - -          | M N              |                           | 4.5                        | 16         | 1 <sup>m</sup>                 | 10fa                          | 70 <sup>f</sup>                         | 0.2 <sup>m</sup>              |
| HA5151   | HA               | 200                                                                  | * * *          | - -              |                           | 2                          | 40         | 2                              | 70                            | 15 <sup>f</sup>                         | 4.5                           |
| TL031    | TI               | 190                                                                  | * * *          | J P              |                           | 10                         | 36         | 0.5                            | 0.002                         | 61                                      | 2.9                           |
| MC33171  | MO               | 180                                                                  | * * *          | - N              |                           | 3                          | 40         | 2                              | 20                            | 32 <sup>f</sup>                         | 2.1                           |
| OP-21    | PM               | 170                                                                  | * * *          | - -              |                           | 5                          | 36         | 0.04                           | 50                            | 20                                      | 0.3                           |
| TL061    | TI+              | 170                                                                  | * * *          | J -              |                           | 4                          | 30         | 3                              | 0.03                          | 85                                      | 3.5                           |
| MAX432   | MA               | 170                                                                  | * - -          | M N              |                           | 6                          | 32         | 0.001                          | 0.01                          | 1.2 <sup>g</sup>                        | 0.13                          |
| TLC25M2A | TI               | 150                                                                  | - * *          | M N              |                           | 1                          | 16         | 5 <sup>m</sup>                 | 0.001                         | 38 <sup>f</sup>                         | 0.6                           |
| LF441    | NS               | 150                                                                  | * * *          | J P <sup>g</sup> |                           | 6                          | 36         | 0.3                            | 0.01                          | 50                                      | 1                             |
| AD548C   | AD               | 150                                                                  | * * -          | J -              |                           | 9                          | 36         | 0.25 <sup>m</sup>              | 0.005                         | 35 <sup>f</sup>                         | 1.8                           |
| TSC900   | TS               | 140                                                                  | * - -          | M N              |                           | 4.5                        | 16         | 0.005 <sup>m</sup>             | 0.07 <sup>m</sup>             | 4 <sup>e</sup>                          | 0.2                           |
| ICL7621  | IL               | 100                                                                  | * * *          | M -              |                           | 2                          | 18         | 5 <sup>m</sup>                 | 0.001                         | 100 <sup>f</sup>                        | 0.2                           |
| LT1006   | LT               | 90                                                                   | * - -          | - N              |                           | 4                          | 40         | 0.03                           | 10                            | 24                                      | -                             |
| TL022    | TI               | 65                                                                   | - * *          | - -              |                           | 4                          | 36         | 1                              | 100                           | 50 <sup>f</sup>                         | 0.5                           |
| HA5141A  | HA               | 45                                                                   | * * *          | - N              |                           | 2                          | 40         | 0.5                            | 45                            | 35                                      | 1.5                           |
| OP-20    | PM               | 40                                                                   | * * *          | - N              |                           | 4                          | 36         | 0.06                           | 15                            | 60                                      | 0.03                          |
| LT1078A  | LT               | 40                                                                   | - * *          | - N              |                           | 2.2                        | 44         | 0.03                           | 6                             | 29                                      | 0.07                          |
| LP324    | NS               | 20                                                                   | - - *          | - N              |                           | 3                          | 32         | 1                              | 1                             | -                                       | 0.05                          |
| TLC1078C | TI               | 15                                                                   | - * *          | M N              |                           | 1.4                        | 16         | 0.18                           | 0.7pA                         | 68 <sup>f</sup>                         | 0.05                          |
| LT1178A  | LT               | 15                                                                   | - * *          | - N              |                           | 2                          | 44         | 0.03                           | 3                             | 50                                      | 0.03                          |
| OP-90    | PM               | 12                                                                   | * * *          | - N              |                           | 1.6                        | 36         | 0.05                           | 4                             | 35 <sup>h</sup>                         | 0.01                          |
| TLC25L2A | TI               | 10                                                                   | - * *          | M N              |                           | 1                          | 16         | 5 <sup>m</sup>                 | 0.001                         | 70 <sup>f</sup>                         | 0.04                          |

(a) see footnote to Table 4.1. (b) J - JFET; M - MOSFET. (c) input operating common-mode range: B - to both rails; N - to negative rail; P - to positive rail. (d) zero when sourcing. (e) μV pp, 0-10Hz (typ). (f) at 1kHz. (g) degraded SR and t<sub>r</sub>. (h) at 30Hz. (i) min/max.

| Type     | $f_T$<br>typ<br>(kHz) | $G_{OL}$<br>typ<br>(dB) | Output<br>current |              | Output $\Delta V$<br>from rails |                  | Comments                          |
|----------|-----------------------|-------------------------|-------------------|--------------|---------------------------------|------------------|-----------------------------------|
|          |                       |                         | source<br>(mA)    | sink<br>(mA) | $V_+$<br>(V)                    | $V_-$<br>(V)     |                                   |
| CA3420   | 500                   | 100                     | 1.5               | 1.5          | 0.1                             | 0.1              | low $I_b$                         |
| AD821B   | 1300                  | 120                     | 10                | 10           | 0                               | 0                | accurate, single-supply           |
| OP-97E   | 900                   | 126                     | 10                | 10           | 1                               | 1                | low-power OP-77                   |
| LT1012C  | 1000                  | 126                     | 10                | 10           | 1                               | 1                | precision, low-noise, low $I_b$   |
| LT1013C  | 1000                  | 137                     | 20                | 20           | 1                               | 0                | greatly improved 358/324          |
| 358/324  | 500                   | 115                     | 20                | 20           | 1.5                             | 0.5 <sup>d</sup> | popular single-supply             |
| TSC911A  | 1500                  | 120                     | 3.5               | 3.5          | 0.7                             | 0                | chopper; int cap                  |
| TSC918   | 700                   | 130                     | -                 | -            | 1                               | 0                | chopper; inexpensive              |
| ALD1701  | 1000                  | 108                     | 0.5               | 0.5          | 0                               | 0                | input, output to both rails       |
| LM10     | 300                   | 120                     | 20                | 20           | 0.01                            | 0.01             | output to rails; int. ref.        |
| 312      | 400                   | 110                     | 5                 | 5            | 1                               | 1                | original low $I_b$ precision      |
| MC34181  | 4000                  | 88                      | 8                 | 11           | 1                               | 0.5              | fast, low distortion              |
| OP-80E   | 300                   | 100                     | 10                | 10           | 1.5                             | 0                | ultra-low $I_b$ , 5pA(max)@85°C   |
| HA5151   | 1300                  | 100                     | 3                 | 3            | 1                               | 0.7 <sup>d</sup> | fast                              |
| TL031    | 1000                  | 83                      | 8                 | 20           | 1                               | 1.1              | improved TL061                    |
| MC33171  | 1800                  | 114                     | 4                 | 15           | 1                               | 1                |                                   |
| OP-21    | 600                   | 120                     | -                 | -            | 1                               | 1                | precision, low-noise              |
| TL061    | 1000                  | 80                      | 15                | 15           | 1.5                             | 1.5              | fast                              |
| MAX432   | 125                   | 150                     | 0.2               | 3            | 0                               | 0                | chopper; low clk noise; $C_{int}$ |
| TLC25M2A | 700                   | 106                     | 10                | 3            | 1.5                             | 0                | LinCMOS series                    |
| LF441    | 1000                  | 100                     | 4                 | 6            | 2                               | 2                | favorite JFET                     |
| AD548C   | 1000                  | 100 <sup>m</sup>        | 5                 | 5            | 2                               | 2                | improved LF441; dual=648          |
| TSC900   | 700                   | 130                     | 2.5               | 2.5          | 1.5                             | 0                | chopper                           |
| ICL7621  | 500                   | 102                     | 20                | 0.6          | 0                               | 0                | output to both rails              |
| LT1006   | -                     | 126                     | 20                | 20           | 1                               | 0                | precision                         |
| TL022    | 500                   | 80                      | 2                 | 2            | 2                               | 2                |                                   |
| HA5141A  | 400                   | 100                     | 3                 | 0.8          | 1                               | 0.5 <sup>d</sup> | fast                              |
| OP-20    | 100                   | 120                     | 0.5               | 0.5          | 0.5                             | 0.5              | precision                         |
| LT1078A  | 200                   | 120                     | 10                | 10           | 1                               | 0                | precision, recommended            |
| LP324    | 100                   | 100                     | 10                | 5            | 1.5                             | 0.5 <sup>d</sup> | favorite bipolar                  |
| TLC1078C | 110                   | 118                     | 15                | 15           | 1                               | 0                | LinCMOS, low offset               |
| LT1178A  | 60                    | 117                     | 5                 | 5            | 1                               | 0                | precision, recommended            |
| OP-90    | 20                    | 122                     | 5                 | 5            | 1                               | 0.7 <sup>d</sup> | precision, recommended            |
| TLC25L2A | 100                   | 110                     | 10                | 3            | 1.5                             | 0                |                                   |

TABLE 14.7. PROGRAMMABLE OP-AMPS

| Type    | Mfg <sup>a</sup> | Supply current |          | Total supply |         | CM to rails <sup>b,c</sup> | #/pkg | I <sub>s</sub> per amp <sup>d</sup> | I <sub>set</sub> (μA) | V <sub>os</sub> typ (mV) | I <sub>b</sub> typ (nA) | e <sub>n</sub> @10Hz typ (nV/√Hz) | Slew rate typ (V/μs) | f <sub>T</sub> typ (kHz) | G <sub>OL</sub> typ (dB) | Output current |           | Output ΔV from rails |                    | Comments                                 |
|---------|------------------|----------------|----------|--------------|---------|----------------------------|-------|-------------------------------------|-----------------------|--------------------------|-------------------------|-----------------------------------|----------------------|--------------------------|--------------------------|----------------|-----------|----------------------|--------------------|------------------------------------------|
|         |                  | min (μA)       | max (μA) | min (V)      | max (V) |                            |       |                                     |                       |                          |                         |                                   |                      |                          |                          | source (mA)    | sink (mA) | V <sub>+</sub> (V)   | V <sub>-</sub> (V) |                                          |
| OP-22   | PM               | 1              | 400      | •            | •       | •                          | •     | •                                   | 1                     | 0.1                      | 3                       | 90                                | 0.008                | 20                       | 125                      | 0.7            | 0.4       | 0.8                  | 0.8                |                                          |
|         |                  |                |          |              |         |                            |       |                                     | 10                    |                          | 20                      | 40                                | 0.08                 | 200                      |                          | 5              | 2         |                      |                    |                                          |
| OP-32   | PM               | 1              | 2000     | •            | •       | •                          | •     | •                                   | 0.6                   | 0.1                      | 2                       | 120                               | 0.03                 | 100                      | 125                      | 0.4            | 0.2       | 0.8                  | 0.8                |                                          |
|         |                  |                |          |              |         |                            |       |                                     | 6                     |                          | 15                      | 50                                | 0.3                  | 1000                     |                          | 2              | 1         |                      |                    |                                          |
| XR094   | XR               | 7000           | •        | •            | •       | •                          | •     | •                                   | 2.3                   | 3                        | 0.08                    | 18 <sup>ef</sup>                  | 0.1                  | 20 <sup>g</sup>          | 105                      | -              | -         | 1.5                  | 1.5                | poor data sheet; I <sub>set</sub> source |
|         |                  |                |          |              |         |                            |       |                                     | 23                    |                          |                         |                                   | 1                    | 200                      |                          |                |           |                      |                    |                                          |
| TLC251B | TI               | 10             | 1000     | •            | •       | •                          | •     | •                                   | -                     | 2 <sup>m</sup>           | 0.001                   | 70 <sup>f</sup>                   | 0.04                 | 100                      | 100                      | 10             | 3         | 1.5                  | 0                  | TLC 25xx, 27xx family                    |
|         |                  |                |          |              |         |                            |       |                                     | -                     |                          |                         | 38 <sup>f</sup>                   | 0.6                  | 700                      |                          |                |           |                      |                    |                                          |
| 346     | NS+              | 4              | 1000     | •            | •       | •                          | •     | •                                   | 0.3                   | 0.5                      | 3                       | 150                               | 0.01                 | 30                       | 120                      | 0.5            | 20        | 1                    | 1                  | 3 + 1; I <sub>set</sub> source           |
|         |                  |                |          |              |         |                            |       |                                     | 3                     |                          | 20                      | 70                                | 0.1                  | 300                      |                          | 5              | 20        |                      |                    |                                          |
| SL562   | PL               | 10             | 3000     | •            | •       | •                          | •     | •                                   | 0.4                   | 1                        | 2                       | 35 <sup>f</sup>                   | 0.01                 | 50                       | 90                       | -              | -         | 0.8                  | 0.8                | I <sub>set</sub> source                  |
|         |                  |                |          |              |         |                            |       |                                     | 4                     |                          | 10                      |                                   | 0.3                  | 400                      |                          |                |           |                      |                    |                                          |
| HA2725  | HA               | 1              | 1500     | •            | •       | •                          | •     | •                                   | 1                     | 2                        | 2                       | 150                               | 0.05                 | 60                       | 92                       | 1              | 1         | 1                    | 1                  |                                          |
|         |                  |                |          |              |         |                            |       |                                     | 9                     |                          | 8                       | 125                               | 0.5                  | 600                      |                          | 5              | 5         |                      |                    |                                          |
| CA3078A | RC               | 0.1            | 1000     | •            | •       | •                          | •     | •                                   | 1                     | 0.7                      | 6                       | 60                                | 0.3 <sup>i</sup>     | 1000 <sup>j</sup>        | 100                      | 10             | 10        | 0.7                  | 0.7                | ext. comp.; I <sub>set</sub> source      |
|         |                  |                |          |              |         |                            |       |                                     | 10                    | 1.3                      | 60                      | 25                                | 1.5 <sup>i</sup>     | 10000 <sup>j</sup>       |                          |                |           |                      |                    |                                          |
| CA3440A | RC               | 0.02           | 10       | •            | •       | •                          | •     | •                                   | 1                     | 2                        | 0.01                    | 250                               | 0.03                 | 50                       | 100                      | 1              | 2         | 2                    | 2                  | lowest I <sub>supply</sub>               |
|         |                  |                |          |              |         |                            |       |                                     | 10                    |                          |                         |                                   | 0.3                  | 300                      |                          |                |           |                      |                    |                                          |
| MC3476  | MO               | 0.1            | 1000     | •            | •       | •                          | •     | •                                   | 1                     | 2                        | 1                       | -                                 | 0.07                 | 200                      | 110                      | -              | -         | 2                    | 2                  |                                          |
|         |                  |                |          |              |         |                            |       |                                     | 10                    |                          | 10                      |                                   | 0.6                  | 700                      |                          | 5              | 5         |                      |                    |                                          |
| XR4202  | XR               | -              | -        | •            | •       | •                          | •     | •                                   | 0.3                   | 0.5                      | 200 <sup>m</sup>        | 300 <sup>l</sup>                  | 0.006                | 15                       | 80                       | -              | -         | 1.4                  | 1.4                | poor data sheet                          |
|         |                  |                |          |              |         |                            |       |                                     | 3                     |                          | -                       | 100 <sup>l</sup>                  | 0.06                 | 150                      |                          |                |           |                      |                    |                                          |
| 4250    | all              | 0.5            | 300      | •            | •       | •                          | •     | •                                   | 1.6                   | 3 <sup>m</sup>           | 6                       | 60                                | 0.03                 | 80                       | 105                      | 5              | 5         | 1                    | 1                  | jellybean; first prog. op-amp            |
|         |                  |                |          |              |         |                            |       |                                     | 16                    |                          | 40                      |                                   | 0.3                  | 250                      |                          |                |           |                      |                    |                                          |
| ICL7612 | IL               | 10             | 1000     | •            | •       | •                          | •     | •                                   | -                     | 2 <sup>m</sup>           | 0.001                   | 100 <sup>l</sup>                  | 0.02                 | 40                       | 102                      | 20             | 0.07      | 0                    | 0                  | 76xx family                              |
|         |                  |                |          |              |         |                            |       |                                     | -                     |                          |                         |                                   | 0.2                  | 500                      |                          | 20             | 0.7       |                      |                    |                                          |
| MC14573 | MO               | 4              | 1500     | •            | •       | •                          | •     | •                                   | 5                     | 8                        | 0.001                   | 1200                              | 0.2                  | 800                      | 95                       | 0.02           | 5         | 0.2                  | 0.1                |                                          |
|         |                  |                |          |              |         |                            |       |                                     | 50                    |                          |                         | 1600                              | 2                    | 2000                     | 80                       | 0.2            | 5         |                      |                    |                                          |

(a) see footnote to Table 4.1.

(b) J - JFET; M - MOSFET.

(c) B - to both rails; N - to negative rail.

(d) specs given at two values of I<sub>s</sub>; I<sub>set</sub> is sinking, unless stated otherwise.

(e) at I<sub>s</sub>=1.5mA per amp.

(f) at 1kHz.

(g) assuming ∞-I<sub>s</sub>.

(h) uncompl/decomp.

(i) 100Hz-10kHz.

(j) min/max.

(a) see footnote to Table 4.1. (b) J - JFET; M - MOSFET. (c) B - to both rails; N - to negative rail. (d) specs given at two values of I<sub>s</sub>; I<sub>set</sub> is sinking, unless stated otherwise. (e) at I<sub>s</sub>=1.5mA per amp. (f) at 1kHz. (g) assuming ∞ I<sub>s</sub>. (h) uncomprod/decamp. (i) 100Hz-10kHz. (m) min/max.

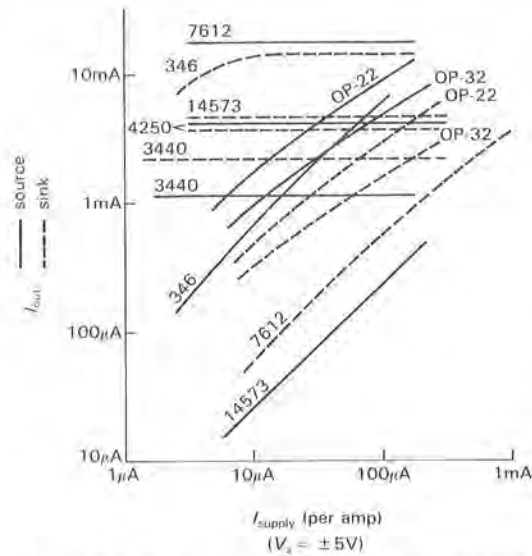


Figure 14.32. Output drive capability (source and sink) versus supply current for various programmable op-amps.

346 is otherwise a nice op-amp, organized as a single plus triple op-amp, with a corresponding pair of programming inputs.

One problem with bipolar programmable op-amps operated at very low currents is that the input bias current does not fall as fast as the supply current (i.e., the input-stage beta drops off at low collector current); thus, for example, the LM346 has a relatively large  $I_b(\text{max})$  of 100 nA when operating at 35  $\mu\text{A}$  per op-amp. This problem is often serious because most bipolar programmable op-amps do not use Darlington inputs or superbeta input transistors. More recent programmable op-amps have emphasized MOSFET design, for example the ICL761x series from Intersil, TI's "LinCMOS" TLC250/270 series, and the RCA CA3440. All have bias currents in the picoamp range, and operating common-mode range down to the negative rail. The ICL7612 has the additional wrinkle of operating with input common-mode swings beyond *both* rails. MOSFET outputs can swing to the supply rails; the 761x-series outputs can saturate all the way to both rails, while the TLC250/270

series are able to saturate only to the negative rail. Only the 3440 is continuously programmable (the others offer 3 supply current choices), and it is the undisputed champ for operation at extremely low currents. You can run it down to a few *nanoamps* supply current, though it won't set any speed records: At 100 nA supply current, the 3440's slew rate is 0.0004 V/ $\mu\text{s}$ , and its  $f_T$  is 200 Hz! However, because of MOS construction, it still delivers good output drive ( $\pm 1\text{ mA}$  at 2 V from the rails). The 3440 is a very good choice for micropower design. (Warning: Note that Figures 7 and 8 in the data sheet should be labeled nA, not  $\mu\text{A}$ ).

TI's LinCMOS line (TLC250/270 series) has some very nice features, including (like the 3440) good output drive at low supply current. It uses phosphorus-doped polysilicon gate technology to give extremely low offset drift with time (0.1  $\mu\text{V}/\text{month}$ ), eliminating a traditional weakness of metal-gate MOSFET op-amps and comparators. TI has a real winner here, unfortunately poorly documented in their traditionally uninformative linear data sheets.

Most CMOS op-amps (including all those mentioned above) share the problem of limited total supply voltage (see Section 4.22), typically 16 volts (e.g.,  $\pm 8\text{ V}$  max). That's the bad news; the good news is that they can run on very low total supply voltages (2 V for the 761x, 1 V for the TLC250, 4 V for the 3440).

We assembled in Tables 14.6 and 14.7 the low-power and programmable op-amps that we know about. If you compare them with Table 3.1, you'll see that micropower design is a specialty subject.

#### Micropower design example: integrating metronome

Figure 14.33 shows a micropower circuit that generates audible ticks at a rate proportional to the light intensity of a photographic-enlarger lamphead. Thus, if you time your enlargement exposures

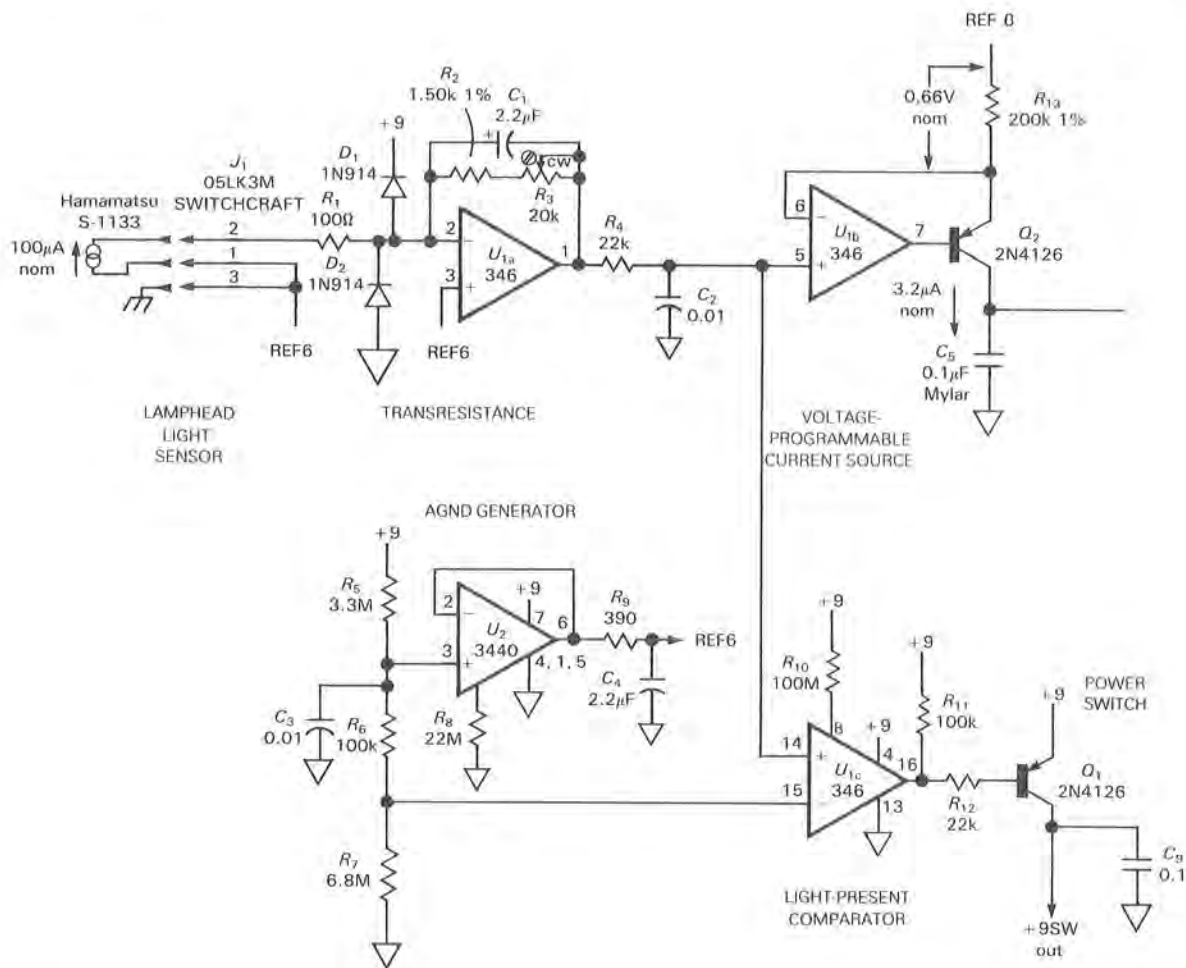
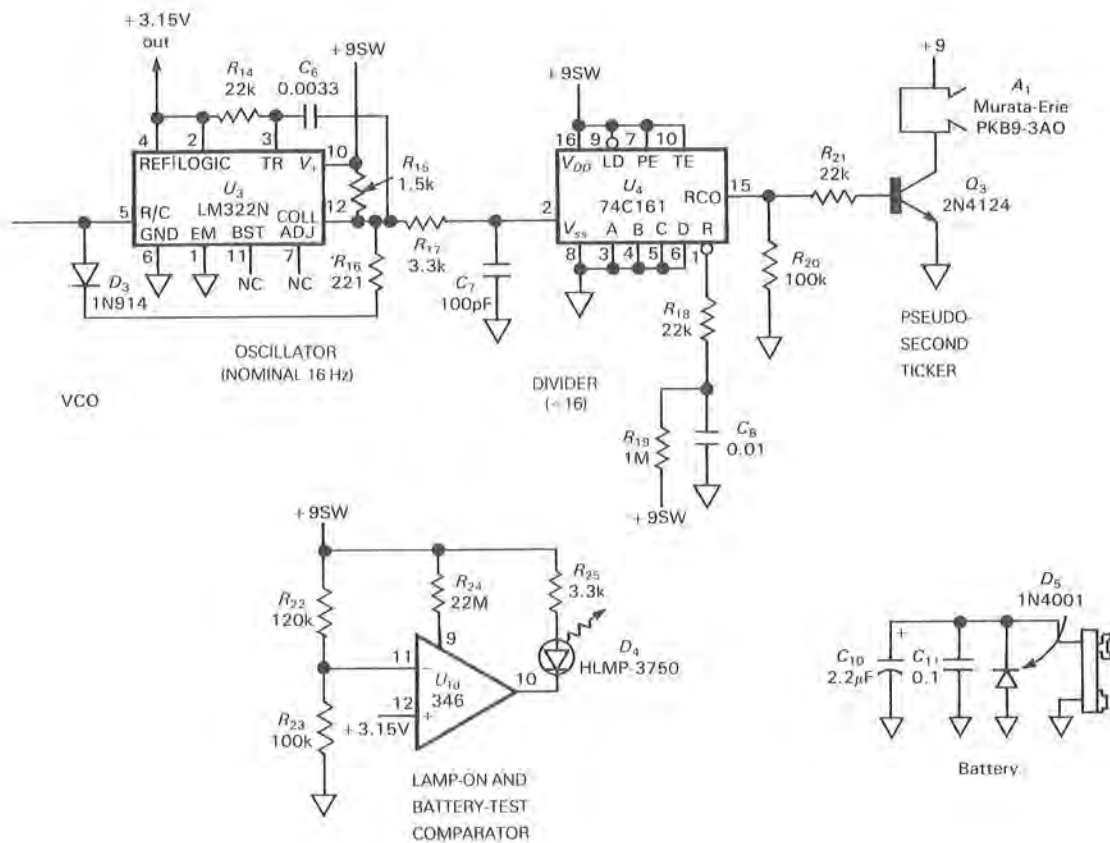


Figure 14.33. Micropower light-integrating darkroom timer.





by the clicks, the prints will be uniformly exposed even if the lamphead brightness is changing (perhaps due to power-line voltage variations, warmup of a fluorescent lamphead, etc.). Design objectives included 9 volt battery operation (simple, cheap), and no ON/OFF switch (people forget to turn it off). It also should signal (with an LED) that the battery is good.

Since battery drain has to be kept below  $20\mu\text{A}$  to give 2 years of life with a 9 volt battery (500mAh, see Table 14.2), and since the LED and piezo beeper require a few milliamps each, the only way to eliminate the ON/OFF switch is to use power switching (see Section 14.17), turning on the power-hungry circuitry only when lamphead light is detected. This is a perfect application for programmable op-amps, which can be shut down via their programming pin. Let's walk through the circuit.

We needed a split supply, and we didn't want to use two batteries, so we began by using micropower follower  $U_2$  to generate a "ground" rail at +6 volts (call it "REF6"). The divider (ignore  $R_6$  for the moment) draws  $1\mu\text{A}$ , and the 3440 is biased to operate at  $2.5\mu\text{A}$  ( $I_Q = 10I_{\text{set}}$ ) via  $R_8$ . The 3440 is a good choice because its CMOS design has negligible input current (50pA max) and maintains several milliamps of output drive capability (both source and sink) even at microamp quiescent currents. In fact, we could have run it at lower current; we picked this value because 22M is the largest standard resistor value, and the resulting current is within our budget! Note the bypass on the divider chain, to suppress capacitively coupled garbage (with megohm impedances you've got to be careful). The  $2.2\mu\text{F}$  capacitor at the output keeps the rail a low impedance even at high frequencies, where  $U_2$  has no gain ( $f_T = 0.01\text{MHz}$  at  $I_Q = 2.5\mu\text{A}$ ), with  $R_9$  for decoupling to prevent oscillation of  $U_2$  into the capacitive load (see Section 7.07).  $U_2$  is always powered.

The light detector is a photodiode, which generates a current (nominally  $100\mu\text{A}$ , but proportional to lamphead intensity) into a short-circuit load, namely the virtual ground at  $U_{1a}$ . We want to generate a frequency accurately proportional to that current, which suggests a capacitor and relaxation oscillator. That won't work, though, because the photodiode, acting as a current source, has very little compliance (0.1V or less). In addition, we need a way to calibrate the instrument, i.e., an adjustment that makes the metronome tick at 1 second intervals when the particular lamphead is at normal brightness (which might be  $50\mu\text{A}$ , or  $200\mu\text{A}$ , rather than the nominal  $100\mu\text{A}$ ). Finally, we need a way to activate the power switching when light is detected.

For these reasons we began with a transresistance (current-to-voltage) stage, with gain adjustable over a 15:1 range via  $R_3$ .  $C_1$  adds some smoothing for fluorescent light sources, which flicker at 120Hz. The photodiode is referenced to REF6, to keep it in the operating common-mode range of  $U_1$ . That makes  $U_{1a}$ 's output a voltage that is below REF6 by an amount proportional to light intensity, nominally 0.66 volt (for 1s ticks) when the calibration has been set.  $U_{1a}$ 's output drives two circuits: a comparator ( $U_{1c}$ ), which controls the power switching, and a current source ( $U_{1b}$ ), which drives the relaxation oscillator whose divided output will be the metronome output.

The comparator ( $U_{1c}$ ) is one of the three sections of  $U_1$  (a, b, and c) that run continually, biased via  $R_{10}$  at about  $9\mu\text{A}$  total. We want to switch power on when  $U_{1a}$ 's output is slightly below REF6, so we pick off a reference from  $U_2$ 's divider, via  $R_6$ , that is 0.1 volt below REF6. The output of comparator  $U_{1c}$  puts  $Q_1$  into saturation to switch on the +9 volt power, called "+9SW," which therefore appears whenever the enlarger lamphead is turned on.

The current source ( $U_{1b}$ ) runs continually. It is the standard op-amp plus *pnp* transistor configuration (see Section 4.07), and it sources  $3.2\mu\text{A}$  into  $C_5$  when its input is the nominal 0.66 volt below REF6. Note that one benefit of not using the photocurrent directly is that we can scale it to a convenient value; in this case,  $3.2\mu\text{A}$  into  $0.1\mu\text{F}$  (i.e.,  $dV/dt = I/C = 32\text{V/s}$ ) generates 16Hz in the 322 (which uses a precise internal reference to set its trigger point at 2.0V). The smorgasbord of resistors, capacitors, and diodes surrounding the 322 is, unfortunately, what it takes to form a relaxation oscillator with this sometimes awkward chip.

The 3.3k and 100pF network at the 322's output is what we have found necessary to cure its genetic predisposition toward double transitions (the 555 often has the same problem, with the same cure). The deglitched output drives a CMOS divide-by-16, whose output – ticks of  $\frac{1}{16}$  second duration at 1 second intervals – saturates  $Q_3$  to drive piezo beeper  $A_1$ .  $R_{19} - C_8$  forces the divider to zero at the beginning of each exposure, so the first beep comes at the end of the first “pseudo-second.”

The last section ( $U_{1d}$ ) of the quad 346 is used, in comparator mode, as a combination “lamphead-on” and “battery-OK” indicator. Its quiescent current is separately programmable via  $R_{24}$ , hence is powered only when +9SW is on.  $D_4$ , a high-efficiency LED that we run at 2mA, is therefore ON when the lamphead is on, provided that voltage from divider  $R_{22} - R_{23}$  is at least 3.15 volts (a stable reference conveniently provided by the 322) above the negative rail. That will be true if the battery voltage is at least 7.0 volts, i.e., not near end of life.

Because of power switching, only  $U_{1a-c}$  and  $U_2$  run continually, with a combined drain of about  $12\mu\text{A}$ . When  $U_{1c}$  senses

lamphead photocurrent, it turns on 9SW, powering the 322 (2.5mA), the LED (2mA), and the piezo beeper ( $\frac{1}{16} \times 8\text{mA}$ , i.e., 0.5mA average current). The battery life works out to about 5 years (or the “shelf life”) in the quiescent state, and 100 hours operating. At 15 seconds exposure per average enlargement, that's 24,000 enlargements.

In the design phase, we chose the 3440 because of its good drive and low input current at low  $I_Q$ . We picked the 346 because of good overall characteristics and low price in a convenient quad package. The 322 was used because its internal reference meant that we didn't need regulated supplies (as we would have with something like a 555) when its timing capacitor was driven with an external current not referenced to the power supply. Its reference voltage output provided a nice bonus, in the form of a “battery-low” indication.

The “ground” rail (REF6) was put (asymmetrically, and as high as possible) at +6 volts to get maximum dynamic range of lamphead brightness: Since  $C_5$  charges to 2 volts above the negative rail, the circuit will stop operating when the current programming voltage across  $R_{13}$  reaches nearly 4 volts (6 times nominal), since the current source will then run out of compliance. At the low end of the dynamic range, voltage offsets in  $U_{1a}$  and  $U_{1b}$  start to produce errors at about  $\frac{1}{8}$  nominal brightness. Thus, the choices of ground-rail voltage (6V) and nominal programming voltage (0.66V) combine to give a dynamic range of  $\frac{1}{8}$  to 6 times nominal, which is far more than any light source should ever fluctuate. For example, a fluorescent lamphead initially at room temperature produces about  $\frac{1}{3}$  of its fully warmed-up brightness. We chose a rate of 16Hz because a single divide-by-16 can then provide the drive signal for the piezo beeper, without monostables.

Note the protective circuitry:  $R_1$  prevents damage to the photodiode from the peak currents that could, under some unusual condition, flow from a charged  $C_1$ . Clamps  $D_1$  and  $D_2$  prevent damage to  $U_{1a}$  from something crazy plugged into the input.  $R_{18}$  prevents charged-up  $C_8$  from putting  $U_4$  into SCR latchup when +9SW turns off. Although these precautions may be unnecessary under most conditions, they were used anyway because this instrument is in commercial production, where a moderate field failure rate can wipe out all your profit (as well as your reputation!).

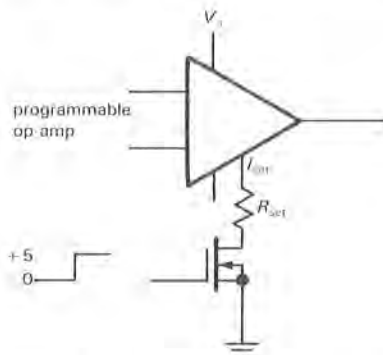


Figure 14.34. Using the programming pin as a power-down control.

#### Miscellaneous ideas: micropower op-amps

The programming pin can be used as a convenient on/off control for power switching, as in Figure 14.34. This is simpler than switching the op-amp's supply voltages, as we were forced to do for the other high-current loads in the metronome example. Most programmable op-amps (3440, 4250) work with an  $I_{set}$  sink, as in the figure, so you can use this simple circuit. It may be a good idea to use a high-value pullup resistor to the op-amp's positive supply, to bypass leakage currents and thus ensure complete turn-off.

Some op-amps are "quasi-programmable," in that they permit a choice of several fixed operating currents (typically  $10\mu\text{A}$ ,  $100\mu\text{A}$ , and  $1\text{mA}$ ). Examples are the ICL7612 and the TLC251/271. The TLC250/270 series also includes multiply packaged op-amps with "low" ( $10\mu\text{A}/\text{section}$ ), "medium" ( $150\mu\text{A}/\text{section}$ ), and "high" ( $1\text{mA}/\text{section}$ ) quiescent current, specified with part numbers like TLC27L2, TLC27M4, and TLC274 (low dual, medium quad, high quad, respectively).

When using CMOS op-amps as comparators, watch out for large drifts of  $V_{os}$  with time, an effect caused by sodium-ion migration in the gate region under long-term differential input voltage; this doesn't much affect their use as op-amps, where feedback enforces zero differential input voltage. The LinCMOS TLC270 series does not suffer from this disease, because of its phosphorus-doped polysilicon gate.

Some CMOS op-amps exhibit a dramatic increase of input (leakage) current when operated at large supply voltage and nonzero input voltage; for example, the LinCMOS line from TI can have  $I_{in}$  as large as  $20\text{nA}$  with  $V_{in} = +2$  volts and  $V_{supply} = \pm 9$  volts. Remember, too, that all FET op-amps (both JFET and MOSFET types) exhibit dramatic increases in input current with rising temperature, typically doubling each  $10^\circ\text{C}$ . At high temperatures, FET op-amps often have higher input currents than do good bipolar types; see Figure 3.30.

It's an unfortunate fact that most micropower op-amps are internally compensated for unity gain. At low quiescent current you need all the slew rate and bandwidth you can get, so it would be nice to have decompensated or uncompensated micropower op-amps for use in high- $G_V$  applications. There is one such op-amp – the OP-32 from PMI – but its slew rate and  $f_T$  are only slightly improved over those for unity-gain compensated jellybeans like the 4250 and 346.



### 14.13 Micropower comparators

The same speed/power trade-off that limits micropower op-amp performance applies to comparators as well. As with conventional comparators and op-amps, however, you'll always get better speed with a comparator than with an op-amp of equivalent power; this is because comparator ICs, not intended for use with negative feedback, do not have speed-robbing frequency compensation. Of course, if you're not concerned with speed, an op-amp will often do the job nicely, as in the metronome circuit above.

As with op-amps, micropower comparators come in two varieties, namely *programmable* and *fixed  $I_Q$* . An example of the former is the LP365, a quad programmable bipolar comparator that is specified as low as  $10\mu\text{A}$  total supply current; it will run from 4 to 36 volts total supply voltage and has a separate output emitter terminal (like the 311) so that you can compare negative voltages while driving logic. An example of the latter is the LP339, a micropower ( $15\mu\text{A}/\text{section}$ ) version of the popular low-power ( $200\mu\text{A}/\text{section}$ ) LM339 quad comparator. TI makes CMOS versions of the 339/393 (TLC339/393), with even lower quiescent current, and with excellent speed/power performance; they also offer them with active pullups (TLC3702/4), so you don't waste precious current (and speed) with an external resistive pullup.

An unusual micropower comparator is the LT1040, which incorporates power switching to achieve an average quiescent current of  $0.1\mu\text{A}$  when strobed externally once per second. Alternatively, you can use the internal-strobing oscillator, which uses an additional  $0.5\mu\text{A}$ . A latency time of 1 second is perfectly adequate if you are monitoring slowly varying quantities, e.g., the level of fluid in a tank. The LT1040 is a dual comparator built with CMOS, and it has latched outputs.

In addition, it provides a "pulsed-power" output pin, active for about  $80\mu\text{s}$  during the conversion time, so you can drive a resistive network (e.g., a bridge with a thermistor in one leg) at the input with switched power also. This chip (or its close cousin the LT1041) would be a good choice simply as a low-frequency micropower oscillator (see next section), since microamp oscillators are not easy to make. Note, however, that it isn't a particularly stable oscillator.

You can use power switching with a conventional (i.e., fast) comparator if you want to do occasional fast comparisons at known times. For example, you might do acoustic ranging by sending short bursts of sound, then measuring the round-trip travel time while listening for the echo. By measuring the *difference* between travel times upstream and downstream, you can even measure a *velocity*. Speed is important here, and with CMOS logic you can do the job, but you probably want the performance of a full-power comparator (see Table 9.3). Power switching is a natural here, since you know when you want to make a measurement.

Table 14.8 lists most currently available low-power comparators.

### 14.14 Micropower timers and oscillators

In battery-operated instrumentation there is often a need to generate intervals of an hour or so. You may be making occasional measurements with various sensors, a power-switched microprocessor, and power-switched communications (or data logging). The whole system needs to "wake up" at predetermined intervals. An hour is far too long for  $RC$  timing, so you need a faster oscillator and a divider (programmable, perhaps, to set the interval to the next wake-up). Since the timer is the only part of the system that runs continuously, it needs to run at low current. What choices are available?



TABLE 14.8. LOW-POWER COMPARATORS

| Type                  | Mfg <sup>a</sup> | #pkg | Total supply |            | $I_s$ /unit<br>typ<br>( $\mu$ A) | CM<br>to<br>$V_s$ | $V_{os}$<br>max<br>(mV) | $I_b$<br>max<br>(nA) | Delay<br>typ, @ $V_s=5V$        |                                 | Output <sup>b</sup> | Emitter <sup>c</sup> | $I_{sink}$<br>typ<br>(mA) @ (V) |
|-----------------------|------------------|------|--------------|------------|----------------------------------|-------------------|-------------------------|----------------------|---------------------------------|---------------------------------|---------------------|----------------------|---------------------------------|
|                       |                  |      | min<br>(V)   | max<br>(V) |                                  |                   |                         |                      | $L \rightarrow H$<br>( $\mu$ s) | $H \rightarrow L$<br>( $\mu$ s) |                     |                      |                                 |
| CMP-04F               | PM               | 4    | 3            | 36         | 200                              | *                 | 1                       | 100                  | 1.4                             | 0.7                             | OC                  | -                    | 12 1                            |
| CMP-404E              | PM               | 4    | 3            | 36         | 55                               | *                 | 1                       | 50                   | 3                               | 4                               | OC                  | -                    | 15 1                            |
| LP311                 | NS               | 1    | 3            | 36         | 150                              | -                 | 7.5                     | 25                   | 1.2                             | 1.2                             | OC                  | *                    | 25 0.4                          |
| LM339                 | NS               | 4    | 2            | 36         | 200                              | *                 | 5                       | 250                  | 1.3                             | 0.75                            | OC                  | -                    | 12 1                            |
| LP339                 | NS               | 4    | 2            | 36         | 15                               | *                 | 5                       | 25                   | 13                              | 7                               | OC                  | -                    | 5 1                             |
| TLC339                | TI               | 4    | 3            | 16         | 10                               | *                 | 5                       | 5pA <sup>t</sup>     | 2.5                             | 2.1                             | OD                  | -                    | 10 0.5                          |
| LP365 <sup>d</sup>    | NS               | 4    | 3            | 36         | 50                               | *                 | 6                       | 20                   | 2                               | 4                               | OC                  | -                    | 2 1                             |
| "                     |                  |      |              |            | 5                                | *                 | 6                       | 5                    | 20                              | 40                              |                     | -                    | 0.2 1                           |
| TLC372C               | TI               | 2    | 2            | 18         | 100                              | *                 | 10                      | 1pA <sup>t</sup>     | 0.65                            | 0.65                            | OD                  | -                    | 10 1                            |
| TLC374C               | TI               | 4    | 2            | 18         | 100                              | *                 | 10                      | 1pA <sup>t</sup>     | 0.9                             | 0.9                             | OD                  | e                    | 10 1                            |
| LM393                 | NS               | 2    | 2            | 36         | 200                              | *                 | 5                       | 250                  | 1.3                             | 0.75                            | OC                  | -                    | 12 1                            |
| TLC393                | TI               | 2    | 3            | 16         | 10                               | *                 | 5                       | 5pA <sup>t</sup>     | 2.5                             | 2.1                             | OD                  | -                    | 10 0.5                          |
| LT1017                | LT               | 2    | 1.1          | 40         | 30                               | *                 | 1                       | 15                   | 18                              | 25                              | TTL                 | -                    | 10 0.15                         |
| LT1018                | LT               | 2    | 1.1          | 40         | 110                              | *                 | 1                       | 75                   | 6                               | 6                               | TTL                 | -                    | 10 0.15                         |
| LT1040                | LT               | 2    | 2.8          | 16         | 0.1 <sup>f</sup> <sub>s</sub>    | g                 | 0.5                     | 0.3 <sup>t</sup>     | -                               | -                               | TTL                 | -                    | 1.6 0.25                        |
| TLC3702               | TI               | 2    | 3            | 16         | 10                               | *                 | 5                       | 5pA <sup>t</sup>     | 2.7                             | 2.3                             | CMOS                | -                    | 10 0.5                          |
| TLC3704               | TI               | 4    | 3            | 16         | 10                               | *                 | 5                       | 5pA <sup>t</sup>     | 2.7                             | 2.3                             | CMOS                | -                    | 10 0.5                          |
| ICL7642C <sup>h</sup> | IL               | 4    | 2            | 16         | 10                               | *                 | 10                      | 0.05                 | 150                             | 300                             | CMOS                | -                    | 0.1 1                           |
| MC14574 <sup>i</sup>  | MO               | 4    | 3            | 15         | 45                               | *                 | 30                      | 0.05                 | 10                              | 5                               | CMOS                | -                    | 5 0.4                           |
| MC14578               | MO               | 1    | 3.5          | 14         | 10 <sup>m</sup>                  | -                 | 50                      | 1pA                  | -                               | -                               | CMOS                | -                    | 1 0.5                           |

(a) see footnote to Table 4.1. (b) CMOS - CMOS output, swings to both rails; OC - open npn collector; OD - open n-channel

drain. TTL - TTL active pullup output, for which the load may return to a more positive supply. (c) npn "open-emitter" pin.

(d) prog op-amp; single prog pin for all 4 sections. (e) common to all 4 sections. (f) sampling comparator.

(g) common-mode range extends to both supply rails. (h) CMOS low-power op-amp used as comp, with  $I_Q$  set to 10 $\mu$ A.

(i) prog op-amp; prog pin for each pair. (j) typical.

### CMOS relaxation oscillators

The first thing to note is that a conventional 4000-series CMOS relaxation oscillator (see Fig. 8.90) running at normal voltages draws a rather large class A current, due mostly to rail-to-rail conduction as the (relaxation) input approaches the CMOS threshold during each half cycle. This average current is in the neighborhood of  $50\mu\text{A}$  when operated at 5 volts (rising rapidly at higher supply voltages) and is relatively independent of oscillation frequency. The problem is not helped by substituting the high-speed 74HC or 74AC series. However, if the oscillator is run from a 3 volt supply, for example from a lithium battery, the supply current will be in the microamp range. Figure 14.35 shows such a micropower oscillator and its current drain at several supply voltages. By substituting 74HC logic you get an oscillator with very low jitter, though the voltage stability of this type of circuit is poor (typically 10% change in frequency going from 1.0V to 1.6V).

#### □ IC oscillators

- **Intersil ICM7242.** This is a CMOS RC oscillator plus 8-bit divider. It runs on 2–16 volts, and draws about  $100\mu\text{A}$  at 5 volts. Supply current is not much reduced at low supply voltage, unfortunately. Typical temperature coefficient of frequency is  $250\text{ppm}/^\circ\text{C}$ .
- **Intersil/Maxim ICM7240/50/60.** These are like the ICL7242, but with program-

mable digital dividers. They draw the same supply current.

- **Intersil ICM7207/A.** These are CMOS crystal oscillators with dividers, intended to provide output frequencies of 100Hz/10Hz and 10Hz/1Hz, respectively; to get these frequencies you have to use 6.5536MHz and 5.24288MHz crystals. These chips draw  $260\mu\text{A}$  at 5 volts, dropping to about  $80\mu\text{A}$  at 3 volts. The data sheet claims they will operate down to 1 volt at lower frequencies, with supply currents of a few microamps.
- **Intersil ICM7555/6 and others.** These are CMOS 555s, with generally improved characteristics (low supply current, higher maximum frequency, much smaller supply current transients). The quiescent current is  $60\mu\text{A}$  at 5 volts, which is also the approximate operating current as an oscillator, if large timing resistors are used. Typical tempco is  $150\text{ppm}/^\circ\text{C}$ . The 7556 is a dual 7555. The similar National LMC555 draws  $100\mu\text{A}$  and has tempco of  $75\text{ppm}/^\circ\text{C}$ . Look at Table 5.3 for other 555 lookalikes from TI, Advanced Linear Devices, and Exar.
- **Op-amps.** An op-amp relaxation oscillator (Fig. 5.29), built with a micropower op-amp, makes a good low-frequency oscillator. Use an op-amp with true CMOS outputs for rail-to-rail swings, particularly at low supply voltages, in order to achieve low tempco and reliable oscillation. The 7611/2 types are good.

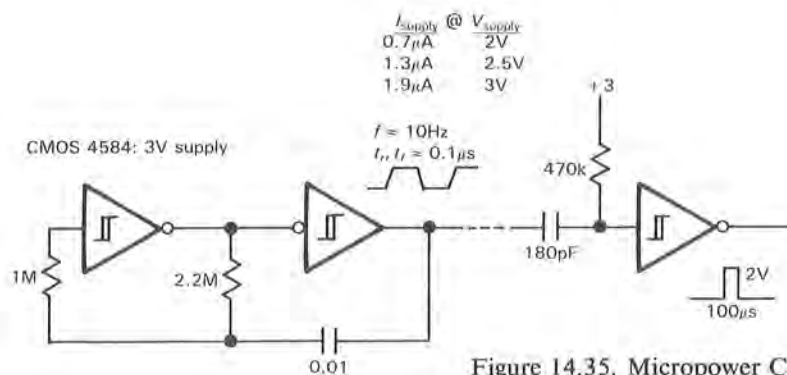


Figure 14.35. Micropower CMOS relaxation oscillator.

- **LT1040.** This is the power-switched comparator from Linear Technology described above, with integral micropower oscillator ( $0.3\mu\text{A}$  at low frequency). The supply current scales linearly with frequency, given by  $I(\mu\text{A}) = f(\text{Hz})/10$ . The oscillator is not particularly stable ( $0.2\%/^{\circ}\text{C}$ ,  $10\%/V$ ), but it certainly is micropower at low frequencies.
- **COPS timer.** The COP498 from National is a member of their COPS series of microcontrollers, with serial “MICRO-WIRE” interface. The COP498 includes a 32.768kHz crystal oscillator circuit that runs at  $20\mu\text{A}$  (max) at 2.4 volts. It must be programmed via the serial line, but then it can generate wake-up pulses at a 1Hz or 16Hz rate. Stabilities of a few ppm/ $^{\circ}\text{C}$  are typical with “tuning-fork”-type crystals.
- **Calendar clocks.** The MM58174/274 from NSC typifies timekeeping chips for microprocessors. They run from battery backup when the computer is off, and they maintain running time and date information, readable (and settable) via the computer’s data bus as an I/O port. The 58174 idles (oscillator running) at  $10\mu\text{A}$  (max) with 2.2 volts supply. Like the COPS chip, it can interrupt at periodic intervals, but only in its full-power 5 volt mode (1mA supply current); also, like the COPS chip, it requires programming via the bus to set up the interrupt interval. Other popular calendar clock chips are the DP8570 (National), the ICM7170 (Intersil), the MC146818 (Motorola), and the MSM5832 (Oki). Some of these allow the use of high-frequency (1MHz or above) crystals for better temperature stability. Epson makes a calendar chip with integral crystal (the RTC58321).
- **Wristwatch circuits.** There are available low-voltage CMOS chips designed to run stepper motors in analog-display wristwatches. For example, the discontinued ICM7245 required a 32.768kHz crystal and ran on 1.5 volts (single-cell battery), drawing only  $0.4\mu\text{A}$ ; various versions provided output frequencies of 1Hz, 0.1Hz, or 0.05Hz. Since it was designed for time-keeping, it had very good stability, typically 0.1ppm over supply voltage. National offers the MM5368, a mini-DIP 32kHz oscillator that draws  $50\mu\text{A}$  (max) at 3 volts and provides outputs at 1Hz, 10Hz, and 50/60Hz. Their MM53107 draws  $75\mu\text{A}$  at 3 volts with a 1MHz crystal, providing  $\approx 30\text{Hz}$  output.
- **Programmable unijunction transistor.** A unijunction transistor (UJT) is a 3-terminal (emitter, base 1, base 2) negative-resistance device that enjoyed some popularity in the 1960s in triggering circuits and free-running oscillators. The device goes into heavy conduction from emitter to base 1 when the emitter terminal is brought more positive than a critical trigger voltage  $V_t = \eta V_{BB} + 0.6$ , i.e., a diode drop above a fixed fraction  $\eta$  (the “intrinsic standoff ratio,” typically around 0.6) of the inter-base voltage; it continues to conduct until the emitter current drops below some minimum (“valley current”) value. The classic UJT oscillator is shown in Figure 14.36A, with the positive spikes at base 1 used to switch an *npn* transistor to generate full logic swings. You hardly see UJT’s used anymore, because op-amps and ICs like the 7555 can do more, and do it better. However, there is an unusual UJT series known as *programmable* UJT’s, in which the trigger parameters ( $\eta$ , peak and valley currents) can be set by an external divider. The 2N6028, in particular, is specified for peak currents as small as  $0.1\mu\text{A}$  and can thus be operated as an oscillator at less than a microamp. Figure 14.36B shows a 10Hz oscillator with CMOS output swing and  $1\mu\text{A}$  operating current; look also at Figure 6.57, where we used a PUJT in a micropower dc-dc converter.

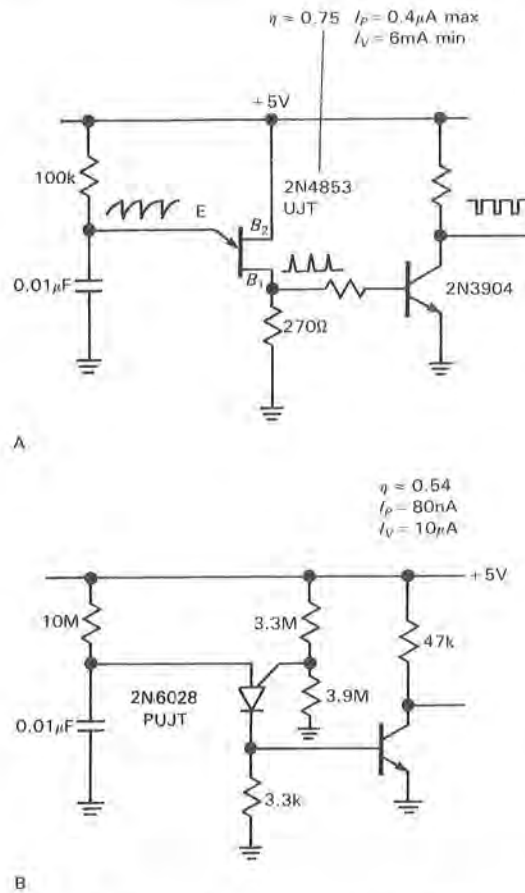


Figure 14.36. Unijunction transistor relaxation oscillators.

## MICROPOWER DIGITAL DESIGN

At first glance, micropower digital design seems simple: Just use CMOS everywhere, including microprocessors and memory, right? Well, almost. CMOS is certainly the family of choice, although bipolar logic can be used with power switching. But there are now several CMOS variants, and there are many pitfalls that can render an unsuspecting CMOS design anything but micropower. In this last section of the chapter we'll review CMOS families and how to design with them so that you really achieve the low power you need for battery operation.

### 14.15 CMOS families

As we discussed at length in Chapter 9, there are several CMOS families; the best one to use depends on the application. They are summarized in Table 9.1.

#### 4000B/74C series

This is the B-series enhancement of the original metal-gate CMOS, which the data sheets say can be operated with 3 to 15 volt supply. Operation is really marginal at 3 volts, however, with high  $Z_{out}$ , poor noise immunity, and low speed. The practical minimum supply voltage is 5 volts. At the high end of the supply voltage range there is considerable class A current during switching, and greater susceptibility to sudden death from spikes on the power supply. These CMOS families are pretty gutless in terms of output drive: 1mA or less when operating from 5 volts. 74C is electrically similar to 4000B, including supply voltage range, but with the functions and pinouts of 74 TTL. Fairchild offered an improved "Isoplanar C" family (faster, same voltage range), as does Philips/Signetics (called "LOCMOS"). These are the only series of CMOS that can be run over a large supply range; they are ideal for operation directly from a 9 volt battery.

#### 74HC and 74AC series

Polysilicon-gate "high-speed" (HC=74LS speed, AC=74F/74AS speed) CMOS operates with 2 to 6 volts (or 1.5V to 5V) supply and is actually *specified* over that range. These devices are quite rugged, with good input protection and virtual freedom from SCR latchup. They have CMOS thresholds (i.e., half  $V_{DD}$ ) and rail-to-rail output swings. The HC series includes many of the popular 4000B-series functions (e.g., 74HC4046), as well as the usual 74LS functions.

### 74HCT and 74ACT series

These are variants of the HC and AC series, with TTL input thresholds for compatibility when coexisting with bipolar TTL devices in the same circuit. Since micropower circuits don't generally have any bipolar TTL, you should always choose HC/AC, with its better noise immunity. HCT and ACT devices require 5 volts  $\pm 10\%$ .

### Low-threshold special function

As we will show below, the dynamic power consumption of CMOS is proportional to the square of the supply voltage. This is a powerful (!) incentive to operate at low supply voltage, and it is the reason for chips like the sub-microwatt (1.5V supply) wristwatch oscillator/dividers. These unusual ICs can be very useful, besides being inexpensive because of their large production runs.

### $M^2L$ (Mickey Mouse logic)

Don't overlook the possibility of using a few discrete components to make or invert a logic level; Figure 14.37 shows the idea. This can be particularly useful if you need to interface between different supply voltages. You can even create impromptu gates, etc., by adding diodes or paralleling transistor outputs.

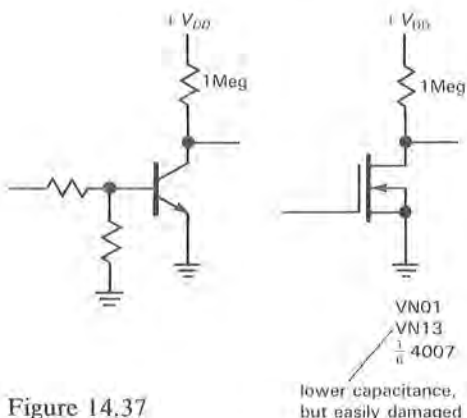


Figure 14.37

### Choice of family

Use the 4000B/74C series if you don't need speed or high output current, or if you want to use unregulated or high supply voltages. Use HC (or AC) if you need the speed or output drive, but note the more limited supply voltage range. The AC (and ACT) families cause more problems from capacitive coupling and reflection (transmission-line effects), and also power-supply transients (crowbarring), because of their fast edge times and robust output drivers. They should be avoided unless absolutely needed. In general, avoid devices with TTL thresholds (HCT, ACT) unless you need them to interface to bipolar TTL, or to NMOS LSI circuitry with TTL logic levels.

### 14.16 Keeping CMOS low power

There are several routine measures you should take to achieve low-current CMOS operation. In addition, it's worthwhile raising your CMOS pathology awareness.

#### Routine design considerations

1. *Keep as few nodes as possible involved with high frequencies.* CMOS has no quiescent current (other than leakage), but current is required to charge internal (and load) capacitances during switching. Since the energy stored in a capacitor is  $\frac{1}{2} CV^2$ , and an equal amount of energy is dissipated by the resistive charging circuit, the power dissipated is

$$P = V_{DD}^2 f C$$

for a switching frequency  $f$ . Thus, CMOS devices consume power proportional to their switching frequency, as shown in Figure 14.38 (compare with Fig. 8.18). At their maximum operating frequency they may use more power than equivalent bipolar TTL logic. The effective capacitance  $C$  is often given on data sheets as the "power dissipation capacitance," to which



you must add the load capacitance  $C_L$  before applying the formula above.

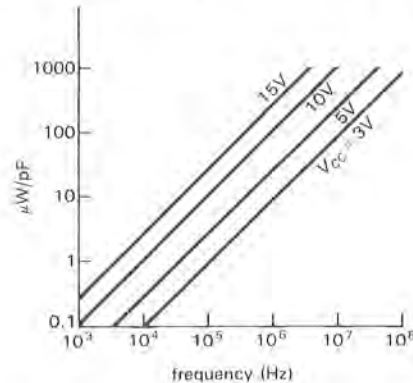


Figure 14.38. CMOS dynamic power consumption.

2. *Within a circuit, keep all  $V_{DD}$ 's and  $V_{SS}$ 's the same.* Otherwise you may have current flowing through input-protection diodes. Even worse, you may force a chip into SCR latchup (see pathologies, below).
3. *Make sure logic swings go all the way to the rails.* CMOS outputs swing rail-to-rail. Outputs from other devices – bipolar TTL, oscillators, NMOS chips – may hover in between, causing class A current and decreased noise immunity.
4. *No open inputs.* Open inputs are the enemy of micropower operation, since there may be considerable class A current (and even oscillation) as the input floats to the logic threshold. Tie all unused inputs to ground (or  $V_{DD}$ , if that disables something you don't want).
5. *Arrange loads to keep normal-state drains low.* Pullups, pulldowns, LEDs, and output drivers should be wired so that current is minimum in the usual state. Thus, for example, use an *npn* (not *pnp*) transistor to switch a high-voltage load from a node that spends most of its time low.
6. *Avoid slow transitions.* Again, class A current is the culprit. A sine-wave input driving a CMOS Schmitt trigger may cause a lot of supply current.

7. *Put current-sensing resistors in the  $V_{DD}$  lead.* In certain failure modes (see below), particularly those caused by static damage, a CMOS chip may draw excessive quiescent current; a 10 ohm resistor in series with  $V_{DD}$  on each board (bypass the load side) makes it easy to see if that is happening. Putting such a resistor on each chip (usually no need to bypass in this case) allows you to locate the bad chip quickly (Figure 14.39).

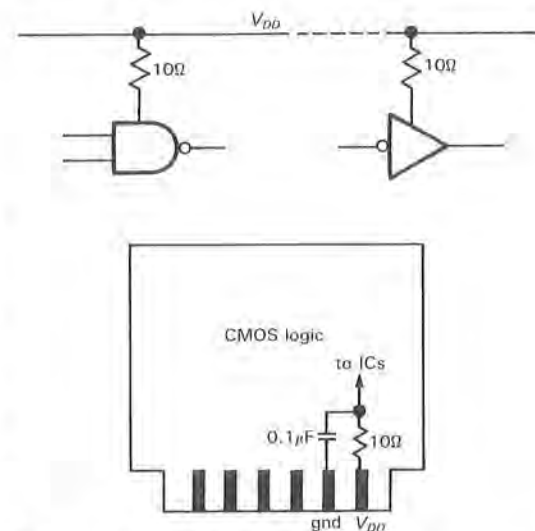


Figure 14.39. Supply current sensing ("current spy").

8. *Quiescent current screening.* A typical HC- or 4000B-series CMOS logic chip has a specified  $I_Q$  of  $0.04\mu\text{A}$  (typ),  $5\mu\text{A}$  (max). Most of the time it is rare to have a quiescent current anywhere near the maximum, but it can happen. If you are operating at low switching frequencies (therefore low dynamic current), and require comparably low quiescent current, you may need to screen incoming chips. The use of small series resistors as recommended above makes the job much easier. We've noticed that in the case of CMOS LSI chips (such as large memories) the typical quiescent current may be close to the manufacturer's maximum leakage specifications – beware!



Fi



Note 339, National Semiconductor Corp.)

A. CMOS IC cross section.

B. Equivalent circuit.

9. *Time-out power switching.* You can save a lot of power by making sure an instrument is turned off when no one is using it. Figure 14.40 shows a straightforward CMOS timeout circuit that turns off the switched +9 volt power an hour after the instrument has been turned on. You might build this into a manual instrument (e.g., a multimeter). It uses a 4536 oscillator/divider/monostable to reset a flip-flop that controls the instrument power. The circuit runs on +3 volts to keep current drain below  $5\mu\text{A}$ . Using the monostable output prevents logic races and runt pulses, and the "8-bypass" line is used to test the circuit by shortening the delay to 15 seconds. MOS switches provide an easy interface with low quiescent power.

#### **CMOS pathologies and failure modes**

CMOS circuits exhibit strange behavior under some circumstances, and they can fail in weird ways. Some of these increase power dissipation dramatically. Here is a rundown.

1. *SCR latchup.* This is a major threat to low-power operation. The silicon substrate forms diode junctions with the elements of the CMOS circuit, producing a parasitic SCR-type circuit (Fig. 14.41) that can be triggered into heavy conduction. It takes typically 20mA–200mA (the larger values for the newer CMOS families) of current through the input-protection (or output-protection) diodes (Fig. 3.50) to turn on the parasitic SCR. Once on, it remains in heavy supply-to-supply conduction, with about 1 volt from supply to ground, often destroying the IC (or even the power supply!). To prevent SCR latchup, design your circuits with series-input current-limiting resistors in places such as external inputs, inputs that can be overdriven, outputs that go off-board, and signals between portions powered by separate supplies. Unfriendly loads that can drive currents into CMOS outputs are potential causes of latchup.

For example, a current-output CMOS D/A converter (running from +5V and ground) driving the summing junction of an op-amp (powered from bipolarity supplies) seems reasonable enough; but on power-up you may get momentary current sinking at the op-amp's input, driving the DAC into latchup. A cure here is to add a Schottky diode to ground. (The newest D/A converters are designed to prevent this; look for a phrase like "protection Schottky not required.") Another place we've seen latchup is when switching large inductive loads with power MOSFETs, whose large feedback capacitances produce high dynamic currents at the gate driver (i.e., the CMOS logic output) during transitions. See "Signal coupling," below, for another latchup scenario.

An easy way to induce latchup is to plug a circuit board into a powered socket, since a signal line may connect before the power-supply line. (With any kind of circuitry it's a bad idea to plug or unplug boards or modules with power on.) When designing with CMOS, it's worth studying the official latchup specifications. The newer polysilicon-gate CMOS types have very effective protection circuits, and some manufacturers (e.g., National) claim that their HC/HCT or AC/ACT lines cannot be triggered into latchup.

2. *Signal coupling.* Because of the high impedances involved, CMOS is prone to capacitive coupling from nearby signals with fast edge times, producing logic spikes. For example, high-impedance pullups or pull-downs allow coupling of spikes through wiring capacitance from nearby lines carrying fast edges; to cure this, use a small ( $0.001\mu\text{F}$ ) bypass capacitor. In general, wiring going off to panels can cause trouble by this mechanism. The relatively high capacitance can even couple to output lines, particularly with 4000B/74C-series CMOS running at 5 volts. In extreme cases (for example, switching high voltages with a relay in the same cable as logic levels)

there can be enough coupling to induce SCR latchup.

3. *Clock skew.* As we mentioned in Chapter 9, the relatively high  $Z_{out}$  of 4000B/74C CMOS can lead to trouble in a synchronous system, particularly if the clock lines are heavily loaded capacitively, delaying the clock relative to the data. The relatively large scatter of CMOS logic thresholds only aggravates the situation. In an unregulated battery-operated system, it's important to check for reliable operation over the full range of supply voltage. Ironically, the problem tends to become worse at higher  $V_{DD}$ , where data delays and transition times become shorter. This is one argument for using regulated supplies in battery-operated CMOS systems.

4. *Failure modes.* Damage at the input can cause input leakage (or a short) to  $V_{SS}$  or to  $V_{DD}$ . Damage to the output stage can cause substantial quiescent current. It may cause one driver to open, so it cannot both sink and source. In such a case there may be quiescent current in one state only. Current-sensing resistors in the  $V_{DD}$  leads, as recommended above, make it easy to track down such problems. When using this method, note that it's easy to be fooled, since the symptom of a damaged input may be nonzero quiescent current in a healthy chip that drives the bad one.

A damaged CMOS chip may work only at very low speeds (faulty driver), or only at very high frequencies (faulty input stage, no dc connection, capacitively coupled). A similar symptom may occur if you've forgotten to connect an input: The circuit may "work" at high speeds due to capacitively coupled edges (Fig. 14.42).

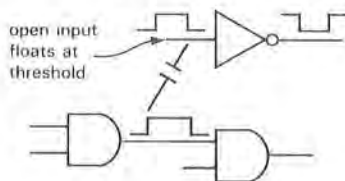


Figure 14.42

A pulldown will reveal this problem, by preventing the input from floating at the transition threshold. As mentioned in Section 8.35, forgetting to connect  $V_{DD}$  leads to weird symptoms, since chips can be powered through their logic inputs (via input-protection diodes); the power goes away, though, if all logic inputs go LOW simultaneously.

### 14.17 Micropower microprocessors and peripherals

There are available CMOS equivalents of many standard microprocessors. Thus, at first glance it would seem easy to design low-power microprocessor circuits. In fact, most of these CMOS microprocessors are simply pin-for-pin replacements of processors originally implemented in NMOS, and in many cases they do not have characteristics tailored to battery operation, e.g., the ability to run with a slow clock. Some aren't even low power—for example the 68020, which dissipates approximately 1 watt.

Since CMOS dissipation is usually proportional to clock frequency, the first question to ask is how much power is used when the clock is shut off. If the chip uses CMOS for its good electrical characteristics, but was not intended for very low power applications, it may include some non-CMOS circuits that result in significant static dissipation. Watch out, also, for chips that use the clock in mysterious ways: An example is a (discontinued) triple 16-bit counter that is microprocessor-compatible (six byte-wide three-state ports, plus control ports). It draws  $10\mu A$  with the clock stopped; however, the surprise is that it needs the clock present even when it isn't talking to the processor. The data sheet doesn't tell you that the clock is used to synchronize the inputs and that its frequency must be at least three times the maximum counting rate. Thus, you need to keep it supplied

with a fast clock for it to function as a counter, resulting in plenty of power dissipation.

The second question to ask is how low a clock frequency you can use while preserving reasonable operation. Some processors may have dynamic registers, mandating a fairly high minimum clock frequency. With a slow clock, some processors (particularly the “controller” types – see below) may respond *very* slowly; you may have a 10ms latency time from interrupts.

### Computer versus controller

Microprocessors generally divide into two categories, according to their intended application. Computation-oriented types have at least a 64K address space and use only a few clock cycles per instruction (to achieve high speed). They require external peripherals for ports, timers, converters, etc. Controller-oriented types, meant for dedicated instrument use, usually have a small address space (2K or 4K is typical) and use many clock cycles per instruction. On the other hand, they tend to have lots of parallel ports built in, many interrupt pins, and an internal clock generator. They often have timers, UARTs, and even A/D converters and nonvolatile RAM.

### Some examples

The 80C85, 80C86, and 80C88 are CMOS equivalents of popular NMOS computing processors. The 80C85 (Oki) draws 2mA with the clock stopped, and 10mA with a 6MHz clock. It may be unsuitable for micropower operation, unless you use power switching. The 80C86 and 80C88 (Harris, Intel) are better, with  $\approx 0.1$ mA static current drain (though it is specified as 0.5mA max). With these you can stop the clock to conserve power, retaining register and program counter contents.

There is a subtlety when restoring clock input to a stopped processor. To keep standby power low, you would like to shut off the crystal oscillator, not just gate its output (with a synchronizer circuit, of course; see Section 8.19). However, because of its high-resonant  $Q$ , a crystal oscillator cannot start up instantaneously, and an oscillator in the megahertz range typically takes 5–20ms to start up; a 32kHz oscillator can take up to a *second* ( $Q = 10^5$ ). Thus, to restore the clock signal, you can either wait for the oscillator to come up to speed before gating it through to the microprocessor, or you can hold the processor reset during that period. The first method is usually better, because you may not want to reboot each time you wake up (see “Power switching,” below, for a “warm-boot” trick). The 82C85 (Harris) is a low-power clock generator intended for the 80C86/88, with internal circuitry supporting stop-clock, stop-oscillator ( $\approx 25\mu$ A), and low-frequency operation, with hooks to the microprocessor for software control.

The 80C31/51 are examples of popular controller-type processors designed with special features for battery operation. These processors include up to 32 programmable I/O lines, a pair of 16-bit timers, 128 bytes of on-chip RAM (they can use ports as multiplexed buses to address 64K of external memory; the 80C51 has in addition 4K bytes of mask-programmed ROM), 5 interrupts, and a programmable serial port. They can be put into a low-power “idle” mode ( $< 1$ mA with 1MHz clock) in which interrupts, serial port, and timers continue to run while the processor clock is halted; all registers and on-chip RAM are preserved. They also can be put into a zero-current “power-down” mode, from which they can be revived only with a full reset, but which preserves on-chip RAM.

The 146805E2 is a similar controller, with the nice capability of waking up from



TABLE 14.9. CMOS MICROPROCESSOR ONE-CHIP CONTROLLERS

| Type      | Mfg <sup>a</sup> | Word size (bits) | Bus width (bits) | Adr bus (bits) | Pins  | NMOS copy? | Number cruncher? | Curr @ f <sub>clk</sub> =0? | Int RAM (words) | Int EPROM (words) | Timer? | Interrupts | Parallel port | Serial port? | Osc stop? | Comments                          |
|-----------|------------------|------------------|------------------|----------------|-------|------------|------------------|-----------------------------|-----------------|-------------------|--------|------------|---------------|--------------|-----------|-----------------------------------|
| 78C05     | NE               | 8                | 8                | 16             | 64    | -          | -                | 0                           | 128             | -                 | -      | 2          | -             | -            | -         | SPI port                          |
| 80C35     | IN               | 8                | 8                | 12             | 40    | -          | -                | 0                           | 64              | -                 | -      | -          | -             | -            | -         | 8048 family (8051 family better)  |
| 80C39     | IN               | 8                | 8                | 12             | 40    | -          | -                | 0                           | 128             | -                 | -      | -          | -             | -            | -         |                                   |
| 80C31     | IN               | 8                | 8                | 16             | 40    | -          | -                | 0                           | 128             | -                 | 2      | 2          | -             | -            | -         | int baud rate gen; 8051 family    |
| 70C42     | TI               | 8                | -                | int            | 40/44 | -          | -                | 0                           | 256             | 4k                | 2      | 6          | -             | -            | -         |                                   |
| 146805E3  | MO               | 8                | 8                | 16             | 40    | -          | -                | 0                           | 112             | -                 | -      | -          | -             | -            | -         |                                   |
| 68HC704P3 | MO               | 8                | -                | int            | 28    | -          | -                | 0                           | 128             | 2k                | -      | -          | -             | -            | -         | can use RC osc                    |
| 1468705G2 | MO               | 8                | -                | int            | 40    | -          | -                | 0                           | 112             | 2k                | -      | -          | -             | -            | -         | piggyback EPROM                   |
| HD63P05Y0 | HI               | 8                | -                | int            | 64    | -          | -                | 0                           | 256             | 8k                | -      | -          | -             | -            | -         | a complete family                 |
| HD6301V1  | HI               | 8                | 8                | 16             | 40    | -          | -                | 0                           | 128             | -                 | -      | 6          | -             | -            | -         | piggyback EPROM                   |
| HD63P01M1 | HI               | 8                | -                | int            | 40    | -          | -                | 0                           | 128             | 8k                | 1      | 6          | -             | -            | -         | lots of I/O                       |
| HD6305Y2  | HI               | 8                | 8                | 14             | 64    | -          | -                | 0                           | 256             | -                 | 1      | 5          | -             | -            | -         |                                   |
| HD6303R   | HI               | 8                | 8                | 16             | 40    | -          | -                | 0                           | 128             | -                 | 1      | 6          | -             | -            | -         | 3 volt version avail              |
| HD6303Y   | HI               | 8                | 8                | 16             | 64    | -          | -                | 0                           | 256             | -                 | 2      | 8          | -             | -            | -         | DMA                               |
| HD647180  | HI               | 8                | -                | int            | 84    | -          | -                | 0                           | 512             | 16k               | 2      | -          | -             | -            | -         |                                   |
| COP8788   | NS               | 8                | -                | int            | 40/44 | -          | -                | 0                           | 192             | 4k                | 2      | 13         | -             | -            | -         | has A/D; 8-bit internal Harv arch |
| 16C54     | GI               | 8/12             | -                | int            | 18    | -          | -                | 0                           | 32              | 512               | 1      | -          | 12            | -            | -         | 12-bit internal Harvard arch      |
| 16C55     | GI               | 8/12             | -                | int            | 28    | -          | -                | 0                           | 32              | 512               | 1      | 1          | 20            | -            | -         | " "                               |
| 87C51     | IN               | 8                | 8                | int            | 40    | -          | -                | 0                           | 128             | 4k                | 2      | 5          | 32            | -            | -         | CMOS vers of most pop controller  |
| 68HC11A8  | MO               | 8                | 8                | int            | 48/52 | -          | -                | 0                           | 256             | 512 <sup>b</sup>  | 4      | 2          | 28/32         | -            | -         | 6801 + 91 instr; A/D, SD1         |
| 80C196    | IN               | 16               | 8/16             | 0/16           | 48/68 | -          | -                | 0                           | 232             | 8k                | 2      | 8          | -             | -            | -         | twice performance of 8096         |
| 16003     | NS               | 16               | 16               | 16             | 68    | -          | -                | 0                           | 256             | -                 | 8      | 8          | -             | -            | -         | complete family                   |
| 16084MH   | NS               | 16               | -                | int            | 68    | -          | -                | 0                           | 256             | 8k                | 8      | 8          | -             | -            | -         | A/D                               |
| 78P312    | NE               | 16               | 8/16             | int            | 64    | -          | -                | -                           | 256             | 8k                | 4      | 15         | 32            | -            | -         | new family, unique instr set      |

(a) see footnote to Table 4.1. (b) EEPROM.

zero-power “stop” mode via an interrupt. There’s internal circuitry to wait for oscillator start-up before resuming processing. The specifications say that start-up time is typically 30ms (300ms max) at  $f_{\text{clk}} = 1\text{MHz}$ ,  $V_{DD} = 3\text{ volts}$ . This is obviously unsuitable for an application where you must wake up frequently, say 10 times per second, but it would be fine in an application where you wake up once each minute. The 146805 series includes versions (specified by suffix) with different port and memory configurations, mask ROM, etc. We’ll use it for a design example in the next section.

Table 14.9 lists most of the interesting low-power microprocessors available as this edition went to press.

### Power switching

You can make any microprocessor operate at low average power, of course, if you switch the power on with a low duty cycle. This is actually easier with NMOS than with CMOS, because inputs such as  $WR'$  can be held high during power-down (with CMOS, that usage would power the chip, through the input-protection diodes!), preventing spurious write cycles, etc. Thus, with CMOS you must use external logic to accomplish orderly shut-down. In either case you’ve got the clock start-up problem mentioned above, which you can solve with external delaying logic, or with a chip like the 82C85.

You usually don’t want to go through the full “cold” reboot software on each restart. The best way to handle this is to let the CPU read a “power-on flag” flip-flop (CMOS, continuously on) at each restart, doing a cold boot the first time only, after which it sets the flip-flop.

With power switching of NMOS processors, you have to store edge-triggered interrupt requests in external CMOS logic during processor hibernation, servicing it when the processor is next restarted.

You may need to use the same trick with some CMOS processors, those that lose edge-triggered interrupts during idle or stop modes.

When power switching NMOS processors, which draw typically 100mA or more of supply current, be sure to use a pretty hefty MOSFET, with  $R_{ON} < 3\text{ ohms}$  at 5 volts of gate drive.

### CMOS peripherals

Many low-power peripheral chips are simply CMOS copies of NMOS parts, for example the 81C55 and 82C55 parallel ports. The data sheets are usually almost exact copies of the original NMOS data sheets, with a few changes. As such, the data sheets are often incorrect! For example, they may specify 2mA output sinking at 0.4 volt and 100 $\mu\text{A}$  sourcing at 2.4 volts, whereas their  $p$ -channel drivers can actually source 2mA, particularly at 2.4 volts. Input thresholds may also be misleading.

Another thing to worry about with CMOS is the disposition of floating inputs from a three-state bus. Harris and Intel have a “bus-hold” circuit that introduces a little positive feedback at the inputs to prevent class A current from floating inputs.

Watch out for CMOS peripherals that draw “quiescent” current because they have a clock. For example, UARTs like the 65C51 and 82C52 draw about 2mA at their recommended oscillator frequency (1.84MHz @ 1.4mA/MHz for the 65C51). You could imagine shutting off the oscillator, but then the UART can’t receive data, for example a command to turn on! Other CMOS peripherals with quiescent currents of 1–5mA are A/D converters, modems, video drivers, EEPROMs, and keyboard encoders. A complex system with several of these devices may have a quiescent current of 25–50mA, giving a 9 volt battery life of only 10 hours. If that’s OK, fine; otherwise, you have to power-switch. But be careful – inputs and outputs may

misbehave; for example, the A/D's three-state bus drivers may pull LOW when the A/D is not powered (in which case use a separate CMOS three-state driver).

RS-232 drivers have traditionally been high-power devices: The classic 1488 quad driver draws  $\pm 20\text{mA}$  quiescent current, not including load current, and the 1489 quad receiver draws  $15\text{mA}$ . Some recent RS-232 chips run at low power. Here are a few good choices:

**Motorola MC145406.** This is a triple CMOS driver/receiver that will work with supply voltages from  $\pm 5$  volts to  $\pm 13$  volts, with less than  $15\text{mW}$  dissipation over that range. CMOS output drivers give rail-to-rail swings, so you can get RS-232 swings even with  $\pm 5$  volts supplies. Various tricks were used in the design to permit CMOS to work with a total supply voltage up to 26 volts, and to permit receiver input swings 20 volts beyond the rails.

**LT1032.** This is a quad bipolar RS-232 driver with  $\pm 5$  to  $\pm 15$  volt supply range, and quiescent current of  $0.5\text{mA}$ . It can be shut down (zero current) with a control pin; during shutdown, the outputs go to a high-impedance state.

**LT1039.** This is a triple bipolar driver/receiver with  $\pm 5$  to  $\pm 15$  volt supply range, and  $4\text{mA}$  quiescent current. Like the LT1032, it has a shut-down pin. It also has a control line that allows one receiver to remain on while the rest of the chip is shut down; you could use that to power up the rest of the chip when something is received. The outputs go to a high-impedance state during shutdown.

**MAX230-239/ICL232 series; LT1080/1.** These devices from Maxim, Intersil, and Linear Technology are dual driver/receivers with on-chip voltage converters, so they run from a single  $+5$  volt supply, producing  $\pm 9$  volt output swings. All except the MAX233 and 235 (which have

built-in capacitors) require four external tantalum capacitors for the voltage converters; their  $\pm 9$  volt outputs are available externally to power low-current loads. Quiescent current is  $5\text{mA}$ . The voltage-conversion portions of these chips are available separately as the MAX680 or LT1026, single  $+5$  volt to dual  $\pm 10$  volt voltage converters; these could be used to power any of the other RS-232 chips above.

**DS14C88/89.** National's CMOS reworkings of the similarly named bipolar classics. The 14C88 driver works with supply voltages from  $\pm 4.5$  volts to  $\pm 12$  volts, producing the usual CMOS rail-to-rail swings. With  $\pm 5$  volt supplies the quad driver draws  $30\mu\text{A}$  max (unloaded), while the receiver draws  $0.9\text{mA}$  max from its single  $+5$  volt supply. (National also makes CMOS RS-422 chips (DS26C31-32).

#### 14.18 Microprocessor design example: degree-day logger

Let's bring all these ideas together with an example. We'll design a small battery-powered data logger whose purpose is to monitor the ambient temperature once per minute, storing the "degree-day" averages in RAM for subsequent readout via serial communication. You might put such an instrument in a remote location, paying it twice-yearly visits during which you read out its data into a portable computer. Or you might prefer to "harvest" the loggers, then read them out when you've got them back home.

The instrument will operate on three alkaline C cells, with battery life of at least a year. To keep power consumption low, we'll use CMOS peripherals and a CMOS controller-type microprocessor with built-in shut-down modes. We'll use power switching of the CPU and front-end circuitry during the momentary data-collection intervals, with wake-up provided by a low-power calendar chip.

Since the serial port will be used only occasionally, we'll do power switching here, too. Our particular circuit is by no means unique; along the way we'll discuss alternative circuit solutions.

### CPU

Figure 14.43 shows our design. We began by choosing the Motorola MC146805 CMOS controller, which is specified for operation down to 3 volts, and includes on-board circuitry to go into WAIT mode (low power, oscillator and timer running) or STOP mode (zero power, oscillator stopped; wake-up via interrupt or reset only). The -E2 suffix version uses external ROM and RAM, but includes 112 bytes of on-board RAM.

When run from 5 volts, the CPU typically uses 7mA running (5MHz clock), 1mA in WAIT mode, and 5 $\mu$ A in STOP mode. Since we need to gather data for only a few milliseconds once per minute, and a wake-up from STOP mode (due to crystal oscillator start-up) takes typically 30ms, we minimize power consumption by using an externally generated interrupt to wake up for each measurement. The alternative – using CPU timer-generated interrupts from WAIT mode – imposes an average CPU current drain of at least 1mA, corresponding to a C-cell battery life of only half a year. That could, of course, be extended to a year by using D cells; another solution would be to run at lower oscillator frequency (say 1MHz), where the WAIT-mode current is significantly lower. Still another possibility is to run at 3 volts, where the WAIT-mode current drain is about 150 $\mu$ A with a 1MHz clock. Any of these solutions is perfectly good. In this example, we'll stick with the power switching, because it illustrates additional techniques. It also provides convenient timekeeping via the calendar chip.

### Calendar clock

For the calendar clock we needed a chip

that not only can keep time at low current drain (all calendar chips do that) but also can make interrupts while in low-power mode. Since the primary use of calendar chips is for ac-line-operated computers, where there is plenty of power available when the CPU is running, many chips cannot interrupt in low-power (battery-backup) mode. We first looked at the ICM7170, a nice calendar chip from Intersil; it can interrupt in low-power mode, but it has an awkward power scheme for single-battery operation. The ever-popular MM58274 from National doesn't interrupt during backup. We finally settled on the Motorola MC146818, a popular part that is made by at least two other manufacturers and is intended for use with chips like the MC146805 CPU. It can be left running at full supply voltage, and it maintains low current drain (50 $\mu$ A max at 32kHz, external oscillator) while not enabled.

Motorola likes memory-mapped I/O, and their MC146805 is no exception. So you don't have the I/O strobes favored by Intel; instead, you decode some portion of memory space as "I/O space." With only a few I/O devices in the system you can be pretty sloppy about I/O decoding, as we explained in Section 10.06. In this case we put the UART at the bottom of external memory (the CPU chip monopolizes the bottom 80H bytes for on-chip memory and ports), and the calendar clock at 80H, using the same 'HC139 decoder that enables memory (see below).

### Memory

For EPROM, we've used a standard 27C64, an 8K $\times$ 8 part of which only the bottom quarter is addressable with the 11 address lines we've connected (we've use the top of address space for I/O). A smaller ROM would be fine, but memory manufacturers have generally been discontinuing the smaller memories as they are able to make larger ones. The 27C64 specifies

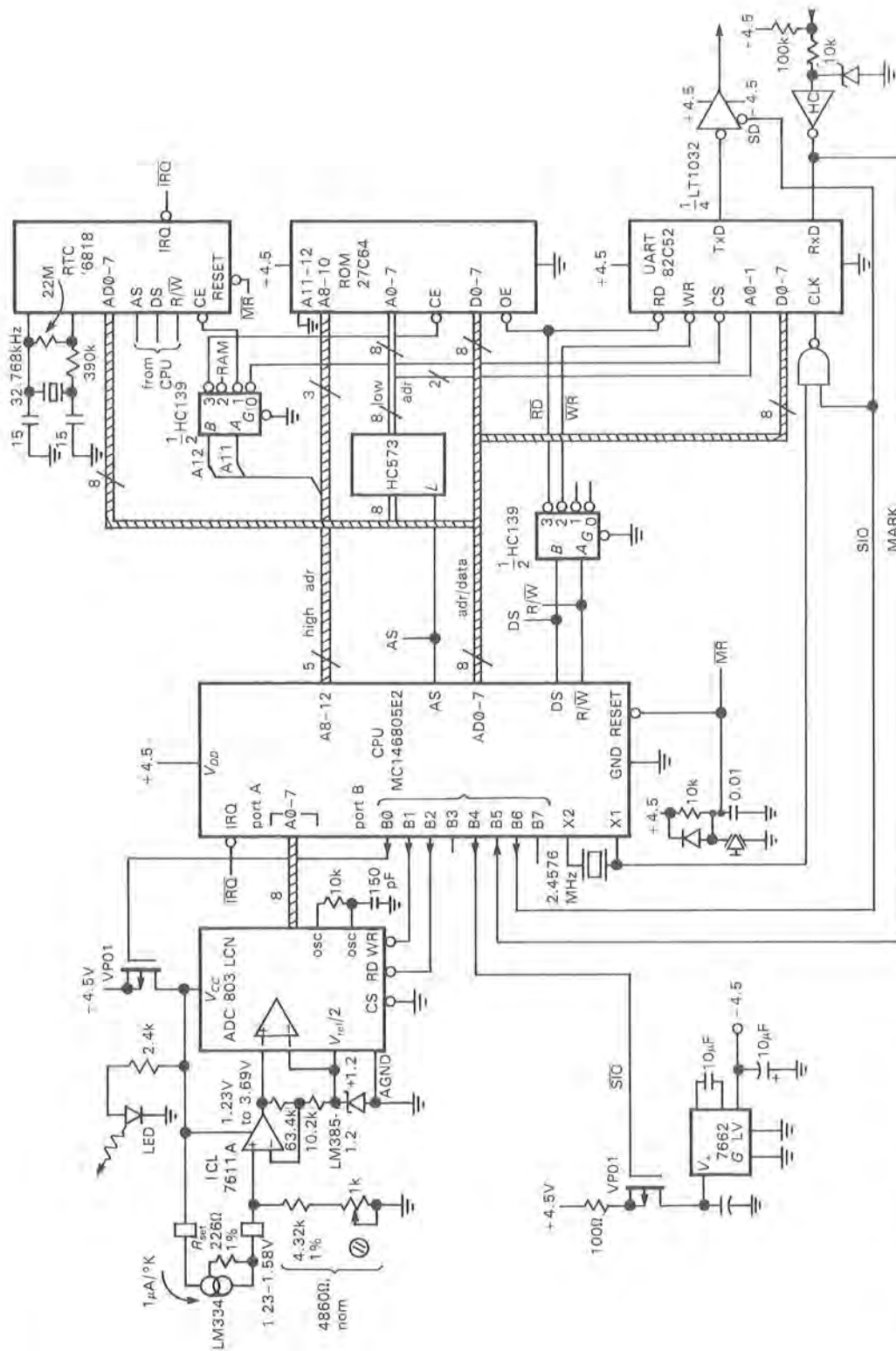


Figure 14.43. General-purpose low-power microprocessor example.



$I_{DD} = 100\mu\text{A}$ , max, when deselected; in fact, the quiescent drain is likely to be less than  $10\mu\text{A}$ . Note the low-order address latching from the multiplexed CPU bus, and the conversion of the Motorola pair (R/W', DS) to the Intel pair (RD', WR'), conveniently done with half of an 'HC139 decoder (the other half does address decoding).

The external CMOS RAM (optional, not shown) sits below ROM in address space (Fig. 14.44) and ties onto the same lines, with the exception of the address decoding. Once again, a smaller RAM would be fine, but the chip manufacturers haven't cooperated.

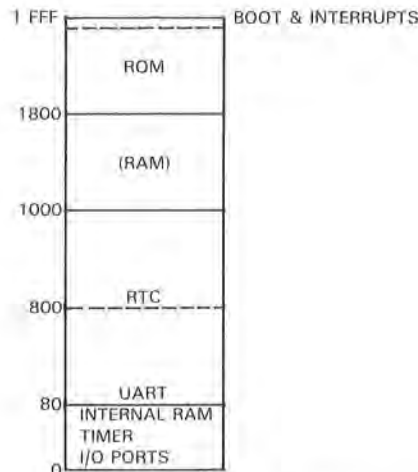


Figure 14.44. Memory map for microprocessor example.

### Serial communication

For the serial communication (Section 10.19) we need a UART and bipolarity RS-232 driver/receiver. Since the serial port will be used only occasionally (during readout and initialization), it can be left in a power-down mode, with the CPU checking to see if something has been connected to it as part of its periodic once-per-minute duties. One possibility, then, is to use a conventional NMOS UART (e.g., an 8251), with power switching controlled by an output port bit of the CPU.

This scheme is fine, although you have to be careful that the powered-down UART doesn't load the bus (or get powered by it, as a CMOS chip would): A good way to solve this problem is to use HC three-state bus buffers between the bus and the power-switched UART, putting them into the high-impedance state during power-down (Fig. 14.45).

In this example we took a simpler solution, namely the use of a CMOS UART that we leave powered at all times, switching the *oscillator* on only when the serial port is used; the resulting quiescent current, though not specified on the data sheet, is typically less than  $20\mu\text{A}$ . For simplicity we also leave most of the RS-232 interface powered. The driver is one section of a low-power quad with a shutdown control; the specified quiescent current in the shutdown state is  $1\mu\text{A}$  (typ),  $10\mu\text{A}$  (max). The negative supply is provided by a 7662 voltage inverter, power-switched under control of one of the CPU's port bits; although the specifications give  $I_Q = 20\mu\text{A}$  (typ), the *maximum*  $I_Q$  is  $150\mu\text{A}$ , enough reason to power-switch a portion of the circuit that might get used 10 minutes each year! Note the 100 ohm current-limiting resistor – the load looks like a short when power is first applied. For the RS-232 receiver, we've cheated by using an HC logic inverter with Schottky clamp and current limiter for the bipolarity input.

By leaving the receiver powered, we can detect when someone has connected to the serial port, since the RS-232 resting state ("mark," at least 3V negative) brings the inverter's output HIGH (note the pullup at the input). That is why the inverter's output, besides driving the UART, is tied to a CPU input port. Of course, the CPU could recognize that fact only during the short ( $<1\text{s}$ ) interval each minute that it is awake. Thus, in practice, the serial-port user would connect to the port, then practice some patience until things begin to happen.

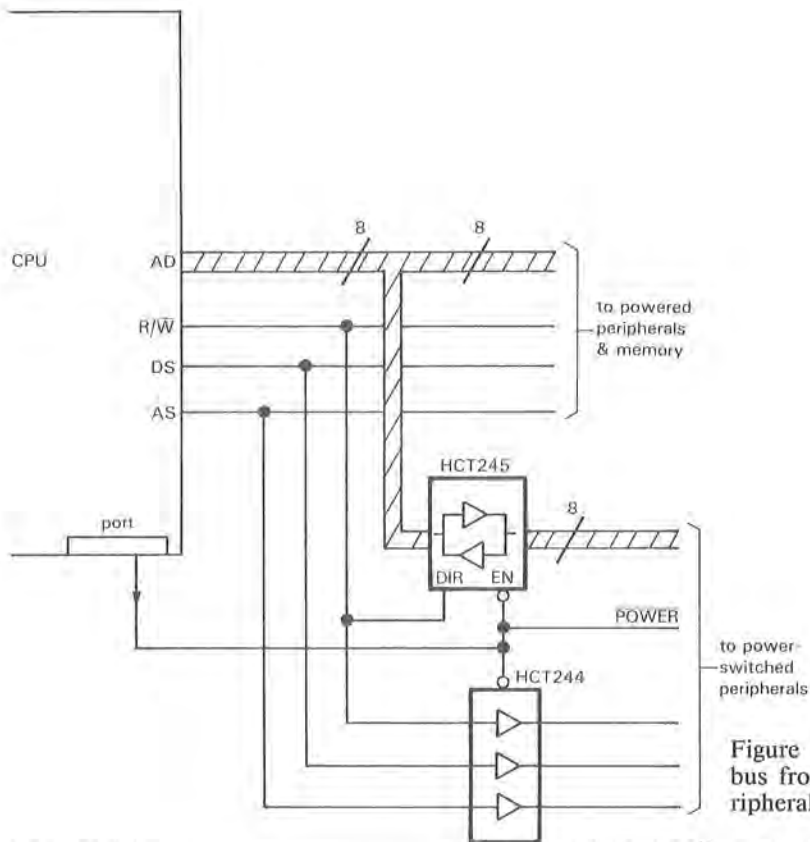


Figure 14.45. Decoupling data bus from a power-switched peripheral.

### Front end

We chose the LM334 temperature-sensing current source as our sensor. The 2-terminal current is proportional to absolute temperature, set by a single resistor  $R_{\text{set}}$ :  $I(\mu\text{A}) = 227 T/R_{\text{set}}$ , where  $T$  is in degrees Kelvin and  $R_{\text{set}}$  is in ohms. It has a voltage compliance from 1 to 40 volts, and initial accuracy of 6%. For the A/D, we picked the simple AD803, an inexpensive (less than \$5) successive-approximation converter with internal clock generator, three-state outputs, and single +5 volt supply. It has a convenient differential input, with circuitry to set the full-scale range. Since it draws 1mA, we've used power switching, controlled by a CPU port bit.

In this implementation we chose  $R_{\text{set}} = 226$  ohms, i.e.,  $1\mu\text{A}/^\circ\text{K}$ . A reasonable temperature range is  $-20^\circ\text{C}$  to  $+50^\circ\text{C}$ , with

corresponding current range  $254\mu\text{A}$  to  $324\mu\text{A}$ . This has to be matched to the A/D input range. This particular A/D has a full-scale analog voltage span of  $2V_{\text{ref}}$ , when an externally supplied reference voltage  $V_{\text{ref}}$  is used; furthermore, the differential input structure lets you add an offset to the input. The simplest configuration is as shown, with an offset equal to  $V_{\text{ref}}$ ; i.e., the analog input range is  $V_{\text{ref}}$  to  $3V_{\text{ref}}$ . With our +4.5 volt supply, the obvious reference is a 1.23 volt bandgap reference, say the micropower LM385-1.2. That makes the analog input span 1.23 to 3.69 volts. The rest is simple, namely choosing a load resistor to put the low end of the sensor output at 1.23 volts, then adding a stage of differential dc gain to put the high end of the sensor output at 3.69 volts. Accordingly, the load resistor for the sensor should be 4.84k, followed by a stage of voltage amplification of  $G_V = 7.26$ , as shown.

The  $\pm 10\%$  trim adjustment is needed on the load resistor to accommodate initial errors (LM334, 6%; LM385, 2%; op-amp offset, 1%). Note the unusual piggyback configuration at the input, which makes the zero point of the dc amplifier equal to the reference, while simultaneously biasing the reference.

## EXERCISE 14.1

Check our arithmetic by calculating for yourself the temperature range produced by the resistor values in Figure 14.43.

Note that we've interfaced the A/D via parallel port lines, rather than the more usual approach of using the CPU's data bus. We did this because the unpowered A/D would otherwise load the bus. Speed is unimportant in this application (especially since the CPU oscillator start-up can take as long as 250ms), and the port lines are available and unused.

**Power consumption**

Table 14.10 shows how the current drain is proportioned, in each of the three operating states. Note the large spread between "typical" and "maximum" quiescent currents. If your ICs are "typical," the average current drain (assuming 500ms of wake-up time per minute) is  $168\mu\text{A}$ , or 3 years with alkaline C cells (4500mAh). The worst-case average drain is  $680\mu\text{A}$  (9 months), marginal for an unattended data logger. Of course, "worst case" means that *every* IC's quiescent current is at the specified limit. There are two solutions: (a) Use much bigger batteries, to guarantee satisfactory worst-case battery life; or (b) pre-screen (by measuring  $I_Q$ ) the ICs that dominate battery life, in this case the CPU. Most of the time you can probably get away with method (c): Live dangerously; most ICs have much lower leakage than the worst-case spec.

TABLE 14.10.  
TEMPERATURE LOGGER CURRENT DRAIN<sup>a</sup>

| Device              | Idle                     |                          | Data in                  |                          | Serial comm              |                          |
|---------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|
|                     | typ<br>( $\mu\text{A}$ ) | max<br>( $\mu\text{A}$ ) | typ<br>( $\mu\text{A}$ ) | max<br>( $\mu\text{A}$ ) | typ<br>( $\mu\text{A}$ ) | max<br>( $\mu\text{A}$ ) |
| CPU                 | 10                       | 175                      | 4000                     | 7500                     | 4000                     | 7500                     |
| ROM <sup>b</sup>    | 10                       | 100                      | 2500                     | 5000                     | 2500                     | 5000                     |
| (RAM)               | 2                        | 10                       | 2500                     | 5000                     | 2500                     | 5000                     |
| Sensor ckt          | 0                        | 0                        | 350                      | 350                      | 0                        | 0                        |
| A/D                 | 0                        | 0                        | 1100                     | 1800                     | 0                        | 0                        |
| LED                 | 0                        | 0                        | 1000                     | 1000                     | 0                        | 0                        |
| UART                | 20                       | 100 <sup>c</sup>         | 20                       | 100 <sup>c</sup>         | 2500                     | 3000                     |
| RS-232 <sup>d</sup> | 1                        | 10                       | 1                        | 10                       | 1600                     | 2000                     |
| 7662                | 0                        | 0                        | 0                        | 0                        | 20                       | 150                      |
| Discrete            | 0                        | 10                       | 0                        | 10                       | 0                        | 10                       |
| RTC                 | 25                       | 100                      | 25                       | 100                      | 25                       | 100                      |
| Totals              | 68 $\mu\text{A}$         | 505 $\mu\text{A}$        | 12mA                     | 21mA                     | 13mA                     | 23mA                     |

(a)  $V_{\text{batt}} = 4.5\text{V}$ ,  $f_{\text{CPU}} = 2.5\text{MHz}$ . (b) assuming 50% ROM, 50% RAM accesses. (c)  $I_Q(\text{max})$  not specified for 82C52; value for 82C51 used. (d) assuming 50% marking, and  $5\text{k}\Omega$  load impedance.

**Programming**

The ROM coding is straightforward and follows the general pattern of firmware coding that we talked about in Chapter 11. There are a few additional wrinkles, however.

When power-switching, be sure to allow time for the powered device to get going. It may include a crystal oscillator, in which case the required delay can be tens of milliseconds. The 7662 voltage inverter in this circuit has large capacitors and needs a few milliseconds to come up to voltage. If the switched device is connected to port bits (as the A/D has been), the port bits should either be set LOW or programmed as inputs before the device is switched off. If three-state drivers are used to isolate the bus, be sure to put them into the high impedance state *before* the device is powered down.

#### EXERCISE 14.2

Make a flow chart for the degree-day logger. Do the right things to the parallel port bits before going to sleep. Don't forget to initialize the calendar clock and UART. Be sure to check for serial port connection during each wake-up.

#### **Design alternatives**

As we mentioned at the outset, by operating at 3 volts instead of 4.5 volts (since battery voltage drops throughout its life, this is best done by regulating down from +4.5V with a micropower low-dropout regulator like the LP2951 or ICL7663), the CPU could be kept in WAIT mode, rather than STOP mode. In wait mode (200 $\mu$ A, max, at 3V and 1MHz), the internal oscillator runs, and supports both interrupts and the internal timer function. Thus, the external calendar clock could be replaced by a simple low-power 32kHz oscillator and divider (e.g., the inexpensive mini-DIP MM5368) to make 1pps interrupts; the internal timer would then wake up the CPU at a programmable interval, with everything else as before. Note that in the present design, most of the battery's energy is used by the calendar clock, a relatively expensive chip of excessive complexity.

#### EXERCISE 14.3

Draw a revised schematic for a temperature logger that sleeps in a WAIT state.

There are some recent chip offerings that allow you to simplify the serial port RS-232 circuitry. The LT1080 and the MAX230 series of RS-232 driver/receivers include on-chip voltage inverters, and, on some models, "shut-down" controls with  $I_Q = 1\mu$ A, typ (10 $\mu$ A, max). We could thus replace the 7662, its discrete power-switching circuitry, and the LT1032 with a single MAX235, which even includes built-in flying capacitors. Unfortunately, the latter's receiver sections go dead during shut-down, so we cannot use it to replace also our 'HC04 receiver. There are RS-232

driver/receivers available that do keep one receiver running ("hibernation"?) during micropower shut-down, for example the LT1039; however, this chip does not provide a solution here, because it requires continuous bipolarity power supplies during hibernation, which would require running the 7662 continuously.

While on the subject of serial port alternatives, it's worth noting that you can actually eliminate the UART, by using a pair of CPU parallel port bits for transmit and receive. To implement such a "software UART" you have to write software to generate and receive serial bit streams. The usual procedure is to use an internal CPU timer function, set to the appropriate baud rate. The transmit function is relatively straightforward, since you just generate 1's and 0's at each timer tick. The receive function turns out to be more of a challenge, since you have to sample the incoming port bit at a much higher clock rate (typically 8 times the baud rate) in order to sample near the middle of the bit cells. In spite of the programming problems, software UARTs are often used in small systems, because they eliminate a large chip.

Instead of the direct RS-232 outputs, the design could incorporate a power-switched modem for connection to the telephone line. The National 74HC943 would be good here, with 8mA quiescent current and single +5 volt supply. A passive "ring-detect" circuit (Fig. 14.46) would then substitute for the negative "mark-detect" circuit of Figure 14.43. It would be best to have the ring-detect trigger an interrupt, since no one likes waiting a minute for a phone to be answered. In the present design, the presence of a marking level (−3V or more) could also be used to trigger an interrupt.

#### EXERCISE 14.4

Show how to make a marking level trigger an interrupt. Be sure to provide a way to clear the interrupt via software.

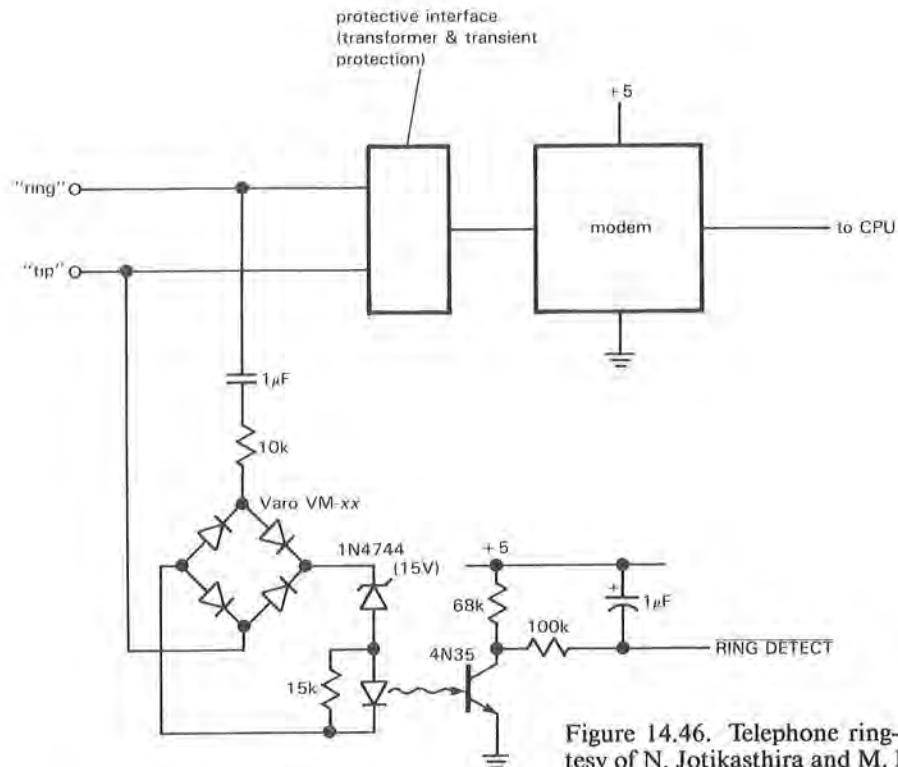


Figure 14.46. Telephone ring-detection. (Courtesy of N. Jotikasthira and M. Bain, Infinet, Inc.)

The A/D front end could tie into the CPU bus (instead of a port), providing that CMOS three-state drivers are used to isolate the powered-down device (Fig. 14.45). The same trick could be used to minimize power drain if a high-performance UART (e.g., an 85C30, as in Fig. 11.13) were substituted for the simple 85C52 used here.

A bit of philosophy: It's always tempting to try to squeeze the last ounce out of micropower design, by making diabolical use of the lowest-power devices, power switching, etc. For this circuit you could probably succeed in a design that used AA cells rather than C cells. But the extra

effort (and cost) would not be worth it, because a reduction of 20% in size and weight wouldn't matter in this application. In fact, it would probably make more sense to simplify the design still further, for example by powering the 7662 from the same switched source as the A/D front end, or even leaving it powered continuously.

## SELF-EXPLANATORY CIRCUITS

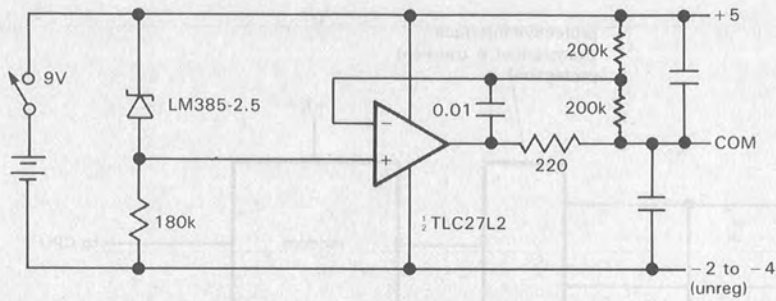
### 14.19 Circuit ideas

Figure 14.47 presents some low-power circuit ideas.

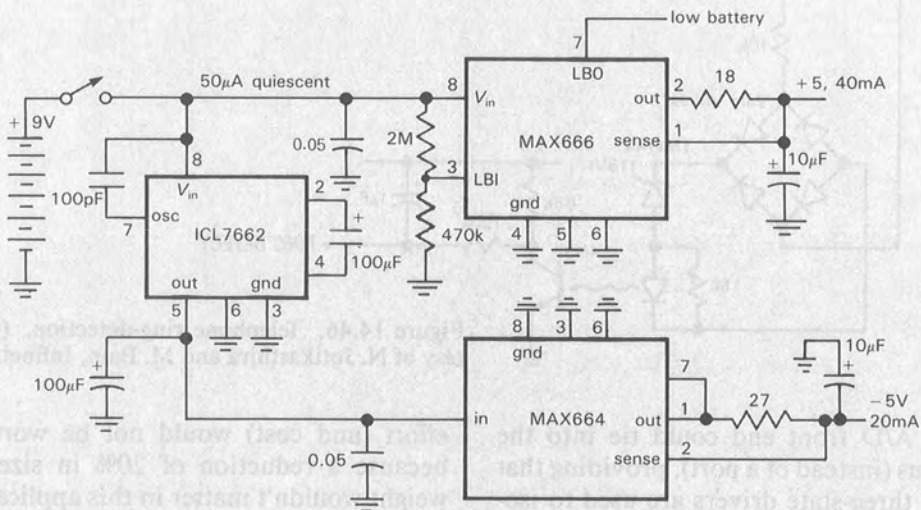




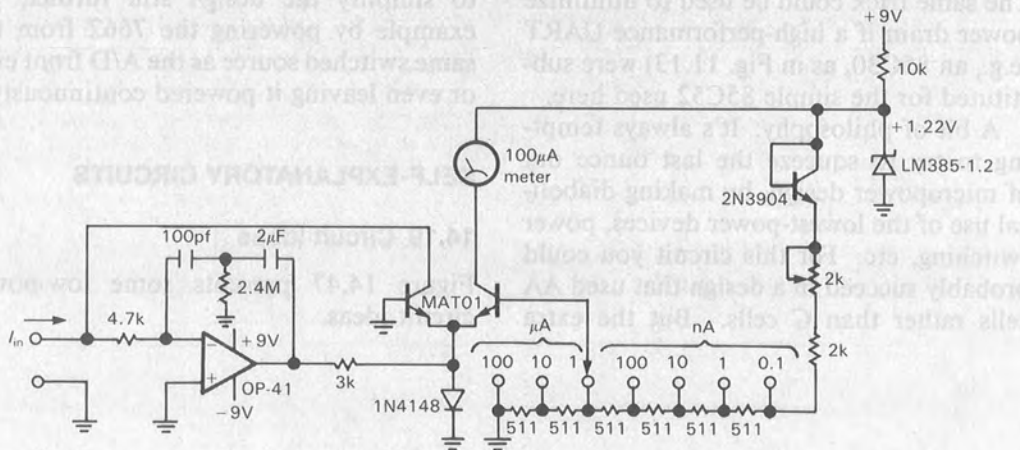
## Circuit ideas



A. generating +5V and -2 to -4V from a single 9V battery



B.  $\pm 5$  volt supply from single 9V battery



C. wide-range nano-ammeter (from PMI App. Note 106)

# MEASUREMENTS AND SIGNAL PROCESSING

CHAPTER

15

## OVERVIEW

Perhaps the most exciting (and most useful) area of electronics involves the gathering and manipulation of data from an industrial process or a scientific experiment. Generally speaking, *transducers* (devices that convert some physical quantity, such as temperature or light level, to a voltage or some other electrical quantity) are used to generate signals that can be manipulated by electronic circuits, quantified by analog-to-digital converters, and logged and analyzed by computers. If the signal you're looking for is masked by noise or interference, powerful "bandwidth-narrowing" techniques such as lock-in detection, signal averaging, multichannel scaling, and correlation and spectral analysis can magically retrieve the sought-after signal. Finally, the results of such physical measurements can be used to control the experiment or process itself, with "on-line" control usually provided by a small computer or micro-processor dedicated to the task. The recent development of powerful and inexpensive microprocessors and support chips has brought about

an explosion in the use of electronics to control and log processes that would not have seemed likely candidates only a decade ago.

In this chapter we will begin with a sampling of quantities that can be measured and the transducers that are normally used for the job. In this area there is plenty of room for ingenuity, and the catalog of transducers we will describe should therefore be considered representative, not exhaustive. We will go into some detail describing the particular problems some of these measurement transducers present and the circuit solutions you might use with them. We will try to cover the most common difficulties, dealing with ultrahigh source impedances (hundreds of megohms in the case of microelectrodes or ion-specific probes), low-level low-impedance transducers (e.g., thermocouples, strain gauges, magnetic pickups), high-impedance ac sensors (capacitance transducers), and others.

The chapter will continue with a look at precision standards (standards of frequency and time, as well as voltage and resistance) and some of the techniques of

precision measurements. We will then describe in some detail the whole business of bandwidth-narrowing, “pulling the signal from the noise.” These techniques are extremely powerful, and they are mysterious to the uninitiated. Finally, we will conclude the chapter with a brief look at spectrum analysis and Fourier techniques. Readers interested primarily in electronic circuit design may wish to skip this chapter.

## MEASUREMENT TRANSDUCERS

In some situations the quantity you want to measure is itself an electrical quantity. Examples might be nerve impulses (voltage), seawater conductivity (resistance), charged-particle fluxes (current), etc. In these cases measurement techniques tend to be relatively straightforward, with most of the difficulties centering around the kind of collection electrode to use and how to handle the signals once they’ve been collected. You might encounter very high impedances (e.g., with microelectrodes) or very small signals (e.g., a current generated from radioactive decay).

More often, a “transducer” of some sort is necessary to convert some physical quantity to a quantifiable electrical quantity. Examples are measurements of temperature, light level, magnetic field, strain, acceleration, sound intensity, etc. In the following sections we will take a look at some of the more common input transducers to give an idea of what can be measured and how accurately. We will go into somewhat greater detail when describing the more common measurements, such as heat and light, but in a book of this scope we can cover only a fraction of the measurement possibilities.

### 15.01 Temperature

Temperature transducers illustrate a nice variety of performance trade-offs. Temper-

ature range, accuracy, repeatability, conformity to a universal curve, size, and price are all involved.

#### Thermocouples

A junction between two dissimilar metals generates a small voltage (with low source impedance!), typically in the millivolt range, with coefficients of about  $50\mu\text{V}/^\circ\text{C}$ . This junction is called a *thermocouple*, and it is useful for measuring temperatures over a broad range. By using various pairs of alloys it is possible to span temperatures from  $-270^\circ\text{C}$  to  $+2500^\circ\text{C}$  with reasonable accuracy ( $0.5$ – $2^\circ\text{C}$ ). The thermoelectric properties of different alloys are well known, so thermocouple probes in different formats (rods, washers, armored probes, etc.) made from the same alloys can be interchanged without affecting calibration.

The classic thermocouple circuit is shown in Figure 15.1. The particular choice of metals in this figure constitutes what is known as a type J thermocouple (look at Table 15.1 for a listing of the standard choices and their properties). Each couple is made by welding the two dissimilar metals together to make a small junction. (People have been known to get away with twisting the wires together, but not for very long!) The reference junction is absolutely necessary, since otherwise you wind up with additional dissimilar thermocouples where the dissimilar metals join the metal terminals. Those extra thermoelectric voltages produced at uncontrolled places in the circuit would result in erratic and inaccurate results. Even with a pair of thermocouple junctions, you still have thermocouples formed where the leads join the metal terminals. However, this seldom causes problems, since those junctions are at the same temperature.

The thermocouple circuit gives you a voltage that depends on the temperatures of both junctions. Roughly speaking, it is proportional to the *difference* between

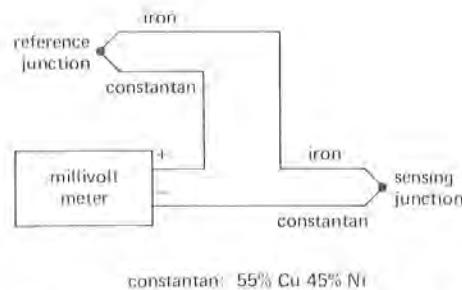


Figure 15.1. Classic thermocouple circuit.

the two junction temperatures. What you actually want is the temperature at the sensing junction. There are two ways to handle the problem of the reference: (a) Classically, you put the reference junction at a fixed temperature, usually  $0^{\circ}\text{C}$ . They used to use ice baths, and you still can, but you can also buy nice little stabilized cold boxes to do the same job. If you are measuring very high temperatures, you may not even care about small errors caused by having the reference junction at “room” temperature. (b) A more modern technique is to build a compensation circuit that corrects for the difference caused by

having the reference junction at a temperature other than  $0^{\circ}\text{C}$ .

Figure 15.2 shows how this is done. The basic idea is to use a temperature-sensing chip and circuitry that adds in a voltage that makes up for the difference between the actual reference junction temperature and the standard  $0^{\circ}\text{C}$ . The AD590 (see the subsequent section on IC temperature sensors) produces an output current in microamps equal to the temperature in degrees Kelvin.  $R_1$  is chosen according to the thermoelectric coefficient, in this case converting  $1\mu\text{A}/^{\circ}\text{C}$  to  $51.5\mu\text{V}/^{\circ}\text{C}$  (see Table 15.1), and the AD580 3-terminal reference (in combination with  $R_2$  and  $R_3$ ) is used to subtract the AD590 offset of  $273\mu\text{A}$  at  $0^{\circ}\text{C}$  ( $273.15^{\circ}\text{K}$ ). Thus, there is no correction made when the reference junction is at  $0^{\circ}\text{C}$ , and  $51.5\mu\text{V}/^{\circ}\text{C}$  (the thermoelectric coefficient of a type J junction at room temperature) is added to the net output voltage of the pair of junctions when the reference junction is at some other temperature.

The metering circuit deserves a few words of comment. The circuit problems

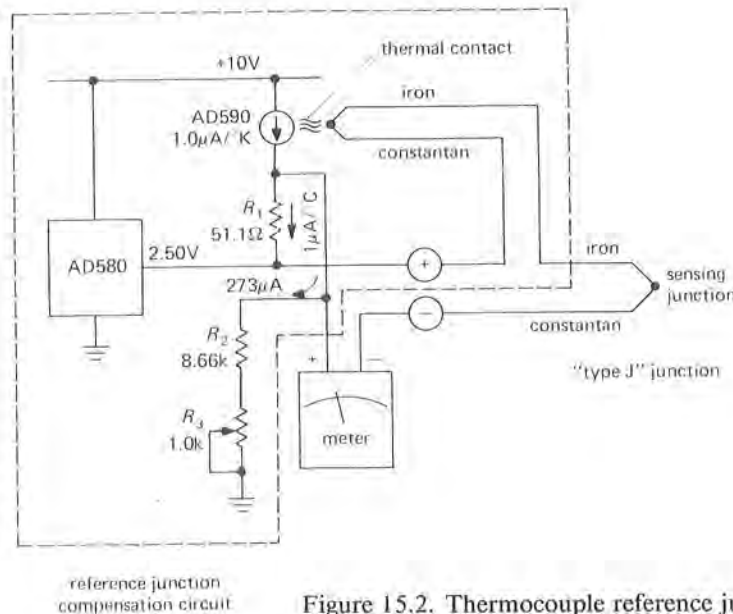


Figure 15.2. Thermocouple reference junction compensation.

TABLE 15.1. THERMOCOUPLES

| Type | Alloy                                            | Max temp <sup>a</sup><br>(°C) | Tempco<br>@20°C<br>( $\mu\text{V}/^\circ\text{C}$ ) | Output voltage <sup>b</sup> |               |                | 30 gauge<br>lead<br>resistance <sup>c</sup><br>( $\Omega$ ) |
|------|--------------------------------------------------|-------------------------------|-----------------------------------------------------|-----------------------------|---------------|----------------|-------------------------------------------------------------|
|      |                                                  |                               |                                                     | 100°C<br>(mV)               | 400°C<br>(mV) | 1000°C<br>(mV) |                                                             |
| J    | {Iron<br>Constantan <sup>d</sup>                 | 760                           | 51.45                                               | 5.268                       | 21.846        | —              | 3.6                                                         |
| K    | {Chromel <sup>e</sup><br>Alumel <sup>f</sup>     | 1370                          | 40.28                                               | 4.095                       | 16.395        | 41.269         | 6.0                                                         |
| T    | {Copper<br>Constantan <sup>d</sup>               | 400                           | 40.28                                               | 4.277                       | 20.869        | —              | 3.0                                                         |
| E    | {Chromel <sup>e</sup><br>Constantan <sup>d</sup> | 1000                          | 60.48                                               | 6.317                       | 28.943        | 76.358         | 7.2                                                         |
| S    | {Platinum<br>90%Pt–10%Rh                         | 1750                          | 5.88                                                | 0.645                       | 3.260         | 9.585          | 1.9                                                         |
| R    | {Platinum<br>87%Pt–13%Rh                         | 1750                          | 5.80                                                | 0.647                       | 3.407         | 10.503         | 1.9                                                         |
| B    | {94%Pt–6%Rh<br>70%Pt–30%Rh                       | 1800                          | 0.00                                                | 0.033                       | 0.786         | 4.833          | 1.9                                                         |

(<sup>a</sup>) thermocouple life is shortened by prolonged operation near maximum temperature. (<sup>b</sup>) reference junction at 0°C. (<sup>c</sup>) per double foot; for 24-gauge, multiply values by 0.25. (<sup>d</sup>) 55%Cu–45%Ni. (<sup>e</sup>) 90%Ni–10%Cr. (<sup>f</sup>) 96%Ni–2%Mn–2%Al.

you have with thermocouples stem from their low output voltage ( $50\mu\text{V}/^\circ\text{C}$  or thereabouts), combined with large common-mode ac and radiofrequency interference. The amplifier must have good common-mode rejection at 60Hz and stable differential gain. In addition, the input impedance must be moderately high (of order 10k or more) in order to prevent error from loading, since the thermocouple leads do have some resistance (5 feet of 30 gauge type K junction wire has a resistance of  $30\Omega$ , for example).

The circuit shown in Figure 15.3 is a good solution. It is just the standard differencing amplifier with the T connection in the feedback path to get high voltage gain (200 in this case) while keeping the input impedance large enough so that loading of the source impedance doesn't contribute error. The op-amp is a precision low-offset type, with drift of less than  $1\mu\text{V}/^\circ\text{C}$  to keep its contribution to the measurement error much less than the  $50\mu\text{V}$  that corresponds to a  $1^\circ\text{C}$  error. The input bypass

capacitors are a good idea to reduce the common-mode interference at 60Hz and at radiofrequencies (thermocouples and their long connecting cables tend to behave like radio antennas). Since thermocouples respond slowly anyway, you can limit the bandwidth with capacitors across the feedback resistors, as shown. In cases of extreme radiofrequency problems, it may be necessary to shield the input leads and add RF chokes before the input bypass capacitors.

Note that the reference junction compensation circuit in Figure 15.3 acts as the *output*, rather than the usual method of compensating the voltage from the thermocouple at the input, as in Figure 15.2. This is done to keep the input truly differential, in order to preserve the advantages of the good common-mode rejection of the differencing amplifier. Since the amplifier has a voltage gain of approximately 200, the compensation circuit has to add  $200 \times 51.5\mu\text{V}/^\circ\text{C}$ , or  $10.3\text{mV}/^\circ\text{C}$  at the output. Note that the OP-97E's input



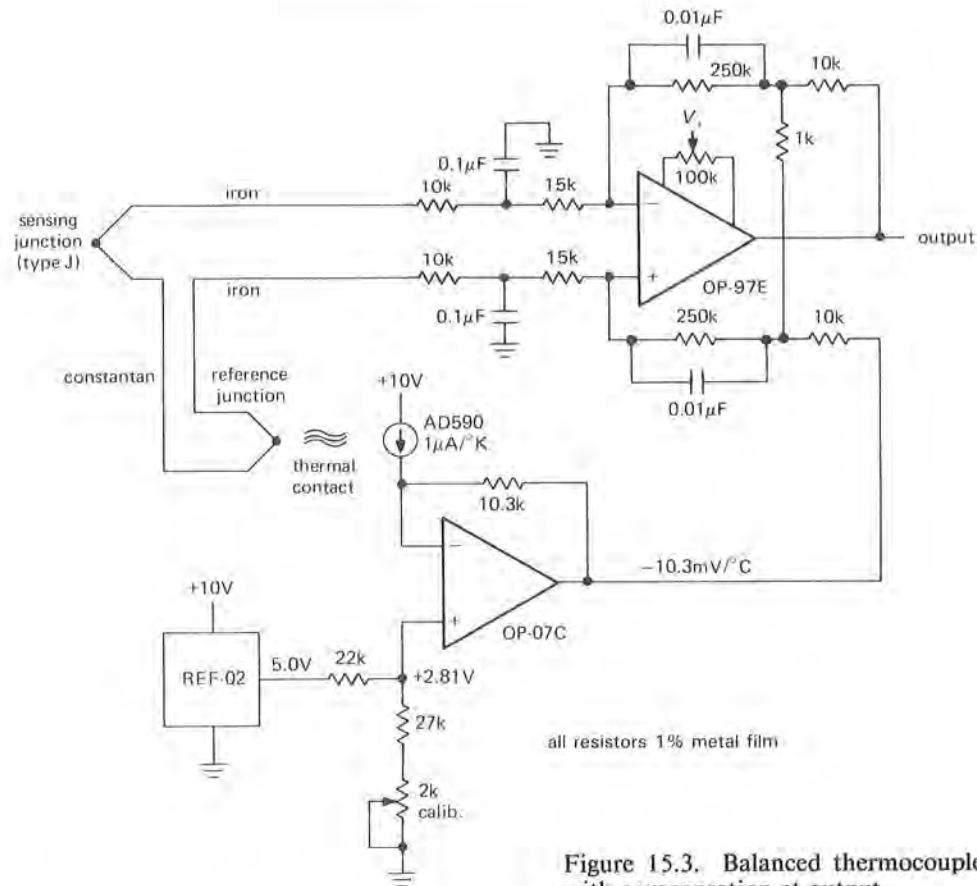


Figure 15.3. Balanced thermocouple amplifier with compensation at output.

offset current of 0.1nA (max) contributes  $25\mu\text{V}$  of input error, which, along with the  $25\mu\text{V}$  (max) of  $V_{os}$ , can be trimmed to zero. Alternatively, one could use a chopper-stabilized op-amp like the 7652 ( $V_{os} = 5\mu\text{V}$  max,  $I_{os} = 40\text{pA}$  max).

An instrumentation amplifier, as in Figure 7.32, could be used instead of the differencing amplifier we've shown; in that case be sure to provide a dc bias path at the input.

Thermocouple users should be aware of Analog Devices' "thermocouple amplifier with cold junction compensation," the AD594 (type J) and AD595 (type K). These monolithic devices have everything you need (including an ice-point reference) to produce a temperature-proportional output voltage, or even programmable trip

point, given a thermocouple input. The best grade is accurate (without trimming) to  $\pm 1^\circ\text{C}$  at room temperature, increasing to  $\pm 3^\circ\text{C}$  at  $-25^\circ\text{C}$  and  $+75^\circ\text{C}$ . Linear Technology makes the LT1025 "micropower thermocouple cold-junction compensator," designed to be used with an external precision op-amp. It includes compensation for all thermocouples in Table 15.1 (except type B), with second-order curvature correction to maintain accuracy over a wider temperature range. The best grade (LT1025A) is accurate (without trimming) to  $\pm 0.5^\circ\text{C}$  at room temperature, increasing to  $\pm 2^\circ\text{C}$  at  $-25^\circ\text{C}$  and  $+80^\circ\text{C}$ .

Complete "smart" temperature-measuring instruments configured for various thermocouple pairs are available commercially. These instruments include computational

circuitry to convert the thermoelectric voltage to temperature. For instance, the digital thermometers manufactured by Analog Devices and Omega Engineering achieve an accuracy of about  $0.4^{\circ}\text{C}$  over a temperature range from  $-200^{\circ}\text{C}$  to  $+1000^{\circ}\text{C}$  and an accuracy of  $1^{\circ}\text{C}$  at temperatures up to  $+2300^{\circ}\text{C}$ .

When compared with other methods of temperature measurement, thermocouples have the advantage of small size and wide temperature range, and they are particularly good for measuring high temperatures.

### Thermistors

Thermistors are semiconductor devices that exhibit a negative coefficient of resistance with temperature, typically in the neighborhood of  $-4\%/^{\circ}\text{C}$ . They are available in all sorts of packages, ranging from tiny glass beads to armored probes. Thermistors intended for accurate temperature measurement (they can also be used as temperature-compensation elements in circuits, for instance) typically have a resistance of a few thousand ohms at room temperature, and they are available with tight conformity ( $0.1$ – $0.2^{\circ}\text{C}$ ) to standard curves. Their large coefficient of resistance change makes them easy to use, and they are inexpensive and stable. Thermistors are a good choice for temperature measurement and control in the range of  $-50^{\circ}\text{C}$  to  $+300^{\circ}\text{C}$ . It is relatively easy to design a simple and effective circuit for “proportional temperature control” using a thermistor sensing element; see, for example, RCA application note ICAN-6158 or the Plessey SL445A data sheet.

Because of their large resistance change with temperature, thermistors make no great demands on the circuitry that follows. Some simple ways to generate an output voltage are shown in Figure 15.4. The circuit in part A expands the low-temperature end of the range because of

the thermistor’s exponential resistance change, whereas the circuit in part B produces a somewhat more linear variation of output voltage with temperature. For example, Figure 15.5 shows resistance versus temperature for two configurations – a bare  $10\text{k}$  thermistor (Fenwal UUA41J1 “uni-curve”), and the same thermistor with a  $10\text{k}$  resistor in series. The series pair is linear to within 3% from  $-10^{\circ}\text{C}$  to  $+50^{\circ}\text{C}$ , and better than 1% from  $0^{\circ}\text{C}$  to  $+40^{\circ}\text{C}$ . The circuit of Figure 15.4B, sensing the drop across  $R$ , produces a voltage output of equivalent linearity.

Parts C and D of Figure 15.4 show elaborations of this linearizing idea, using composite matched “thermlinear” thermistors (and corresponding resistor pairs) from Yellow Springs Instrument Company. These 2-thermistor configurations achieve 0.2% linearity from  $0^{\circ}\text{C}$  to  $100^{\circ}\text{C}$ . YSI also makes modules with *three* thermistors (and three resistors), for even better linearity. The circuit in part E is the classic Wheatstone bridge, balanced when  $R_T/R_2 = R_1/R_3$ ; since it is ratiometric, the null doesn’t shift with variations in supply voltage. The bridge circuit, with a high-gain amplifier, is particularly good for detecting small changes about some reference temperature; for small deviations the (differential) output voltage is linear in the unbalance. With all the thermistor circuits you have to be careful about self-heating effects. A typical small thermistor probe might have a dissipation constant of  $1\text{mW}/^{\circ}\text{C}$ , meaning that  $I^2R$  heating should be kept well below  $1\text{mW}$  if you would like your reading accurate to better than 1 degree.

Complete “smart” temperature-measuring instruments using curve-conforming thermistors are available commercially. These devices include internal computational circuitry to convert resistance readings directly to temperature. As an example, the Omega model 5800 digital thermometer covers  $-30^{\circ}\text{C}$  to  $+100^{\circ}\text{C}$ .

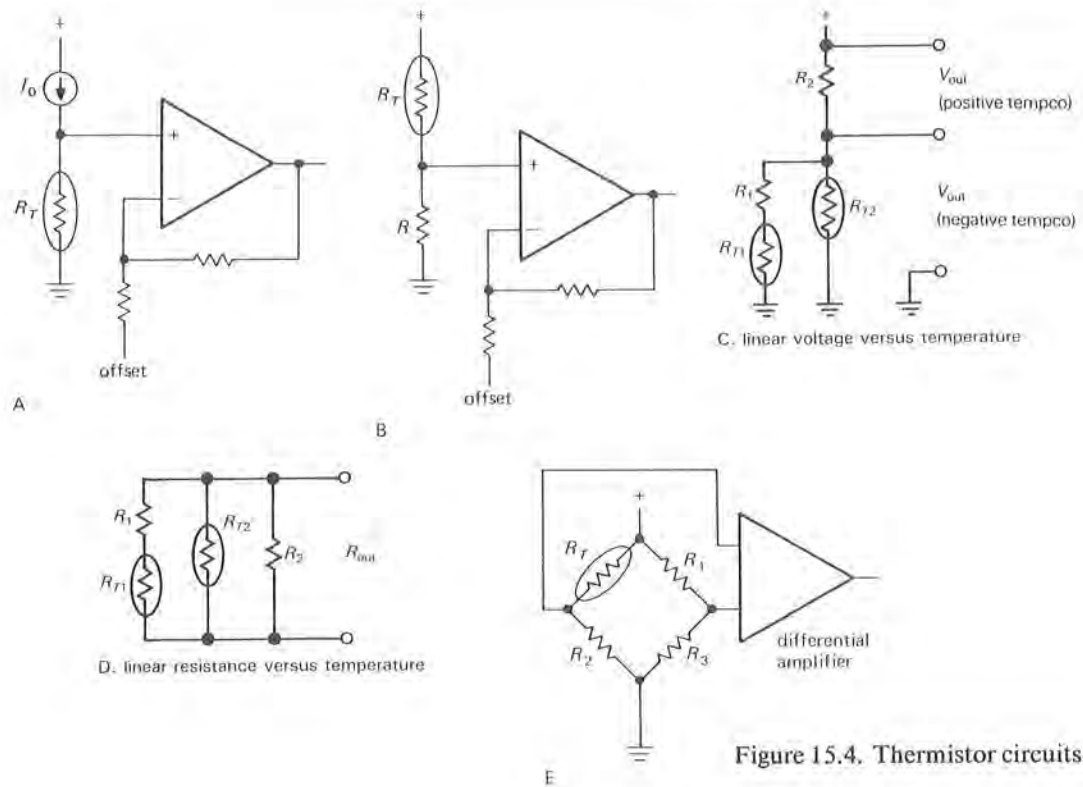


Figure 15.4. Thermistor circuits.

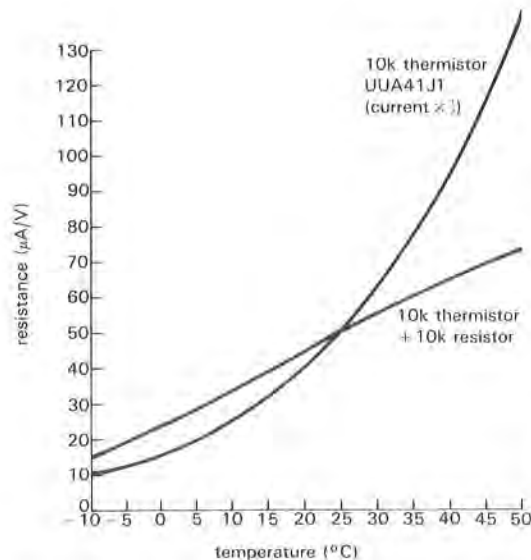


Figure 15.5 Resistance versus temperature for thermistor and thermistor-resistor pair.

in two ranges, reading out in either centigrade or Fahrenheit on a 4-digit LED display. It has  $0.5^{\circ}C$  accuracy over the full temperature range, with  $0.1^{\circ}C$  resolution.

When compared with other methods of temperature measurement, thermistors provide simplicity and accuracy, but they suffer from self-heating effects, fragility, and a narrow temperature range.

#### Platinum resistance thermometers

These devices consist simply of a coil of very pure platinum wire, which has a positive temperature coefficient of about  $0.4\%/^{\circ}C$ . Platinum thermometers are extremely stable with time and conform very closely ( $0.02$ – $0.2^{\circ}C$ ) to a standard curve. They are usable over a temperature range of  $-200^{\circ}C$  to  $+1000^{\circ}C$ . They aren't terribly cheap.

### IC temperature sensors

As we remarked in Section 6.15, a bandgap voltage reference can be used to generate a temperature-sensing output voltage proportional to absolute temperature, as well as its usual stable zero-tempco reference output. The REF-02, for example, provides a “temp” output with a linear coefficient of  $+2.1\text{mV}/^\circ\text{C}$ . If you buffer this output with an amplifier of adjustable gain and offset for calibration, you can achieve accuracies of about  $0.5^\circ\text{C}$  over the range  $-55^\circ\text{C}$  to  $+125^\circ\text{C}$ . The LM335 is a convenient 2-terminal temperature sensor that behaves like a zener diode with a voltage of  $+10\text{mV}/^\circ\text{K}$ ; e.g., at  $25^\circ\text{C}$  ( $298.2^\circ\text{K}$ ) it acts like a  $2.982$  volt zener (Fig. 15.6). It comes with an initial accuracy as good as  $\pm 1^\circ\text{C}$  and it can be externally trimmed. A single point calibration can typically improve its accuracy to  $\pm 0.5^\circ\text{C}$  max over a  $-55^\circ\text{C}$  to  $+125^\circ\text{C}$  range. After trimming, the output should be accurate to  $0.1^\circ\text{C}$  at the test temperature, with an accuracy budget increasing to  $\pm 0.5^\circ\text{C}$  at the temperature extremes (Fig. 15.7).

The LM35 also provides a voltage output with  $+10\text{mV}/^\circ\text{C}$  slope, but it behaves like a 3-terminal reference (rather than a 2-terminal zener), powered by  $+4$  volts to  $+30$  volts applied to a third terminal; it has an internal offset such that the output is  $0$  volts at  $0^\circ\text{C}$ . To operate near or below  $0^\circ\text{C}$ , you must use a pulldown resistor, as shown. The best grade (LM35A) has a maximum error of  $0.5^\circ\text{C}$ , but cannot be trimmed. Its cousin, the LM34A, works similarly, with readout in Fahrenheit ( $0$  volts at  $0^\circ\text{F}$ ).

Another approach to IC temperature sensors is the AD590, a two-terminal device that acts as a constant-current element, passing a current in microamps equal to the absolute temperature; e.g., at  $25^\circ\text{C}$  ( $298.2^\circ\text{K}$ ) it behaves like a constant-current regulator of  $298.2\mu\text{A}$  ( $\pm 0.5\mu\text{A}$ ). With this simple device you get  $1^\circ\text{C}$  accuracy (best grade) over the range  $-55^\circ\text{C}$

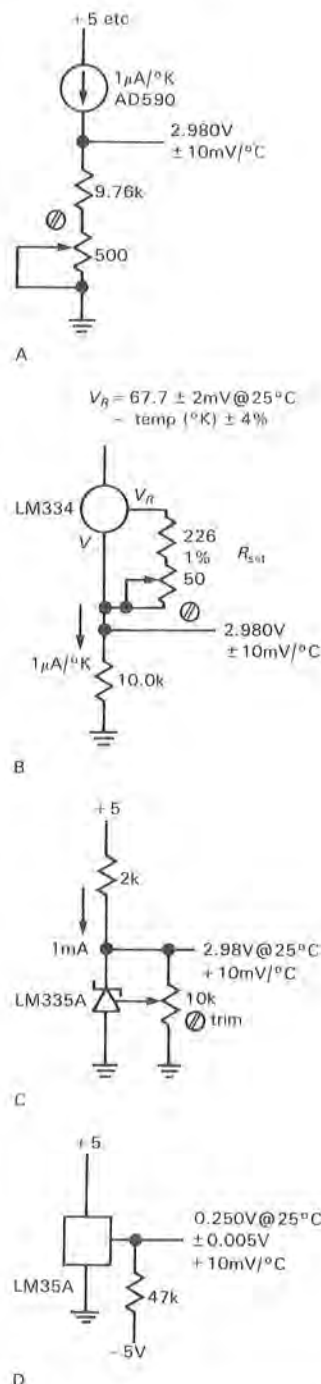


Figure 15.6 IC temperature sensors. Voltages are at  $25^\circ\text{C}$ . Methods A and B incur an additional  $1\%$  error per  $33\text{ppm}/^\circ\text{C}$  resistor tempco.

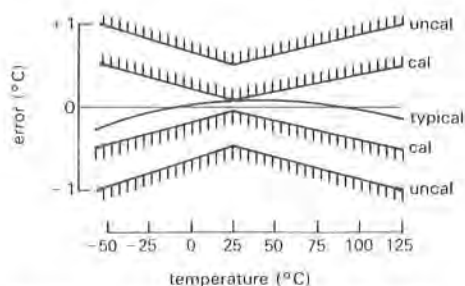


Figure 15.7. LM335 temperature error.

to  $+150^{\circ}\text{C}$ . The simplicity of external circuitry required makes this a very attractive device. A plastic version, the AD592, has comparable performance over a smaller temperature range ( $-25^{\circ}\text{C}$  to  $+105^{\circ}\text{C}$ ). The LM334 current-source IC (see Section 6.18) also has an output proportional to absolute temperature, settable with a single resistor according to  $I_{\text{out}}(\mu\text{A}) = 227T(^{\circ}\text{K})/R_{\text{set}}(\Omega)$  (see Fig. 15.6); this formula includes  $\approx 5\%$  correction for  $V_{-}$  current.

### Quartz thermometer

The change of resonant frequency of a quartz crystal with temperature can be exploited to make an accurate and repeatable thermometer. Although the usual objective in quartz-crystal oscillator design is lowest possible temperature coefficient, in this case you choose a crystal cut with a large coefficient and take advantage of the high precision possible in frequency measurements. A good example of a commercial instrument is the Hewlett-Packard 2804A, a microprocessor-controlled thermometer with an absolute accuracy of 40 millidegrees over the range  $-50^{\circ}\text{C}$  to  $+150^{\circ}\text{C}$  (reduced accuracy over a wider range) and temperature resolution of 100 microdegrees. To get this kind of performance, the instrument contains calibration data for the individual sensor that it uses in the temperature calculation.

### Pyrometers and thermographs

An interesting method of “noncontacting” temperature measurement is exemplified by the classic pyrometer, a gadget that lets you sight through a telescope at an incandescent object, comparing its glowing color with that of a filament inside the pyrometer. You adjust the filament’s current until its brightness matches the object’s, both being viewed through a red filter, then read off the temperature. This is a handy method of measuring the temperatures of very hot objects, objects in inaccessible places like ovens or vacuum chambers, or objects in oxidizing or reducing atmospheres where thermocouples cannot be used. Typical optical pyrometers cover a range of  $750^{\circ}\text{C}$  to  $3000^{\circ}\text{C}$ , with an accuracy of about  $4^{\circ}\text{C}$  at the low end and  $20^{\circ}\text{C}$  at the high end.

The development of good infrared detectors has extended this sort of measurement technique down to ordinary temperatures and for ordinary use. For example, Omega offers a line of digital-readout infrared pyrometers with a temperature range of  $-30^{\circ}\text{C}$  to  $+5400^{\circ}\text{C}$ . By measuring the intensity of infrared radiation, perhaps at several infrared wavelengths, you can determine with good accuracy the temperature of a remote object. Such “thermography” has recently become popular in quite diverse fields: in medicine, for the detection of tumors, and in the energy business, where a thermograph of your house can tell you where your energy dollars are evaporating.

### Low-temperature measurements

Cryogenic (very cold) systems pose special problems when it comes to accurate temperature measurement. What matters there is how close to absolute zero ( $0^{\circ}\text{K} = 273.15^{\circ}\text{C}$ ) you are. Two popular methods involve measuring the resistance of ordinary carbon-composition resistors,



which soars at low temperatures, and measuring the degree of paramagnetism of some salt. These are really specialty measurement techniques that will not be dealt with here.

### Measurement allows control

If you have a way of adjusting some quantity, then the availability of a good measurement technique lets you control that quantity accurately. Thermistors, in particular, provide a nice method for controlling the temperature of a bath or oven.

## 15.02 Light level

The measurement, timing, and imaging of low light levels are parts of a well-developed field, thanks to the existence of amplification methods that do not rely on conventional circuit techniques. Photomultipliers, channel-plate intensifiers, CCDs (charge-coupled devices), and ISITs (intensified silicon intensifier target) are included in the catalog of high-performance optical detection devices. We will begin with the simplest detectors (photodiodes and phototransistors) and then go on to discuss the exotic and the wonderful.

### Photodiodes and phototransistors

A diode junction acts as a photodetector: Light creates electron-hole pairs, and therefore a current through the external circuit. Diodes intended as photodetectors (photodiodes and PIN diodes) are packaged in a transparent case and are designed for high speed, high efficiency, low noise, and low leakage current. In the simplest mode of operation, a photodiode can be connected directly across a resistive load, or current/voltage converter, as shown in Figure 15.8. You get faster response (and the same photocurrent) by reverse-biasing

the junction, as in Figure 15.9. High-speed PIN diodes have response times of a nanosecond or less (1GHz bandwidth) when loaded into a low impedance. It should be noted that the leakage current of good PIN diodes is so low (less than a nanoamp) that the Johnson noise in the load resistance dominates for load resistances less than 100M $\Omega$  or so, implying a speed/noise trade-off. An additional problem to be aware of is the error caused by the amplifier's input offset voltage, or the applied bias voltage, in combination with the photodiode's "dark resistance," when working at low light levels.

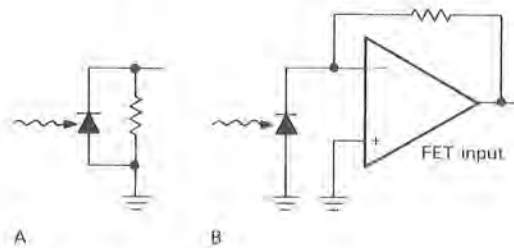


Figure 15.8

Photodiodes are pretty good light detectors when there is plenty of light around, but the output signal can be inconveniently small at low light levels. Typical sensitivities are of the order of 1 microamp per microwatt of incident light. A flux of 1000 photons per second, quite visible with the unaided eye, would cause a photocurrent of  $4 \times 10^{-16}$  amps when focused onto a PIN diode, totally undetectable when compared with the leakage current and noise. No silicon photodetectors are sensitive at the photon level (see the subsequent section on photomultipliers for that), but a device known as a *phototransistor* has considerably more output current than a photodiode at comparable light levels, bought at the expense of speed. It works like an ordinary transistor, with the base current provided by the photocurrent produced in the base-collector junction.

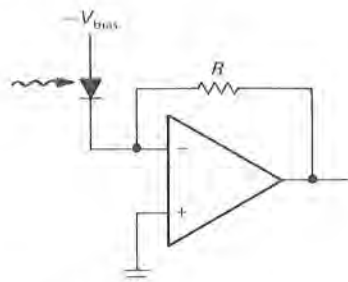


Figure 15.9

Inexpensive phototransistors like the MRD701 have output currents of a milliamp or so at an illumination of  $1\text{ mW/cm}^2$ , with rise and fall times of tens of microseconds, and photo-Darlingtonts like the MRD711 have photocurrents roughly 50 times higher, but with rise times of  $100\mu\text{s}$  or more. Note, however, that the additional current gain of a phototransistor or photo-Darlington doesn't improve its ability to detect extremely low light levels (its "detectivity"), since the ultimate limit is set by the detector diode's "dark current."

We have had excellent results with the photodiodes from UDT, Siemens, and Hamamatsu. The latter's catalog lists an impressive variety of detectors, including silicon *pn*-junction, silicon PIN diodes, GaAs (both diffusion and Schottky types) and GaP photodiodes, and avalanche detectors. It includes single detectors of many sizes and shapes, as well as linear photodiode arrays. UDT makes a series of detectors with integral MBC connectors.

### Photomultipliers

For low-light-level detection and measurement (and, incidentally, for nanosecond resolution), you can't beat the photomultiplier. This clever device allows a photon (the smallest unit of light) to eject an electron from a photosensitive alkali metal "photocathode." The photomultiplier then amplifies this feeble photocur-

rent by accelerating the electron onto successive surfaces (dynodes), from which additional electrons are easily ejected. Figure 15.10 suggests the process. This use of "electron multiplication" yields extremely low noise amplification of the initial photocurrent signal. Typically, you use a voltage divider to put about 100 volts between successive dynodes, for a gain of about 10 per stage, or 1 million overall. The final current is collected by the anode, usually run near ground potential (look at Fig. 15.11), and is large enough so that subsequent amplifier noise is negligible.

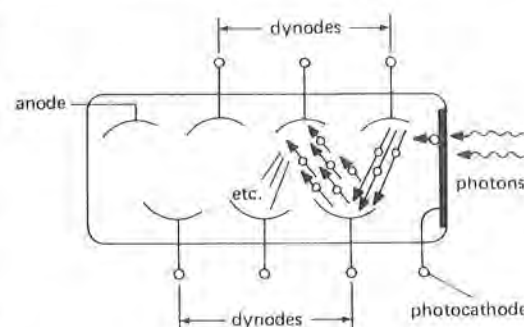


Figure 15.10. Photomultiplier multiplication process.

The most efficient photocathode materials have quantum efficiencies exceeding 25%, and with the large gain provided by the dynodes, individual photoelectron events are easily seen. At low light levels that's how you would use it, following the PMT (photomultiplier tube) with charge-integrating pulse amplifiers, discriminators (as detailed in Fig. 13.60), and counters. At higher levels the individual photoelectron count rate becomes too high, so you measure the anode *current* as a macroscopic quantity, instead. PMTs have typical sensitivities of an ampere per microwatt, although you could never operate them at such a high current; maximum PMT anode currents are limited to a milliamp or less. It should be noted that the practical limit to photon counting,

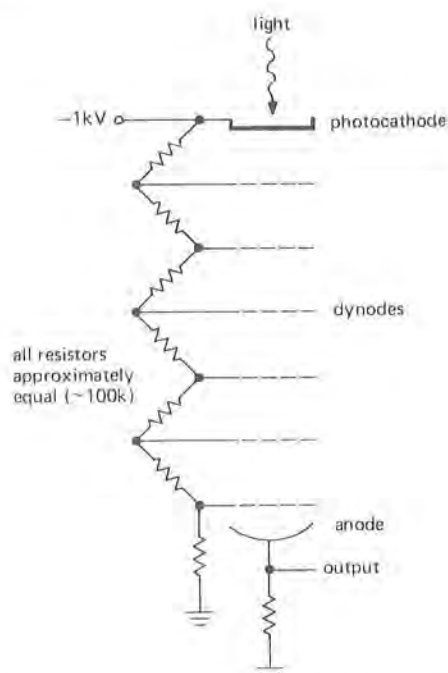


Figure 15.11. Photomultiplier biasing.

something like 1 million counts per second, corresponds to roughly 2 micro-micro-watts incident power!

Convenient electronic packages are available for both pulse-counting and output-current-measurement modes. An example is the “quantum photometer” from PAR, a handy gadget with built-in high-voltage supply and both pulse and current electronics. It has 11 ranges of pulse counting (10 pps to  $10^6$  pps, full scale) and 11 ranges of anode-current read-out (10nA to 1mA, full scale).

Even in total darkness you get a small anode current from a photomultiplier. This is caused by electrons thermally excited from the photocathode and dynodes, and it can be reduced by cooling the PMT down to  $-25^\circ\text{C}$  or so. Typical dark currents for a sensitive “bialkali” cathode PMT are in the neighborhood of 30 counts per second per square centimeter of cathode area, at room temperature. A cooled

PMT with a small cathode can have dark currents of less than a count per second. It should be pointed out that a powered PMT should never be exposed to ordinary light levels; a PMT that has seen the light of day, even without power applied, may require 24 hours or more to “cool down” to normal dark-current levels. In some applications (e.g., fluorescence measurements) a PMT may be exposed to bright flashes of light at known times. In that case you can minimize the overload recovery time by disabling the accelerating voltage on the first few dynodes during the flash (some manufacturers offer this capability in their PMT/socket combinations).

When compared with photodiodes, PMTs have the advantage of high quantum efficiency while operating at high speed (2ns rise time, typically). They are bulky, though, and they require a stable source of high voltage, since the tube’s gain rises exponentially with applied voltage.

It should be emphasized that PMTs are to be used with extremely low light levels. You run them with typical anode currents of a microamp or less, and they can easily see light that you cannot. Photomultipliers are used not only for the detection of light directly, as in astronomy (photometry) and biology (bioluminescence, fluorescence), but also in conjunction with scintillators as particle detectors and x-ray/gamma-ray detectors, as we will discuss in Section 15.07. Photomultipliers find wide use in spectrophotometry, where they are combined with prisms, gratings, or interferometers to make precise measurements of optical spectra. PMTs are manufactured by RCA (Burle), Hamamatsu, EMI, and EG&G.

#### **CCDs, intensifiers, SITs, ISITs, and image dissectors**

It is possible to do *imaging* at the light quantum level, thanks to some clever recent technology; i.e., you can form an

image with the same sort of sensitivity to low light levels that you get with the (non-imaging) photomultiplier. These recent inventions are amazing to see. You can sit in what appears to be a completely dark room, then peer into a television monitor in which are imaged, albeit with plenty of "snow," all the objects in the room.

The key to all this is the image intensifier, an incredible device that produces as its output a brightened replica of an input image. You begin with either an ordinary silicon target vidicon (TV camera) or a CCD array. These are light-sensitive two-dimensional targets that accumulate an image and can be read out electronically by scanning with an electron beam or by shifting the image along as an analog shift register, respectively. At this point, all you have is a television camera whose sensitivity is far below the individual photon level; it is the two-dimensional analog of a photodiode. To bring about the miraculous, you simply put an imaging intensifier tube in front. Figure 15.12 suggests the process schematically.

Intensifiers come in two varieties. The first-generation type consists of a sensitive photocathode surface of the type used in photomultipliers, with electron-focusing optics and a phosphor screen arranged behind so that photoelectrons from the cathode are accelerated by high applied

voltages and hit the phosphor with enough energy to give off a bright flash of light. With this kind of intensifier you can get single-stage light amplification of about 50, with resolution of about 50 lines/mm. Popular types cascade two, three, or four such stages of light amplification to achieve overall light amplifications of a million or more. The input and output may simply be glass surfaces with their internal photosensitive and phosphor coatings, or they may be paved with a dense fiber-optic bundle. Fiber-optics are nice because they let you match a perfectly flat entrance or exit surface to a curved tube surface, and they simplify the external optical system, since you can cascade these devices by just sticking things together, without any lenses.

Second-generation intensifiers using "microchannel plates" allow you to achieve much higher values of single-stage light amplification, and they are better at really low light levels because of fewer "ion events," the result of positive ions being ejected from the phosphor and returning to the cathode, where they make a big splash. In these channel-plate intensifiers the space from cathode to phosphor contains a bundle of microscopic hollow tubes whose insides are coated with a dynode-type multiplication surface. Photoelectrons from the cathode bounce their way down these channels, ejecting secondary

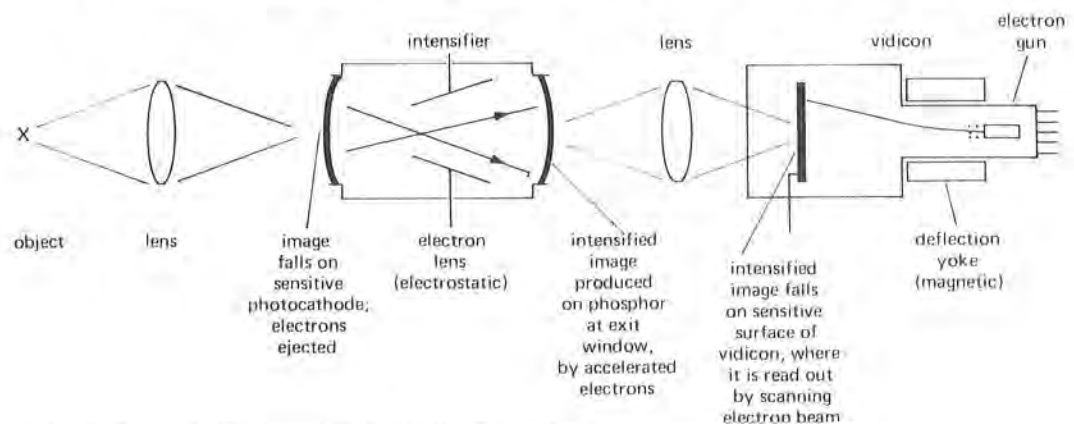


Figure 15.12. Vidicon with single-stage intensifier.



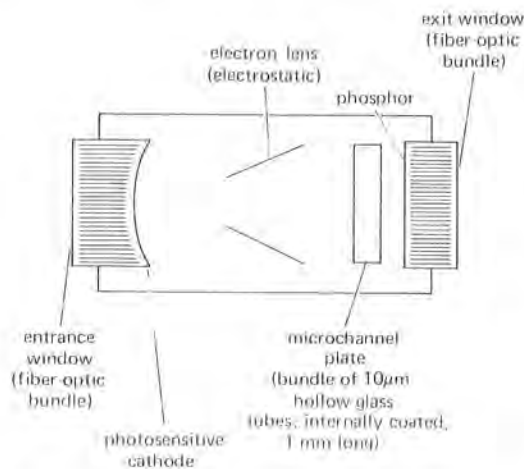


Figure 15.13. Electrostatically focused channel-plate intensifier.

electrons to give light amplification of about 10,000 (Fig. 15.13). You can get resolutions of about 20 lines/mm, and with special configurations ("J-channel," "C-channel," "chevron") the ion-event problem can be eliminated almost entirely. The result is an imaging intensifier with the same sort of quantum efficiency as photomultipliers (20%–30%). The use

of nearly noiseless electron multiplication results in light amplification to a level that the vidicon or CCD can see.

Such an intensifier combined with a silicon-target vidicon in a single tube is called a "SIT" (silicon intensifier target). An ISIT is a SIT with an additional intensifier placed externally in front (Fig. 15.14); this is the sort of gadget that lets you see in the dark. These things are very popular with astronomers and with night-warfare people.

An interesting variation of the imaging intensifier is the so-called image dissector, a clever device that actually preceded the devices just described. It consists of a sensitive photocathode area, followed by the usual photomultiplier dynode chain. In between is a small aperture and some deflection electrodes, so that any spot on the photocathode can become the active area for electron multiplication by the dynode system. You can think of an image dissector as a photomultiplier with an electronically movable photocathode area. It has the quantum efficiency and gain of a conventional PMT, but it differs from the intensified vidicons, CCDs, and SITs

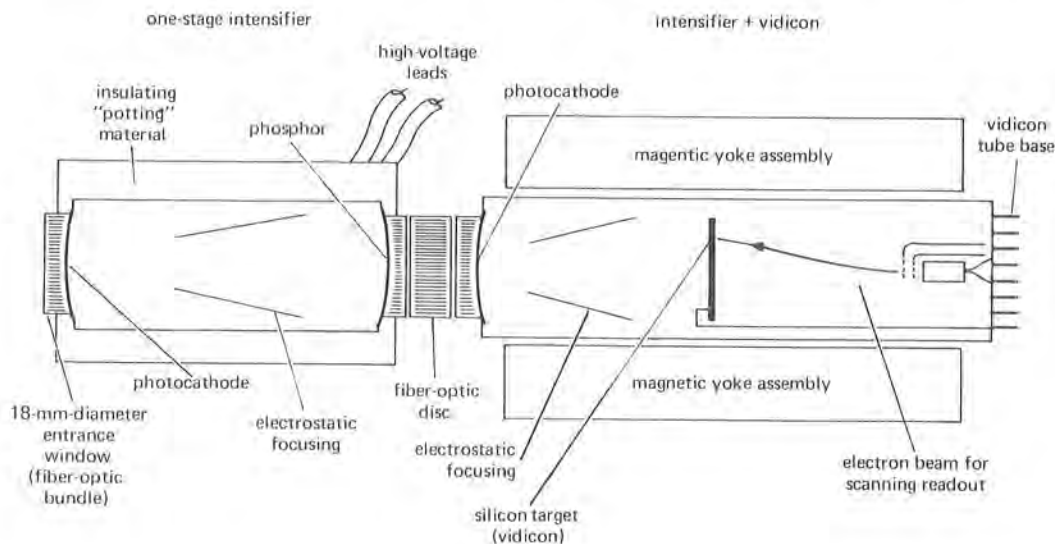


Figure 15.14. Intensified silicon intensifier target (ISIT).



(which are all image-integrating devices) in that it does not accumulate the image over the entire field in between readouts.

CCD detector arrays can, of course, be used without intensification. You can buy them from companies like EGG Reticon, Kodak, TI, Tektronix, Thomson, and Toshiba. They are available as 1-dimensional ("linear") arrays containing up to 4096 cells, or as 2-dimensional ("area") arrays containing 256K ( $512 \times 512$ ) or more cells (called "pixels," for "picture elements"); Toshiba has an area detector with an impressive 2 million pixels on a single chip. Linear arrays are handy as detectors for spectroscopy; area detectors, of course, are used for full 2-dimensional images, for example in television cameras.

All CCDs are light-*integrating* devices, which accumulate charge in each pixel until the array is read out. During readout the CCD becomes an analog shift register, with the image emerging, raster style, as a serial analog waveform at the single output line.

### 15.03 Strain and displacement

The field of measurements of physical variables such as position and force has its own bag of tricks, and any accomplished measurer should be aware of things like strain gauges, LVDTs, and the like. The key to all of these measurements is the measurement of displacement.

There are several nice ways to measure position, displacement (changes in position), and strain (relative elongation).

**LVDT.** A popular method is the LVDT (linear variable differential transformer), which is almost self-explanatory. You construct a transformer with a movable core, excite one set of coils with ac, and measure the induced voltage in the second set. The secondary is center-tapped (or brought out as two separate windings) and arranged

symmetrically with respect to the primary, as shown in Figure 15.15. LVDTs come in an enormous variety of sizes, with full-scale displacements ranging from 0.005 inch to 25 inches, excitation frequencies from 50Hz to 30kHz, and accuracies of 1% down to 0.1% or better. A leader in this field is Schaevitz, whose catalog lists a broad selection of linear and angular ("RVDT") transducers, and measurement transducers utilizing LVDT sensors (e.g., pressure, force, acceleration, etc.), as well as LVDT readout electronics. If you become involved with LVDTs, you may want to build your own instrumentation, perhaps using the special ICs designed for the purpose. For example, the monolithic Signetics NE5520/1 "LVDT Signal Conditioner" provides sine-wave excitation signals and includes a synchronous demodulator to provide a voltage output proportional to the LVDT displacement. The 2S54/6 from Analog Devices are LVDT synchronous demodulators with excellent linearity (0.01%) and built-in A/D converters providing direct digital outputs (14 and 16 bits, respectively).

**Strain gauge.** A strain gauge measures elongation or flexure by subjecting an array of four metal thin-film resistors to deformation. They come as complete assemblies, in sizes from 1/64 of an inch to several inches, and they generally have impedances in the neighborhood of 350 ohms/leg. Electrically they look like a Wheatstone bridge; you apply dc across two of the terminals and look at the voltage difference between the other two, as discussed in Section 7.09. The output voltages are very small, typically 2mV per volt of excitation for full-scale deformation, with accuracies ranging from 1% down to 0.1% of full scale (see Fig. 15.15D).

It is not easy to measure small relative elongations, and strain-gauge specifications are notoriously unreliable. Small differences in the temperature coefficients

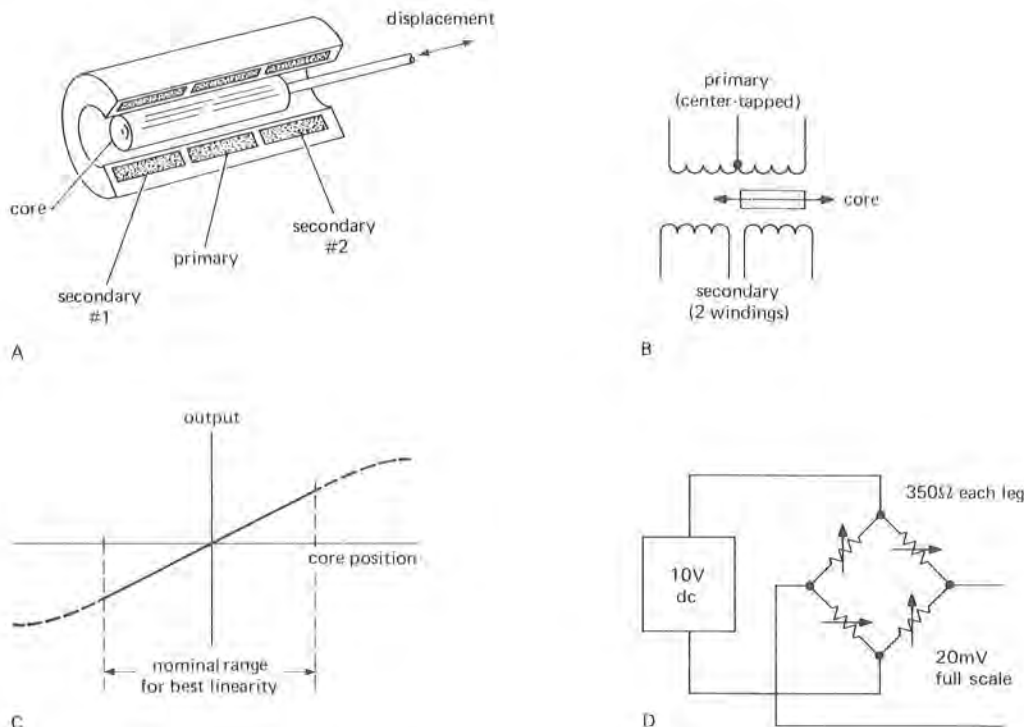


Figure 15.15. Displacement transducers.

- A. Linear variable differential transformer (LVDT) cutaway view.  
 B. LVDT schematic.  
 C. LVDT output versus displacement.  
 D. Strain gauge schematic.

of the bridge elements are responsible for the temperature sensitivity, which limits the performance of the strain gauge. This is a problem, even in controlled-temperature environments, because of self-heating. For example, 10 volts of dc excitation on a 350 ohm bridge will produce 300mW dissipation in the sensor, with a temperature rise of 10°C (or more), causing errors corresponding to a real signal of 0.1% to 0.5% of full scale.

Recently, semiconductor strain gauges have become popular. They have outputs that are 10 times higher than those of the metal-film variety and impedances of a few thousand ohms. It is often necessary to use a current source as excitation, rather than

a voltage source, to minimize temperature sensitivities.

**Capacitance transducers.** Very sensitive measurements of displacement can be made with a transducer consisting simply of two closely spaced plates, or a plate suspended between a pair of outer plates. By making the capacitor part of a resonant circuit, or by using a high-frequency ac bridge, you can sense or control very small changes in position. Capacitor microphones use this principle to convert acoustic pressure or velocity to an audio signal.

The amplifiers used with capacitor microphones illustrate some interesting circuit ideas, and they are of practical

importance, since many of the best recording microphones are capacitive position transducers, made by supporting a thin metallized plastic foil in close proximity to a fixed plate. You charge the capacitor through a large resistor with a bias of 50 to 100 volts, and you look at the changes of voltage as the diaphragm moves in the sound field.

Capacitor microphones have enormously high source impedances (a typical capsule has about 20pF of capacitance, or a reactance of about 400M at 20Hz), which means that you don't have a chance of running the signal through any length of cable whatsoever without putting a preamp right at the capsule. Figure 15.16 shows

two ways to buffer the voltage from the capsule, which might have an amplitude of 1mV to 100mV (rms) for typical ranges of program material. In the first circuit, a low-noise FET op-amp provides 20dB of gain and the low impedance necessary to drive a single-ended shielded line. Since the amplifier has to be located close to the microphone capsule (within a few inches), it is necessary to supply the operating voltages (bias for the capsule, as well as op-amp power) through the microphone cable, in this case on additional wires. Note the trick of floating the microphone capsule in order to simplify biasing of the op-amp.  $R_1$  and  $C_1$  filter the bias supply, and  $R_2$  and  $C_1$  must be chosen to have a high impedance

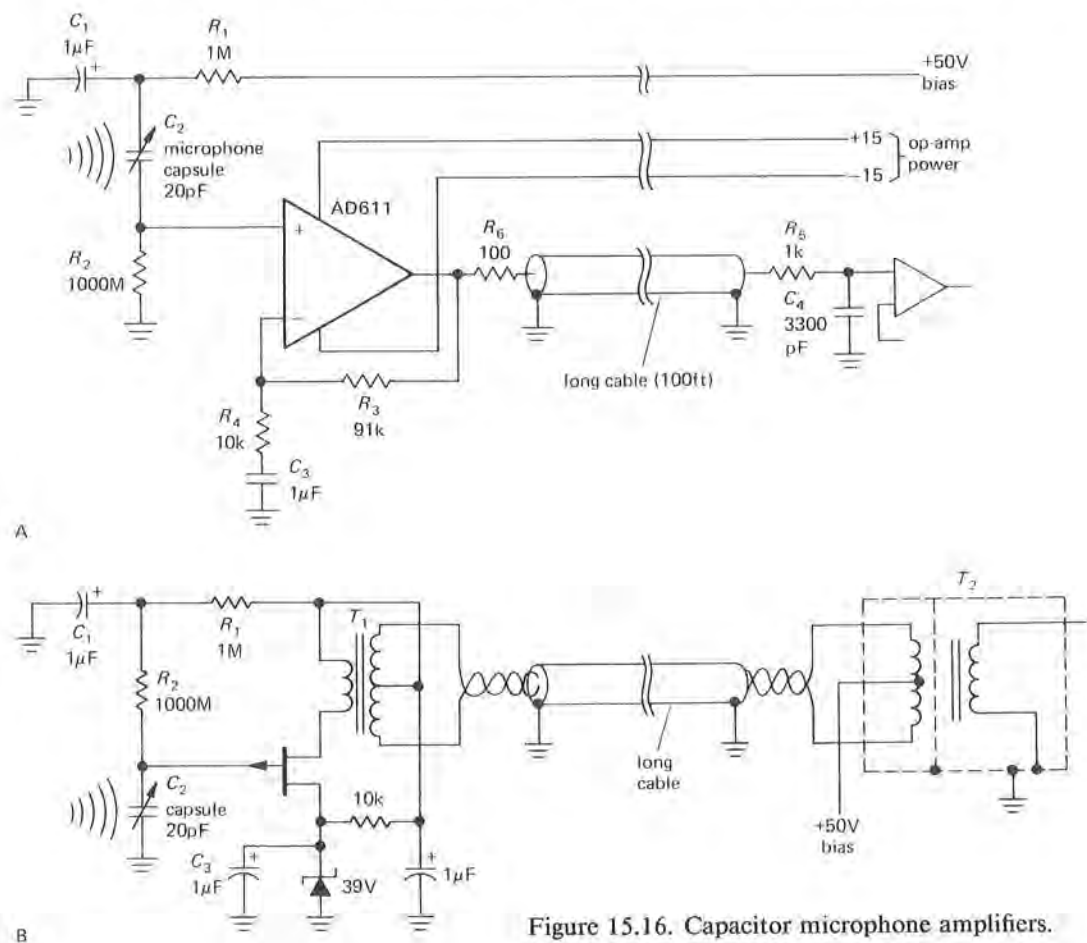


Figure 15.16. Capacitor microphone amplifiers.

compared with the capsule at all audio frequencies.  $R_5$  and  $C_4$  form an RF filter, since the line is unbalanced and therefore prone to radiofrequency interference.

There are a few bad features about this circuit. It requires a 4-wire cable, rather than the industry-standard shielded pair. Also, the floating capsule can create mechanical problems. These drawbacks are remedied in the second circuit, in which the capsule bias is sent on the same lines as the audio itself, which happens to be a balanced 200 ohm pair. One side of the capsule is grounded, and a  $p$ -channel JFET is used as a source follower to drive a small audio transformer. A single-ended output is recovered at the far end, where bias is applied via the transformer's center tap. Some would complain that the proliferation of transformers is a poor idea, but in practice they perform admirably.

- *Angle.* It is possible to convert angles to electronic signals with pretty good precision. There are angular versions of the LVDT, for instance, and popular devices known as resolvers. In both cases you use an ac excitation, and you can easily measure angular position down to an arc-minute. With great care it is possible to measure angles at the arc-second level. There are other techniques, e.g., using light beams looking at a glass disc with gray coded radial stripes.

- *Interferometry.* Highly accurate position measurements can be made by bouncing laser beams off mirrors attached to the object and counting interference fringes. The ultimate accuracy of such methods is set by the wavelength of light, so you have to work hard to do much better than a half micron (1 micron, or micrometer, is 1/1000 of a millimeter, or 1/25000 of an inch). An example of a commercial laser measurement instrument is the 5527A from Hewlett-Packard, which claims resolution of a microinch or better. Laser

interferometer systems are now used routinely in surveying, in flatness measurements, and in various tasks around research laboratories.

The most precise distance measurements have been made interferometrically by Deslattes at the National Bureau of Standards. Deslattes is a real wizard when it comes to precise physical measurement, routinely measuring spacings to milliångstroms (one ten-millionth of a micron) and angles to milliseconds of arc.

- *Quartz oscillators.* A quartz crystal responds to deformation with a change in its resonant frequency, thus providing a very accurate method for measuring small displacements or changes in pressure. Quartz-oscillator pressure transducers provide the highest resolution presently available (more about this later).

#### 15.04 Acceleration, pressure, force, velocity

The techniques just mentioned allow you to measure acceleration, pressure, and force. Accelerometers consist of strain gauges attached to a test mass, or capacitance-sensing transducers that sense the change of position of the test mass. There are various tricks to damp the system to prevent oscillation, in accelerometers that simply measure the displacement of the test mass to provide an output signal; alternatively, some systems use feedback to prevent the test mass from being displaced relative to the body of the accelerometer, the amount of applied feedback force then being the accelerometer's output signal.

LVDT, strain gauges, capacitance transducers, and quartz oscillators are used for pressure measurements, along with special devices such as a Bourdon gauge, a spiral hollow quartz tube that unwinds when inflated. LVDT transducers, for example, are available with full-scale ranges going



from 1 psi to 100,000 psi or more. Quartz-crystal oscillator types provide the highest resolution and accuracy. The types available from Paroscientific, for example, will deliver accuracies of 0.01% and stabilities of 0.001%. Hewlett-Packard has a quartz pressure gauge with 11,000 psi full-scale sensitivity and claimed resolution of 0.01 psi.

LVDI transducers are often used to measure force or weight, although any of the displacement techniques can be used. Full-scale sensitivities go from 10 grams to 250 tons for one popular series, with accuracies of 0.1%. For highly precise laboratory measurements of small forces, you will find quartz-fiber torsion balances, electrostatic balances, and the like. An interesting example of the latter is the clever gravimeter developed by Goodkind and Warburton. It uses a superconducting sphere levitated approximately to zero weight by a persistent magnetic field, then balanced the rest of the way by electrostatic sensing and levitating plates. It can measure changes in gravitational field of one part in a billion, and it easily sees barometric pressure variations because of the effect of the changing overhead air mass on local gravity!

#### ***Magnetic velocity transducers***

The position transducers we have been talking about can also be used to keep track of velocity, which is just the time derivative of position. However, it is possible to make a direct velocity measurement by exploiting the fact that the voltage induced in a loop of wire moving through a magnetic field is proportional to the rate of change of magnetic flux linked by the loop. There are velocity-measurement gadgets available that consist of long coils of wire with magnetic rods moving through the central bore.

Much more prevalent are the magnetic velocity transducers used in the audio industry: microphones (and their inverse,

loudspeakers), phono cartridges, and analog tape recorders. These devices typically generate signals at very low levels (a few millivolts is typical), and they present unique and interesting circuit challenges. For high-quality sound you have to keep noise and interference down 60dB or more, i.e., at the microvolt level. Since these signals get piped around over large distances in recording studios and radio stations, the problem can become serious.

Figure 15.17 shows how low-level signals from microphones and phono cartridges are usually handled. A dynamic microphone is a loudspeaker in reverse: A coil moves in a magnetic field, propelled by the sound pressure. Typically these things have output impedances of 200 ohms, with signals of  $50\mu\text{V}$  to 5mV (rms) for quiet speech and concert-hall sound levels, respectively. For any significant length of connecting cable, you always use a balanced and shielded twisted-pair, terminated in the industry-standard Cannon XLR 3-pin audio connector.

At the far end you transform to a terminated impedance level of about 50k with a high-quality audio matching transformer, as indicated. Signal levels are then in the range of 1mV to 100mV (rms) and should be amplified by a low-noise preamp, as shown. Although you will see preamps with 40dB front-end gain, for good overload performance it is best to stick with a gain of 20dB. This is especially true for popular-music recording, where singers often wind up hollering into the microphone at close range.

The use of balanced 200 ohm microphone cable pretty much eliminates interference because of its good common-mode rejection. Good audio transformers for this kind of application have electrostatic shielding between the windings, which further reduces sensitivity to RF pickup. If radiofrequency interference is not suppressed enough with this scheme, as may be the case near transmitting



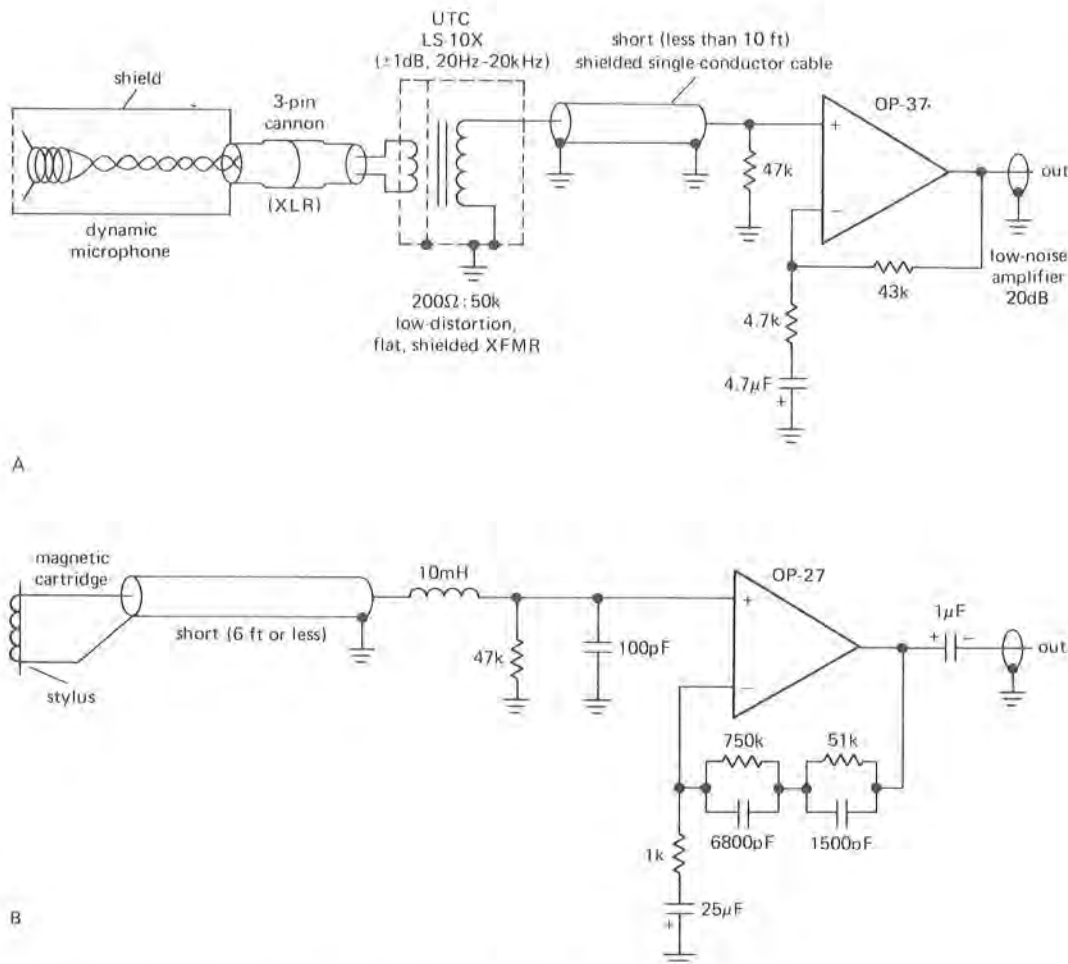


Figure 15.17. Dynamic microphone and phono cartridge amplifiers.

stations, you can add a low-pass filter at the preamp input. A 1k resistor or small RF choke in series at the input, followed by a 100pF capacitor to ground, will usually tame the beast.

Phono cartridges don't require balanced lines, because the cable run to the amplifier is usually very short. The standard method is simply to use single-conductor shielded cable, terminated with the 47k to ground that the cartridge requires for proper frequency response (Fig. 15.17B). We have also shown an input filter to reduce RF interference, because that is such

a common problem in urban areas. RF signals at the inputs of audio equipment pose a particularly insidious problem, because the audio amplifier's nonlinearities at radiofrequencies produce rectification, with consequent interference (audio detection) and distortion. When designing RF filters, be sure to keep the load capacitance small (a maximum of 300pF, including cable capacitance), since otherwise the cartridge's frequency response is changed. The series impedance should not have a resistance greater than a few hundred ohms, to keep the noise low. Quite large values of

inductance can be safely used, since the cartridge's inductance is typically 0.5 henry. The amplifier circuit shown has the standard RIAA response used for recording in the United States.

### 15.05 Magnetic field

Accurate magnetic-field measurements are important in the physical sciences, in connection with instrumentation that uses a magnetic field (magnetic resonance, magnetrons, magnetically focused electron devices, etc.), and in geology and prospecting. For measurements at the 1% level, a Hall-effect probe is adequate. The Hall effect is the production of a transverse voltage in a current-carrying conductor (usually a semiconductor) in a magnetic field, and commercial Hall-effect magnetometers cover a range of about 1 gauss to 10kG full-scale. To give an idea of scale, the earth's field is about 0.5 gauss, whereas the field of a strong permanent magnet is a few thousand gauss. Hall magnetometers are inexpensive, simple, small, and reliable. For example, the inexpensive TL173 from TI is a complete linear Hall-effect sensor in a 3-terminal plastic TO-92 package. You power it from +12 volts, and out comes a dc voltage that increases 1.5mV per gauss of applied magnetic field. Sprague also markets a line of linear Hall sensors (the UG3500/3600 series). The Hall effect is also used to make noncontacting keyboard and panel switches, as we remarked in Section 9.04.

A method with considerably greater roots into the past is the flip coil, a multi-turn coil of wire that is either rotated in the magnetic field at some fixed speed or simply pulled out; you measure the induced ac voltage or the integrated current, respectively. A flip coil is simplicity itself, and it has the elegance of pure electromagnetic theory, but it tends to be a bit bulky and old-fashioned looking.

For measurements of minuscule magnetic fields you can't beat the exotic SQUID (superconducting quantum interference device), a clever arrangement of superconducting junctions that can easily measure a single quantum of magnetic flux ( $0.2 \text{ microgauss-cm}^2$ ). A SQUID can be used to measure the magnetic fields set up in your body when you drink a glass of water, for whatever that's worth. These are fancy devices that require a considerable investment in cryogenic hardware, liquid helium, etc., and shouldn't be considered ordinary circuit items.

For precision magnetic-field measurements in the kilogauss range you can't do better than an NMR (nuclear magnetic resonance) magnetometer, a device that exploits the precession of nuclear (usually hydrogen) spins in an external magnetic field. This is the physicist's magnetometer, and it effortlessly yields values of magnetic field accurate to a part in a million or better. Since the output is a *frequency*, all the precision of frequency/time measurements can be used (more about this later).

Devices known as flux-gate magnetometers and transducers provide yet another way to measure magnetic fields. They work by exciting a piece of ferrite with an ac excitation field, with the response, as modified by the ambient field, being observed.

### 15.06 Vacuum gauges

The measurement of vacuum presents no great obstacles, which is fortunate, since it is a quantity of great importance in processes such as transistor and IC manufacture, thin-film evaporation, and the preparation of freeze-dried coffee. The basic device here is the Bayard/Alpert ionization gauge, which looks like an inside-out vacuum tube (Fig. 15.18). A hot filament emits electrons that are collected at a positive wire anode. Along the way the electron beam scatters from residual gas molecules, creating positive ions that are

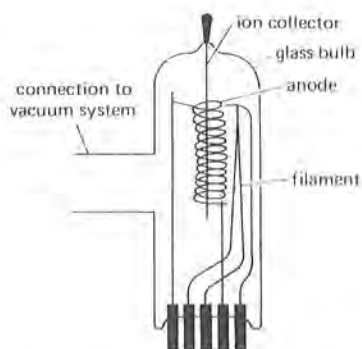


Figure 15.18. Ionization gauge.

collected at a central wire electrode held near ground. The ion current is accurately proportional to the density of gas molecules, i.e., the pressure. Ion gauges are usable at pressures (vacuums!) from about  $10^{-3}$  to  $10^{-11}$  mm Hg (1 mm Hg is also known as 1 torr; atmospheric pressure is 760 mm Hg). It takes great care to maintain a vacuum of  $10^{-10}$  mm Hg; even a fingerprint on the side of the chamber will frustrate your efforts.

At more mundane levels of vacuum (1 mm Hg down to  $1\mu\text{m}$  Hg, which you get with mechanical “roughing” pumps) the popular choice of measurement device is the thermocouple bonded to a small heater; you run some current through the heater, then measure the temperature with the thermocouple. Residual gas cools the contraption, lowering the thermocouple’s output voltage. Thermocouple gauges are usually used so that you know when it’s safe to turn on the high-vacuum (diffusion or ion) pumps. Granville-Phillips has an improved heat-loss gauge that works on similar principles. It’s called a “convec-tron” gauge and measures from  $1\mu\text{m}$  Hg all the way up to atmospheric pressure.

### 15.07 Particle detectors

The detection, identification, spectroscopy, and imaging of charged particles and energetic photons (x rays, gamma rays)

constitute an essential part of the fields of nuclear and particle physics, as well as numerous fields that make use of radioactivity (medical radiography tracers, forensic science, industrial inspection, etc.). We will treat x-ray and gamma-ray detectors first, then charged-particle detectors.

#### *X-ray and gamma-ray detectors*

The classic uranium prospector was a slightly grizzled and shriveled character who went poking around the desert with clicking Geiger counter in hand. The detector situation has now improved considerably. These detectors all have in common the property that they use the energy of an incoming photon to ionize an atom of something, giving off an electron via the photoelectric effect. What they do with the electron depends on the particular detector.

*Ionization chamber, proportional counter, Geiger counter.* These detectors consist simply of a cylindrical (usually) chamber, typically a few inches in size, with a thin wire running down the center. They are filled with some gas or mixture of gases. There’s a thin “window” on one side, made of some material that the desired x rays can penetrate (plastic, beryllium, etc.). The central wire is held at a positive potential and connected to some electronics. Figure 15.19 shows a typical configuration.

When an x ray enters, it ionizes an atom by ejecting a photoelectron, which then loses energy by ionizing gas atoms until it is brought to rest. It turns out that the electron loses about 20 volts of energy per electron-ion pair it creates, so the total free charge left after the photoelectron is brought to rest is proportional to the x ray’s initial energy. In an ionization chamber that charge is collected and amplified by a charge-sensitive (integrating) amplifier, just as with a photomultiplier. Thus, the output pulse is proportional to

the x ray's energy. The proportional counter works the same way but with the central wire held at a higher voltage, so that electrons drawn toward it cause additional ionization, resulting in a larger signal. This charge-multiplication effect makes proportional counters useful at low x-ray energies (down to a kilovolt, or less) where an ionization counter would be useless. In a Geiger counter the central wire is at a high enough voltage that *any* amount of initial ionization causes a single large (fixed-size) output pulse. This gives a nice large output pulse, but in the process you lose all information about the x ray's energy.

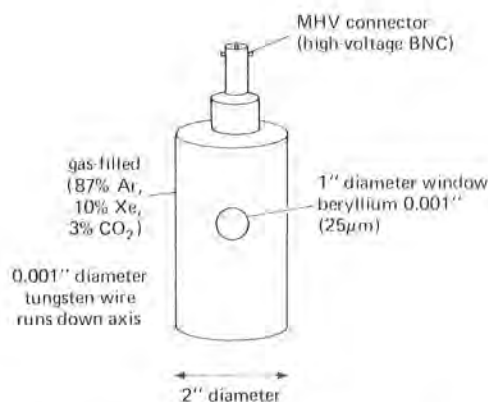


Figure 15.19. Proportional counter.

As you will see in Section 15.16, a clever device known as a pulse-height analyzer lets you convert an input stream of pulses of assorted heights into a histogram. If the pulse heights are a measure of particle energy, you wind up with an energy spectrum! Thus, with a proportional counter (but not with a Geiger counter) you are doing x-ray energy spectroscopy.

These gas-filled counters are usable in the energy range from about 1keV to about 100keV. Proportional counters have an energy resolution of about 15% at 5.9keV (a popular x-ray calibration energy provided by decay of iron 55). They're inexpensive and can be made in very large

or very small sizes, but they require a well-regulated power supply (the multiplication rises exponentially with voltage) and are not terribly fast (25,000 count/s is a rough practical maximum counting rate).

**Scintillators.** Scintillators work by converting the energy of the photoelectron, Compton electron, or electron-positron pair to a pulse of light, which is then detected by an attached photomultiplier. A popular scintillator is crystalline sodium iodide (NaI) doped with thallium. As with proportional counters, the output pulse is proportional to the incoming x-ray (or gamma-ray) energy, which means that you can do spectroscopy, with the help of a pulse-height analyzer (see Section 15.16). Typically, an NaI crystal will give an energy resolution of about 6% at 1.3MeV (a popular gamma-ray calibration energy provided by decay of cobalt 60) and is usable in the energy range of 10keV to many GeV. The light pulse is about  $1\mu\text{s}$  long, making these detectors reasonably fast. NaI crystals come in various sizes up to a few inches; they absorb water, though, so they have to be sealed. Since you must keep light out anyway, they're usually supplied in a metal package with thin aluminum or beryllium entrance window and integral photomultiplier tube.

Plastic (organic) scintillator materials are also popular, being very inexpensive. They have poorer resolution than sodium iodide and are used primarily at energies above 1MeV. Their light pulses are very short, roughly 10ns. Liquid scintillation "cocktails" are routinely used in biological studies. In such applications the material being examined for radioactivity is mixed into the scintillator cocktail, and the whole works is put into a dark chamber with a photomultiplier. You'll see handsome instruments in biology labs that automate the whole process, passing one vial after another through the counting chamber and recording the results.



**Solid-state detectors.** As with the rest of electronics, the great revolution in x-ray and gamma-ray detection has come about through advances in silicon and germanium technology. “Solid-state” detectors work just like the classic ionization chamber, but with the active volume filled with a nonconducting (intrinsic) semiconductor. An applied potential of about 1000 volts sweeps the ionization out, generating a pulse of charge. In silicon, an electron loses only about 2eV per electron-ion pair created, so many more ions are created for the same incident x-ray energy, as compared with a gas-filled proportional detector, giving better energy resolution through improved statistics. Other subtle effects also contribute to improved performance.

Solid-state detectors come in several varieties, Si(Li), Ge(Li), and intrinsic germanium, or IG (the first two are pronounced “silly” and “jelly”), according to the semiconductor material and dopants used to make it insulating. They are all operated at liquid-nitrogen temperature ( $-196^{\circ}\text{C}$ ), and the lithium-drifted types must be kept cold at all times (if allowed to warm up, they decay, permanently, with about the same time constant as fresh fish). Typical Si(Li) detectors come in diameters from 4mm to 16mm and are usable for x-ray energies from about 1keV to 50keV. Ge(Li) and IG detectors are used at higher energies, 10keV to 10MeV. Good Si(Li) detectors have energy resolutions of 150eV at 5.9keV (2.5%, six to eight times better than proportional counters), and the germanium detectors have energy resolutions of about 1.8keV at 1.3MeV (0.14%).

In order to illustrate what that extra resolution buys you, we bombarded a random hunk of stainless steel with 2MeV protons and measured the x-ray spectrum produced. This is called PIXE (proton-induced x-ray emission), and it is a powerful technique for determining spatially resolved trace-element distribution. Figure 15.20 shows the energy spectrum (made

with a pulse-height analyzer), with two x-ray lines visible for each element, at least with the Si(Li) detector. You can see iron, chromium, and nickel. A few additional elements are visible if you expand the lower part of the graph. With the proportional counter, all you get is mush.

Figure 15.21 shows the same kind of comparison for gamma detectors. This time it's an NaI scintillator versus a Ge(Li). (We ran out of steam, so we cribbed this one from the friendly folks at Canberra Industries. Many thanks, Mr. Tench.) As before, solid-state detectors win, hands down, for resolution.

Solid-state detectors have the best energy resolution of all the x-ray and gamma-ray detectors, but they have the disadvantage of a small active area in a large clumsy package (see Fig. 15.22 for an example), relatively slow speed ( $50\mu\text{s}$  or longer recovery time), high price, and high nuisance value (unless you enjoy being a full-time baby-sitter to a liquid-nitrogen guzzler).

### ***Charged-particle detectors***

The detectors we've just described are intended for energetic *photons* (x rays and gamma rays), not particles. Particle detectors have somewhat different incarnations; in addition, charged particles are deflected by electric and magnetic fields according to their charge, mass, and energy, making it much easier to measure particle energies.

**Surface-barrier detectors.** These germanium and silicon detectors are the analogs of the Ge(Li) and Si(Li) detectors. They don't have to be cooled, which simplifies the packaging enormously. (It also lets you take an occasional vacation!) Surface-barrier detectors are available in diameters from 3mm to 50mm. They are usable at particle energies of 1MeV to hundreds of MeV, and they have energy resolutions of 0.2% to 1% for 5.5MeV alpha particles (a



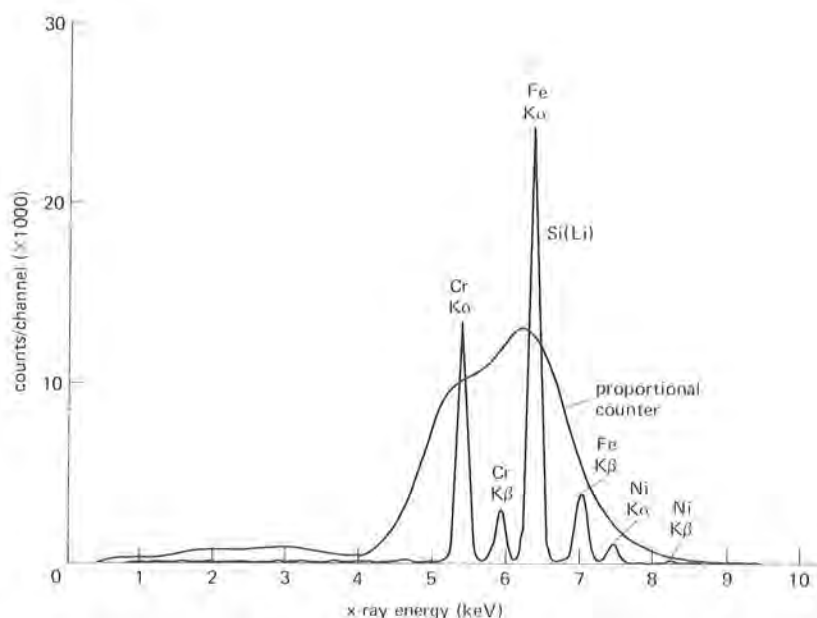


Figure 15.20. An x-ray spectrum from a piece of stainless steel, as seen by an argon proportional counter and a Si(Li) detector.

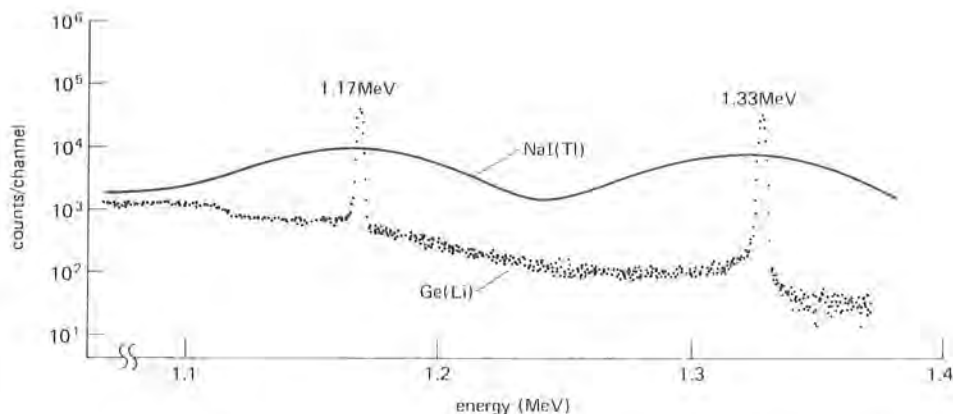


Figure 15.21. Cobalt-60 gamma-ray spectrum, as seen by a sodium iodide scintillator and a Ge(Li) detector. [From Canberra Ge(Li) Detector Systems brochure, Canberra Industries, Inc.]

popular alpha calibration energy provided by decay of americium 241).

- *Cerenkov detectors.* At very high energies (1GeV and above) even a heavy

charged particle can outrun light in material media, giving rise to Cerenkov radiation, a “visible sonic boom.” They are used extensively in high-energy physics experiments.

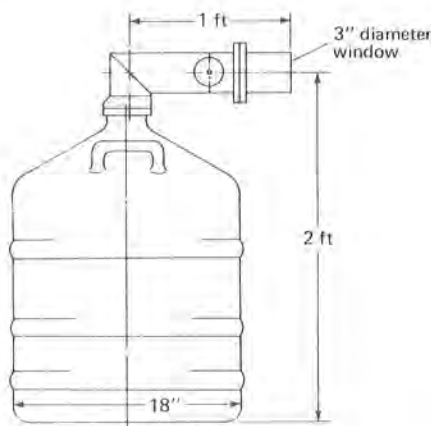


Figure 15.22. Ge(Li) cryostat. (Courtesy of Canberra Industries, Inc.)

- **Ionization chambers.** The classic gas-filled ionization chamber previously described in connection with x-ray detection can also be used as a detector of energetic charged particles. In its simplest form it consists of a single collecting wire running the length of an argon-filled chamber. Depending on the particle energies involved, the chamber may range from inches to feet in size; variations include the use of multiple collecting wires or plates, and other filling gases.
- **Shower chambers.** A shower chamber is the electron equivalent of an ionization chamber. An energetic electron enters a box of liquid argon, where it generates a "shower" of charged particles that are subsequently collected at charged plates. High-energy physicists like to call these things "calorimeters."
- **Scintillation chambers.** A charged particle can be detected with very good energy resolution by using photomultipliers to detect the ultraviolet-rich scintillations caused by the particle's ionized path in a chamber filled with argon or xenon in gas or liquid form. Scintillation chambers are

delightfully fast, in contrast with the more leisurely response of ionization and shower chambers.

- **Drift chambers.** These are the latest rage in high-energy physics, and they are made possible by advances in high-speed on-line computing. They're simple in conception: a box filled with gas at atmospheric pressure (an argon-ethane mixture is typical) and crisscrossed by hundreds of wires with an applied voltage. The box is full of electric fields, and when a charged particle goes in and ionizes the gas, the ions are swept out by the array of wires. You keep track of the signal amplitudes and timing on all the wires (that's where the computer comes in), and from that information you deduce the particle's path. With an applied magnetic field, that also tells you the momentum.

The drift chamber has become the universal imaging charged-particle detector for high-energy physics. It can deliver spatial resolution of 0.2mm or better over a volume large enough for you to climb into.

### 15.08 Biological and chemical voltage probes

In the biological and chemical sciences there are many examples of measurement wizardry: electrochemical methods such as electrochemistry with ion-specific electrodes, electrophoresis, voltametry, and polarography, as well as techniques like chromatography, IR and visible spectroscopy, NMR, mass spectroscopy, x-ray spectroscopy, nuclear quadrupole spectroscopy, ESCA, etc. It is hopeless in a volume this size to attempt any kind of comprehensive catalog of these sophisticated techniques. Furthermore, these techniques can be characterized as less fundamental than the direct physical measurements cataloged earlier in this chapter.

In order to give an idea of the special problems that arise in chemical and

biological measurements, we will describe only the simplest sort of measurement: the determination of the potentials generated by a microelectrode (used to explore nerve and muscle signals in biological systems), by an ion-specific electrode (used to measure the concentration of some specific ionic species in solution), and by a voltametric electrochemical probe. As usual, there are some interesting electronic challenges you face, if you want to get anything meaningful out of your measurement.

### Microelectrodes

In order to look at the voltage on nerves or in the interiors of cells, it is standard practice to make electrodes that are just a few hundred angstroms in tip diameter (1 angstrom =  $10^{-8}$  cm, approximately the size of a hydrogen atom). That turns out to be easily done by drawing a glass capillary, then filling it with a conductive solution. You wind up with a nice probe, but with interesting circuit problems arising from the electrode's source impedance of  $100\text{M}\Omega$  or more. Interference pickup, loading by the circuit, and high-frequency rolloffs of a few hertz due to cable and stray capacitances plague the unwary.

In order to see nerve or muscle signals you want to have decent high-frequency performance, at least out to a few kilohertz, or so (this isn't exactly high frequency in the sense of Chapter 13!). The amplifier must have very high input impedance, and preferably low input noise. In addition, it must be insensitive to common-mode interference.

The circuit in Figure 15.23 represents a good solution. The use of a reference electrode connected near the point of actual measurement keeps interference from appearing as normal (differential) mode signals. The inputs are buffered as close as possible to the actual microelectrode by low-noise FET-input op-amps  $\text{IC}_1$  and  $\text{IC}_2$ , which also bootstrap the guard elec-

trode in order to reduce the effective cable capacitance. Note that the guard is itself shielded. You've got to use FET amplifiers in order to get high input impedance and low input current noise; the particular types shown were chosen for their low input noise voltage ( $2.5\mu\text{V}$  pp max 0.1–10Hz), often a problem with FET or MOSFET amplifiers. The pair of buffered signals is applied to the standard differencing amplifier configuration with a low-noise-voltage low-drift op-amp, with 100mV of stable adjustable output offset added via  $\text{IC}_6$ .

At this point you have an amplifier with a differential gain of 10, suitable noise performance, good common-mode rejection, and low input current ( $<1\text{pA}$ ). However, even with input guarding, the residual input capacitance at the input buffers and at the microelectrode tip will result in poor speed performance. For example, a  $100\text{M}\Omega$  source impedance driving 20pF has a high-frequency 3dB point of only 80Hz. The solution is active compensation via positive feedback, provided by  $\text{IC}_3$  and  $\text{IC}_4$  through  $C_1$  and  $C_2$ . In practice, you adjust the voltage gain of amplifiers  $\text{IC}_3$  and  $\text{IC}_4$  for decent high-frequency performance (or transient response), with response to several kilohertz possible.

### Ion-specific electrodes

The classic example of an ion-specific electrode is the pH meter, which measures the voltage developed between a reference electrode and a thin-walled glass electrode through which hydrogen ions can diffuse. Once again you're dealing with very high source impedances, although the problems here are less severe than with microelectrodes, because you don't often care about frequency response.

There are more than 20 kinds of ion-specific electrode systems available, e.g., to measure activities of  $\text{K}^+$ ,  $\text{Na}^+$ ,  $\text{NH}_4^+$ ,

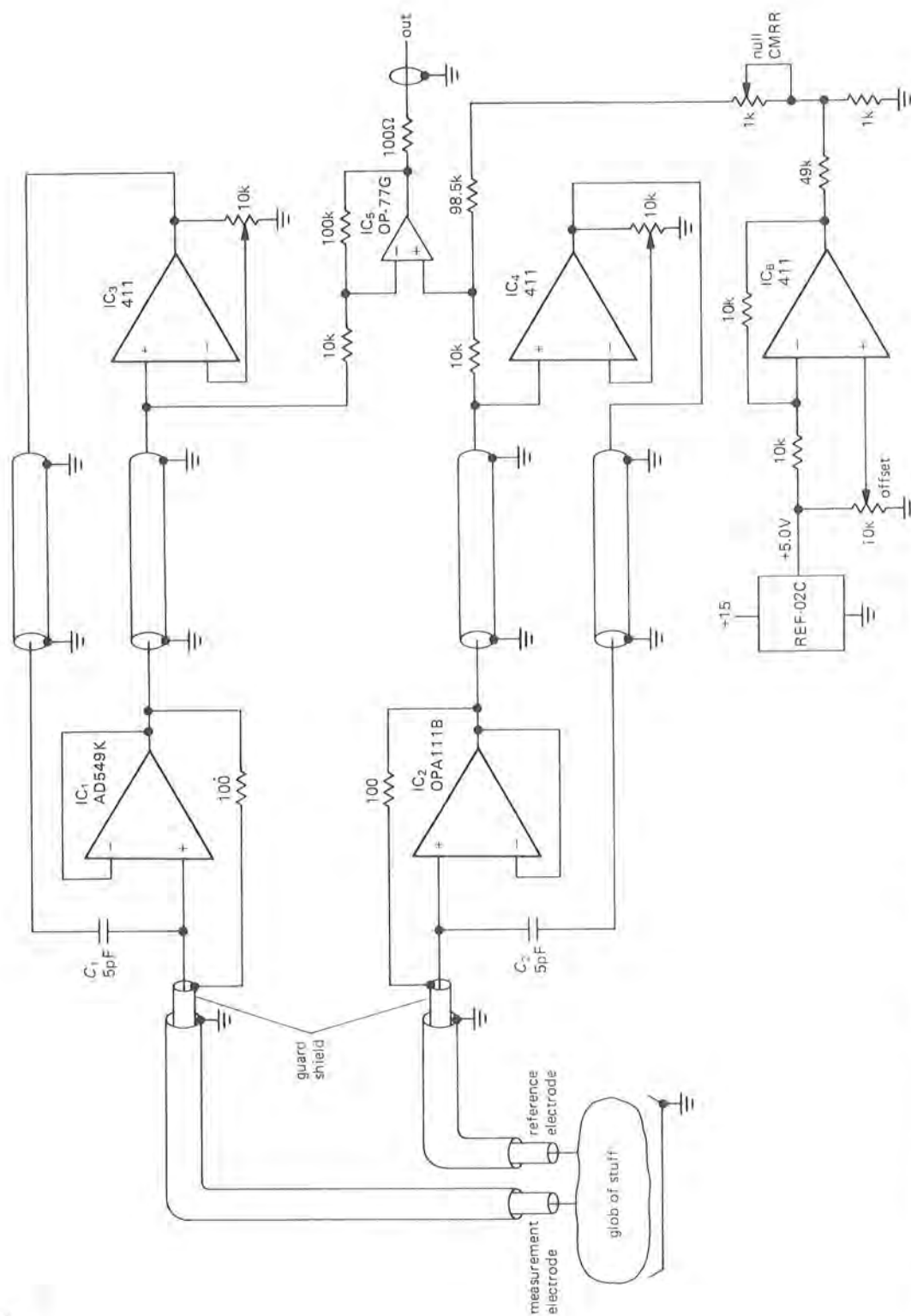


Figure 15.23. Compensating microelectrode amplifier with guarding and reference channel.

$\text{CN}^-$ ,  $\text{Hg}^{++}$ ,  $\text{SCN}^-$ ,  $\text{Br}^-$ ,  $\text{Cl}^-$ ,  $\text{F}^-$ ,  $\text{I}^-$ ,  $\text{Ca}^{++}$ , or  $\text{Cu}^{++}$ . In general, you have two electrodes: a reference electrode, typically silver-coated with silver chloride and immersed in a concentrated solution of potassium chloride that communicates with the solution you want to measure via a porous plug or gel, and an ion-specific electrode, typically consisting of an electrode immersed in a concentrated solution of the ion you're interested in and separated from the solution under test by a membrane that is selectively permeable to the ion of interest. The membrane is commonly an ion-selective glass or an organic liquid containing mobile ion-transporting organic molecules. Your task is to measure a voltage that is in the range of 0 to 2 volts, with an accuracy of a millivolt, while drawing less than 100pA. The situation is complicated by a temperature coefficient of as much as a few percent change in voltage per degree centigrade, which you can attempt to cancel automatically with thermistor-driven compensation circuitry. Conversion from measured ionic activity to concentration requires attention to the total ionic strength of the sample and to the crossover sensitivity of the ion-specific electrode to other ions present. In any case, chemists say you get best results with this sort of black art if you calibrate on some standard solutions just before and after making your measurements. With care, you can see concentrations of 0.1ppm and achieve measurement accuracies of about 1% in solutions of moderate concentration.

### **Electrochemical measurements**

In the area of electrochemistry, it is possible to make very sensitive analytical measurements of the concentrations of specific ions by measuring electrode currents (reaction rates) versus applied voltage in a solution. By scanning the applied voltage, you pass through the potentials at which specific reactions occur, giving rise to steps or

peaks. Terms such as cyclic voltametry, polarography, and anodic stripping voltametry (ASV) are used to describe various ways of doing such analytical measurements. Among the most sensitive of these techniques is ASV, which uses a hanging drop mercury (hdm) electrode, a renewable electrode onto which you electroplate at a relatively high potential for a while, then reverse the current and strip off each element sequentially. This technique can detect elements like lead and cadmium at the parts-per-billion level, and it should be considered on a par with other trace-element techniques such as neutron activation, flame spectroscopy, and x-ray and ion microprobes.

The technique of measuring a small current while subjecting a system to a fixed voltage is called a "voltage clamp," and it finds application also in nerve and cell physiology. Nerve membranes have voltage-dependent channels through which specific ions can diffuse, and nerve physiologists like to measure the voltages at which such channels open. Again, voltage clamps are used, this time with microelectrodes.

In preparative electrochemistry, the same techniques are used, but with currents measured in amperes rather than microamperes. Once again the idea is to drive a specific reaction product by applying the right voltage.

Figure 15.24 shows a simple potentiostat (or voltage clamp) circuit. The electrolytic cell consists of an electrode to inject current (the counter electrode), a common return electrode (the working electrode), and a small probe to measure the voltage in the solution near the working electrode (the reference electrode). IC<sub>1</sub> maintains a voltage equal to  $V_{\text{ref}}$  between the reference and working electrodes by varying the current into the counter electrode appropriately (in measurements of membrane potential, the upper two electrodes would be inside the cell, and the working electrode outside). IC<sub>2</sub> holds the



working electrode at virtual ground, converting the current to an output voltage. The range of voltages encountered is typically  $\pm 1$  volt; currents in the range of  $1\text{ nA}$  to  $1\text{ mA}$  are typical of analytical measurements, whereas currents of  $1\text{ mA}$  to  $10\text{ amps}$  are used in preparative electrochemistry.

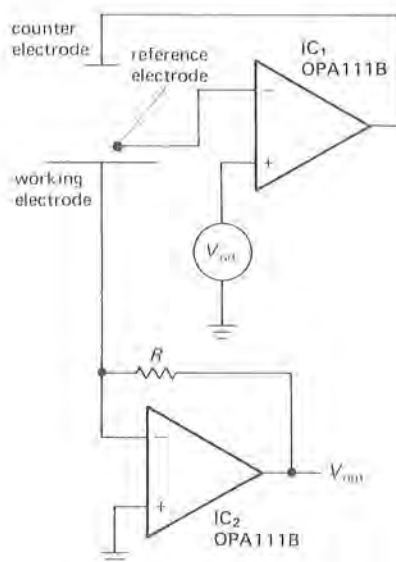


Figure 15.24. Potentiostatic electrochemistry circuit (voltage clamp).

In order to do scanning measurements,  $V_{\text{ref}}$  would be replaced with a ramp generator. For low-current membrane measurement you would have to shield the input leads carefully, and perhaps use guarding and positive feedback with a capacitor in the manner of Figure 15.23 to maintain some frequency response.

## PRECISION STANDARDS AND PRECISION MEASUREMENTS

In Chapter 7 we talked about the circuit methods that are necessary in high-accuracy instruments to maintain small voltage offsets and drifts, e.g., when amplifying very small voltages. We dealt there

only with analog electronics, the amplification of continuously varying voltages and currents. For a number of reasons, it turns out that measurements of digital quantities such as frequency, period, and time intervals can be made to far greater precision than any analog measurement. In the following sections, we will explore the accuracy of electronic standards (time, voltage, resistance), and you will see how to make analog measurements of high precision using such standards as references. We will devote the bulk of the discussion to time/frequency measurements, because of their greater inherent precision and because we have already treated precision analog circuitry in some detail in Chapter 7.

### 15.09 Frequency standards

Let's take a look at the high-stability frequency standards you can get, and then discuss how you set and then maintain their frequency.

### Quartz-crystal oscillators

Back in Sections 5.12–5.19, we described briefly the stabilities you can expect from frequency standards, going from the simplest  $RC$  relaxation oscillator to the atomic standards based on rubidium and cesium. For any serious timing you wouldn't consider anything less stable than a quartz-crystal oscillator. Fortunately, the cheapest crystal oscillators cost only a few dollars and can deliver stabilities of a few parts per million. For about \$50 you can buy a good TCXO (temperature-compensated crystal oscillator), stable to one part per million from  $0^\circ\text{C}$  to  $50^\circ\text{C}$ . For better performance you need ovenized crystals, with price tags from a couple of hundred dollars to more than \$1000. Once you begin talking about stabilities of a few parts per billion, you have to worry also about "aging," the tendency of crystal oscillators to

drift in frequency at a more or less constant rate once they are initially "broken in." The 105B from Hewlett-Packard typifies good crystal oscillators with stability of better than 2 parts per billion over the full temperature range and aging rates of less than 0.5 part per billion per day.

Uncompensated crystal oscillators, and even TCXOs, are logical choices as part of a small instrument. The fancier ovenized oscillators are usually rack-mounted standards with an identity all their own.

□ **Atomic standards**

There are three standards in use today: rubidium, cesium, and hydrogen. Rubidium has a microwave absorption at 6,834, 682,608Hz, cesium has an absorption at 9,192,631,770Hz, and hydrogen has an absorption at 1,420,405,751.768Hz. A frequency standard based on one of these is considerably more complicated (and expensive) than a good crystal oscillator.

- **Rubidium.** In the rubidium standard you have a glass bulb with rubidium vapor, heated and contained in a microwave cavity with glass end windows. A rubidium lamp shines through the cavity, with a photocell detecting the transmitted light. Meanwhile, a modulated microwave signal referenced to a stable crystal oscillator is introduced into the cavity. By using lock-in detection (see Section 15.15) of the transmitted light, you can bring the microwave signal exactly to the rubidium resonance frequency, since the optical absorption of the rubidium gas is altered when its microwave resonance is excited. The crystal's frequency is then related in a known way to the rubidium resonance, so it is straightforward to generate a standard frequency like 10MHz. (There are actually several additional complications that we have glossed over.)

Rubidium standards have better stability than ovenized crystal oscillators, although they do exhibit a form of aging.

Commercial units are available with stabilities of a few parts in  $10^{11}$  over the full temperature range and long-term stabilities of 1 part in  $10^{11}$  per month. Rubidium standards make sense in a laboratory situation, and you find them at observatories and other places where extremely accurate observations are made. It should be pointed out that a rubidium standard, just like a crystal oscillator, must be calibrated, because changing conditions within the resonance cell affect the frequency at the part-per-billion level.

- **Cesium.** A cesium standard is practically a small atomic-beam laboratory, in which cesium atoms are launched from an oven into a vacuum chamber, where they pass through spin-state selector magnets and oscillatory electric fields before being detected with a hot-wire ionization detector. As with the rubidium standard, a microwave signal referenced to a stable crystal oscillator is locked to the resonance with feedback from a phase-sensitive detector, and the output frequency is synthesized from the crystal.

Cesium standards aren't small, and they aren't cheap. But they are *primary* standards; you don't have to calibrate them. In fact, by international agreement, cesium *defines* the second: "the duration of exactly 9192631770 periods of the radiation corresponding to the transition between the two hyperfine levels of the ground state of the cesium-133 atom." Cesium clocks are used to keep official time in this country and to calibrate time transmissions (more on this shortly). The cesium clocks used to keep time are elaborate devices, but even commercially built cesium standards keep exceptional time: long-term stability and reproducibility of 3 parts in  $10^{12}$  for the model 5061B from Hewlett-Packard (priced at \$32,500).

- **Hydrogen.** Neutral hydrogen atoms have a hyperfine resonance at about 1420MHz, and in contrast to the situation

with the other atomic standards, it is possible to make an actual oscillator with them. As with cesium, you make an atomic beam and run it through magnetic state selectors, then into a Teflon-coated quartz bulb in a microwave cavity. The atoms bounce around inside this "storage bulb" for about 1 second and give off enough radiofrequency energy to sustain an oscillation in the cavity. That makes it easy to lock a crystal oscillator, using PLLs and mixers. You call this object a hydrogen *maser* (micro-wave amplification by stimulated emission of radiation).

Hydrogen masers are extraordinarily stable over short times (up to a few hours), with stabilities of 1 part in  $10^{15}$ . They have not replaced cesium-beam apparatus for primary timekeeping, however, because the problem of determining the frequency-pulling effect of the cavity has not been solved, and because of long-term drifts caused by the changing properties of the storage-bulb wall surface. R. Vessot at the Smithsonian Astrophysical Observatory (Cambridge, MA) is the world's leader in hydrogen clocks, having built more than two dozen; his timekeepers cost \$0.5M apiece.

□ *Methane laser.* A fourth atomic standard is used at infrared wavelengths, namely the methane-stabilized helium/neon laser. It has a frequency stability comparable to that of the other atomic standards, but at its frequency of  $8.85 \times 10^{13}$  Hz ( $3.39\mu\text{m}$  wavelength) it is not a usable radiofrequency standard.

□ *Recent developments.* The latest research in stable frequency standards involves two promising areas: "trapped ions" and cryogenic hydrogen masers. Proponents of each talk about ultimate stabilities of parts in  $10^{18}$ , if all goes well.

#### □ *Calibrating a clock*

Unless you happen to own a cesium-beam standard, you've got to have access to a stable calibration signal to keep your oscillator on frequency. In addition, you may wish to keep accurate absolute time as well as frequency, i.e., you have to set your clock, after you have it running at the right rate. There are several services to help you keep time. On the East Coast of the United States, and in several other areas, you can receive Loran-C, a navigational signal at 100kHz, from which you can determine frequency and time. Loran-C is generated by cesium clocks and is compared with the cesium-beam master clock at the Naval Observatory, which publishes corrections each month. Another time service is WWVB, from the National Institute of Standards and Technology (formerly NBS) in Colorado. This is a 60kHz signal that you can receive most anywhere in the United States. For both these low-frequency transmissions you can synchronize to  $1\mu\text{s}$  or better if you are within range of the "ground-wave" signal (a few hundred miles), but ionospheric effects (day/night shifts, solar winds, etc.) make synchronization via the "sky wave" less accurate ( $10\text{--}50\mu\text{s}$ ). A more recent network known as Omega transmits at very low frequencies (around 10kHz) and can be received anywhere, although with an accuracy of only  $10\mu\text{s}$  or so. The geostationary weather satellites known as "GOES" transmit UHF timing signals (at 469MHz) that you can use for time synchronization to a millisecond, if you are in the "footprint" of one of them (all of North and South America).

If you can receive one of these time services, you can compare your oscillator frequency with the real thing. There are nice commercial gadgets that will take care of all the fuss and bother and even generate pretty graphs of the results. It is a bit more difficult to set your clock's time. The

most reliable way is to carry it (or some portable clock) to one of the standards, set it, then carry it back. As soon as you get home you make observations of Loran-C, or whatever, to determine the time delay from the transmitter to you. Save that number! (We still remember the magic number  $53,211\mu\text{s}$  for the delay from Loran-C in Nantucket to Harvard's 60 inch telescope dome.) As long as no one builds a new mountain between you and the transmitter, you're all set to tell time.

The recent Global Positioning System (GPS, or "NAVSTAR") is a constellation of 21 satellites, in 12 hour orbits of high inclination, containing on-board atomic clocks. When fully operational, it will permit the determination of time (to 20 nanoseconds) and location (to 10 meters) anywhere on earth, using a self-contained "smart" GPS receiver with a small L-band (1.2GHz and 1.6GHz) doorknob-shaped antenna. The incomplete GPS system has already been used to synchronize clocks worldwide to better than 50ns. The complete GPS system could ultimately be used to achieve time transfer at the 2ns level, if a set of proposed upgrades is adopted.

### 15.10 Frequency, period, and time-interval measurements

With an accurate reference oscillator and just a small amount of digital electronics, it is disarmingly easy to make frequency and period measurements of high precision.

#### Frequency

Figure 15.25 shows the basic circuit of a frequency counter. A Schmitt trigger converts the analog input signal to logic levels, at which point it is gated by an accurate 1 second pulse derived from a crystal oscillator. The frequency in hertz is the number of pulses counted by the multidigit BCD counter. It is best to latch

the count and reset the counter between counting intervals.

In practice, you would arrange the clock circuit so that shorter or longer intervals can be selected, with a choice of 0.1 second, 1 second, and 10 seconds as a minimum. Also, you can eliminate the 1 second interval between measurements. Additional features might include the following: an adjustable preamp, with selectable trigger point and hysteresis, and perhaps a front-panel output from the discriminator so you can see the trigger point on an oscilloscope; BCD output for readout into a computer or logger; provision for an external oscillator, when a precision standard is available; a manual start/stop input for simple counting (totalization).

□ *Microwave counting.* You can go to frequencies of 3GHz with the digital ICs available today. In particular, GigaBit Logic manufactures a series of astoundingly fast GsAs ripple counters with guaranteed 3GHz clock rate. For higher frequencies you can use heterodyne techniques to mix the microwave input signal down to a directly countable frequency, or you can use a so-called transfer oscillator technique, in which you phase-lock the  $n$ th harmonic of a VCO to the input signal, then measure the VCO frequency and multiply the result by  $n$ .

□ *The  $\pm 1$  count ambiguity.* One disadvantage of this simple frequency counting scheme is that low frequencies cannot be measured to high precision, because of the  $\pm 1$  count error. For example, if you were to measure a signal near 10Hz with a gate time of 1 second, your answer would be accurate only to 10%, since the result would be either 9, 10, or 11. You could measure for a longer interval, but it would take a whole day's counting to get the relative accuracy (1 part per million) that you would get in 1 second when measuring



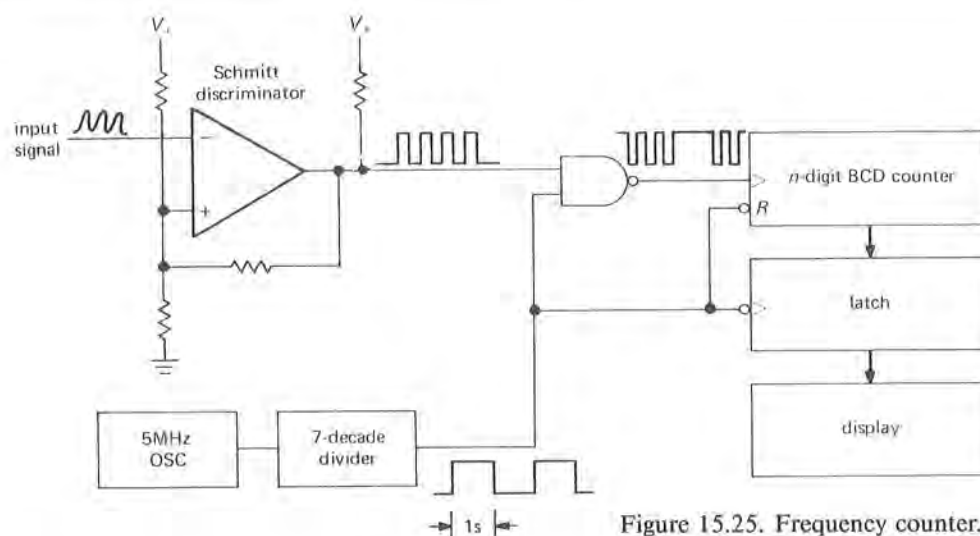


Figure 15.25. Frequency counter.

a 1MHz signal, for example. There are several solutions to this problem: period (or reciprocal) counting, interpolation methods, and phase-locked-loop frequency multiplication techniques. We will deal with the first two in the next sections, since they aren't really direct frequency measurements.

Figure 15.26 shows the PLL "resolution multiplication" technique. A standard

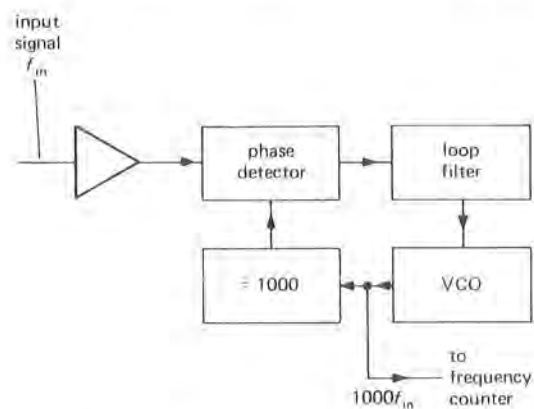


Figure 15.26. PLL resolution multiplication for low-frequency counting.

phase-locked loop is used to synthesize a frequency of 1000 times the input signal,

say, which is then counted as described earlier. The accuracy of this technique is limited by the phase jitter in the PLL phase detector and the loop compensation parameters. For example, if a 100Hz signal is multiplied by 1000 and counted for 1 second, and the jitter in the phase detector is 1% of a cycle ( $3.6^\circ$ ), or  $100\mu\text{s}$ , then the *accuracy* of the measurement will be 1 part in 10,000, even though the *resolution* is 1 part in 100,000.

We will now mention two other ways to improve frequency-measurement accuracy: period measurement and interpolated time-interval measurement.

### Period (reciprocal counting)

A good way to handle the problem of resolution when measuring low frequencies is to turn things around and use the input signal (or some subdivision of it) to gate the clock. Figure 15.27 shows the standard configuration for such a period counter. The number of periods measured is normally switch-selectable to some power of 10 (1, 10, 100, etc.). You will usually pick a number of periods such that the measurement takes a convenient length of time, typically a second, giving an answer to about seven significant figures. Of course,



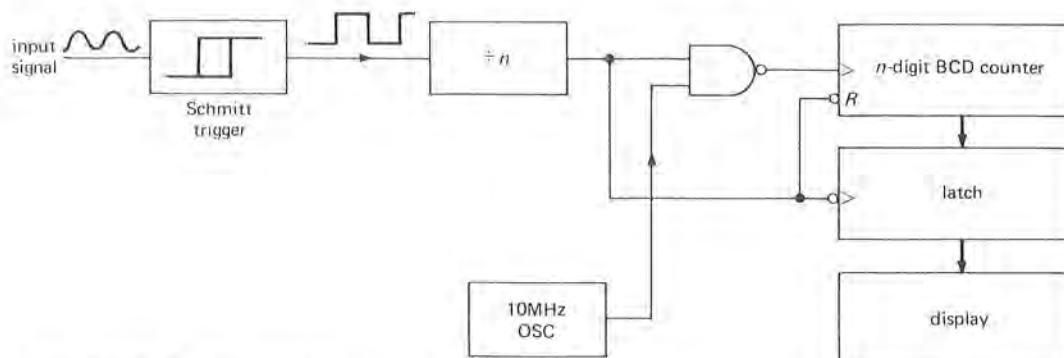


Figure 15.27. Period counter.

that answer is in units of time, not frequency, so you have to take the reciprocal to recover the frequency. Luckily, you soon won't even have to know how to divide, since modern counters use dedicated microprocessors to do the period-to-frequency conversion.

Note that the accuracy of period measurement is critically dependent on stable triggering and requires good signal/noise ratios. Figure 15.28 indicates the problem here.

The main advantage of reciprocal counting is that you get a constant resolution  $\Delta f/f$  for a given length of measurement, independent of the input frequency. The graph in Figure 15.29 compares the resolution of frequency and reciprocal frequency (period) measurements of duration 1 second, using a 10MHz clock. The period graph should actually be somewhat jagged, since you normally have to live with the closest power of 10 for the number of periods averaged. Even this restriction is evaporating, with the advent of "smart" microprocessor counters (e.g., the low-cost Hewlett-Packard 5315A) that have *continuous* adjustment of gate time; they know how many periods were averaged, and they divide the answer accordingly. They also oblige you by switching from period to frequency mode for input frequencies greater than the clock frequency, in order to get optimum resolution at any input frequency.

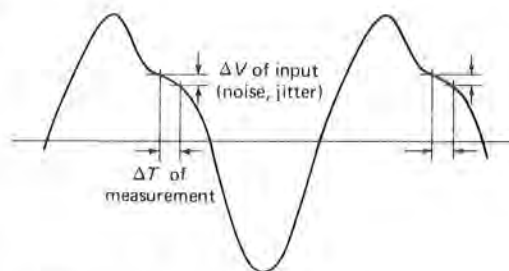


Figure 15.28

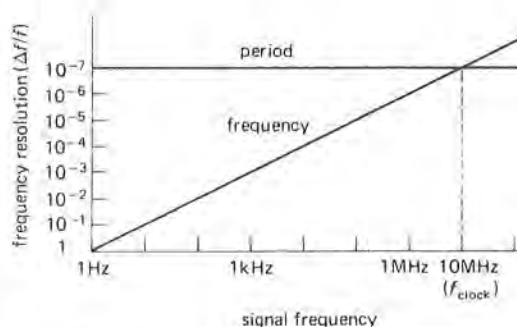


Figure 15.29. Fractional resolution of frequency and period counting.

A second advantage of reciprocal frequency measurement is the ability to control externally the time at which the gating occurs. This is advantageous if you wish to measure the frequency of a tone burst, for example, a situation in which a simple frequency counter would give incorrect

results, since its internally controlled gating interval might not coincide with the burst. With period counting you can gate the measurement externally and can even make a set of measurements at various points along the burst, given the generally superior resolution of period measurement.

You might wonder if it is possible to do better than the “uncertainty principle” resolution limit of  $\Delta f/f \approx 1/f_{\text{clock}}T$  (period measurement) or  $1/f_{\text{input}}T$  (frequency counting), for the relative error  $\Delta f/f$  of a frequency measurement made by counting for time  $T$ . The answer is yes. In fact, several clever schemes have been invented. We will discuss them in the next subsection (time-interval measurement), but just to show that it can be done, we’ve drawn in Figure 15.30 a method of measuring the frequency of a 1MHz oscillator to a resolution of 1 part in  $10^{12}$  in 1 second of measurement time. The unknown oscillator is mixed with a stable reference offset slightly from 1.0MHz, say 1.000001MHz (this could be synthesized with a PLL). The mixer output contains the sum and difference frequencies. After low-pass filtering, you’ve got a 1Hz signal that is the difference between the two oscillators, and that can be easily measured with a period counter to one part per million in 1 second. In other words, you’ve measured 1MHz to 1 $\mu$ Hz in 1 second.

This technique assumes that you have extremely good signal/noise ratios, and in practice you would have to worry about low-frequency noise, settling time of the filter, etc., so you might not do better than 1 part in  $10^{10}$  in 1 second. Still, this is considerably better than simple frequency (or period) counting. In addition, the *accuracy* will be less than the resolution unless the reference oscillator is also accurate to 1 part in  $10^{12}$  (possible, but not easy, with today’s technology). You can think of this scheme as a way of comparing the relative frequencies of two oscillators, if you prefer.

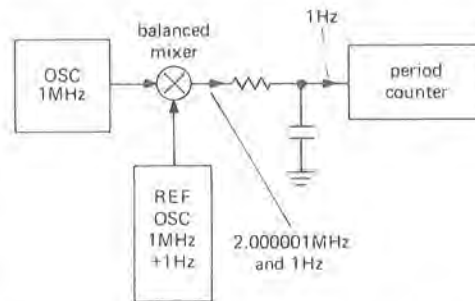


Figure 15.30. High-resolution frequency comparison.

### Time-interval measurement

With a trivial change in the circuitry of the period counter, you can measure the time interval between two events. Figure 15.31 shows how. In practice it may be better to add a synchronizer, as shown in the second circuit, to prevent the generation of runt pulses. The best resolution is obviously obtained by running the oscillator at the highest possible frequency, and commercial counters use local oscillator references as high as 500MHz, phase-locked from a stable crystal at 5MHz or 10MHz. With a 500MHz reference, you have a resolution of 2ns.

As we hinted earlier, there are ways to beat the reciprocal frequency resolution limit when making time-interval measurements, essentially by exploiting the extra information you have about the position of zero crossings of the input signal relative to the reference. The oscillator comparison scheme we showed earlier really exploited that same information, but in a more subtle way. For these schemes you must have a clean signal with very low noise level. There are two interpolation methods in use in commercial instruments: linear interpolation and vernier interpolation.

- **Linear interpolation.** Suppose you wish to measure the time interval between the start and stop pulses in Figure 15.32. You begin by measuring the number of

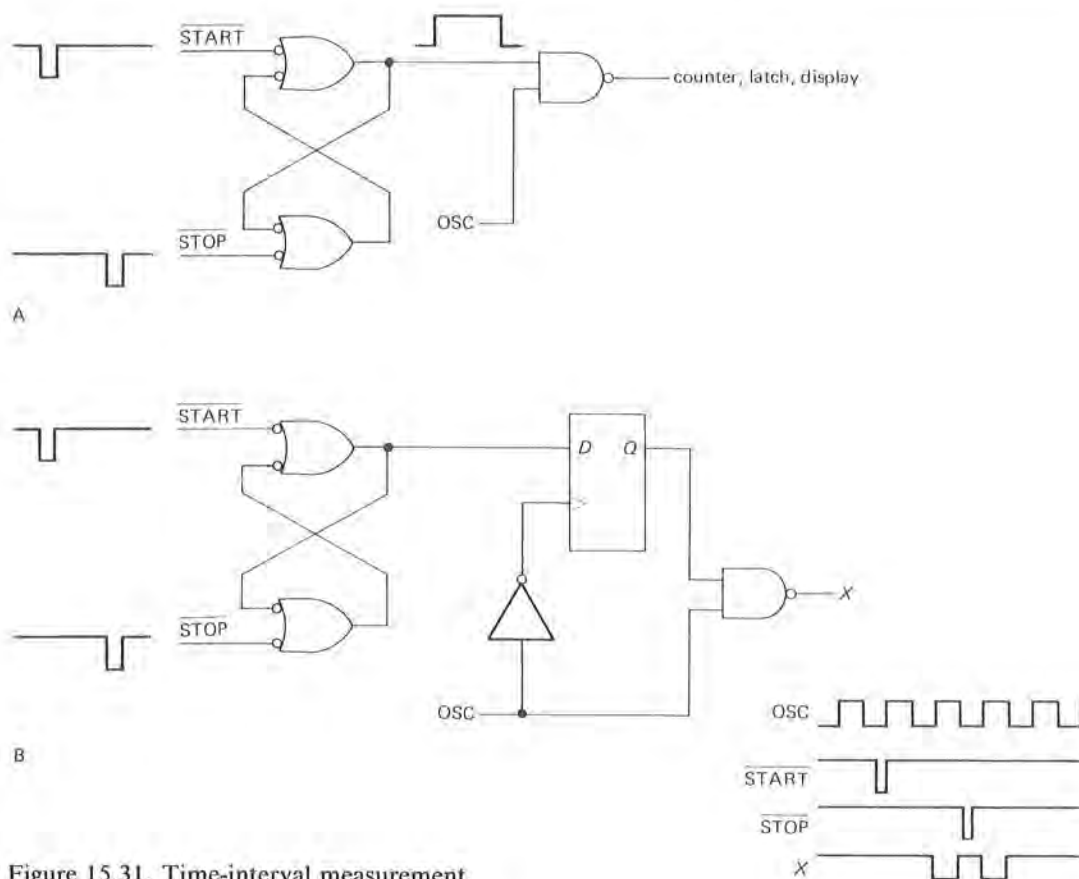


Figure 15.31. Time-interval measurement.

clock impulses,  $n$ , during the interval  $\tau$ , as shown (with a synchronizer you would start and stop with the first clock pulse after the respective input signal, as shown). To improve the resolution of the measurement, all you need to know are the time intervals  $T_0$  and  $T_1$ , the time elapsed from the occurrence of each input pulse to the next clock pulse. Assuming you're already running the system clock at the highest rate you can conveniently count, you have to expand those unknown intervals in order to measure them. A variation of the dual-slope principle works here: Integrate charge onto a capacitor during those intervals, then ramp down at a small fraction (say 1/1000th) of the charging rate, thus expanding the unknown time intervals by a

factor of 1000. During those expanded intervals, you count the system clock, generating counts of  $n_0$  and  $n_1$ . The unknown time interval is therefore given by

$$\tau = T_{\text{clock}} \times (n + n_0/1000 - n_1/1000)$$

with the obvious improvement in resolution. The ultimate accuracy of this method is limited by the accuracy of the interpolators and the system clock. An example of this kind of instrument is the Hewlett-Packard 5334B counter, which displays 9 digits (frequency or time) per second of gate time.

□ *Vernier interpolation.* Vernier interpolation is a digital technique that lets you

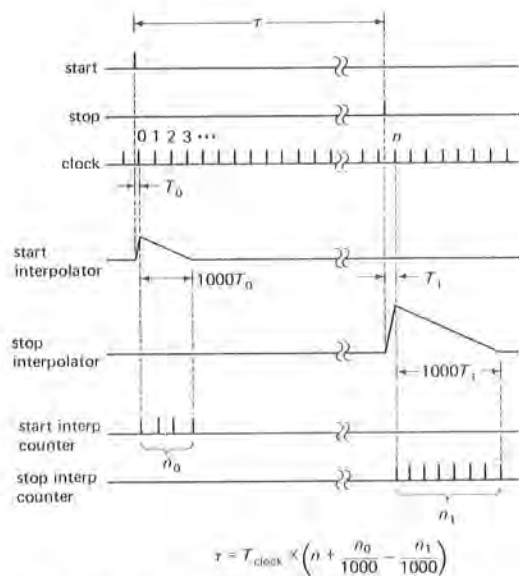


Figure 15.32. Linear interpolation (time-interval measurement).

find out where in the clock cycle the input pulse occurred. Figure 15.33 shows the method schematically. There are three clocks involved: The master reference clock runs continuously, with a period  $T_0$  of, say, 5ns; the input START pulse triggers a second oscillator with a period greater than the reference by a factor  $1 + 1/n$  (we've set  $n = 16$  for this example); the input STOP pulse triggers a third oscillator of the same period as the other triggered oscillator. Fast circuitry then looks for coincidences between the triggered oscillators and the master clock, while counting the number ( $n_1, n_2$ ) of ticks of each before coincidence. The arithmetic is shown in the figure; the net result is to determine to within  $1/n$ th of a master clock pulse the duration between START and STOP.

The Hewlett-Packard 5370B uses this technique, with  $T_0 = 5\text{ns}$  and  $n = 256$ . The result is time-interval resolution of 20ps. The technique can clearly be used for period measurements, since period is simply the time interval for one cycle of an input

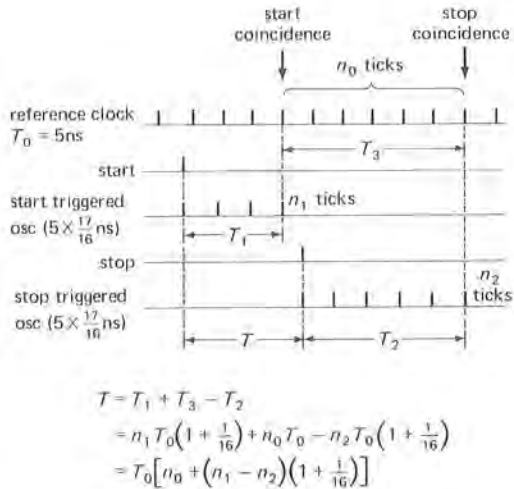


Figure 15.33. Vernier interpolation (time-interval measurement).

wave. When used this way, the counter just mentioned determines frequency to 11-digit resolution in 1 second!

- **Time-interval averaging.** There is a third way to improve the resolution of a time-interval measurement, namely by repeating the measurement many times and taking the average. The  $\pm 1$  count ambiguity gets averaged that way, and the result converts to the true interval, provided only that the repetition rate of START pulses is not commensurate with the master clock. Some counters include a "jittered clock" to make sure this doesn't happen.

### Spectrum analysis

A powerful technique that must be mentioned in connection with the measurement of frequency is *spectrum analysis*, looking at signals in the frequency domain. Spectrum analyzers can measure frequency (and in fact they are very useful when you need to know the frequency of a weak signal in the presence of other stronger signals), but in addition they can do a lot more. We will talk about them in Section 15.18.

□ 15.11 Voltage and resistance standards and measurements

As we hinted earlier, analog standards and measurements do not have anything like the precision we have just been talking about. Here you're lucky to get accuracy of a part per million. The analog standards are voltage and resistance; from them you can determine current, if need be.

The traditional standard of voltage is the Weston cell, an electrochemical device with reproducible output voltage, intended for use as a reference only (no more than  $10\mu\text{A}$ , and preferably no current at all, should be drawn from it). Its terminal voltage is 1.018636 volts at  $20^\circ\text{C}$ . Unfortunately, Weston cells are fussy gadgets. They have to be maintained at a precise temperature because of their large temperature coefficient ( $40\mu\text{V}/^\circ\text{C}$ , far worse than good IC voltage references) and even larger sensitivity to temperature gradients (the individual "limbs" of the cell have tempcos of about  $350\mu\text{V}/^\circ\text{C}$ ). Standard cells are carefully maintained by the National Institute of Standards and Technology for comparison with secondary standards. Nowadays there are very stable solid-state references with controllable output voltage. They can be used to transfer a measurement from a finicky standard cell to an actual measurement situation. Typical specifications are 10 ppm stability for a month after calibration and 30 ppm stability in a year.

To make an accurate voltage measurement, you use precision voltage dividers (known as Kelvin-Varley dividers), available with linearities in the 0.1ppm range. The divider is used to generate a precise fraction of the unknown voltage, for comparison with the voltage standard. Accurate null detectors and instruments for compensation of wiring resistance are available for this task. Routine calibrations at accuracies of just a few parts per million are possible.

Recently a measurement based on a superconducting Josephson junction has replaced the standard cell as the definition of voltage. With care it is possible to measure voltages reproducibly to a few parts in  $10^{10}$ . The method has the pleasant simplicity of requiring only a measurement of a frequency and knowledge of the physical constants  $h$  (Planck's constant) and  $e$  (the electron charge). Although Josephson junction technology has traditionally been considered too complicated for use as a voltage standard, the situation is changing: The National Institute of Standards and Technology (NIST, formerly NBS) has developed an affordable chip containing 19,000 series junctions, which can go to 10 volts and beyond. Serious users of voltage standards can now have their own precision Josephson standard, although at a cost of \$100k. If recent breakthroughs in high-temperature superconductivity can be harnessed to make voltage standards, Josephson voltage references could become accessible to every laboratory.

As with voltage, standards of resistance are carefully maintained by NIST. By using such standards in a Wheatstone bridge circuit, you can calibrate a secondary standard and maintain accuracies of a few parts per million.

We should point out some of the limitations that prevent analog measurements from having the same high accuracy as time measurements. Analog measurements rely on physical properties such as electrochemical potentials, breakdown voltages, and resistances, and these all vary with temperature and time. Interfering effects such as Johnson and  $1/f$  noise, leakage currents, and thermoelectric potentials (thermocouple effect) complicate any measurement. To measure a voltage with precision comparable to state-of-the-art time or frequency measurements would require a measurement accuracy of a picovolt at a voltage of 1 volt. Think of this not as an indictment of analog methods but merely



as a celebration of the incredible precision attainable in the time/frequency domain. And, in practice, choose time/frequency transducers and measurements, rather than voltage/resistance measurements, whenever possible.

## BANDWIDTH-NARROWING TECHNIQUES

### 15.12 The problem of signal-to-noise ratio

Up to this point we have been talking about the various experimental quantities that can be detected, how you might measure them, and what sort of trade-offs you face. As luck would have it, the signals you often want to measure are buried in noise or interference, frequently to the extent that you can't even see them on an oscilloscope. Even when external noise isn't a problem, the statistics of the signal itself may make detection difficult, as, for example, when counting nuclear disintegrations from a weak source, with only a few counts detected per minute. Finally, even when the signal is detectable, you may wish to improve the detected signal strength in order to make a more accurate measurement. In all these cases some tricks are needed to improve the signal/noise ratio; as you will see, they all amount to a narrowing of the detection bandwidth in order to preserve the desired signal while reducing the total amount of (broadband) noise accepted.

The first thing you might be tempted to try when thinking of reducing the bandwidth of a measurement is to hang a simple low-pass filter on the output, in order to average out the noise. There are cases where that therapy will work, but most of the time it will do very little good, for a couple of reasons. First, the signal itself may have some high frequencies in it, or it may be centered at some high frequency. Second, even if the signal is in fact slowly

varying or static, you invariably have to contend with the reality that the density of noise signal usually has a  $1/f$  character, so as you squeeze the bandwidth down toward dc you gain very little. Electronic and physical systems are twitchy, so to speak.

In practice, there are a few basic techniques of bandwidth narrowing that are in widespread use. They go under names like signal averaging, transient averaging, boxcar integration, multichannel scaling, pulse-height analysis, lock-in detection, and phase-sensitive detection. All these methods assume that you have a repetitive signal; that's no real problem, since there is almost always a way to force the signal to be periodic, assuming it isn't already. Let's see what is going on.

### 15.13 Signal averaging and multichannel averaging

By forming a cumulative sum of a repetitive signal versus time, you can improve the signal/noise ratio enormously. This usually goes under the heading of "signal averaging," and it is often applied to analog signals. We will consider first what may seem to be an artificial situation, namely a signal consisting of pulses whose rate is proportional to the amplitude of some sought-after waveform versus time. We begin with this example because it makes our calculations easier. In reality, it isn't even an artificial situation, since it is the rule when using pulse-counting electronics such as particle detectors or photomultipliers at low light levels.

#### *Multichannel scalars*

We begin with multichannel scaling because it typifies all these techniques and, in addition, is easy to understand and quantify. The multichannel scaler (MCS) is a piece of hardware that contains a set of memory registers (typically 1024 or more), each of which can store a number up to 1 million (20 bits binary or 24 bits BCD) or

so. The MCS accepts pulses (or continuous voltages, as will be described later) as its input; in addition, it accepts either a channel-advance signal (a pulse) or a parallel multibit channel address. Each time there is an input pulse, the MCS increments the count in the memory channel currently being addressed. Additional inputs let you reset the address to 0, clear the memory, etc.

To use an MCS you need a signal that repeats itself at some interval. Let's suppose for the time being that the phenomenon you're observing is itself periodic, with period  $T$ ; although this is not the case most of the time (you usually have to coax the experiment into periodicity), there are good examples in the real world of strictly periodic phenomena, e.g., the light output of a pulsar. Let's suppose that the input consists of pulses, with rate proportional to the signal plus a large background rate of noise pulses, i.e., pulses randomly distributed in time (again, realistic for pulsars, where the actual signal is swamped by light from the night sky). By sending timing pulses to the channel advance and reset inputs, we arrange to sweep the MCS repetitively through its 1024 channels once every  $T$  seconds, accumulating additional input (signal plus background) counts into the memory channels each sweep. As time goes on, the signal will keep adding counts to the same subgroup of channels, with the background noise adding counts in all channels, because the sweep through the entire set of channels is timed to coincide with the signal's periodicity. Thus the signal keeps adding on top of itself, the accumulated sum getting larger after each repetition.

#### Signal-to-noise computation

Let's see what happens. To be specific, let the background pulse rate have an average value that contributes  $n_b$  pulses per channel each sweep, with the signal contribut-

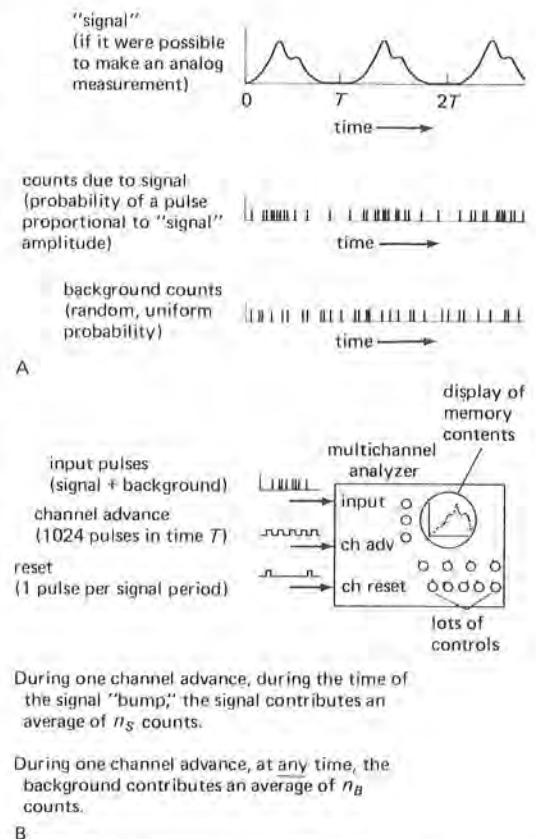


Figure 15.34. Multichannel signal averaging (pulse input).

ing an additional  $n_s$  pulses into the channel where its peak lies (Fig. 15.34). Let's give ourselves a poor signal/background ratio, i.e.,  $n_s \ll n_b$ , meaning that most of the counts added during each sweep through the memory are contributed by background, rather than signal. Now, when the memory contents are graphed, the signal should be recognizable as a bump above the background. You might think the criterion is that the number of signal counts in a channel with signal should be comparable with the number of counts contributed to that channel by the background noise. That would be wrong, since the *average* value contributed by noise is

quite irrelevant; all that matters is the level of *fluctuations* of that average value about the mean.

Thus, a poor input signal/noise ratio is actually characterized by  $n_s \ll \sqrt{n_b}$ , meaning that in one sweep the signal will not be recognizable above the “noise” consisting of an undulating graph of accumulated random background pulses. For purposes of computation, let’s let  $n_s = 10$  and  $n_b = 1000$ . Therefore, in one sweep an initially cleared MCS will acquire an average of 1000 counts in each channel, with an additional 10 counts in the channels where the signal peaked. Since the fluctuations in the channel totals equal about 31 (square root of 1000), the actual signal bump is left pretty much buried in the noise after only one sweep. But after 1000 sweeps, say, the average count in any channel is about 1,000,000, with fluctuations of 1000. The channels where the signal peaks have an additional 10,000 counts (1000 sweeps  $\times$  10 counts/sweep), for a signal/noise ratio of 10. In other words, the signal has emerged from the background.

### Example: Mössbauer resonance

Figure 15.35 shows the results of just such an analysis, in this case a Mössbauer resonance signal consisting of six dips in the transmission of an enriched iron-57 foil to gamma radiation from a cobalt-57 radioactive source. In this case  $n_b = 0.4$  and  $n_s = 0.1$ , approximately, for a situation of poor signal/noise ratio. The Mössbauer signal is totally swamped by noise even after 10 or 100 sweeps; it becomes visible only after 1000 sweeps or so. The results are shown after 1000, 10,000, and 100,000 sweeps, with each graph scaled to keep this signal size the same. Note the rise of the “baseline” caused by the steady background, as well as the nice enhancement of SNR with time.

It is easy to see by what factor the ratio of signal amplitude to background fluctuation (“noise”) increases as time goes on. The signal amplitude increases proportional to  $t$ ; the average background count (“baseline”) also increases proportional

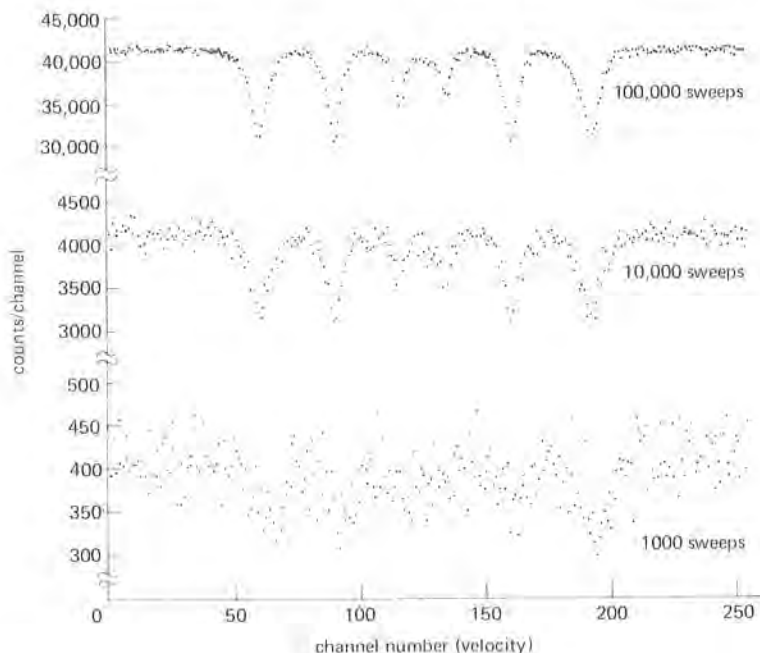


Figure 15.35. Mössbauer absorption spectrum, showing effect of signal averaging.

to  $t$ , but the *fluctuations* in the background count (“noise”) rise only proportional to the square root of  $t$ . Therefore, the ratio between signal and fluctuations in background increases as  $t$  divided by the square root of  $t$ . In other words, the signal-to-noise ratio improves in proportion to the square root of time.

#### **Multichannel analysis of analog signals (signal averaging)**

You can play the same game with analog signals by simply using a voltage-to-frequency converter at the input. Commercial MCSs often provide the electronics for you, giving you a choice of analog or pulse input modes. In this form you often hear these gadgets called signal averagers or transient averagers. One company (TMC) called theirs a “CAT” (computer of averaged transients), and the name has stuck, in some circles at least.

It is possible to make a completely analog MCS by using a set of integrators to store the accumulated signal. A simpler device, known as a boxcar integrator, is an analog signal averager with a single “sliding channel.” With the enormous reductions in digital memory prices that have taken place in the last decade, such analog signal averagers are becoming impractical, except perhaps for specialized applications.

#### **Multichannel analysis as bandwidth narrowing**

We suggested at the beginning of this discussion that there was an equivalence between the magical SNR-reduction methods and a reduction in effective measurement bandwidth. It is not hard to see how that goes in this case. Imagine another (interfering) signal added into the input, but with periodicity  $T'$  slightly different from the desired signal of period  $T$ . After just a few sweeps, its signal will also begin to

accumulate, causing trouble. But wait – as time goes on, its “bump” will gradually drift along through the channels, successively contributing counts through all the channels. It will have drifted all the way around through all the channels once after a time,

$$t = 1/\Delta f$$

where  $\Delta f$  is the frequency difference  $1/T - 1/T'$  between the desired signal and the interfering signal.

#### **EXERCISE 15.1**

Derive this result.

In other words, by accumulating data for a time  $t$  (as given in the preceding equation), the interfering signal has been spread equally through all the channels. Another way to say the same thing is that the measurement’s bandwidth is reduced roughly to

$$\Delta f = 1/t$$

after accumulating data for time  $t$ . By running for a long time, you reduce the bandwidth and exclude nearby interfering signals! In fact, you also exclude most of the noise, since it is spread evenly in frequency. Viewed in this light, the effect of multichannel analysis is to narrow the accepted bandwidth, thereby accepting the signal power but squeezing down the amount of noise power.

Let’s see how the calculation goes. After time  $t$ , the bandwidth is narrowed to  $\Delta f = 1/t$ . If the noise power density is  $p_n$  watts per hertz, and the signal power  $P_s$  stays within the measurement bandwidth, then the SNR after time  $t$  is

$$\text{SNR} = 10 \log(P_s t / p_n)$$

The signal amplitude improves proportionally to the square root of  $t$  (3dB for each doubling of  $t$ ), just as we found in the analysis we did earlier by considering the number of counts per channel and its fluctuations.



### 15.14 Making a signal periodic

We mentioned initially that all signal-averaging schemes require a signal that repeats many times in order to realize significant reduction in signal/noise ratio. Since most measurements don't involve intrinsically periodic quantities, it is usually necessary to force the signal to repeat. There are many ways to do this, depending on the particular measurement. It is probably easiest to give a few examples, rather than attempt to set down rules.

A measurable quantity that depends on some external parameter can easily be made periodic – just vary the external parameter. In NMR (nuclear magnetic resonance) the resonance frequency varies linearly with the applied field, so it is standard to modulate the current in a small additional magnet winding. In Mössbauer studies you vary the source velocity. In quadrupole resonance you can sweep the oscillator.

In other cases an effect may have its own well-defined transient, but allow external triggering. A classic example is the pulse

of depolarization in a nerve fiber. In order to generate a clean graph of the waveform of such a pulse, you can simply trigger the nerve with an externally applied voltage pulse, starting the MCS sweep at the same time (or even “anticipating” the trigger by starting the sweep, then triggering the nerve with a delayed pulse); in this case you would pick a repetition period long enough so that the nerve has fully recovered before the next pulse. This last case illustrates graphically the importance of a repeatable phenomenon as fodder for signal averaging; if the frog whose leg is twitching chances to expire, your experiment is over, whatever the signal/noise ratio!

It should be pointed out that cases where the phenomenon you're measuring has its own well-defined periodicity may in fact be the most difficult to work with, since you have to know the periodicity precisely. The graph of the “light curve” (brightness versus time) in Figure 15.36 is an example. We made this curve by using an MCS on the output of a photomultiplier stationed at the focus of a 60 inch telescope, run

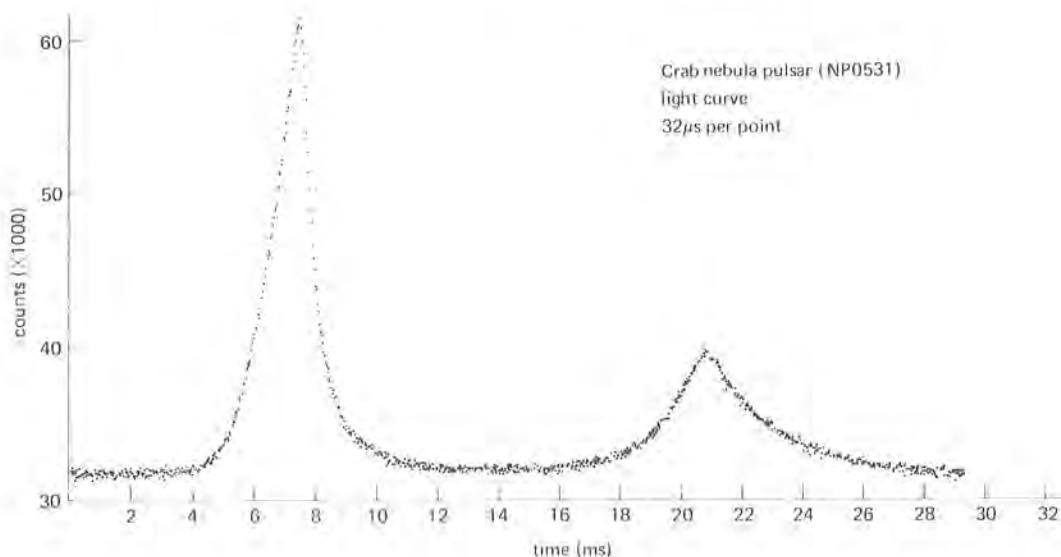


Figure 15.36. Crab nebula pulsar brightness versus time (light curve).

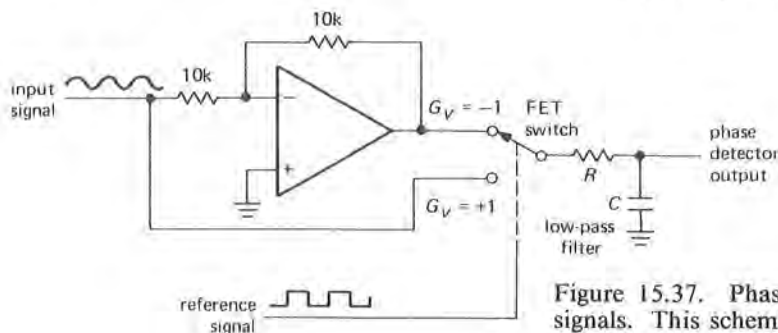


exactly in synchronism with the pulsar's rotation. Even with that size telescope it required an average of approximately 5 million sweeps to generate such a clean curve, since the average number of detected photons for each entire pulsar pulse was about 1. With such a short period, that puts enormous accuracy requirements on the MCS channel-advance circuitry, in this case requiring clocks of part-per-billion stability and frequent adjustment of the clock rate to compensate for the earth's motion.

It is worth saying again that the essence of signal averaging is a reduction in bandwidth, gained by running an experiment for a long period of time. The bottom line here is the total length of the experiment; the particular rate of scanning, or modulation, is usually not important, as long as it takes you far enough from the  $1/f$  noise present near dc. You can think of the modulation as simply shifting the signal you wish to measure from dc up to the modulating frequency. The effect of the long data accumulation is then to center an effective bandwidth  $\Delta f = 1/T$  at  $f_{\text{mod}}$ , rather than at dc.

### 15.15 Lock-in detection

This is a method of considerable subtlety. In order to understand the method, it is necessary to take a short detour into the phase detector, a subject we first took up in Section 9.27.



### Phase detectors

In Section 9.27 we described phase detectors that produce an output voltage proportional to the phase difference between two digital (logic-level) signals. For purposes of lock-in detection, you need to know about linear phase detectors, since you are nearly always dealing with analog voltage levels.

The basic circuit is shown in Figure 15.37. An analog signal passes through a linear amplifier whose gain is reversed by a square-wave "reference" signal controlling a FET switch. The output signal passes through a low-pass filter,  $RC$ . That's all there is to it. Let's see what you can do with it.

- *Phase-detector output.* To analyze the phase-detector operation, let's assume we apply a signal

$$E_s \cos(\omega t + \phi)$$

to such a phase detector, whose reference signal is a square wave with transitions at the zeros of  $\sin \omega t$ , i.e., at  $t = 0, \pi/\omega, 2\pi/\omega$ , etc. Let us further assume that we average the output,  $V_{\text{out}}$ , by passing it through a low-pass filter whose time constant is longer than one period:

$$\tau = RC \gg T = 2\pi/\omega$$

Then the low-pass filter output is

$$\langle E_s \cos(\omega t + \phi) \rangle_0^{\pi/\omega} - \langle E_s \cos(\omega t + \phi) \rangle_{\pi/\omega}^{2\pi/\omega}$$

Figure 15.37. Phase detector for linear input signals. This scheme is used in the monolithic AD630.

where the brackets represent averages, and the minus sign comes from the gain reversal over alternate half cycles of  $V_{ref}$ . As an exercise, you can show that

$$\langle V_{out} \rangle = -(2E_s/\pi) \sin \phi$$

#### EXERCISE 15.2

Perform the indicated averages by explicit integration to obtain the preceding result for unity gain.

Our result shows that the averaged output, for an input signal of the same frequency as the reference signal, is proportional to the amplitude of  $V_s$  and sinusoidal in the relative phase.

We need one more result before going on: What is the output voltage for an input signal whose frequency is close to (but not equal to) the reference signal? This is easy, since in the preceding equations the quantity  $\phi$  now varies slowly, at the difference frequency:

$$\cos(\omega + \Delta\omega)t = \cos(\omega t + \phi)$$

with  $\phi = t\Delta\omega$

giving an output signal that is a slow sinusoid:

$$V_{out} = (2E_s/\pi) \sin(\Delta\omega)t$$

which will pass through the low-pass filter relatively unscathed if  $\Delta\omega < 1/\tau = 1/RC$  and will be heavily attenuated if  $\Delta\omega > 1/\tau$ .

#### The lock-in method

Now the so-called lock-in (or phase-sensitive) amplifier should make sense. First you make a weak signal periodic, as we've discussed, typically at a frequency in the neighborhood of 100Hz. The weak signal, contaminated by noise, is amplified and phase-detected relative to the modulating signal. Look at Figure 15.38. You need an experiment with two "knobs" on it, one for fast modulation in order to do phase detection and one for a slow sweep through the interesting features of the signal (in NMR, for example, the fast modulation might be a small 100Hz modulation of the magnetic field, and the slow modulation might be a frequency sweep 10 minutes in duration through the resonance). The phase shifter is adjusted to give maximum output signal, and the low-pass filter is set for a time constant long enough to give good signal/noise ratio. The low-pass-filter rolloff sets the bandwidth, so a 1Hz rolloff, for example, gives you sensitivity to spurious signals and noise only within 1Hz of the desired signal. The bandwidth also determines how fast you can adjust the "slow modulation," since now you must not sweep through any features of the signal faster than the filter can respond. People use time constants of fractions of a second up to tens of seconds and often do the slow modulation with a geared-down

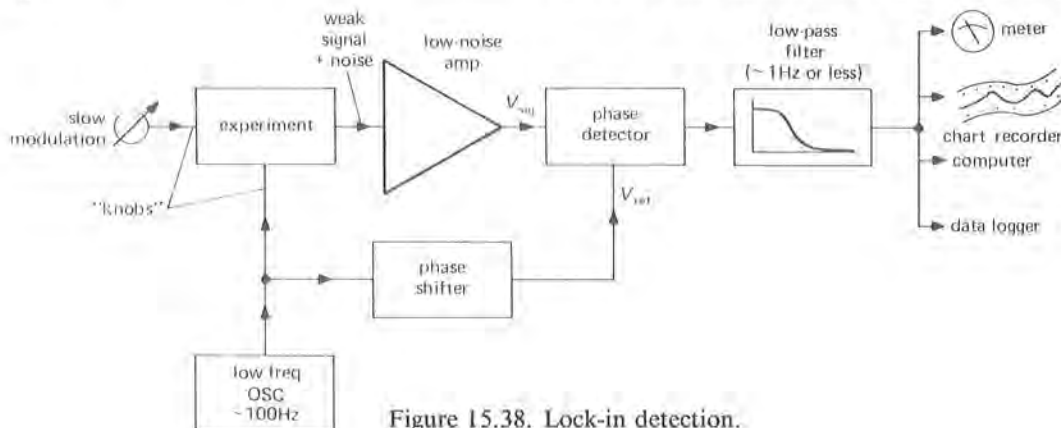


Figure 15.38. Lock-in detection.

clock motor turning an actual knob on something!

Note that lock-in detection amounts to bandwidth narrowing again, with the bandwidth set by the post-detection low-pass filter. As with signal averaging, the effect of the modulation is to center the signal at the fast modulation frequency, rather than at dc, in order to get away from  $1/f$  noise (flicker noise, drifts, and the like).

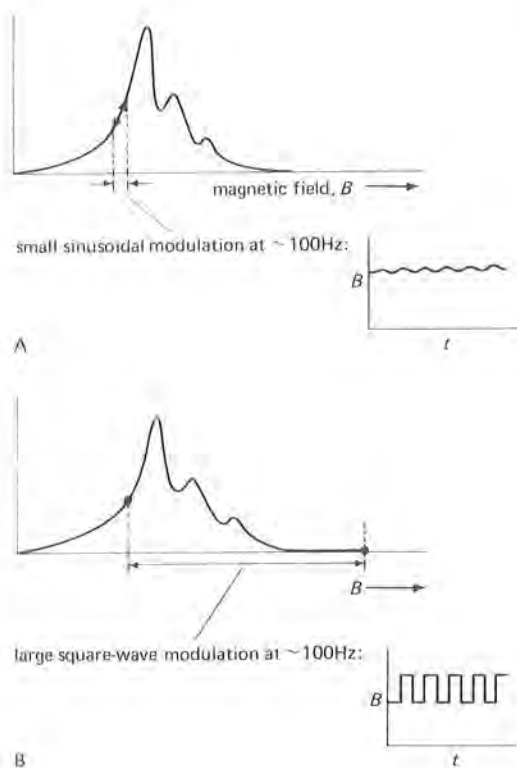


Figure 15.39. Lock-in modulation methods.  
A. Small sinusoid.  
B. Large square wave.

#### Two methods of "fast modulation"

There are some ways to do the fast modulation: The modulation waveform can be either a very small sine wave or a very large square wave compared with the features of

the sought-after signal (line shape versus magnetic field, for example, in NMR), as sketched in Figure 15.39. In the first case the output signal from the phase-sensitive detector is proportional to the *slope* of the line shape (i.e., its derivative), whereas in the second case it is proportional to the line shape itself (providing there aren't any other lines out at the other endpoint of the modulation waveform). This is the reason all those simple NMR resonance lines come out looking like dispersion curves (Fig. 15.40).

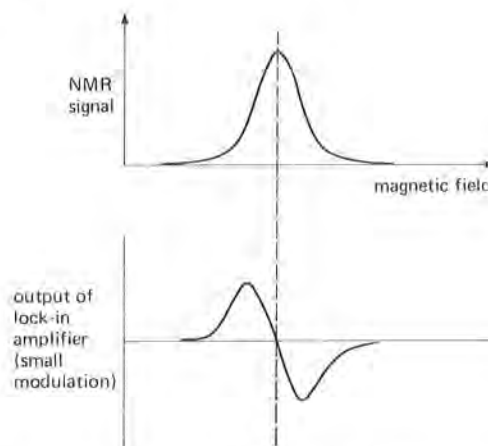


Figure 15.40. Line shape differentiation resulting from lock-in detection.

For large-shift square-wave modulation there's a clever method for suppressing modulation feedthrough, in cases where that is a problem. Figure 15.41 shows the modulation waveform. The offsets above and below the central value kill the signal, causing an on/off modulation of the signal at *twice* the fundamental of the modulating waveform. This is a method for use in special cases only; don't get carried away by the beauty of it all!

Large-amplitude square-wave modulation is a favorite with those dealing in infrared astronomy, where the telescope secondary mirrors are rocked to switch the image back and forth on an infrared source.

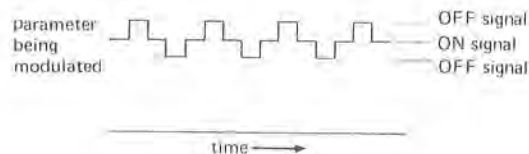


Figure 15.41. Modulation scheme for suppressing modulation feedthrough.

It is also popular in radioastronomy, where it's called a Dicke switch.

Commercial lock-in amplifiers have a variable-frequency modulating source and tracking filter, a switchable time-constant post-detection filter, a good low-noise wide-dynamic-range amplifier (you wouldn't be using lock-in detection if you weren't having noise problems), and a nice linear phase detector. They also let you use an external source of modulation. There's a knob that adjusts the phase shift, so you can maximize the detected signal. The whole item comes packed in a handsome cabinet, with a meter to read output signal. Typically these things cost a few thousand dollars and are manufactured by companies like EG&G Princeton Applied Research, Ithaco, and Stanford Research Systems. Board-level components are made by Evans Electronics, among others.

In order to illustrate the power of lock-in detection, we usually set up a small demonstration for our students. We use a lock-in to modulate a small LED of the kind used for panel indicators, with a modulation rate of a kilohertz or so. The current is very low, and you can hardly see the LED glowing in normal room light. Six feet away a phototransistor looks in the general direction of the LED, with its output fed to the lock-in. With the room lights out, there's a tiny signal from the phototransistor at the modulating frequency (mixed with plenty of noise), and the lock-in easily detects it, using a time constant of a few seconds. Then we turn the room lights on (fluorescent), at which point the

signal from the phototransistor becomes just a huge messy 120Hz waveform, jumping in amplitude by 50dB or more. The situation looks hopeless on the oscilloscope, but the lock-in just sits there, unperturbed, calmly detecting the same LED signal at the same level. You can check that it's really working by sticking your hand in between the LED and the detector. It's darned impressive.

### 15.16 Pulse-height analysis

A pulse-height analyzer (PHA) is a simple extension of the multichannel scaler principle, and it is a very important instrument in nuclear and radiation physics. The idea is simplicity itself: Pulses with a range of amplitudes are input to a peak-detector/ADC circuit that converts the relative pulse height to a channel address. A multichannel scaler then increments the contents of the selected address. The result is a graph that is a histogram of pulse heights. That's all there is to it.

The enormous utility of pulse-height analyzers stems from the fact that many detectors of charged particles, x rays, and gamma rays have output pulse sizes proportional to the energy of the radiation detected (e.g., proportional counters, solid-state detectors, surface-barrier detectors, and scintillators, as we discussed in Section 15.07). Thus a pulse-height analyzer converts the detector's output to an energy spectrum.

Pulse-height analyzers used to be designed as dedicated hardware devices, with buckets of ICs and discrete components. Nowadays the standard method is to use an off-the-shelf microcomputer, preceded by a fast pulse-input ADC. That way you can build in all sorts of useful computational routines, e.g., background subtraction, energy calibration and line identification, disk and tape storage, and on-line control of the experiment. We have an apparatus that scans a proton microbeam

over a specimen in a two-dimensional raster pattern, detects the emitted x rays, sorts them by chemical element, and stores a picture of the distribution of each element in the sample, all the while letting you view the x-ray spectrum and images as the picture accumulates. The whole operation is handled by a pulse-height analyzer that doesn't realize that it's really a computer.

There is an interesting subtlety involving the ADC front end of a pulse-height analyzer. It turns out that you can't use something like a successive-approximation A/D converter, in spite of its superior speed, because you wouldn't get exact equality of channel widths, with the disastrous effect of producing a lumpy baseline from a smooth continuum of input radiation. All PHAs use a so-called Wilkinson converter, a variation on single-slope conversion whereby an input pulse charges a capacitor, which is then discharged by a constant current while a fast counter (200MHz is typical) counts up the address. This has the disadvantage of giving an analyzer "dead time" that depends on the height of the last pulse, but it gives absolute equality of channel widths.

Most pulse-height analyzers provide inputs so that you can use them as multi-channel scalars. Why shouldn't they? All the electronics are already there. Some big names in pulse-height analyzers are Canberra, EG&G, Nuclear Data, and Tracor-Northern.

### 15.17 Time-to-amplitude converters

In nuclear physics it is often important to know the distribution of decay times of some short-lived particle. This turns out to be easy to measure, by simply hooking a time-to-amplitude converter (TAC) in front of a pulse-height analyzer. The TAC starts a ramp when it receives a pulse at one input and stops it when it receives a pulse at a second input, discharging the ramp and generating an output pulse

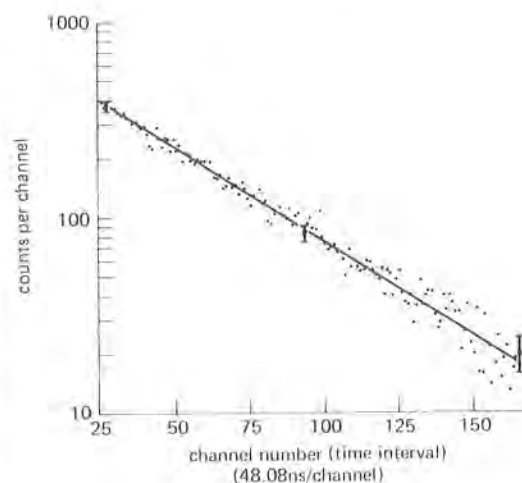


Figure 15.42. Muon lifetime measurement from time-interval spectrum (TAC + PHA).

proportional to the time interval between pulses. It is possible to build these things with resolution down in the picoseconds. Figure 15.42 shows a measurement of the muon lifetime made by a student by timing the delay between the capture of a cosmic-ray muon in a scintillator and its subsequent decay. Each event creates a flash of light, and a TAC is used to convert the intervals into pulses. A cosmic-ray muon decayed in this student's apparatus once a minute on the average, so he accumulated data for 18 days to determine a lifetime of  $2.198 \pm 0.02 \mu\text{s}$  (accepted value is  $2.197134 \pm 0.00008 \mu\text{s}$ ). Note the use of log-lin axes to plot data that should be an exponential, and the systematic shift of  $n^{1/2}$  (counting) error bars. The line plotted is the decay according to the accepted value,  $n(t) = n_0 \exp(-t/\tau)$ .

## SPECTRUM ANALYSIS AND FOURIER TRANSFORMS

### 15.18 Spectrum analyzers

An instrument of considerable utility, particularly in radiofrequency work, is the spectrum analyzer. These devices gen-



erate an  $xy$  oscilloscope display, with  $y$  representing signal strength (usually logarithmic, i.e., in decibels), but with  $x$  representing frequency. In other words, a spectrum analyzer lets you look in the *frequency domain*, plotting the amount of input signal versus its frequency. You can think of it as a Fourier decomposition of the input waveform (if you know about such things), or as the response you would get as you tuned the dial of a broadrange high-performance (wide dynamic range, stable, sensitive) receiver through its frequency range. This ability can be very handy when analyzing modulated signals, looking for intermodulation products or distortion, analyzing noise and drift, trying to make accurate frequency measurements on weak signals in the presence of stronger signals, and making a host of other measurements.

Spectrum analyzers come in two basic varieties: swept-tuned and real-time. Swept analyzers are the most common variety, and they work as shown in Figure 15.43. What you have is basically a superheterodyne receiver (see Section 13.16), with a local oscillator (LO) that can be swept by an internally generated ramp waveform. As the LO is swept through its range of frequencies, different input frequencies are successively mixed to pass through the IF amplifier and filter. For example, suppose you have a spectrum analyzer with an IF of 200MHz and an LO

that can sweep from 200MHz to 300MHz. When the LO is at 210MHz, input signals at 10MHz ( $\pm$  the IF filter bandwidth) pass through to the detector and produce vertical deflection on the scope. Signals at 410MHz (an “image” frequency) would also pass through, which is the reason for the low-pass filter at the input. At any given time, input frequencies 200MHz lower than the LO are detected.

Real spectrum analyzers allow lots of flexibility as to sweep range, center frequency, filter bandwidth, display scales, etc. Typical input frequency ranges go from hertz to gigahertz, with selectable bandwidths ranging from hertz to megahertz. A range of 10MHz to 22GHz is popular, with resolution bandwidths of 10Hz to 3MHz. In addition, sophisticated spectrum analyzers have convenience features such as absolute amplitude calibration, storage of spectra to prevent flicker during sweeping, additional storage for comparison and normalization, and display of digital information on the screen. Fancy spectrum analyzers let you analyze phase versus frequency, generate frequency markers, program the operation via the IEEE-488 bus, include tracking oscillators (for increased dynamic range), make precise frequency measurements of features in the spectrum, generate tracking noise voltages for system stimulus, and even do signal averaging (particularly useful for noisy signals).

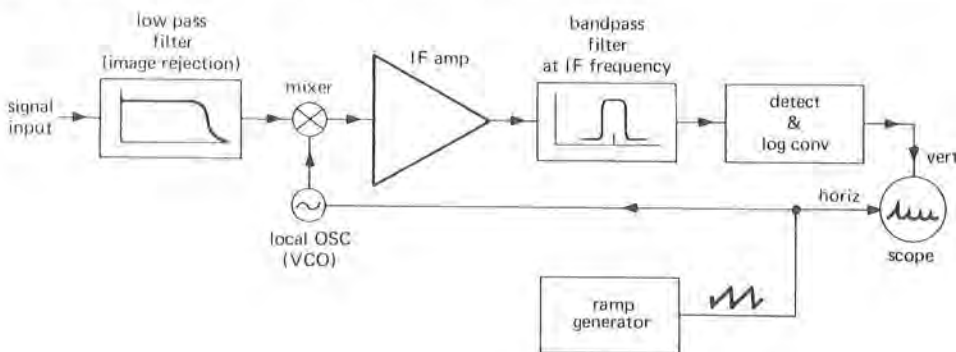


Figure 15.43. Swept-LO spectrum analyzer.

Note that this sort of swept spectrum analyzer looks at only one frequency at a time and generates a complete spectrum by sweeping in time. This can be a real disadvantage, since you can't look at transient events. In addition, when scanning with narrow bandwidth, the sweep rate must be kept slow. Finally, only a small portion of the input signal is being used at any one time.

These disadvantages of swept spectrum analysis are remedied in real-time spectrum analyzers. Again, there are several approaches. The clumsy method employs a set of narrow filters to look at a range of frequencies simultaneously. More recently, sophisticated analyzers based on digital Fourier analysis (in particular, the famous Cooley-Tukey fast Fourier transform, FFT for short) are becoming popular. These instruments convert the analog input signal (after mixing, etc.) to numbers, using a fast analog/digital converter. Then a special-purpose computer turns the crank on the FFT, generating a digital frequency spectrum. Since this method looks at all frequencies simultaneously, it has excellent sensitivity and speed, and it can be used for analysis of transients. It is particularly good for low-frequency signals, where swept analyzers are too slow. In addition, it can perform correlations between signals. Since the data comes out in digital form, it is natural to apply the full power of signal averaging, a feature available in some commercial instruments.

Note that these digital spectrum analyzers, being limited by computational speed, have *much* less bandwidth than the radiofrequency analog types (swept-LO or filter-bank). For example, the popular HP 3561A goes from 125  $\mu$ Hz to 100kHz. You can, of course, use it to look at a 100kHz band centered at some higher frequency, by translating that band down in frequency with heterodyne techniques.

A clever real-time spectrum analyzer can also be constructed using the so-called

chirp/Z transform. In this method a dispersive filter (delay time proportional to frequency) replaces the IF bandpass filter in the swept-LO analyzer (Fig. 15.43). By matching the LO sweep rate to the filter's dispersion, you get an output that superficially resembles the swept analyzer output, namely a linear scan of frequency versus time during each sweep. However, in contrast to the swept-LO analyzer, this scheme gathers signals from the entire band of frequencies continuously. Another interesting technique for real-time spectral analysis is the Bragg cell (or "acousto-optic spectrometer"), in which the IF signal is used to generate acoustic waves in a transparent crystal. These deformations diffract a laser beam, generating a real-time display of the frequency spectrum as light intensity versus position. An array of photodetectors completes the analyzer output. Bragg-cell spectrometers are used in radio astronomy. A typical unit has 2GHz instantaneous bandwidth, analyzed into 16,000 channels of 125kHz bandwidth each. When choosing a spectrum analyzer type, be sure to consider trade-offs among bandwidth, resolution, linearity, and dynamic range.

Figure 15.44 shows the sort of radiofrequency spectra that endear spectrum analyzers to people who earn their living above 1MHz. The first four spectra show oscillators: A is just a pure sine-wave oscillator, B is distorted (as indicated by its harmonics), C has noise sidebands, and D has some frequency instability (drifting or residual FM). You can measure amplifier intermodulation products, as in E, where second-, third-, and fourth-order intermodulation frequencies are visible in the output of an amplifier driven by a "two-tone" test signal consisting of pure sine waves at frequencies  $f_1$  and  $f_2$ . Finally, in F you can see the uncouth behavior of a double-balanced mixer; there is feedthrough of both the LO and input signal, as well as distortion terms ( $f_{LO} \pm 2f_{sig}$ ,  $f_{LO} \pm 3f_{sig}$ ). This

last spectrum may actually indicate quite respectable mixer performance, depending on the vertical scale shown. Spectrum analyzers are designed with enormous dynamic range (internally generated distortion products are typically down by 70dB or more; with a “tracking preselector” they’re down by 100dB) so that you can see the failings of even a very good circuit.

The last graph G in Figure 15.44 shows what happens when you sweep the LO too fast in a swept analyzer. If the sweep causes a signal to pass through the filter bandwidth  $\Delta f$  in a time shorter than  $\Delta t \approx 1/\Delta f$ , it will be broadened, roughly to  $\Delta f' \approx 1/\Delta t$ .

For instance, we have used the FFT to search for pulsars, perform audio analysis, enhance the resolution of astronomical images (speckle imaging), and look for signals from intelligent life in space (SETI). In the last experiment, a GaAs FET amplifier connected to a receiving dish 84 feet in diameter drives a heterodyne receiver, with 400kHz of bandwidth analyzed (in real time) into 8 million simultaneous 0.05Hz channels. Our digital spectrum analyzer has 20,000 ICs and a half million solder joints (all done by hand!) and can detect narrowband signals 60dB below receiver noise in a 20 second integration. This corresponds to a radio flux of less than 1 nanowatt total over the entire earth’s disk!

### 15.19 Off-line spectrum analysis

The fast Fourier transform applied to digitized data from an experiment provides a very powerful method of signal analysis, particularly the recognition of weak signals of well-defined periodicity buried in interfering signals or noise, or the recognition of vibrations or oscillatory modes.

### SELF-EXPLANATORY CIRCUITS

#### 15.20 Circuit ideas

In Figure 15.45 we’ve collected some circuits that are useful in measurement and control applications.

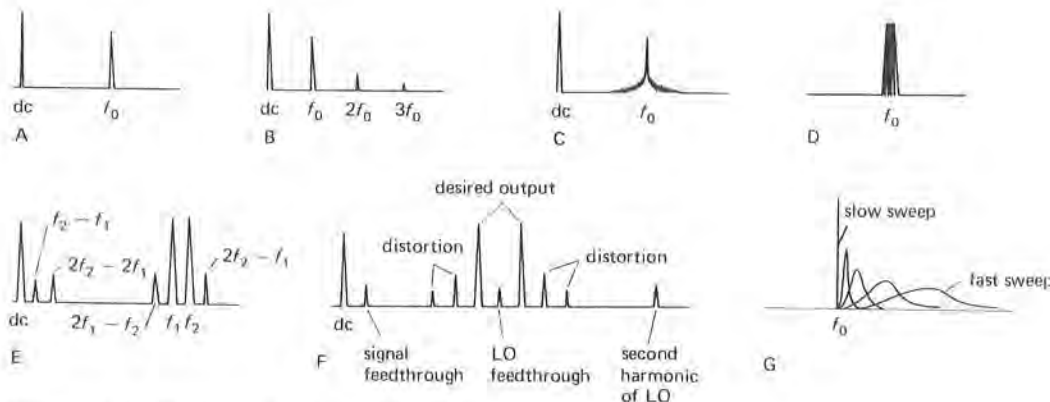
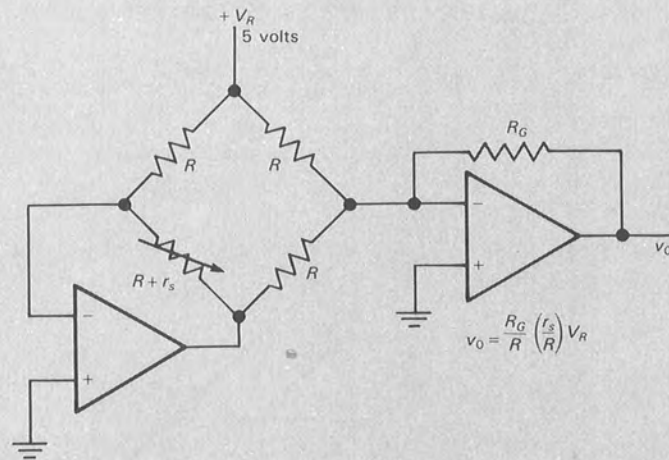
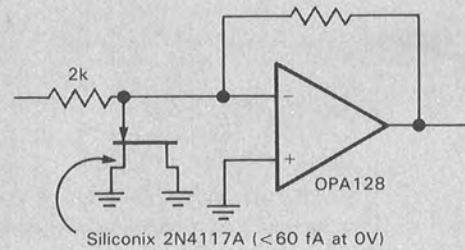


Figure 15.44. Spectrum analyzer displays.

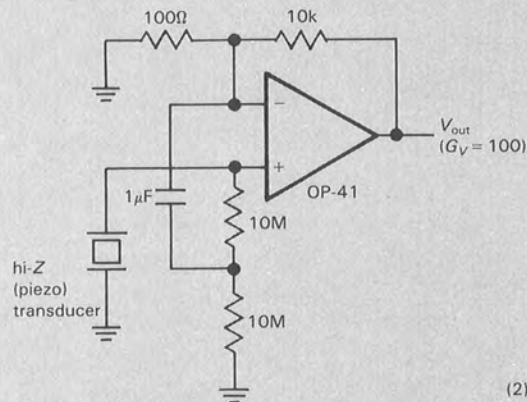
## Circuit ideas



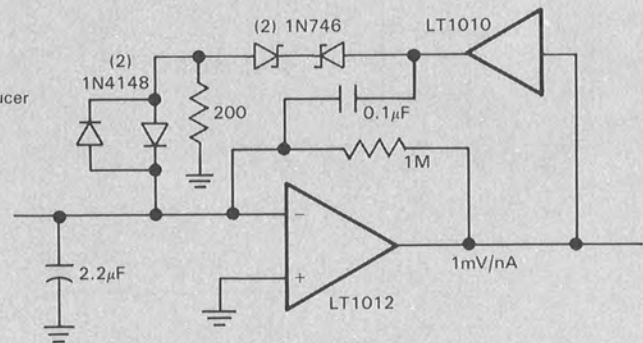
A. linearizes bridge response, with minimum  $V_{OS}$  effect; note that many bridges (e.g., semiconductor strain gauge) have high TCs of  $R$ .



C. charge amplifier input protection with low leakage



B. high- $Z_{in}$  bootstrapped amplifier for piezo transducer



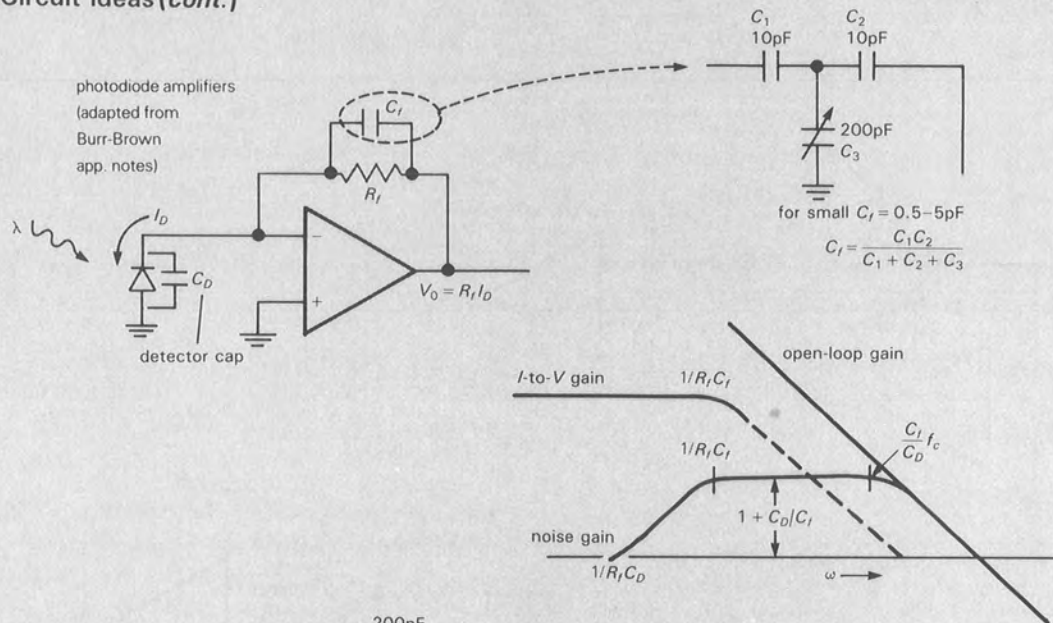
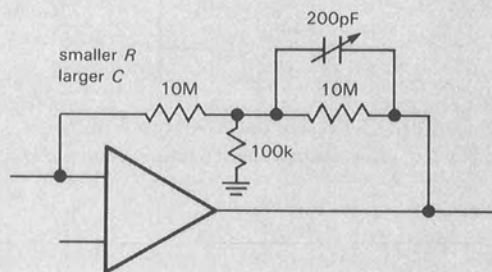
D. I-to-V converter, sensitive to 35pA, but able to keep summing junction under control to  $\pm 150\text{mA}$

Figure 15.45



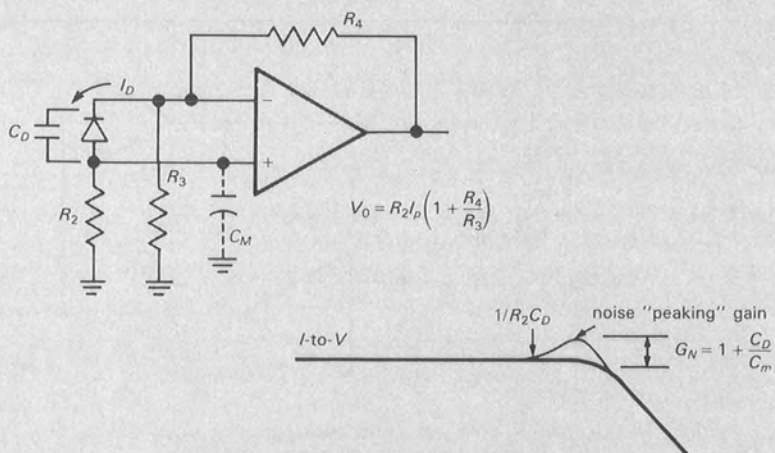


## Circuit ideas (cont.)



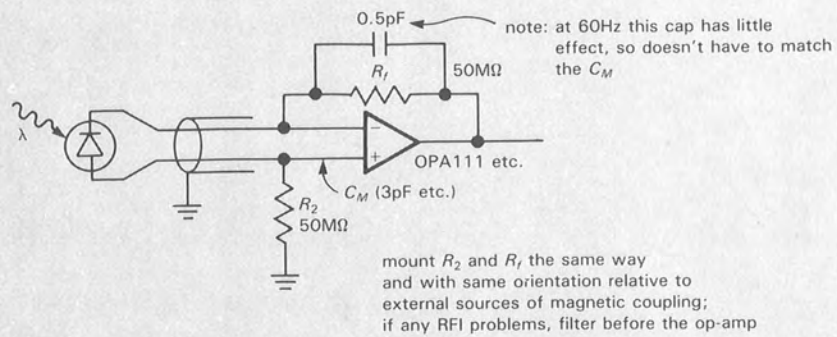
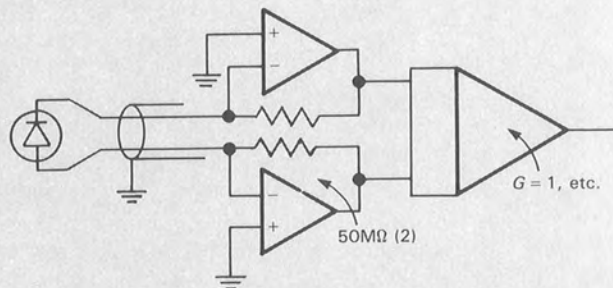
E. stabilizing photodiode amplifiers: contending with the necessary small  $C$  and large  $R$

one op-amp with moderate noise gain

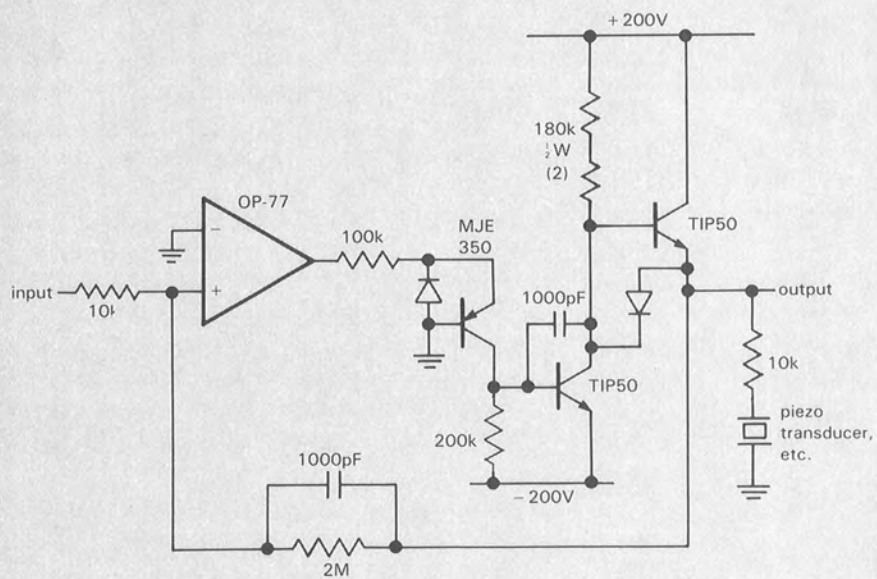


F. alternate photodiode amplifier configuration





G. balanced photodiode amplifiers for long detector cables





# APPENDIXES



## THE OSCILLOSCOPE

### APPENDIX A

The oscilloscope (scope for short) is the most useful and versatile electronic test instrument. As usually used, it lets you “see” voltages in a circuit as a function of time, triggering on a particular point of the waveform so that a stationary display results. We’ve drawn a block diagram (Fig. A1) and typical front panel (Fig. A2) to help explain how it works. The scope we will describe is usually called a decoupled dual-trace triggered scope. There are special-purpose scopes used for TV servicing and the like, and there are scopes of an older vintage that don’t have the features needed for circuit testing.

#### VERTICAL

Beginning with the signal inputs, most scopes have two channels; that’s very useful, since you often need to see the relationship between signals. Each channel has a calibrated gain switch, which sets the scale of VOLTS/DIVISION *on the screen*. There’s also a VARIABLE gain knob (concentric with the gain switch) in case you want to set a given signal to a certain number of divisions. Warning: Be sure the variable gain knob is in the “calibrated” position when making voltage measurements! It’s easy to forget. The better scopes have indicator lights to warn

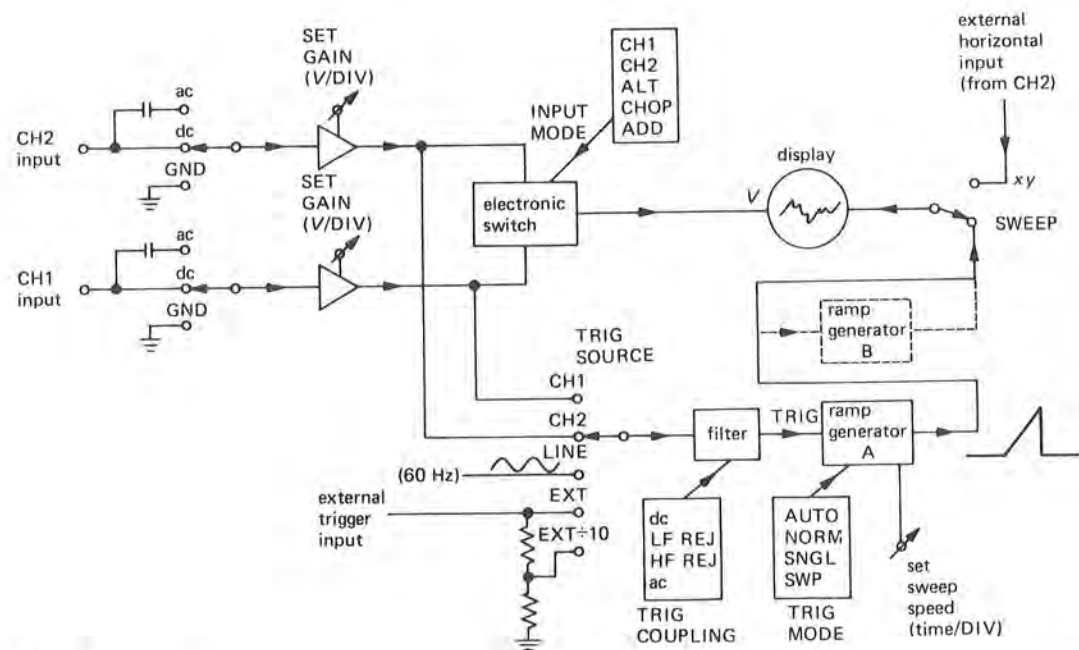


Figure A1



you if the variable gain knob is out of the calibrated position.

The scope is dc-coupled, an essential feature: What you see on the screen is the signal voltage, dc value and all. Sometimes you may want to see a small signal riding on a large dc voltage, though; in that case you can switch the input to ac coupling, which capacitively couples the input with a time constant of about 0.1 second.

Most scopes also have a grounded input position, which lets you see where zero volts is on the screen. (In GND position the signal isn't shorted to ground, just disconnected from the scope, whose input is grounded.) Scope inputs are usually high-impedance (a megohm in parallel with about 20pF), as any good voltage-measuring instrument should be. The input resistance of 1.0 megohm is an

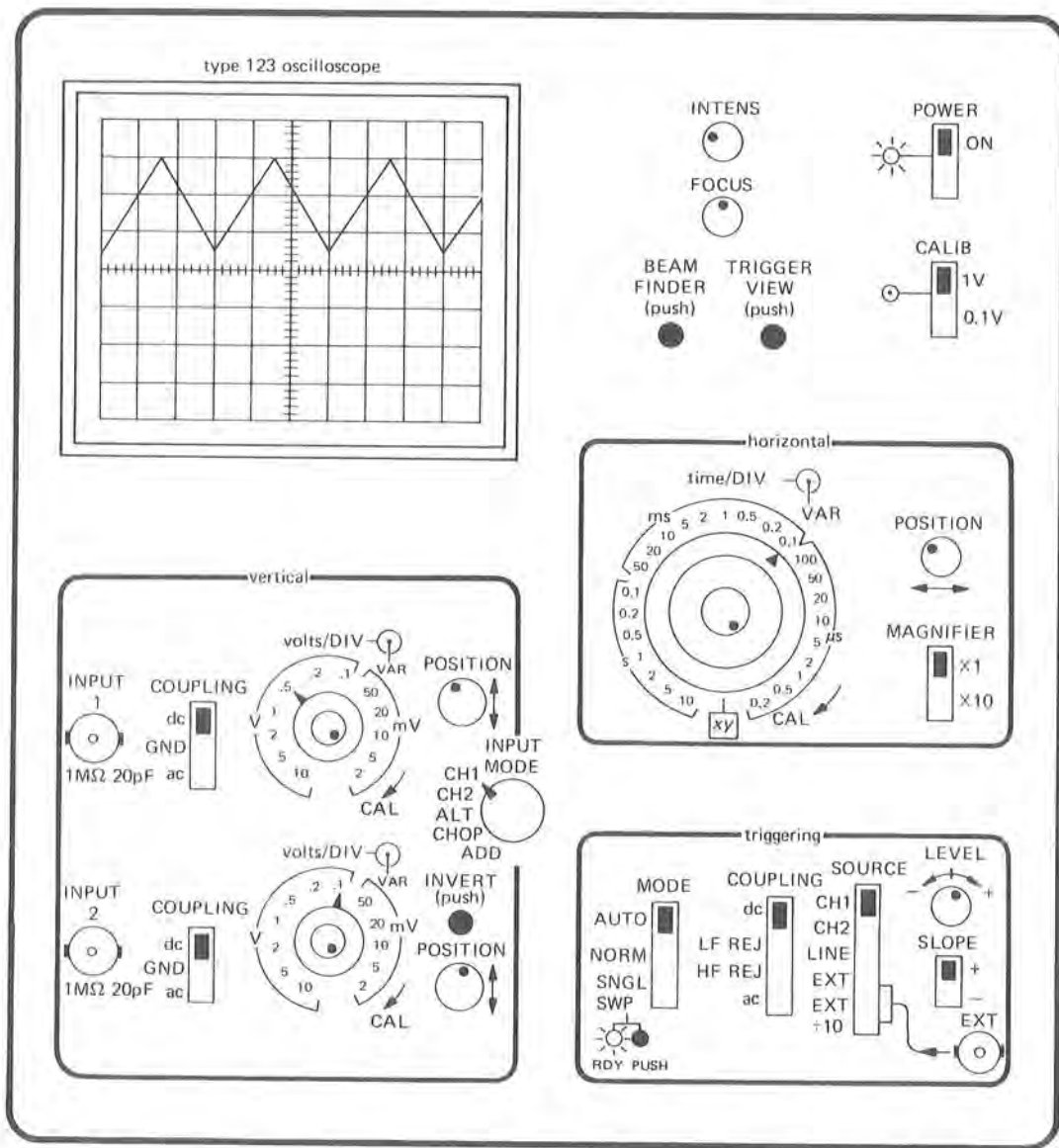


Figure A2

accurate and universal value, so that high-impedance attenuating probes can be used (as will be described later); unfortunately, the parallel capacitance is not standardized, which is a bit of a nuisance when changing probes.

The vertical amplifiers include a vertical POSITION control, an INVERT control on at least one of the channels, and an INPUT MODE switch. The latter lets you look at either channel, their sum (their difference, when one channel is inverted), or both. There are two ways to see both: ALTERNATE, in which alternate inputs are displayed on successive sweeps of the trace; CHOPPED, in which the trace jumps back and forth rapidly (0.1–1MHz) between the two signals. ALTERNATE mode is generally better, except for slow signals. It is often useful to view signals both ways, to make sure you're not being deceived.

## HORIZONTAL

The vertical signal is applied to the vertical deflection electronics, moving the dot up and down on the screen. The horizontal sweep signal is generated by an internal ramp generator, giving deflection proportional to time. As with the vertical amplifiers, there's a calibrated TIME/DIVISION switch and a VARIABLE concentric knob; the same warning stated earlier applies here. Most scopes have a  $10 \times$  MAGNIFIER and also allow you to use one of the input channels for horizontal deflection (this lets you generate those beloved but generally useless "Lissajous figures" featured in elementary books and science fiction movies).

## TRIGGERING

Now comes the trickiest part: triggering. We've got vertical signals and horizontal

sweep; that's what's needed for a graph of voltage versus time. But if the horizontal sweep doesn't catch the input signal at the same point in its waveform each time (assuming the signal is repetitive), the display will be a mess – a picture of the input waveform superimposed over itself at different times. The trigger circuitry lets you select a LEVEL and SLOPE (+ or –) on the waveform at which to begin the sweep. You can see from the front panel that you have a number of choices about trigger sources and mode. NORMAL mode produces a sweep only when the source selected crosses through the trigger point you have set, moving in the direction (SLOPE) you have selected. In practice, you adjust the level control for a stable display. In AUTO the sweep will "free run" if no signal is present; this is good if the signal sometimes drops to small values, since the display won't disappear and make you think the signal has gone away. It's the best mode to use if you are looking at a bunch of different signals and don't want to bother setting the trigger each time. SINGLE SWEEP is used for nonrepetitive signals. LINE causes the sweep to trigger on the ac power line, handy if you're looking at hum or ripple in a circuit. The EXTERNAL trigger inputs are used if you have a clean signal available at the same rate as some "dirty" signal you're trying to see; it's often used in situations where you are driving some circuit with a test signal, or in digital circuits where some "clock" signal synchronizes circuit operations. The various coupling modes are useful when viewing composite signals; for instance, you may want to look at an audio signal of a few kilohertz that has some spikes on it. The HF REJ position (high-frequency reject) puts a low-pass filter in front of the trigger circuitry, preventing false triggering on the spikes. If the spikes happen to be of interest, you can trigger on them instead in LF REJ position.

Many scopes now have BEAM FINDER

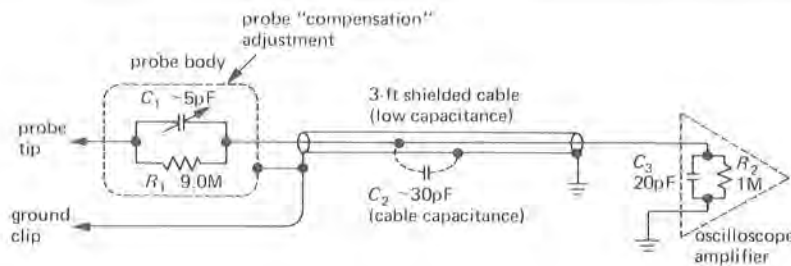


Figure A3

and TRIGGER VIEW controls. The beam finder is handy if you're lost and can't find the trace; it's a favorite of beginners. Trigger view displays the trigger signal; it's especially handy when triggering from external sources.

### HINTS FOR BEGINNERS

Sometimes it's hard to get *anything* to show on the scope. Begin by turning the scope on; set triggering for AUTO, DC COUPLING, CH 1. Set sweep speed at 1ms/div, cal, and the magnifier off ( $\times 1$ ). Ground the vertical inputs, turn up the intensity, and wiggle the vertical position control until a horizontal line appears (if you have trouble at this point, try the beam finder). Warning: Some scopes (the popular Tektronix 400 series, for example) don't sweep on AUTO unless the trigger level is adjusted correctly. Now you can apply a signal, unground the input, and fiddle with the trigger. Become familiar with the way things look when the vertical gain is far too high, when the sweep speed is too fast or slow, and when the trigger is adjusted incorrectly.

### PROBES

The oscilloscope input capacitance seen by a circuit under test can be quite high, especially when the necessary shielded connecting cable is included. The resulting

input impedance (1 megohm in parallel with 100 picofarads or so) is often too low for sensitive circuits and loads it by the usual voltage divider action. Worse yet, the capacitance may cause some circuits to misbehave, even to the point of going into oscillation! In such cases the scope obviously is not acting like the "low-profile" measurement instrument we expect; it's more like a bull in a china shop.

The usual solution is the use of high-impedance "probes." The popular 10 $\times$  probe works as shown in Figure A3. At dc it's just a 10 $\times$  voltage divider. By adjusting  $C_1$  to 1/9th the parallel capacitance of  $C_2$  and  $C_3$ , the circuit becomes a 10 $\times$  divider at all frequencies, with input impedance of 10 megohms in parallel with a few picofarads. In practice, you adjust the probe by looking at a square wave of about 1kHz, available on all scopes as CALIB, or PROBE ADJ, setting the capacitor on the probe for a clean square wave without overshoot. Sometimes the adjustment is cleverly hidden; on some probes you twist the body of the probe and lock it by tightening a second threaded part. One drawback: A 10 $\times$  probe makes it difficult to look at signals of only a few millivolts; for these situations use a "1 $\times$  probe," which is simply a length of low-capacitance shielded cable with the usual probe hardware (wire "grabber," ground clip, handsome knurled handle, etc.). The 10 $\times$  probe should be the standard probe, left connected to the scope, with the 1 $\times$  probe used when necessary. Some probes feature a convenient choice of 1 $\times$  or

10× attenuation, switchable at the probe tip.

## GROUNDING

As with most test instruments, the oscilloscope input is referred to the instrument ground (the outer connection of the input BNC connectors), which is usually tied electrically to the case. That, in turn, connects to the ground lead of the ac power line, via the 3-wire power cord. This means that you cannot measure voltages between the two arbitrary points in a circuit, but are forced to measure signals relative to this universal ground.

An important caution is in order here: If you try to connect the ground clip of an oscilloscope probe to a point in the circuit that is at some voltage relative to ground, you will end up shorting it to ground. This can have disastrous consequences to the circuit under test; in addition, it can be downright dangerous with circuits that are “hot to ground” (transformerless consumer electronics like television sets, for example). If it is imperative to look at the signal between two points, you can either “float” the scope by lifting the ground lead (not recommended, unless you know what you’re doing) or make a differential measurement by inverting one input channel and switching to ADD (some plug-in modules permit direct differential measurements).

Another caution about grounds when you’re measuring weak signals or high frequencies: Be sure the oscilloscope ground is the same as the circuit ground where you’re measuring. The best way to do this is by connecting the short ground wire on the probe body directly to the circuit ground, then checking by measuring the voltage of “ground” with the probe, observing no signal. One problem with this scheme is that those short ground clips are usually missing, lost! Keep your probe accessories in a drawer somewhere.

## OTHER SCOPE FEATURES

Many scopes have a DELAYED SWEEP that lets you see a segment of a waveform occurring some time after the trigger point. You can dial the delay accurately with a multiturn adjustment and a second sweep-speed switch. A delay mode known as A INTENSIFIED BY B lets you display the whole waveform at the first sweep speed, with the delayed segment brightened; this is handy during setup. Scopes with delayed sweep sometimes have “mixed sweep,” in which the trace begins at one sweep speed, then switches to a second (usually faster) speed after the selected delay. Another option is to begin the delayed sweep either immediately after the selected delay or at the next trigger point after the delay; there are two sets of trigger controls, so the two trigger points can be set individually. (Don’t confuse delayed sweep with “signal delay.” All good scopes have a delay in the signal channel, so you can display the event that caused the trigger; it lets you look a little bit backward in time!) Many scopes now have a TRIGGER HOLDOFF control; it inhibits triggering for an adjustable interval after each sweep, and it is very useful when viewing complicated waveforms without the simple periodicity of a sine wave, say. The usual case is a digital waveform with a complicated sequence of 1’s and 0’s, which won’t generate a stable display otherwise (except by adjustment of the sweep-speed vernier, which means you don’t get a calibrated sweep). There are also scopes with a “storage” that let you see a nonrepetitive event, and scopes that accept plug-in modules. These let you do just about anything, including display of eight simultaneous traces, spectrum analysis, accurate (digital) voltage and time measurements on waveforms, etc. Digital-storage analog oscilloscopes of a new generation are becoming popular; they let you catch a one-shot waveform, and even let you look backward in time (before the trigger event).

## MATH REVIEW

### APPENDIX B

Some knowledge of algebra and trigonometry is essential to understand this book. In addition, a limited ability to deal with complex numbers and derivatives (a part of calculus) is helpful, although not entirely essential. This appendix is meant as the briefest of summaries of complex numbers and differentiation. It is not meant as a textbook substitute. For a highly readable self-help book on calculus, we recommend *Quick Calculus*, by D. Kleppner and N. Ramsey (John Wiley & Sons, 1972).

#### COMPLEX NUMBERS

A complex number is an object of the form

$$N = a + bi$$

where  $a$  and  $b$  are real numbers and  $i$  (called  $j$  in the rest of the book, to avoid confusion with small-signal currents) is the square root of  $-1$ ;  $a$  is called the real part, and  $b$  is called the imaginary part. Boldface letters or squiggly underlines are sometimes used to denote complex numbers. At other times you're just supposed to *know*!

Complex numbers can be added, subtracted, multiplied, etc., just as real numbers:

$$\begin{aligned}(a + bi) + (c + di) &= (a + c) + (b + d)i \\(a + bi) - (c + di) &= (a - c) + (b - d)i \\(a + bi)(c + di) &= (ac - bd) + (bc + ad)i\end{aligned}$$

$$\begin{aligned}\frac{a + bi}{c + di} &= \frac{(a + bi)(c - di)}{(c + di)(c - di)} \\&= \frac{ac + bd}{c^2 + d^2} + \frac{bc - ad}{c^2 + d^2}i\end{aligned}$$

All these operations are natural, in the sense that you just treat  $i$  as something that multiplies the imaginary part, and go ahead with ordinary arithmetic. Note that  $i^2 = -1$  (used in the multiplication example) and that division is simplified by multiplying top and bottom by the *complex conjugate*, the number you get by changing the sign of the imaginary part. The complex conjugate is sometimes indicated with an asterisk. If

$$N = a + bi$$

then

$$N^* = a - bi$$

The magnitude (or *modulus*) of a complex number is

$$\begin{aligned}|N| = |a + bi| &= [(a + bi)(a - bi)]^{\frac{1}{2}} \\&= (a^2 + b^2)^{\frac{1}{2}}\end{aligned}$$

i.e.,

$$|N| = (NN^*)^{\frac{1}{2}}$$

simply obtained by multiplying by the complex conjugate and taking the square root. The magnitude of the product (or quotient) of two complex numbers is simply the product (or quotient) of their magnitudes.

The real (or imaginary) part of a complex number is sometimes written

real part of  $N = \text{Re}(N)$

imaginary part of  $N = \text{Im}(N)$



You get them by writing out the number in the form  $a + bi$ , then taking either  $a$  or  $b$ . This may involve some multiplication or division, since the complex number may be a real mess.

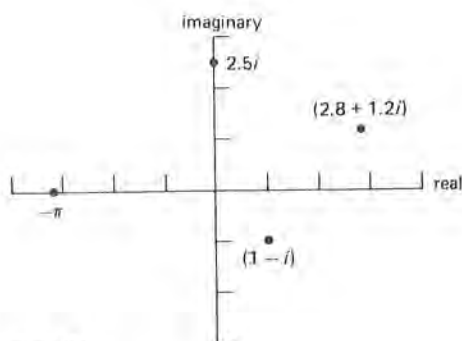


Figure B1

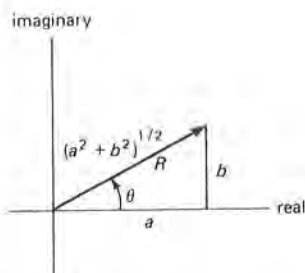


Figure B2

Complex numbers are sometimes represented on the complex plane. It looks just like an ordinary  $x, y$  graph, except that a complex number is plotted by taking its real part as  $x$  and its imaginary part as  $y$ ; i.e., the axes represent REAL ( $x$ ) and IMAGINARY ( $y$ ), as shown in Figure B1. In keeping with this analogy, you sometimes see complex numbers written just like  $x, y$  coordinates:

$$a + bi \leftrightarrow (a, b)$$

Just as with ordinary  $x, y$  pairs, complex numbers can be represented in polar coordinates; that's known as "magnitude, angle" representation. For example, the number  $a + bi$  can also be written (Fig. B2)

$$a + bi = (R, \theta)$$

where  $R = (a^2 + b^2)^{\frac{1}{2}}$  and  $\theta = \tan^{-1}(b/a)$ . This is usually written in a different way, using the fact that

$$e^{ix} = \cos x + i \sin x$$

(You can easily derive the preceding result, known as Euler's formula, by expanding the exponential in a Taylor series.) Thus we have the following equivalents:

$$N = a + bi = Re^{i\theta}$$

$$R = |N| = (NN^*)^{\frac{1}{2}} = (a^2 + b^2)^{\frac{1}{2}}$$

$$\theta = \tan^{-1}(b/a)$$

i.e., the modulus  $R$  and angle  $\theta$  are simply the polar coordinates of the point that represents the number in the complex plane. Polar form is handy when complex numbers have to be multiplied (or divided); you just multiply (divide) their magnitudes and add (subtract) their angles:

$$(ae^{ib})(ce^{id}) = ace^{i(b+d)}$$

Finally, to convert from polar to rectangular form, just use Euler's formula:

$$ae^{ib} = a \cos b + ia \sin b$$

i.e.,

$$\operatorname{Re}(ae^{ib}) = a \cos b$$

$$\operatorname{Im}(ae^{ib}) = a \sin b$$

If you have a complex number multiplying a complex exponential, just do the necessary multiplications. If

$$N = a + bi$$

$$\begin{aligned} Ne^{ix} &= (a + bi)(\cos x + i \sin x) \\ &= (a \cos x - b \sin x) \\ &\quad + i(b \cos x + a \sin x) \end{aligned}$$

## DIFFERENTIATION (CALCULUS)

We start with the concept of a *function*  $f(x)$ , i.e., a formula that gives a value  $y = f(x)$  for each  $x$ . The function  $f(x)$  should be *single-valued* i.e., it should give

a single value of  $y$  for each  $x$ . You can think of  $y = f(x)$  as a graph, as in Figure B3. The derivative of  $y$  with respect to  $x$ , written  $dy/dx$  ("dee  $y$  dee  $x$ "), is the *slope* of the graph of  $y$  versus  $x$ . If you draw a tangent to the curve at some point, its slope is  $dy/dx$  at that point; i.e., the derivative is itself a function, since it has a value at each point. In Figure B3, the slope at the point (1, 1) happens to be 2, whereas the slope at the origin is zero (you will see shortly how to compute the derivative).

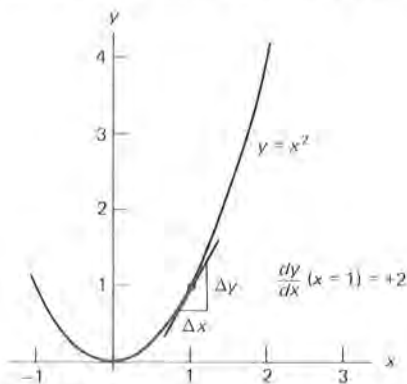


Figure B3

In mathematical terms, the derivative is the limiting value of the ratio of the change in  $y$  ( $\Delta y$ ) to the change in  $x$  ( $\Delta x$ ) corresponding to a small change in  $x$  ( $\Delta x$ ), as  $\Delta x$  goes to zero. To quote a song once sung in the hallowed halls of Harvard (by Tom Lehrer and Lewis Branscomb),

You take a function of  $x$ , and you call it  $y$   
Take any  $x$ -nought that you care to try  
Make a little change and call it delta  $x$   
The corresponding change in  $y$  is what  
you find nex'  
And then you take the quotient, and now,  
carefully  
Send delta  $x$  to zero, and I think you'll  
see  
That what the limit gives us (if our work  
all checks)  
Is what you call  $dy/dx$   
It's just  $dy/dx$   
(to the tune of "There'll Be Some Changes  
Made," W. Benton Overstreet).

Differentiation is a straightforward art, and the derivatives of many common functions are tabulated in standard tables. Here are some rules ( $u$  and  $v$  are arbitrary functions  $x$ ):

#### Some derivatives

$$\frac{d}{dx} x^n = nx^{n-1}$$

$$\frac{d}{dx} \sin x = \cos x$$

$$\frac{d}{dx} e^x = e^x$$

$$\frac{d}{dx} au(x) = a \frac{d}{dx} u(x) \quad (a = \text{constant})$$

$$\frac{d}{dx} (u + v) = \frac{du}{dx} + \frac{dv}{dx}$$

$$\frac{d}{dx} \left( \frac{u}{v} \right) = \frac{v \frac{du}{dx} - u \frac{dv}{dx}}{v^2}$$

$$\frac{d}{dx} \{u[v(x)]\} = \frac{du}{dv} \frac{dv}{dx}$$

The last one is very useful and is called the chain rule.

Once you have differentiated a function, you often want to evaluate the value of the derivative at some point. Other times you may want to find a minimum or maximum of the function; that's the same thing as having a zero derivative, so you can just set the derivative to zero and solve for  $x$ . Here are some examples:

$$\frac{d}{dx} x^2 = 2x \quad \left( \text{Fig. B3: slope} = 2 \text{ at } x = 1, \right. \\ \left. \text{slope} = 0 \text{ at } x = 0 \right)$$

$$\frac{d}{dx} xe^x = xe^x + e^x \quad (\text{product rule})$$

$$\frac{d}{dx} \sin(ax) = a \cos(ax) \quad (\text{chain rule})$$

$$\frac{d}{dx} a^x = \frac{d}{dx} (e^{x \log a}) = a^x \log a \quad (\text{chain rule})$$

$$\frac{d}{dx} \left( \frac{1}{x^{\frac{1}{2}}} \right) = -\frac{1}{2} x^{-\frac{3}{2}}$$

## THE 5% RESISTOR COLOR CODE

### APPENDIX C

Low-power axial-lead carbon-composition and film resistors with 2% to 20% tolerances have a standard set of values and a standard color-band marking scheme. Although it may seem diabolical to the beginner, the practice of color banding makes it easy to recognize resistor values in a circuit or parts bin, without having to search for a printed legend. The standard resistor values are chosen so that adjacent values have relative ratios of about 10% for the 2% and 5% tolerance types and 20% for the 10% and 20% tolerance types. Thus there are many values that could be described by the color code but that are not available.

Two digits and a multiplier digit determine the resistor value, and resistors are color-banded in that order starting from one end of the resistor (Fig. C1). A fourth

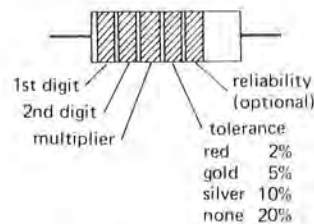
tolerance band is usually present, and occasionally you'll see a fifth band for other parameters (such as yellow or orange band for MIL spec reliability rating).

Here is the set of standard values for the first two digits (lightface type indicates 2% and 5% only):

|           |           |           |           |            |
|-----------|-----------|-----------|-----------|------------|
| <b>10</b> | 16        | 27        | 43        | <b>68</b>  |
| <b>11</b> | <b>18</b> | 30        | <b>47</b> | 75         |
| <b>12</b> | 20        | <b>33</b> | 51        | <b>82</b>  |
| 13        | <b>22</b> | 36        | <b>56</b> | 91         |
| <b>15</b> | 24        | <b>39</b> | 62        | <b>100</b> |

Carbon-composition resistors range in price from 3 cents each (in quantities of 1000) to 15 cents (quantities of 25). Distributors may be unwilling to sell less than 25 to 50 pieces of one value; thus an assortment box (made by Stackpole or Ohmite) may be a wise purchase.

| digit | color  | multiplier | number of zeros |
|-------|--------|------------|-----------------|
|       | silver | 0.01       | -2              |
|       | gold   | 0.1        | -1              |
| 0     | black  | 1          | 0               |
| 1     | brown  | 10         | 1               |
| 2     | red    | 100        | 2               |
| 3     | orange | 1k         | 3               |
| 4     | yellow | 10k        | 4               |
| 5     | green  | 100k       | 5               |
| 6     | blue   | 1 M        | 6               |
| 7     | violet | 10 M       | 7               |
| 8     | gray   |            |                 |
| 9     | white  |            |                 |



example: red-yellow-orange-gold is a 2, 4, and 3 zeros, or 24k 5%, resistor

Figure C1

## 1% PRECISION RESISTORS

### APPENDIX D

Metal-film precision resistors with  $\frac{1}{2}\%$  and 1% tolerance ratings have seen sufficient use in the industry to have attractively low prices. In particular, the RN55D and RN60D resistors are often available for as little as 5 cents each in quantities of 100, and a distributor may be willing to sell an assortment of mixed values at a quantity discount. The RN55D resistors are the same size as ordinary 1/4 watt "composition" resistors (although their military rating will be 1/10 or 1/8 watt at 70°C ambient temperature), whereas the RN60D are the size of 1/2 watt composition resistors. The RN55D resistors have a temperature coefficient of 100ppm/°C, and RN55C resistors (same size) have a 50ppm/°C rating.

Metal-film precision resistors use a four-digit code printed on the resistor body, rather than the ordinary color-banding scheme. The first three digits denote a value, and the last digit is the "number-of-zeros" multiplier. For example, 1693 denotes a 169k resistor, and 1000 denotes a 100 ohm resistor. (Note that the color bands work the same way, but with only three digits altogether. Many capacitor types use this same printed number scheme.) If the resistor's value is too small to be described this way, an R will be used to indicate the decimal point; for example, 49R9 is a 49.9 ohm resistor, and 10R0 is 10.0 ohms.

The standard values range from 10.0 ohms to 301kΩ by approximately 2% ratios, although some companies may offer similar (non-MIL-spec) resistors with values from 4.99 ohms to 2.00 MΩ. Standard

values in each decade are given in the list that follows.

One percent resistors are often used in applications that require excellent stability and accuracy; a small adjustable "trimmer" resistor may be connected in series to set a precise resistance value. But it's important to realize that from a worst-case standpoint, 1% resistors are only guaranteed to be within 1% of their rated value under a specified set of conditions. Resistance variation due to temperature change, high humidity, and operation at full rated power can easily exceed 1%. Resistance drift with time can approach 0.5%, particularly if the resistors are used at rated power. Circuits that require extremely accurate or stable performance (good to 0.1% or better, say) should use precision wire-wound resistors or some of the special metal-film resistors designed for such stability (for example, the Mepco 5023Z). This advice goes for composition resistors, as well. Resist, if you will, the temptation to regard the manufacturer's specifications as being overly conservative.

|     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|
| 100 | 140 | 196 | 274 | 383 | 536 | 750 |
| 102 | 143 | 200 | 280 | 392 | 549 | 768 |
| 105 | 147 | 205 | 287 | 402 | 562 | 787 |
| 107 | 150 | 210 | 294 | 412 | 576 | 806 |
| 110 | 154 | 215 | 301 | 422 | 590 | 825 |
| 113 | 158 | 221 | 309 | 432 | 604 | 845 |
| 115 | 162 | 226 | 316 | 442 | 619 | 866 |
| 118 | 165 | 232 | 324 | 453 | 634 | 887 |
| 121 | 169 | 237 | 332 | 464 | 649 | 909 |
| 124 | 174 | 243 | 340 | 475 | 665 | 931 |
| 127 | 178 | 249 | 348 | 487 | 681 | 953 |
| 130 | 182 | 255 | 357 | 499 | 698 | 976 |
| 133 | 187 | 261 | 365 | 511 | 715 |     |
| 137 | 191 | 267 | 374 | 523 | 732 |     |

**TABLE D.1. SELECTED RESISTOR TYPES**

| Property                             | Carbon comp<br>(RCR-07) | Metal film           |                            |                            |
|--------------------------------------|-------------------------|----------------------|----------------------------|----------------------------|
|                                      |                         | standard<br>(RN-55D) | precision<br>(Mepco 5023Z) | miniature<br>(Mepco 5063J) |
| <b>Load life</b><br>1000hrs @85°C    | 10%                     | 0.5%                 | 0.01%                      | 0.15%                      |
| <b>Moisture</b><br>Mil std 202       | 15%                     | 0.5%                 | 0.04%                      | 1%                         |
| <b>Temp cycle</b><br>-65°C to +150°C | 4%                      | 0.25%                | 0.005%                     | 0.25%                      |
| <b>Low-temp operation</b><br>-65°C   | 3%                      | 0.25%                | 0.01%                      | 0.25%                      |
| <b>Short-term overload</b>           | 2.5%                    | 0.25%                | 0.01%                      | 0.25%                      |
| <b>Soldering</b><br>@ 350°C          | 3%                      | 0.25%                | 0.01%                      | 0.25%                      |
| <b>Shock</b><br>50G, 11ms            | 2%                      | 0.25%                | 0.01%                      | 0.25%                      |
| <b>Vibration</b><br>10-2000Hz        | 2%                      | 0.25%                | 0.01%                      | 0.25%                      |
| <b>Storage</b><br>1 year             | —                       | —                    | 0.003%                     | —                          |
| <b>Tolerances avail</b>              | 5%, 10%                 | 0.1%—1%              | 0.025%—1%                  | 1%, 5%                     |
| <b>Tempco avail</b><br>ppm/°C        | 5000                    | 25—100               | 5—25                       | 100                        |
| <b>Voltage coef</b>                  | —                       | 5ppm/V               | 0.1ppm/V                   | —                          |
| <b>Thermal emf</b>                   | —                       | —                    | 2μV/°C                     | —                          |
| <b>Insulation resistance</b>         | —                       | —                    | 10,000MΩ                   | 1000MΩ                     |



## HOW TO DRAW SCHEMATIC DIAGRAMS

### APPENDIX E

A well-drawn schematic makes it easy to understand how a circuit works and aids in troubleshooting; a poor schematic only creates confusion. By keeping a few rules and suggestions in mind, you can draw a good schematic in no more time than it takes to draw a poor one. In this appendix we dispense advice of three varieties: general principles, rules, and hints. We have also drawn some real knee-slappers to illustrate habits to avoid.

#### GENERAL PRINCIPLES

1. Schematics should be unambiguous. Therefore, pin numbers, parts values, polarities, etc., should be clearly labeled to avoid confusion.
2. A good schematic makes circuit functions clear. Therefore, keep functional areas distinct; don't be afraid to leave blank areas on the page, and don't try to fill the page. There are conventional ways to draw functional subunits; for instance, don't draw a differential amplifier as in Figure E1, because the function won't be easily recognized. Likewise, flip-flops are usually drawn with clock and inputs on the left, set and clear on top and bottom, and outputs on the right.

#### RULES

1. Wires connecting are indicated by a heavy black dot; wires crossing, but not connecting, have no dot (don't use a little half-circular "jog"; it went out in the 1950s).

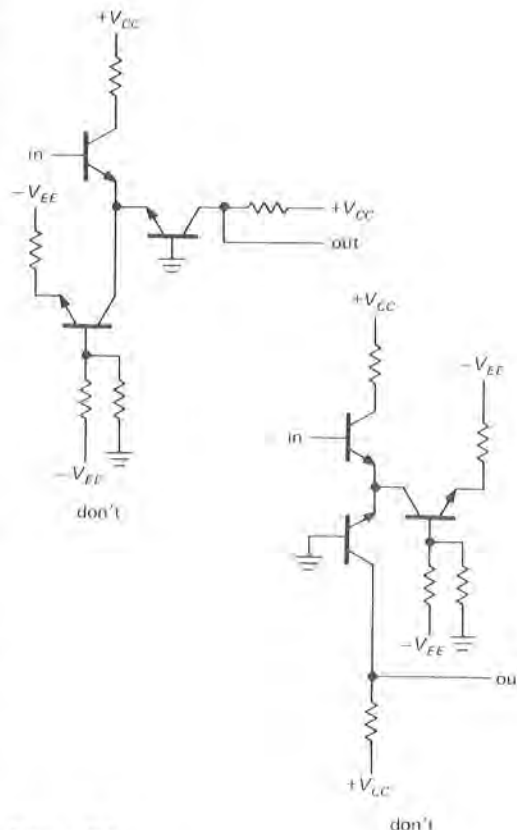


Figure E1

2. Four wires must not connect at a point; i.e., wires must not cross *and* connect.
3. Always use the same symbol for the same device; e.g., don't draw flip-flops in two different ways (exception: assertion-level logic symbols show each gate in two possible ways).
4. Wires and components are aligned horizontally or vertically, unless there's a good reason to do otherwise.

5. Label pin numbers on the outside of a symbol, signal names on the inside.
6. All parts should have values or types indicated; it's best to give all parts a label, too, e.g.,  $R_7$  or  $IC_3$ .

### HINTS

1. Identify parts immediately adjacent to the symbol, forming a distinct group giving symbol, label, and type or value.
2. In general, signals go from left to right; don't be dogmatic about this, though, if clarity is sacrificed.
3. Put positive supply voltages at the top of the page, negative at the bottom. Thus, *npn* transistors will usually have their emitter at the bottom, whereas *pnp*'s will have their emitter topmost.
4. Don't attempt to bring all wires around to the supply rails, or to a common ground wire. Instead, use the ground symbol(s) and labels like  $+V_{CC}$  to indicate those voltages where needed.
5. It is helpful to label signals and functional blocks and show waveforms; in logic diagrams it is especially important to label signal lines, e.g., RESET' or CLK.
6. It is helpful to bring leads away from components a short distance before making connections or jogs. For example, draw transistors as in Figure E2.
7. Leave some space around circuit symbols; e.g., don't draw components or wires

too close to an op-amp symbol. This keeps the drawing uncluttered and leaves room for labels, pin numbers, etc.

8. Label all boxes that aren't obvious: comparator versus op-amp, shift register versus counter, etc. Don't be afraid to invent a new symbol.

9. Use small rectangles, ovals, or circles to indicate card-edge connections, connector pins, etc. Be consistent.

10. The signal path through switches should be clear. Don't force the reader to follow wires all over the page to find out how a signal is switched.

11. Power-supply connections are normally assumed for op-amps and logic devices. However, show any unusual connections (e.g., an op-amp run from a single supply, where  $V_- = \text{ground}$ ) and the disposition of unused inputs.

12. It is very helpful to include a small table of IC numbers, types, and power-supply connections (pin numbers for  $V_{CC}$  and ground, for instance).

13. Include a title area near the bottom of the page, with name of circuit, name of instrument, by whom drawn, by whom designed or checked, date, and assembly number. Also include a revision area, with columns for revision number, date, and subject.

14. We recommend drawing schematics freehand on coarse graph paper (nonreproducing blue, 4 to 8 lines per inch) or on plain paper on top of graph paper. This is fast, and it gives very pleasing results. Use dark pencil or ink; avoid ball-point pen.

As an illustration, we've drawn a humble example (Fig. E3) showing "awful" and "good" schematics of the same circuit; the former violates nearly every rule and is almost impossible to understand. See how many bad habits you can find illustrated. We've seen all of them in professionally drawn schematics! (Drawing the "bad" schematic was an occasion of great hilarity; we laughed ourselves silly.)

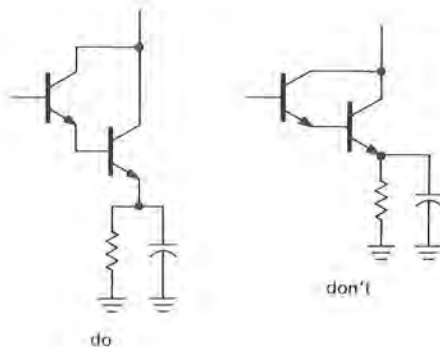


Figure E2

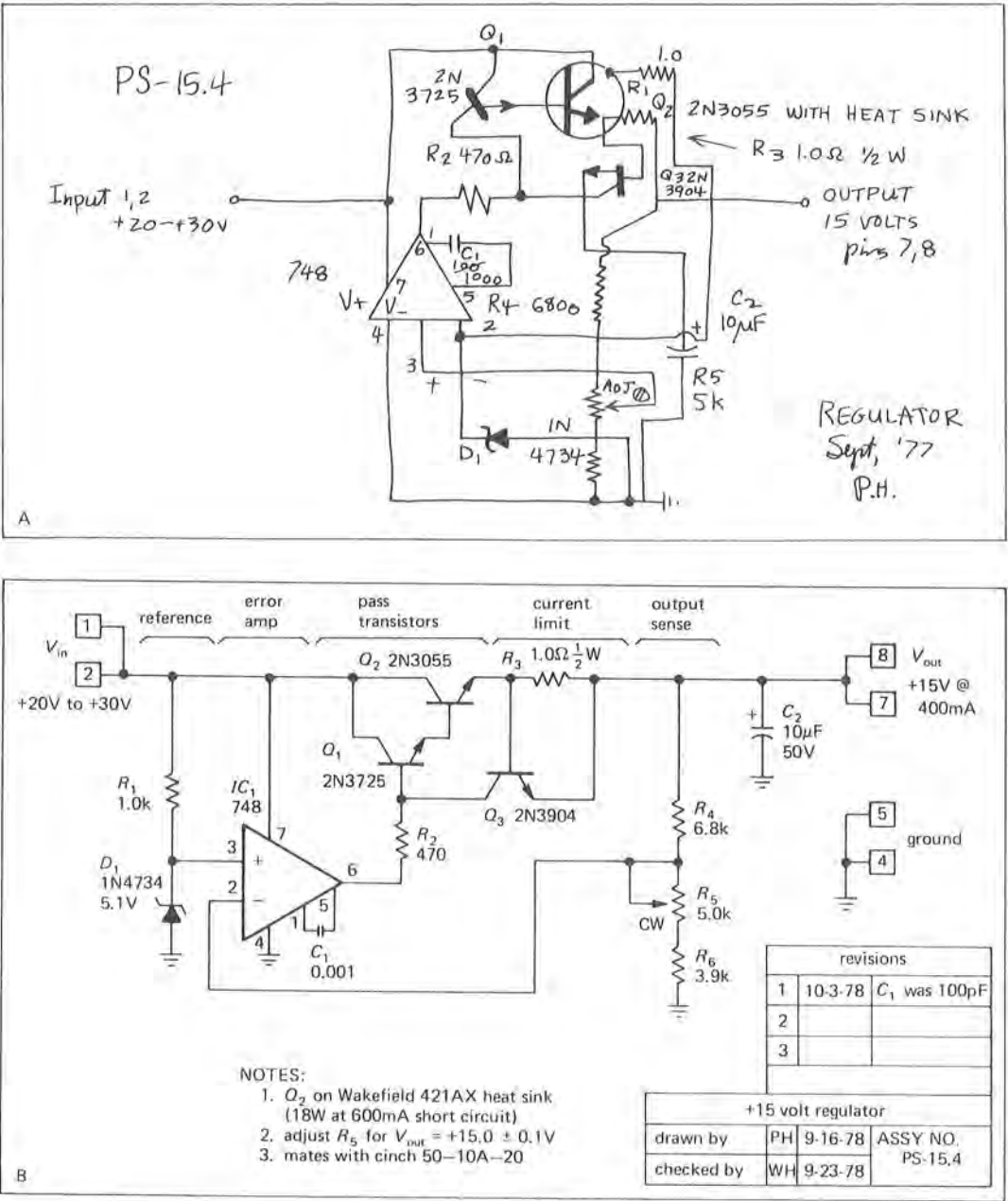


Figure E3

## LOAD LINES

### APPENDIX F

The graphic method of “load lines” usually makes an early appearance in electronic textbooks. We have avoided it because, well, it just isn’t useful in transistor design, the way it was in vacuum-tube circuit design. However, it is of use in dealing with some nonlinear devices (tunnel diodes, for example), and in any case it is a useful conceptual tool.

Let’s start with an example. Suppose you want to know the voltage across the diode in Figure F1. Assume that you know the voltage-versus-current ( $V$ - $I$ ) curve of the particular diode (of course, it would have a manufacturing “spread,” as well as depending on ambient temperature); it might look something like the curve drawn. How would you figure out the quiescent point?

One method might be to guess a rough value of current, say 0.6mA, then use the curve to get the drop across the resistor, from which you get a new guess for the current (in this case, 0.48mA). This

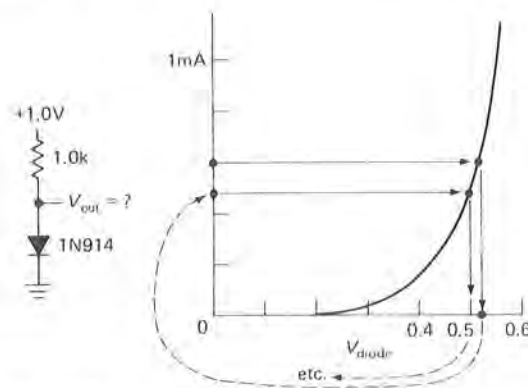


Figure F1

iterative method is suggested in Figure F1. After a few iterations, this method will get you an answer, but it leaves a lot to be desired.

The method of load lines gets you the answer to this sort of problem immediately. Imagine *any* device connected in place of the diode; the 1.0k resistor is still the load. Now plot, on a  $V$ - $I$  graph, the curve of resistor current versus device voltage. This turns out to be easy: at zero volts the current is just  $V_+/R$  (full drop across the resistor); at  $V_+$  volts the current is zero; points in between fall on a straight line between the two. Now, on the same graph, plot the  $V$ - $I$  curve of the device. The operating point lies on both curves, i.e., at the intersection, as shown in Figure F2.

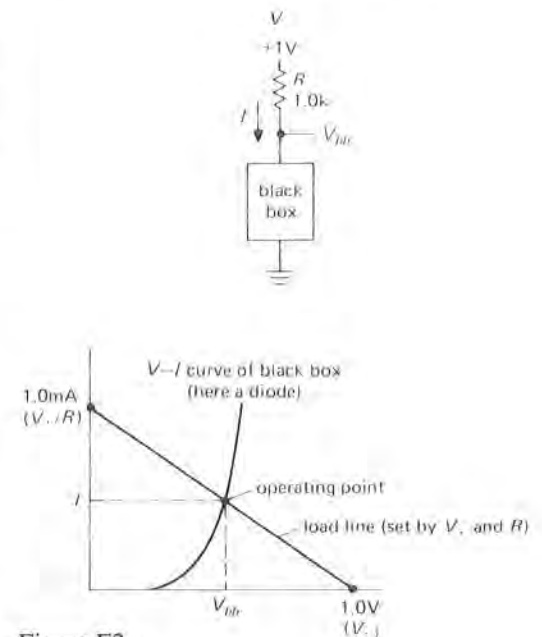


Figure F2

Load lines can be used with a 3-terminal device (tube or transistor, for example) by plotting a family of curves for the device. Figure F3 shows what such a thing would look like for a depletion-mode FET, with the curve family parameterized by gate-source voltage. You can read off the output for a given input by sliding along the load line between appropriate curves corresponding to the input you've got, then projecting onto the voltage axis. In this example we've done this, showing the drain voltage (output) for a gate swing (input) between ground and  $-2$  volts.

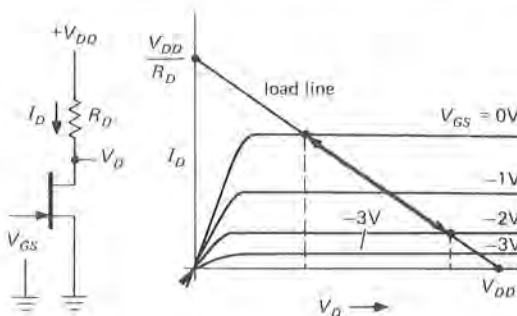


Figure F3

As nice as this method seems, it has very limited use for transistor or FET design, for a couple of reasons. For one thing, the curves published for semiconductor devices are "typical," with manufacturing spread that can be as large as a factor of 5. Imagine what would happen to those nice load-line solutions if all the curves shrank by a factor of 4! Another reason is that for an inherently logarithmic device like a diode junction, a linear load-line graph can be used to give accurate results only over a narrow region. Finally, the nongraphic methods we've used in this book are adequate to handle solid-state design. In particular, these methods emphasize the parameters you can count on ( $r_e$ ,  $I_C$  vs.  $V_{BE}$  and  $T$ , etc.), rather than the ones that are highly variable ( $h_{FE}$ ,  $V_P$ , etc.). If anything, the use of load lines on published

curves for transistors only gives you a false sense of security, since the device spread isn't also shown.

Load lines turn out to be very useful in understanding the circuit behavior of highly nonlinear devices. The example of tunnel diodes illustrates a couple of interesting points. Let's analyze the circuit in Figure F4. Note that in this case,  $V_{in}$  takes the place of the supply voltage in the previous examples. So a signal swing will generate a family of parallel load lines intersecting with a single device  $V$ - $I$  curve (Fig. F5A). The values shown are for a 100 ohm load resistor. As can be seen, the output varies most rapidly as the input swing takes the load line across the negative-resistance portion of the tunnel-diode curve. By reading off values of  $V_{out}$  (projection on the  $x$  axis) for various values of  $V_{in}$  (individual load lines), you get the "transfer" characteristics shown. This particular circuit has some voltage gain for input voltages near 0.2 volt.

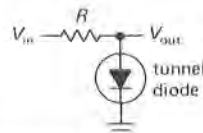


Figure F4

An interesting thing happens if the load lines become flatter than the middle section of the diode curve. That happens when the load resistance exceeds the magnitude of the diode's negative resistance. It is then possible to have *two* intersection points, as in Figure F6. A rising input signal carries the load lines up until the intersection point has nowhere to go and has to jump across to a higher  $V_{out}$  value. On returning, the load lines similarly carry the intersection point down until it must again jump back. The overall transfer characteristic has hysteresis, as shown. Tunnel diodes are used in this manner as fast switching devices (triggers).



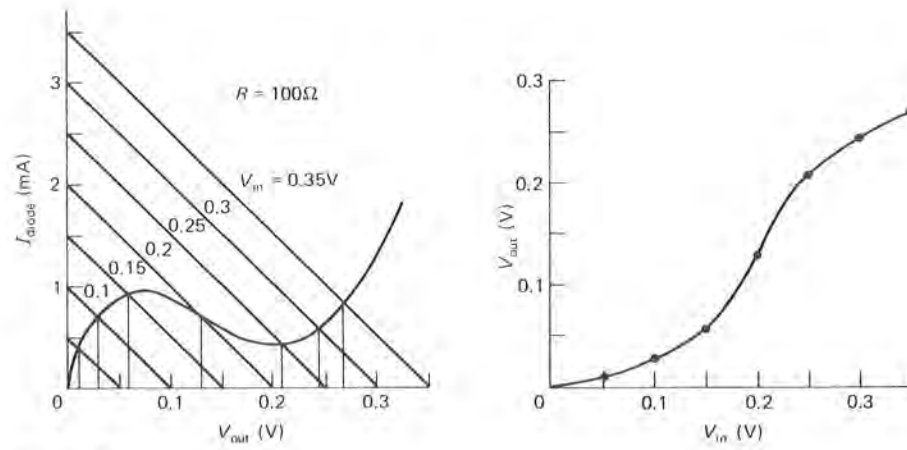


Figure F5

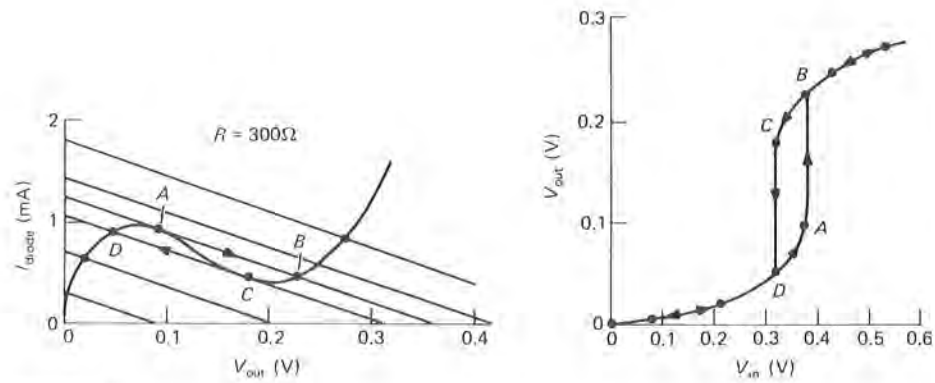


Figure F6

## TRANSISTOR SATURATION APPENDIX G

The subtitle of this appendix might be “transistor man defeated by the collector-base diode.” With a simple model we can see the reason for the finite “saturation voltage” exhibited by bipolar transistors. The basic idea is that the collector-base junction is a big diode, with a high  $I_S$  (Ebers-Moll equation), so that it has a lower ON voltage for a given current than the base-emitter diode. Therefore, at small values of collector-to-emitter voltage (typically 0.25V or less), some of the base current will be “robbed” by conduction of the collector-base diode (Fig. G1). This lowers the effective  $h_{FE}$  and makes it necessary to supply relatively large base currents to bring the collector close to the emitter, as shown in the measured data of Figure G2.

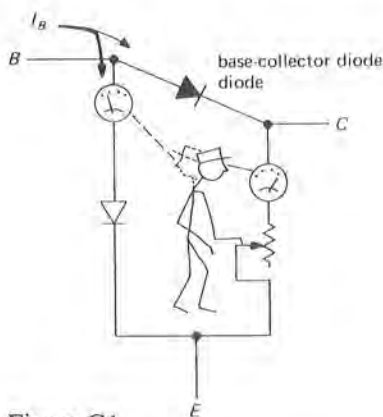


Figure G1

$V_{CE(sat)}$ , the collector saturation voltage at a particular value of base current and collector current, is also relatively

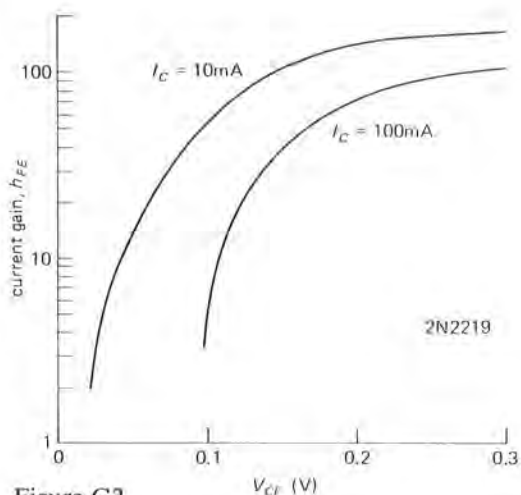


Figure G2

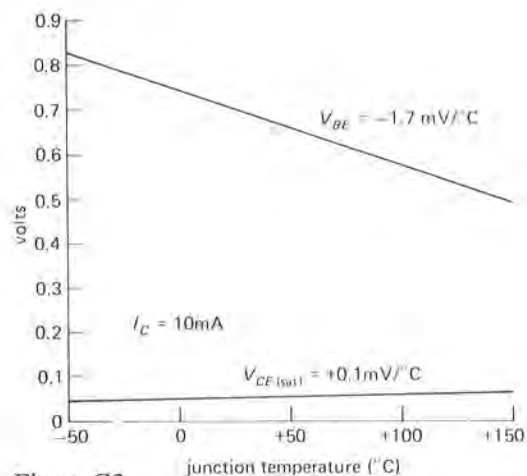


Figure G3

independent of temperature because of cancellation of the temperature coefficients of the two diodes (Fig. G3). This is of interest because a saturated transistor is

frequently used to switch large currents and may get hot (e.g., 10A at a saturation voltage of 0.5V is 5W, enough to bring the junction of a small power transistor to 100°C or more).

In saturated switching applications you usually provide generous amounts of base current (typically 1/10 or 1/20 of the collector current) to achieve values of  $V_{CE(sat)}$  of 0.05 to 0.2 volt. If the load inadvertently demands much greater collector currents, the transistor will go out of saturation, with greatly increased power dissipation. The measured data in Figure G4 show that it is hard to define exactly when a transistor is “saturated”;

you might use some arbitrary criterion such as  $I_C = 10I_B$ .

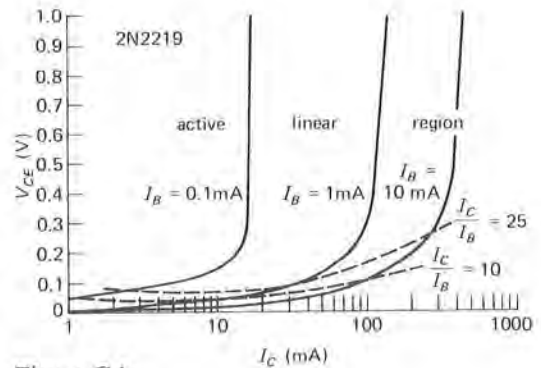


Figure G4

## LC BUTTERWORTH FILTERS

### APPENDIX H

Active filters, as discussed in Chapter 4, are very convenient at low frequencies, but they are impractical at radiofrequencies because of the slew-rate and bandwidth requirements they impose on the operational amplifiers. At frequencies of 100kHz and above (and often at lower frequencies), the best approach is to design a passive filter with inductors and capacitors. (Of course, at UHF and microwave frequencies these “lumped-component” filters are replaced by stripline and cavity filters.)

As with active filters, there are many methods and filter characteristics possible with LC filters. For example, you can design the classic Butterworth, Chebyshev, and Bessel filters, each in low-pass, band-pass, high-pass, and band-reject varieties. It turns out that the Butterworth filter is

particularly easy to design, and we can present in just a page or two all the essential design information for low-pass and high-pass Butterworth LC filters, and even a few examples. For further information we recommend the excellent handbook by Zverev cited in the Bibliography.

Table H1 gives the values of normalized inductances and capacitances for low-pass filters of various orders, from which actual circuit values are obtained by the frequency and impedance scaling rules:

*Low-pass scaling rules:*

$$L_n(\text{actual}) = \frac{R_L L_n(\text{table})}{\omega}$$

$$C_n(\text{actual}) = \frac{C_n(\text{table})}{\omega R_L}$$

TABLE H1. BUTTERWORTH LOW-PASS FILTERS<sup>a</sup> ( $R_L = 1\Omega$ )

| $\pi$<br>$T$ | $R_S$<br>$1/R_S$ | $C_1$<br>$L_1$ | $L_2$<br>$C_2$ | $C_3$<br>$L_3$ | $L_4$<br>$C_4$ | $C_5$<br>$L_5$ | $L_6$<br>$C_6$ | $C_7$<br>$L_7$ | $L_8$<br>$C_8$ |
|--------------|------------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|
| $n = 2$      | 1                | 1.4142         | 1.4142         |                |                |                |                |                |                |
|              | $\infty$         | 1.4142         | 0.7071         |                |                |                |                |                |                |
| $n = 3$      | 1                | 1.0            | 2.0            | 1.0            |                |                |                |                |                |
|              | $\infty$         | 1.5            | 1.3333         | 0.5            |                |                |                |                |                |
| $n = 4$      | 1                | 0.7654         | 1.8478         | 1.8478         | 0.7654         |                |                |                |                |
|              | $\infty$         | 1.5307         | 1.5772         | 1.0824         | 0.3827         |                |                |                |                |
| $n = 5$      | 1                | 0.6180         | 1.6180         | 2.0            | 1.6180         | 0.6180         |                |                |                |
|              | $\infty$         | 1.5451         | 1.6944         | 1.3820         | 0.8944         | 0.3090         |                |                |                |
| $n = 6$      | 1                | 0.5176         | 1.4142         | 1.9319         | 1.9319         | 1.4142         | 0.5176         |                |                |
|              | $\infty$         | 1.5529         | 1.7593         | 1.5529         | 1.2016         | 0.7579         | 0.2588         |                |                |
| $n = 7$      | 1                | 0.4450         | 1.2470         | 1.8019         | 2.0            | 1.8019         | 1.2470         | 0.4450         |                |
|              | $\infty$         | 1.5576         | 1.7988         | 1.6588         | 1.3972         | 1.0550         | 0.6560         | 0.2225         |                |
| $n = 8$      | 1                | 0.3902         | 1.1111         | 1.6629         | 1.9616         | 1.9616         | 1.6629         | 1.1111         | 0.3902         |
|              | $\infty$         | 1.5607         | 1.8246         | 1.7287         | 1.5283         | 1.2588         | 0.9371         | 0.5776         | 0.1951         |

<sup>a</sup>Values of  $L_n$ ,  $C_n$  for  $1\Omega$  load resistance and cutoff frequency ( $-3\text{dB}$ ) of 1 rad/s. See text for scaling rules.

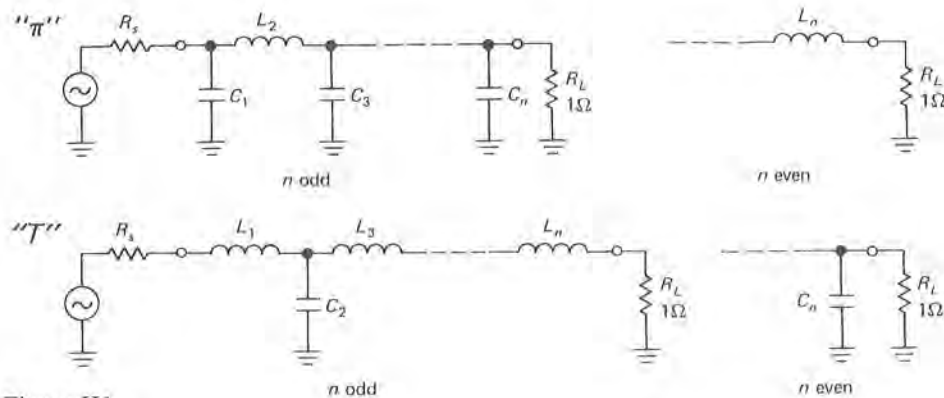


Figure H1

where  $R_L$  is the load impedance and  $\omega$  is the angular frequency ( $\omega = 2\pi f$ ).

Table H1 gives normalized values for 2-pole through 8-pole low-pass filters, for the two most common cases, namely (a) equal source and load impedances and (b) either source or load impedance much larger than the other. To use the table, first decide how many poles you need, based on the Butterworth response (graphs are plotted in Sections 5.05 and 5.07). Then use the preceding equations to determine the filter configuration ( $T$  or  $\pi$ ; see Fig. H1) and component values. For equal source and load impedances, either configuration is OK; the  $\pi$  configuration may be preferable because it requires fewer inductors. For load impedance much higher (lower) than the source impedance, use the  $T$  ( $\pi$ ) configuration.

To design a high-pass filter, follow the procedure outlined to determine which filter configuration to use and how many poles are necessary. Then do the universal low-pass to high-pass transformation shown in Figure H2, which consists simply of replacing inductors by capacitors, and vice versa. The actual component values are determined from the normalized values in Table H1 by the following frequency and impedance scaling rules:

*High-pass scaling rules:*

$$L_n(\text{actual}) = \frac{R_L}{\omega C_n(\text{table})}$$

$$C_n(\text{actual}) = \frac{1}{R_L \omega L_n(\text{table})}$$

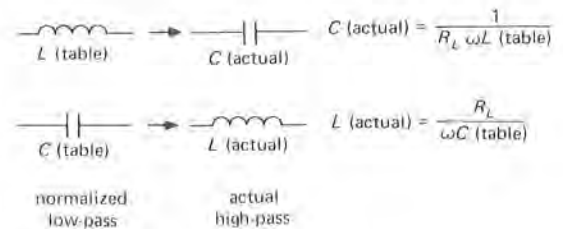


Figure H2

We will show how to use the table to design both low-pass and high-pass filters with a few examples.

### EXAMPLE I

Design a 5-pole low-pass filter for source and load impedance of 75 ohms, with a cutoff frequency ( $-3\text{dB}$ ) of 1MHz.

We use the  $\pi$  configuration to minimize the number of required inductors. The scaling rules give us



$$C_1 = C_5 = \frac{0.618}{2\pi \times 10^6 \times 75} = 1310 \text{ pF}$$

$$L_2 = L_4 = \frac{75 \times 1.618}{2\pi \times 10^6} = 19.3 \mu\text{H}$$

$$C_3 = \frac{2}{2\pi \times 10^6 \times 75} = 4240 \text{ pF}$$

The complete filter is shown in Figure H3. Note that all filters with equal source and load impedances will be symmetrical.

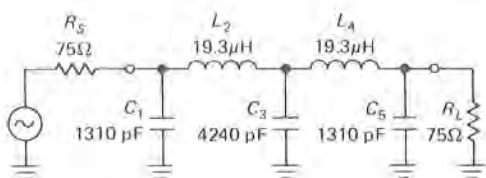


Figure H3

#### EXAMPLE II

Design a 3-pole low-pass filter for a source impedance of 50 ohms and a load impedance of 10k, with a cutoff frequency of 100kHz.

We use the  $T$  configuration, because  $R_S \ll R_L$ . For  $R_L = 10\text{k}$ , the scaling rules give

$$L_1 = \frac{10^4 \times 1.5}{2\pi \times 10^5} = 23.9 \text{ mH}$$

$$C_2 = \frac{1.3333}{2\pi \times 10^5 \times 10^4} = 212 \text{ pF}$$

$$L_3 = \frac{10^4 \times 0.5}{2\pi \times 10^5} = 7.96 \text{ mH}$$

The complete filter is shown in Figure H4.

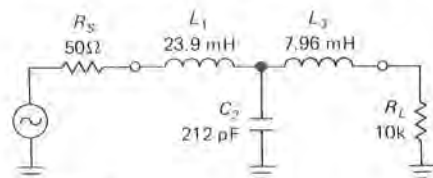


Figure H4

#### EXAMPLE III

Design a 4-pole low-pass filter for a zero-impedance source (voltage source) and a 75 ohm load, with cutoff frequency of 10MHz.

We use the  $T$  configuration, as in the previous example, because  $R_S \ll R_L$ . The scaling rules give

$$L_1 = \frac{75 \times 1.5307}{2\pi \times 10^7} = 1.83 \mu\text{H}$$

$$C_2 = \frac{1.5772}{2\pi \times 10^7 \times 75} = 335 \text{ pF}$$

$$L_3 = \frac{75 \times 1.0824}{2\pi \times 10^7} = 1.29 \mu\text{H}$$

$$C_4 = \frac{0.3827}{2\pi \times 10^7 \times 75} = 81.2 \text{ pF}$$

The complete filter is shown in Figure H5.

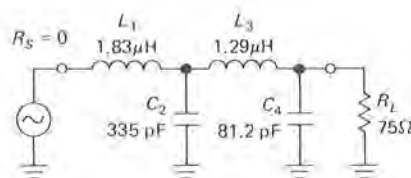


Figure H5

#### EXAMPLE IV

Design a 2-pole low-pass filter for current-source drive and 1k load impedance, with cutoff frequency of 10kHz.

We use the  $\pi$  configuration, because  $R_S \gg R_L$ . The scaling rules give

$$C_1 = \frac{1.4142}{2\pi \times 10^4 \times 10^3} = 0.0225 \mu\text{F}$$

$$L_2 = \frac{10^3 \times 0.7071}{2\pi \times 10^4} = 11.3 \text{ mH}$$

The complete filter is shown in Figure H6.

### EXAMPLE V

Design a 3-pole high-pass filter for 52 ohm source and load impedances, with cutoff frequency of 6MHz.

We begin with the  $T$  configuration, then transform inductors to capacitors, and vice versa, giving

$$C_1 = C_3 = \frac{1}{52 \times 2\pi \times 6 \times 10^6 \times 1.0} = 510\text{pF}$$

$$L_2 = \frac{52}{2\pi \times 6 \times 10^6 \times 2.0} = 0.690\mu\text{H}$$

The complete filter is shown in Figure H7.

We could like to emphasize that the field of passive filter design is rich and varied and that this simple table of Butterworth

filters doesn't even begin to scratch the surface.

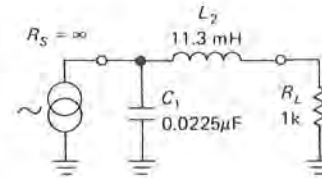


Figure H6

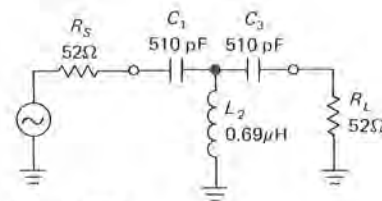


Figure H7

## ELECTRONICS MAGAZINES AND JOURNALS

### APPENDIX I

In this appendix we've gathered together a selection of magazines that are worth knowing about. Most are well supplied with advertisements proclaiming the specifications and virtues of new ICs, instruments, computers, etc., and these ads are a good source of information about new products; they don't have the nuisance value of ordinary newspaper and magazine ads. There is often a "reader information card" (or "bingo" card) at the back of the magazine on which you can circle the numbers corresponding to advertisements that you're interested in. More information then arrives in the mail within a few weeks. The system works well.

#### ELECTRONICS MAGAZINES

**EDN; Electronic Design; Electronics; Electronic Products.** At least one of these magazines should be considered required reading to keep up with new components and design ideas. The advertisements are as important as the articles.

**EE Times; Electronic News.** Newspapers for the electronics industry.

**Electronics and Wireless World.** British all-around electronics magazine for hobbyists and professionals.

**Ham Radio.** The most technical of the amateur radio publications.

**Journal of Solid State Circuits** (IEEE). Circuit design and new ICs.

**QST.** Amateur radio magazine from the official ARRL.

**Spectrum** (IEEE). General-interest electronics magazine, put out by IEEE. Good review articles covering broad range of subjects.

#### COMPUTER MAGAZINES

**Byte.** The first large-circulation personal computing magazine. Good tutorials on wide range of topics to which computers can be applied.

**Computer Design.** Definitive magazine on digital hardware and software techniques for large computer systems.

**Computers in Physics.** The title says it all.

**Dr. Dobbs Journal.** For programmers. Emphasis on software systems and design.

**MacWorld; MacUser.** End-user oriented evaluation of hardware and software; product reviews.

**PC Magazine.** Helpful magazine for the PC owner.

**PC Tech Journal.** Semi-technical magazine of PC computing.

**PC Week; Infoworld; Macintosh Today.** Newspaper-format trade weeklies, with the latest scuttlebutt.

#### OTHER MAGAZINES

**Measurement and Control News.** Includes biomedical and chemical instrumentation.

**Nuclear Instruments and Methods; Review of Scientific Instruments.** Scientific instrumentation.

## IC PREFIXES

### APPENDIX J

Typical problem: You need to replace an integrated circuit, or at least find some data on it. It says

DM8095N  
7410 NS

and lives in a 16-pin DIP. What is it? The 7410 has a familiar sound, so you order a few. A week later they arrive, in 14-pin DIPS! Banging your head on the nearest wall (since you should have known this all along), you realize you're back where you were a week ago, but with a handful of spare 3-input NANDs for consolation.

What's needed is a master list of IC prefixes, from which you can quickly identify the manufacturer. This appendix is our attempt to bring some order out of chaos. We make no pretense of accuracy or completeness, especially since the list is proliferating daily. (That mystery chip is a National Semiconductor 8095 hex three-state TTL buffer, by the way, manufactured in the 10th week of 1974.)

#### PREFIXES

The various semiconductor manufacturers use distinctive (usually) prefixes in front of the IC number, even if it is an IC type made by many different companies. The DM in the preceding example indicates a digital monolithic IC made by National Semiconductor (also indicated by the NS logo). Here is a list of most of the prefixes now in use:

| Prefixes                              | Manufacturer                    |
|---------------------------------------|---------------------------------|
| ACF, AY, GIC,<br>GP, SPR              | General Instrument (GI)         |
| AD, CAV, HAS,<br>HDM                  | Analog Devices                  |
| ADC, DM, DS, LF<br>LFT, LH,<br>LM, NH | National<br>Semiconductor (NSC) |
| AH                                    | Optical Electronics Inc.        |
| Am                                    | Advanced Micro<br>Devices (AMD) |
| AM                                    | Datel                           |
| AN                                    | Panasonic                       |
| Bt                                    | Brooktree                       |
| BX, CX                                | Sony                            |
| C, I, i                               | Intel                           |
| CA, CD, CDP                           | GE/RCA                          |
| CA, TDC, MPY,<br>THC, TMC             | TRW                             |
| CM, HV                                | Supertex                        |
| CLC                                   | Comlinear                       |
| CMP, DAC, MAT,<br>OP, PM, REF,<br>SSS | Precision<br>Monolithics        |
| CY                                    | Cypress                         |
| D, DE, DG, SI                         | Siliconix                       |
| DS                                    | Dallas<br>Semiconductor         |
| EF, ET, MK,<br>SFC, TDF, TS,          | Thomson/Mostek                  |
| EP, EPM, PL                           | Altera                          |
| F, $\mu$ A, $\mu$ L,<br>Unx           | Fairchild/NSC                   |
| FSS, ZLD                              | Ferranti                        |
| GA                                    | Gazelle                         |
| GAL                                   | Lattice                         |
| GEL                                   | GE                              |
| HA, HI                                | Harris                          |

|                                       |                                         |                                                                               |                                                         |
|---------------------------------------|-----------------------------------------|-------------------------------------------------------------------------------|---------------------------------------------------------|
| HA, HD, HG,<br>HL, HM, HN             | Hitachi                                 | RD, RF, RM,<br>RT, RU                                                         | EG&G Reticon                                            |
| HADC, HDAC                            | Honeywell                               | S                                                                             | AMI                                                     |
| HEP, MC, MCC,<br>MCM, MEC,<br>MM, MWM | Motorola                                | SFC                                                                           | ESMF                                                    |
| ICH, ICL, ICM,<br>IM                  | GE/Intersil                             | SG                                                                            | Silicon General                                         |
| IDT                                   | Integrated Device<br>Technology         | SN, TL, TLC,<br>TMS                                                           | Texas Instruments (TI)                                  |
| IMS                                   | Siemens                                 | SS                                                                            | Silicon Systems                                         |
| INA, ISO, OPA,<br>PWR                 | Inmos                                   | T, TA, TC, TD<br>TMM, TMP                                                     | Toshiba                                                 |
| IR                                    | Burr-Brown                              | OM, PCD, PCF,<br>SAA, SAB, SAF,<br>SCB, SCN, TAA,<br>TBA, TCA, TDA,<br>TEA, U | AEG, Amperex, SGS,<br>Siemens, Signetics,<br>Telefunken |
| ITT, MIC                              | Sharp                                   | TML                                                                           | Telmos                                                  |
| KA                                    | ITT                                     | TP                                                                            | Teledyne Philbrick                                      |
| L                                     | Samsung                                 | TPQ, UCN, UCS,<br>UDN, UDS, UHP,<br>ULN, ULS                                  | Sprague                                                 |
| L, LD                                 | SGS                                     | TSC                                                                           | Teledyne Semiconductor                                  |
| L, UC                                 | Siliconix, Siltronics                   | $\mu$ PB, $\mu$ PC, $\mu$ PD                                                  | NEC                                                     |
| LA, LC                                | Unitrode                                | V                                                                             | Amtel                                                   |
| LS                                    | Sanyo                                   | VA, VC                                                                        | VTC                                                     |
| LT, LTC, LTZ                          | LSI Computer Systems                    | VT                                                                            | VLSI Technology Inc.<br>(VTI)                           |
| M                                     | Linear Technology Corp.                 | X                                                                             | Xicor                                                   |
| MA                                    | Mitsubishi                              | XC                                                                            | Xilinx                                                  |
| MAX                                   | Analog Systems,<br>Marconi              | XR                                                                            | Exar                                                    |
| MB                                    | Maxim                                   | Z                                                                             | Zilog                                                   |
| MCS                                   | Fujitsu                                 | ZN                                                                            | Ferranti                                                |
| MIL                                   | MOS Technology                          | 5082- <i>nnnn</i>                                                             | Hewlett-Packard (HP)                                    |
| ML, MN, SL,<br>SP, TAB                | Microsystems<br>International           |                                                                               |                                                         |
| ML, MT                                | Plessey                                 |                                                                               |                                                         |
| MM                                    | Mitel                                   |                                                                               |                                                         |
| MN                                    | Teledyne-Amelco,<br>Monolithic Memories |                                                                               |                                                         |
| MP                                    | Micro Networks                          |                                                                               |                                                         |
| MSM                                   | Micro Power Systems                     |                                                                               |                                                         |
| N, NE, PLS,<br>S, SE, SP              | Oki                                     |                                                                               |                                                         |
| <i>nnG</i>                            | Signetics                               |                                                                               |                                                         |
| NC                                    | Gigabit Logic                           |                                                                               |                                                         |
| PA                                    | Nitron                                  |                                                                               |                                                         |
| PAL                                   | Apex                                    |                                                                               |                                                         |
| R                                     | AMD/MMI                                 |                                                                               |                                                         |
| R, Ray, RC, RM                        | Rockwell                                |                                                                               |                                                         |
|                                       | Raytheon                                |                                                                               |                                                         |

#### Suffixes

Suffix letters indicate package type and temperature range. There are three standard temperature ranges: "Military" ( $-55^{\circ}\text{C}$  to  $+125^{\circ}\text{C}$ ), "Industrial" ( $-25^{\circ}\text{C}$  to  $+85^{\circ}\text{C}$ ), and "Commercial" ( $0^{\circ}\text{C}$  to  $+70^{\circ}\text{C}$ ). Commercial is adequate for anything intended for use in normal indoor environments. As luck would have it, each manufacturer has its own set of suffixes, subject to frequent modification. Be sure to look up the correct suffix before you order, or ask the distributor for assistance.



**DATE CODES**

Most ICs and transistors, and many other electronic components, are stamped with a simple four-digit code giving date of manufacture: the first two digits are the year, the last two are the week of the year. In the example given earlier, 7410 means the second week of March 1974. They're sometimes useful, for example to estimate the age of components that have a finite useful life (electrolytic capacitors, for instance); unfortunately the components with the

shortest life (batteries) are often purposely coded so you can't figure out the date. If you get a batch of ICs with an abnormally high failure rate (most manufacturers test only a sample of each batch; typically 0.01% to 0.1% of the ICs you buy will not meet specifications), avoid replacements with the same date code. Date codes can also help you estimate the date of manufacture of commercial electronic equipment. Since ICs don't become stale, there's no reason to avoid an IC with an old date code.

## DATA SHEETS

### APPENDIX K

In this appendix we have reproduced three data sheets just as they were printed by the manufacturer. We chose representative or popular devices, looking especially for data sheets that were comprehensive and clear.

On the following pages you will find data sheets for these devices:

2N4400-4401 A popular signal transistor (from the *Motorola Semiconductor Library*,

*Vol. 1*, 1974. (Courtesy of Motorola Semiconductor Products Inc.)

LF411-412 A popular series of JFET operational amplifiers (from the *National Semiconductor Linear Data Book*, Vol. 1, 1988. (Courtesy of National Semiconductor Corp.)

LM317 A popular 3-terminal adjustable positive voltage regulator (from the *National Semiconductor Linear Data Book*, 1978). (Courtesy of National Semiconductor Corp.)

**Semiconductors**  
BOX 955 • PHOENIX, ARIZONA 85001

**NPN SILICON ANNULAR® TRANSISTORS**

... designed for general purpose switching and amplifier applications and for complementary circuitry with PNP types 2N4402 and 2N4403.

- High Voltage Ratings —  $BV_{CEO} = 40$  V minimum
- Current Gain Specified from 0.1 mA to 500 mA
- Low Saturation Voltage  
 $V_{CE(sat)} = 0.4$  V maximum @  $I_C = 150$  mA
- Complete Switching and Amplifier Specifications
- One-Piece, Injection-Molded Unibloc® Package

**MAXIMUM RATINGS**

| Characteristic                                    | Symbol         | Rating      | Unit                 |
|---------------------------------------------------|----------------|-------------|----------------------|
| Collector-Emitter Voltage                         | $V_{CEO}$      | 40          | Vdc                  |
| Collector-Base Voltage                            | $V_{CB}$       | 60          | Vdc                  |
| Emitter-Base Voltage                              | $V_{EB}$       | 6           | Vdc                  |
| Collector Current - Continuous                    | $I_C$          | 600         | mA dc                |
| Total Device Dissipation $T_A = 25^\circ\text{C}$ | $P_D$          | 310         | mW                   |
| Derate above $25^\circ\text{C}$                   |                | 2.81        | mW/ $^\circ\text{C}$ |
| Operating & Storage Junction Temperature Range    | $T_J, T_{stg}$ | -55 to +135 | $^\circ\text{C}$     |

**THERMAL CHARACTERISTICS**

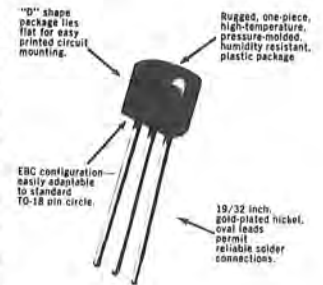
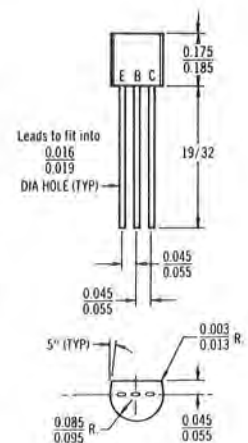
| Characteristic                          | Symbol        | Max   | Unit                       |
|-----------------------------------------|---------------|-------|----------------------------|
| Thermal Resistance, Junction to Case    | $\theta_{JC}$ | 0.137 | $^\circ\text{C}/\text{mW}$ |
| Thermal Resistance, Junction to Ambient | $\theta_{JA}$ | 0.357 | $^\circ\text{C}/\text{mW}$ |

\*Annular Semiconductors patented by Motorola Inc.  
†Trademark of Motorola Inc.

**2N4400**  
**2N4401**

**NPN SILICON  
SWITCHING &  
AMPLIFIER TRANSISTORS**

AUGUST 1966 — DS 5198



TO-92 OUTLINE

ELECTRICAL CHARACTERISTICS ( $T_A = -25^\circ\text{C}$  unless otherwise noted)

| Characteristic                                                                                   | Fig. No. | Symbol       | Min | Max | Unit            |
|--------------------------------------------------------------------------------------------------|----------|--------------|-----|-----|-----------------|
| <b>OFF CHARACTERISTICS</b>                                                                       |          |              |     |     |                 |
| Collector-Emitter Breakdown Voltage*<br>( $I_C = 1\text{ mAdc}$ , $I_B = 0$ )                    |          | $BV_{CEO}^*$ | 40  | —   | Vdc             |
| Collector-Base Breakdown Voltage<br>( $I_C = 0.1\text{ mAdc}$ , $I_E = 0$ )                      |          | $BV_{CBO}$   | 60  | —   | Vdc             |
| Emitter-Base Breakdown Voltage<br>( $I_E = 0.1\text{ mAdc}$ , $I_C = 0$ )                        |          | $BV_{EBO}$   | 6   | —   | Vdc             |
| Collector Cutoff Current<br>( $V_{CE} = 35\text{ Vdc}$ , $V_{EB(\text{off})} = 0.4\text{ Vdc}$ ) |          | $I_{CEX}$    | —   | 0.1 | $\mu\text{Ade}$ |
| Base Cutoff Current<br>( $V_{CE} = 35\text{ Vdc}$ , $V_{EB(\text{off})} = 0.4\text{ Vdc}$ )      |          | $I_{BL}$     | —   | 0.1 | $\mu\text{Ade}$ |

**ON CHARACTERISTICS**

|                                                                                               |        |            |                      |      |      |     |
|-----------------------------------------------------------------------------------------------|--------|------------|----------------------|------|------|-----|
| DC Current Gain<br>( $I_C = 0.1\text{ mAde}$ , $V_{CE} = 1\text{ Vdc}$ )                      | 2N4401 | 15         | $h_{FE}$             | 20   | —    | —   |
| ( $I_C = 1\text{ mAde}$ , $V_{CE} = 1\text{ Vdc}$ )                                           | 2N4400 |            |                      | 20   | —    | —   |
|                                                                                               | 2N4401 |            |                      | 40   | —    | —   |
| ( $I_C = 10\text{ mAde}$ , $V_{CE} = 1\text{ Vdc}$ )                                          | 2N4400 |            |                      | 40   | —    | —   |
|                                                                                               | 2N4401 |            |                      | 80   | —    | —   |
| ( $I_C = 150\text{ mAde}$ , $V_{CE} = 1\text{ Vdc}$ )*                                        | 2N4400 | 10, 17, 18 | $V_{CE(\text{sat})}$ | 50   | 150  | Vdc |
|                                                                                               | 2N4401 |            |                      | 100  | 300  |     |
| ( $I_C = 500\text{ mAde}$ , $V_{CE} = 2\text{ Vdc}$ )*                                        | 2N4400 |            |                      | 20   | —    |     |
|                                                                                               | 2N4401 |            |                      | 40   | —    |     |
| Collector-Emitter Saturation Voltage*<br>( $I_C = 150\text{ mAde}$ , $I_B = 15\text{ mAde}$ ) |        | 17, 18     | $V_{BE(\text{sat})}$ | —    | 0.4  | Vdc |
| ( $I_C = 500\text{ mAde}$ , $I_B = 50\text{ mAde}$ )                                          |        |            |                      | —    | 0.75 |     |
| Base-Emitter Saturation Voltage*<br>( $I_C = 150\text{ mAde}$ , $I_B = 15\text{ mAde}$ )      |        |            |                      | 0.75 | 0.95 |     |
| ( $I_C = 500\text{ mAde}$ , $I_B = 50\text{ mAde}$ )                                          |        |            |                      | —    | 1.2  |     |

**SMALL-SIGNAL CHARACTERISTICS**

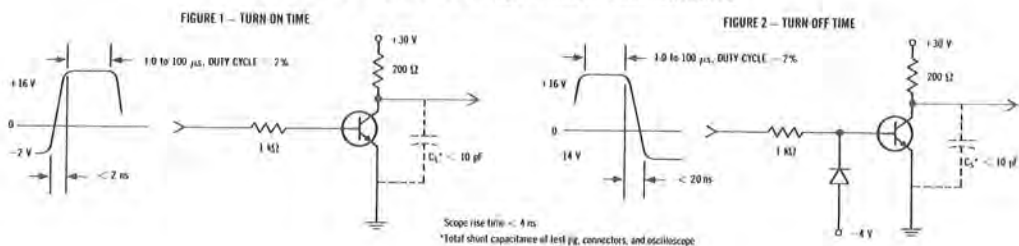
|                                                                                                                  |                  |    |          |             |             |                  |
|------------------------------------------------------------------------------------------------------------------|------------------|----|----------|-------------|-------------|------------------|
| Current-Gain — Bandwidth Product<br>( $I_C = 20\text{ mAde}$ , $V_{CE} = 10\text{ Vdc}$ , $f = 100\text{ MHz}$ ) | 2N4400<br>2N4401 |    | $f_T$    | 200<br>250  | —           | MHz              |
| Collector-Base Capacitance<br>( $V_{CB} = 5\text{ Vdc}$ , $I_E = 0$ , $f = 100\text{ kHz}$ , emitter guarded)    |                  | 3  | $C_{cb}$ | —           | 6.5         | pF               |
| Emitter-Base Capacitance<br>( $V_{BE} = 0.5\text{ Vdc}$ , $I_C = 0$ , $f = 100\text{ kHz}$ , collector guarded)  |                  | 3  | $C_{eb}$ | —           | 30          | pF               |
| Input Impedance<br>( $I_C = 1\text{ mAde}$ , $V_{CE} = 10\text{ Vdc}$ , $f = 1\text{ kHz}$ )                     | 2N4400<br>2N4401 | 12 | $h_{ie}$ | 500<br>1.0k | 7.5k<br>15k | ohms             |
| Voltage Feedback Ratio<br>( $I_C = 1\text{ mAde}$ , $V_{CE} = 10\text{ Vdc}$ , $f = 1\text{ kHz}$ )              |                  | 13 | $h_{re}$ | 0.1         | 8           | $\times 10^{-4}$ |
| Small-Signal Current Gain<br>( $I_C = 1\text{ mAde}$ , $V_{CE} = 10\text{ Vdc}$ , $f = 1\text{ kHz}$ )           | 2N4400<br>2N4401 | 11 | $h_{fe}$ | 20<br>40    | 250<br>500  | —                |
| Output Admittance<br>( $I_C = 1\text{ mAde}$ , $V_{CE} = 10\text{ Vdc}$ , $f = 1\text{ kHz}$ )                   |                  | 14 | $h_{oe}$ | 1           | 30          | $\mu\text{mhos}$ |

**SWITCHING CHARACTERISTICS**

|              |                                                                                                                         |         |       |   |     |    |
|--------------|-------------------------------------------------------------------------------------------------------------------------|---------|-------|---|-----|----|
| Delay Time   | $V_{CC} = 30\text{ Vdc}$ , $V_{EB(\text{off})} = 2\text{ Vdc}$ ,<br>$I_C = 150\text{ mAde}$ , $I_{B1} = 15\text{ mAde}$ | 1, 5    | $t_d$ | — | 15  | ns |
| Rise Time    |                                                                                                                         | 1, 5, 6 | $t_r$ | — | 20  | ns |
| Storage Time | $V_{CC} = 30\text{ Vdc}$ , $I_C = 150\text{ mAde}$ ,<br>$I_{B1} = I_{B2} = 15\text{ mAde}$                              | 2, 7    | $t_s$ | — | 225 | ns |
| Fall Time    |                                                                                                                         | 2, 8    | $t_f$ | — | 30  | ns |

\*Pulse Test: Pulse Width  $\leq 300\text{ }\mu\text{s}$ , Duty Cycle  $\leq 2\%$

**SWITCHING TIME EQUIVALENT TEST CIRCUITS**



TRANSIENT CHARACTERISTICS

— 25°C — — 100°C

FIGURE 3 — CAPACITANCES

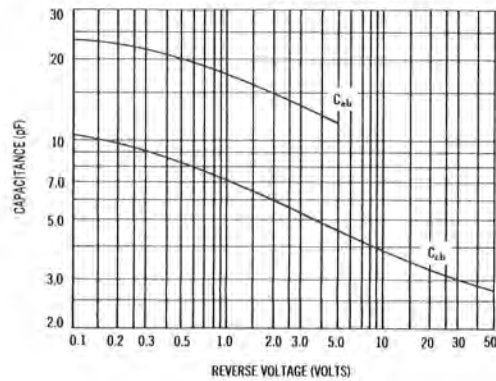


FIGURE 4 — CHARGE DATA

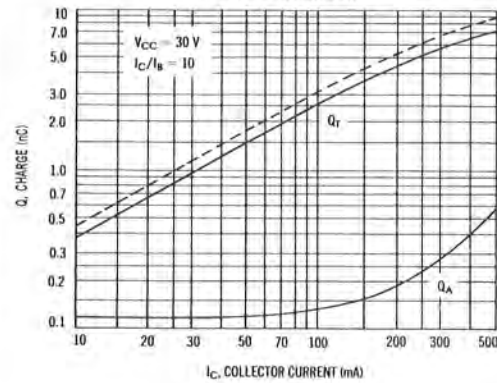


FIGURE 5 — TURN-ON TIME

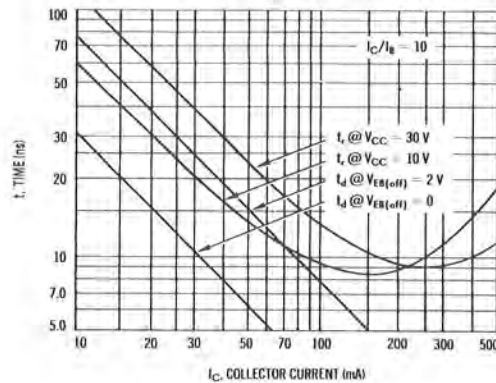


FIGURE 6 — RISE AND FALL TIMES

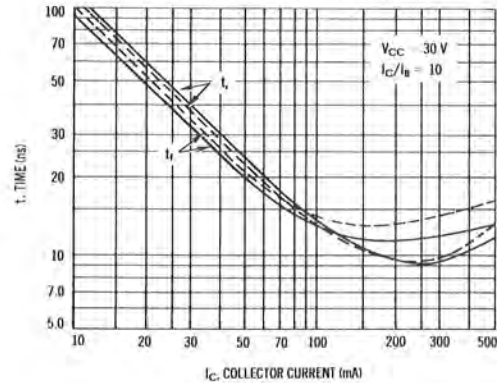


FIGURE 7 — STORAGE TIME

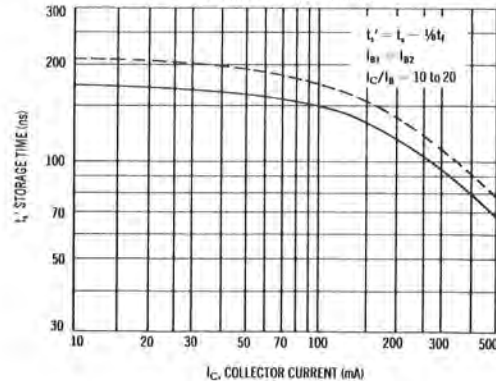
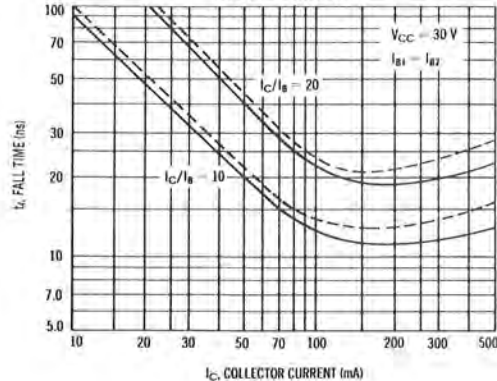


FIGURE 8 — FALL TIME





### SMALL-SIGNAL CHARACTERISTICS

#### NOISE FIGURE

$V_{CE} = 10 \text{ Vdc}$ ,  $T_A = 25^\circ\text{C}$

FIGURE 9 — FREQUENCY EFFECTS

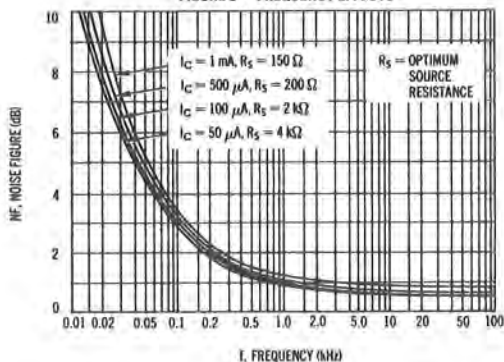
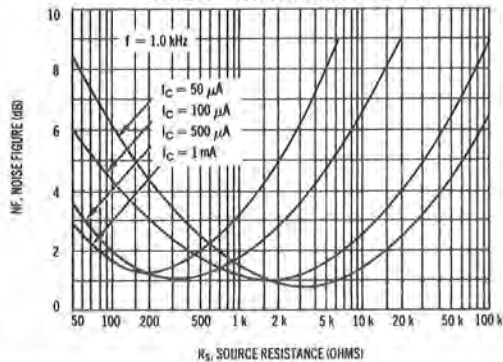


FIGURE 10 — SOURCE RESISTANCE EFFECTS



#### h PARAMETERS

$V_{CE} = 10 \text{ Vdc}$ ,  $f = 1 \text{ kHz}$ ,  $T_A = 25^\circ\text{C}$

This group of graphs illustrates the relationship between  $h_{fe}$  and other "h" parameters for this series of transistors. To obtain these curves, a high-gain and a low-gain unit were selected from both the

2N4400 and 2N4401 lines, and the same units were used to develop the correspondingly numbered curves on each graph.

FIGURE 11 — CURRENT GAIN

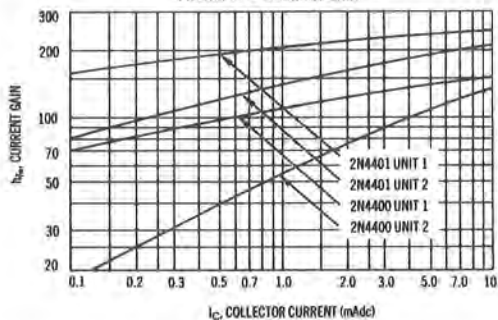


FIGURE 12 — INPUT IMPEDANCE

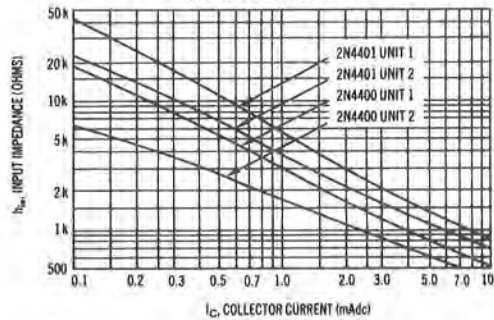


FIGURE 13 — VOLTAGE FEEDBACK RATIO

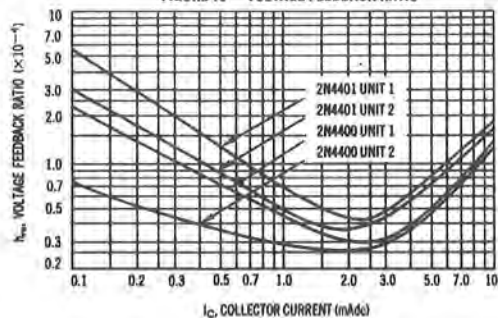
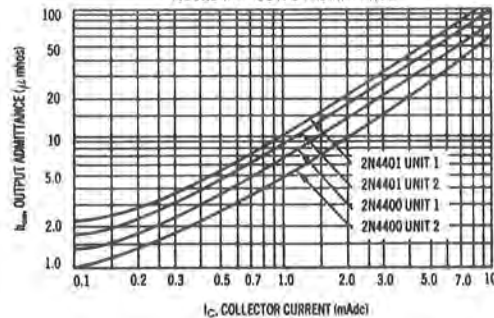


FIGURE 14 — OUTPUT ADMITTANCE



STATIC CHARACTERISTICS

FIGURE 15 — DC CURRENT GAIN

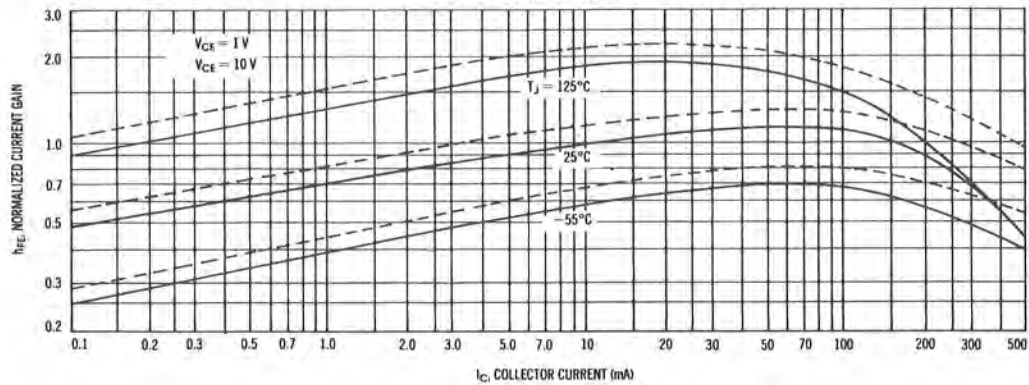


FIGURE 16 — COLLECTOR SATURATION REGION

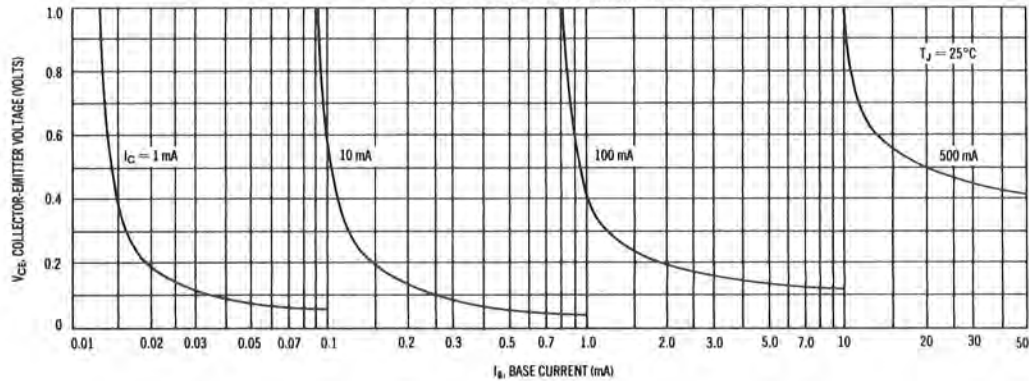


FIGURE 17 — "ON" VOLTAGES

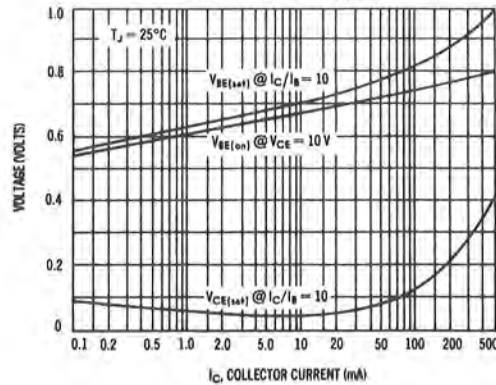
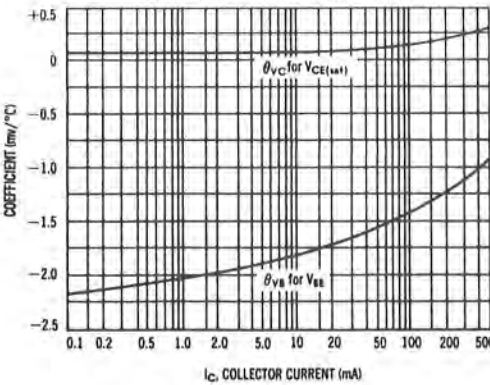
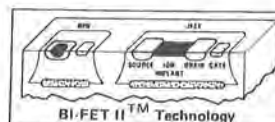


FIGURE 18 — TEMPERATURE COEFFICIENTS





## LF411A/LF411 Low Offset, Low Drift JFET Input Operational Amplifier



July 1987

### General Description

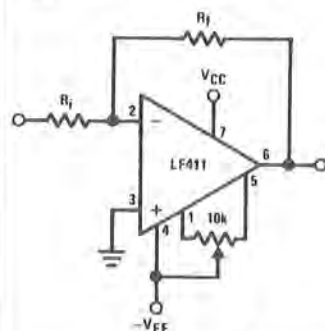
These devices are low cost, high speed, JFET input operational amplifiers with very low input offset voltage and guaranteed input offset voltage drift. They require low supply current yet maintain a large gain bandwidth product and fast slew rate. In addition, well matched high voltage JFET input devices provide very low input bias and offset currents. The LF411 is pin compatible with the standard LM741 allowing designers to immediately upgrade the overall performance of existing designs.

These amplifiers may be used in applications such as high speed integrators, fast D/A converters, sample and hold circuits and many other circuits requiring low input offset voltage and drift, low input bias current, high input impedance, high slew rate and wide bandwidth.

### Features

|                                                                                                                                          |                                             |
|------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------|
| ■ Internally trimmed offset voltage                                                                                                      | 0.5 mV(max)                                 |
| ■ Input offset voltage drift                                                                                                             | 10 $\mu\text{V}/^\circ\text{C}(\text{max})$ |
| ■ Low input bias current                                                                                                                 | 50 pA                                       |
| ■ Low input noise current                                                                                                                | 0.01 pA/ $\sqrt{\text{Hz}}$                 |
| ■ Wide gain bandwidth                                                                                                                    | 3 MHz(min)                                  |
| ■ High slew rate                                                                                                                         | 10V/ $\mu\text{s}(\text{min})$              |
| ■ Low supply current                                                                                                                     | 1.8 mA                                      |
| ■ High input impedance                                                                                                                   | 10 <sup>12</sup> $\Omega$                   |
| ■ Low total harmonic distortion $A_V = 10$ ,<br>$R_L = 10\text{k}$ , $V_O = 20\text{ Vp-p}$ , $\text{BW} = 20\text{ Hz} - 20\text{ kHz}$ | <0.02%                                      |
| ■ Low 1/f noise corner                                                                                                                   | 50 Hz                                       |
| ■ Fast settling time to 0.01%                                                                                                            | 2 $\mu\text{s}$                             |

### Typical Connection



TL/H/5655-1

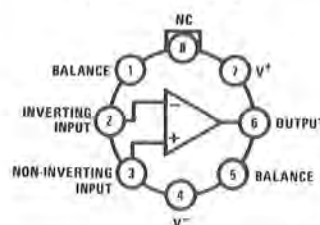
### Ordering Information

#### LF411XYZ

- X indicates electrical grade
- Y indicates temperature range
- "M" for military
- "C" for commercial
- Z indicates package type
- "H" or "N"

### Connection Diagrams

#### Metal Can Package

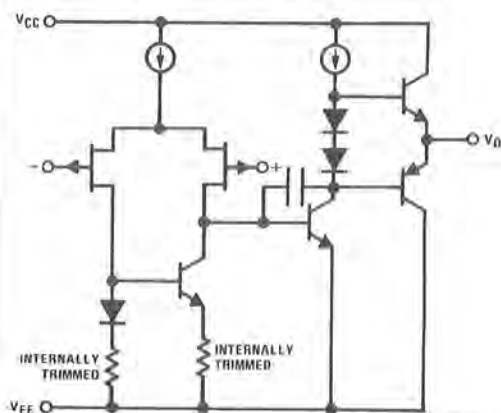


TL/H/5655-6

#### Top View

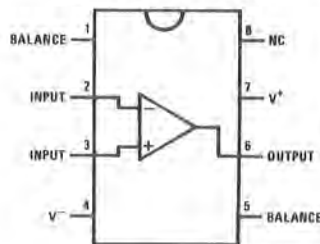
Note: Pin 4 connected to case.  
Order Number LF411AMH,  
LF411MH, LF411ACH or LF411CH  
See NS Package Number H08B

### Simplified Schematic



TL/H/5655-6

#### Dual-In-Line Package



TL/H/5655-7

#### Top View

Order Number  
LF411ACN or LF411CN  
See NS Package Number N08E

BI-FET II™ is a trademark of National Semiconductor Corporation.

### Absolute Maximum Ratings

If Military/Aerospace specified devices are required, please contact the National Semiconductor Sales Office/Distributors for availability and specifications. (Note 8)

|                               | LF411A     | LF411      |                                   | H Package                                            | N Package             |
|-------------------------------|------------|------------|-----------------------------------|------------------------------------------------------|-----------------------|
| Supply Voltage                | ±22V       | ±18V       | Power Dissipation (Notes 2 and 9) | 670 mW                                               | 670 mW                |
| Differential Input Voltage    | ±38V       | ±30V       | $T_{jmax}$                        | 150°C                                                | 115°C                 |
| Input Voltage Range (Note 1)  | ±19V       | ±15V       | $\theta_{JA}$                     | 225°C/W (Still Air)<br>160°C/W (400 LF/min Air Flow) | 120°C/W               |
| Output Short Circuit Duration | Continuous | Continuous | $\theta_{JC}$                     | 25°C/W                                               |                       |
|                               |            |            | Operating Temp. Range             | (Note 3)                                             | (Note 3)              |
|                               |            |            | Storage Temp. Range               | −65°C ≤ $T_A$ ≤ 150°C                                | −65°C ≤ $T_A$ ≤ 150°C |
|                               |            |            | Lead Temp. (Soldering, 10 sec.)   | 260°C                                                | 260°C                 |
|                               |            |            | ESD rating to be determined.      |                                                      |                       |

### DC Electrical Characteristics (Note 4)

| Symbol                   | Parameter                          | Conditions                                                                                        | LF411A                    |                  |     | LF411 |                  |             | Units                        |
|--------------------------|------------------------------------|---------------------------------------------------------------------------------------------------|---------------------------|------------------|-----|-------|------------------|-------------|------------------------------|
|                          |                                    |                                                                                                   | Min                       | Typ              | Max | Min   | Typ              | Max         |                              |
| $V_{OS}$                 | Input Offset Voltage               | $R_S = 10\text{ k}\Omega$ , $T_A = 25^\circ\text{C}$                                              |                           | 0.3              | 0.5 |       | 0.8              | 2.0         | mV                           |
| $\Delta V_{OS}/\Delta T$ | Average TC of Input Offset Voltage | $R_S = 10\text{ k}\Omega$ (Note 5)                                                                |                           | 7                | 10  |       | 7                | 20 (Note 5) | $\mu\text{V}/^\circ\text{C}$ |
| $I_{OS}$                 | Input Offset Current               | $V_S = \pm 15\text{V}$<br>(Notes 4, 6)                                                            | $T_J = 25^\circ\text{C}$  | 25               | 100 |       | 25               | 100         | pA                           |
|                          |                                    |                                                                                                   |                           |                  | 2   |       |                  | 2           | nA                           |
|                          |                                    |                                                                                                   |                           |                  | 25  |       |                  | 25          | nA                           |
| $I_B$                    | Input Bias Current                 | $V_S = \pm 15\text{V}$<br>(Notes 4, 6)                                                            | $T_J = 25^\circ\text{C}$  | 50               | 200 |       | 50               | 200         | pA                           |
|                          |                                    |                                                                                                   | $T_J = 70^\circ\text{C}$  |                  | 4   |       |                  | 4           | nA                           |
|                          |                                    |                                                                                                   | $T_J = 125^\circ\text{C}$ |                  | 50  |       |                  | 50          | nA                           |
| $R_{IN}$                 | Input Resistance                   | $T_J = 25^\circ\text{C}$                                                                          |                           | 10 <sup>12</sup> |     |       | 10 <sup>12</sup> |             | $\Omega$                     |
| $A_{VOL}$                | Large Signal Voltage Gain          | $V_S = \pm 15\text{V}$ , $V_O = \pm 10\text{V}$ ,<br>$R_L = 2\text{k}$ , $T_A = 25^\circ\text{C}$ | 50                        | 200              |     | 25    | 200              |             | V/mV                         |
|                          |                                    | Over Temperature                                                                                  | 25                        | 200              |     | 15    | 200              |             | V/mV                         |
| $V_O$                    | Output Voltage Swing               | $V_S = \pm 15\text{V}$ , $R_L = 10\text{k}$                                                       | ±12                       | ±13.5            |     | ±12   | ±13.5            |             | V                            |
| $V_{CM}$                 | Input Common-Mode Voltage Range    |                                                                                                   | ±16                       | ±19.5            |     | ±11   | ±14.5            |             | V                            |
|                          |                                    |                                                                                                   |                           | −16.5            |     |       | −11.5            |             | V                            |
| CMRR                     | Common-Mode Rejection Ratio        | $R_S \leq 10\text{k}$                                                                             | 80                        | 100              |     | 70    | 100              |             | dB                           |
| PSRR                     | Supply Voltage Rejection Ratio     | (Note 7)                                                                                          | 80                        | 100              |     | 70    | 100              |             | dB                           |
| $I_S$                    | Supply Current                     |                                                                                                   |                           | 1.8              | 2.8 |       | 1.8              | 3.4         | mA                           |

### AC Electrical Characteristics (Note 4)

| Symbol | Parameter                      | Conditions                                                           | LF411A |      |     | LF411 |      |     | Units                  |
|--------|--------------------------------|----------------------------------------------------------------------|--------|------|-----|-------|------|-----|------------------------|
|        |                                |                                                                      | Min    | Typ  | Max | Min   | Typ  | Max |                        |
| SR     | Slew Rate                      | $V_S = \pm 15\text{V}$ , $T_A = 25^\circ\text{C}$                    | 10     | 15   |     | 8     | 15   |     | V/ $\mu\text{s}$       |
| GBW    | Gain-Bandwidth Product         | $V_S = \pm 15\text{V}$ , $T_A = 25^\circ\text{C}$                    | 3      | 4    |     | 2.7   | 4    |     | MHz                    |
| $e_n$  | Equivalent Input Noise Voltage | $T_A = 25^\circ\text{C}$ , $R_S = 100\Omega$ ,<br>$f = 1\text{ kHz}$ |        | 25   |     |       | 25   |     | nV/ $\sqrt{\text{Hz}}$ |
| $i_n$  | Equivalent Input Noise Current | $T_A = 25^\circ\text{C}$ , $f = 1\text{ kHz}$                        |        | 0.01 |     |       | 0.01 |     | pA/ $\sqrt{\text{Hz}}$ |

**Note 1:** Unless otherwise specified the absolute maximum negative input voltage is equal to the negative power supply voltage.

**Note 2:** For operating at elevated temperature, these devices must be derated based on a thermal resistance of  $\theta_{JA}$ .

**Note 3:** These devices are available in both the commercial temperature range  $0^{\circ}\text{C} \leq T_A \leq 70^{\circ}\text{C}$  and the military temperature range  $-55^{\circ}\text{C} \leq T_A \leq 125^{\circ}\text{C}$ . The temperature range is designated by the position just before the package type in the device number. A "C" indicates the commercial temperature range and an "M" indicates the military temperature range. The military temperature range is available in "H" package only.

**Note 4:** Unless otherwise specified, the specifications apply over the full temperature range and for  $V_S = \pm 20\text{V}$  for the LF411A and for  $V_S = \pm 15\text{V}$  for the LF411.  $V_{OS}$ ,  $I_B$ , and  $I_{OS}$  are measured at  $V_{CM} = 0$ .

**Note 5:** The LF411A is 100% tested to this specification. The LF411 is sample tested to insure at least 90% of the units meet this specification.

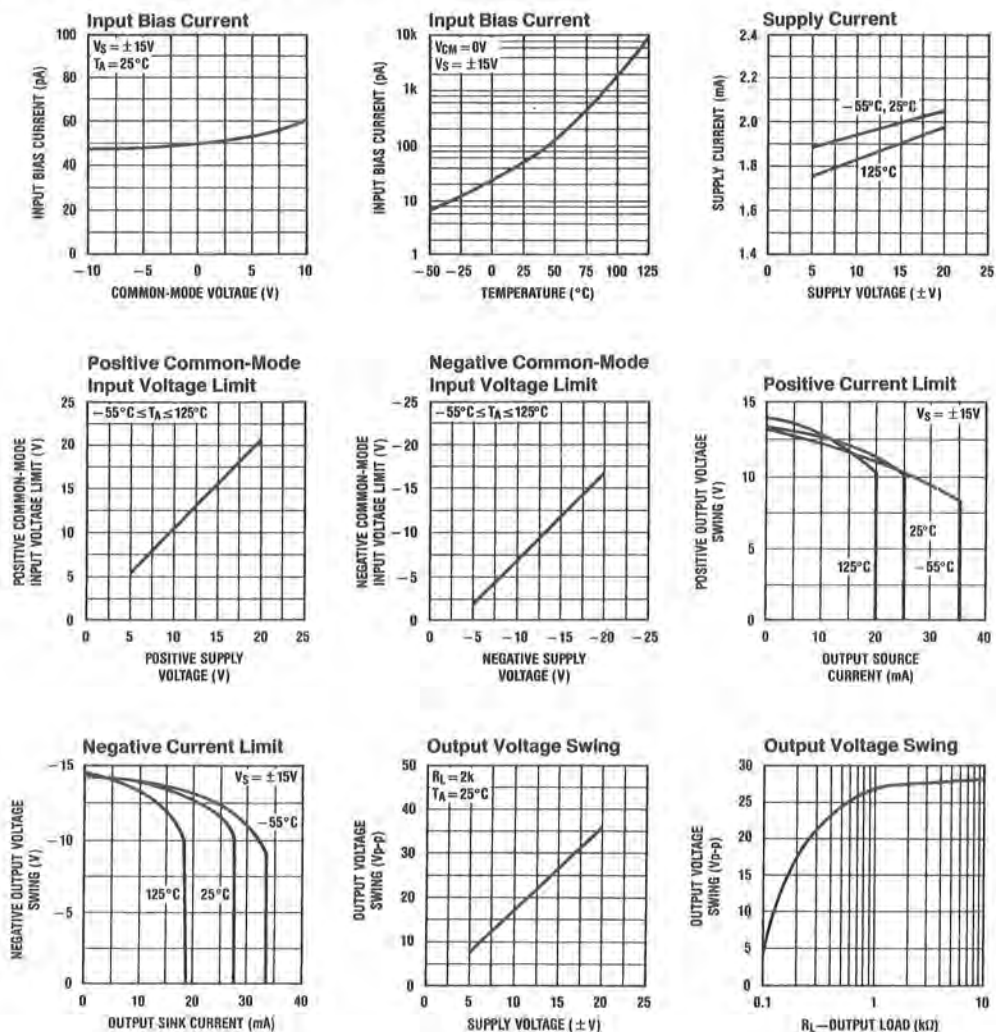
**Note 6:** The input bias currents are junction leakage currents which approximately double for every  $10^{\circ}\text{C}$  increase in the junction temperature,  $T_J$ . Due to limited production test time, the input bias currents measured are correlated to junction temperature. In normal operation the junction temperature rises above the ambient temperature as a result of internal power dissipation,  $P_D$ .  $T_J = T_A + \theta_{JA} P_D$  where  $\theta_{JA}$  is the thermal resistance from junction to ambient. Use of a heat sink is recommended if input bias current is to be kept to a minimum.

**Note 7:** Supply voltage rejection ratio is measured for both supply magnitudes increasing or decreasing simultaneously in accordance with common practice, from  $\pm 15\text{V}$  to  $\pm 5\text{V}$  for the LF411 and from  $\pm 20\text{V}$  to  $\pm 5\text{V}$  for the LF411A.

**Note 8:** Refer to RETS 411AX for LF411AMH military specifications and to RETS 411X for LF411MH military specifications.

**Note 9:** Max. Power Dissipation is defined by the package characteristics. Operating the part near the Max. Power Dissipation may cause the part to operate outside guaranteed limits.

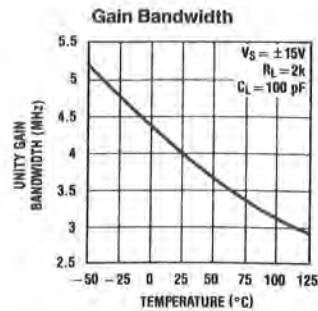
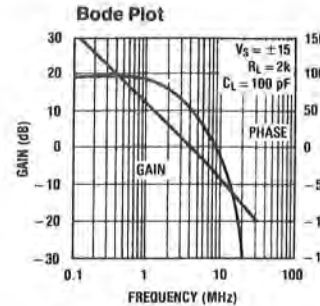
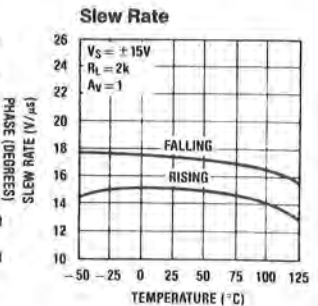
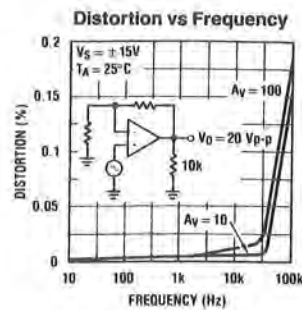
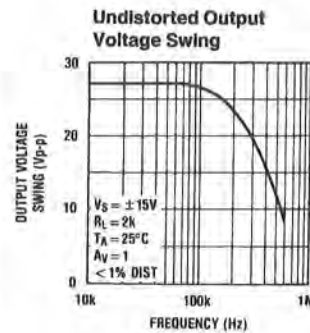
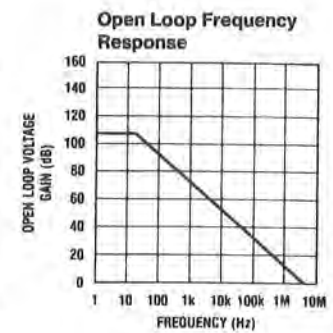
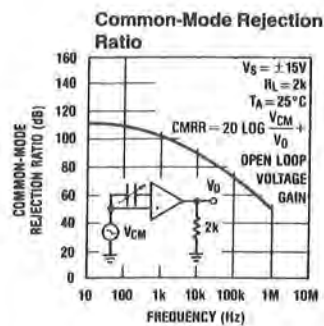
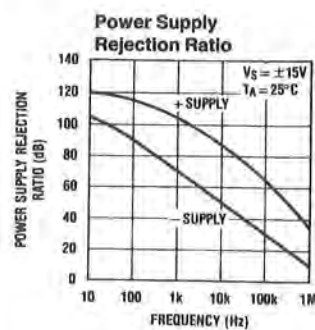
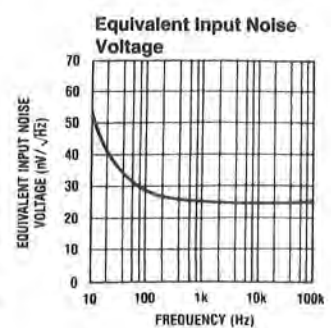
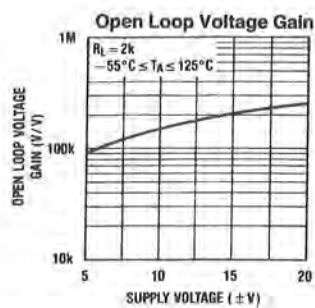
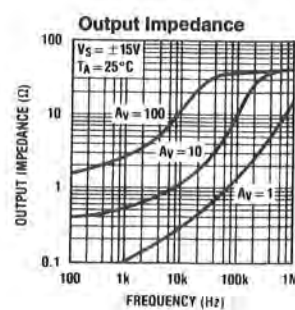
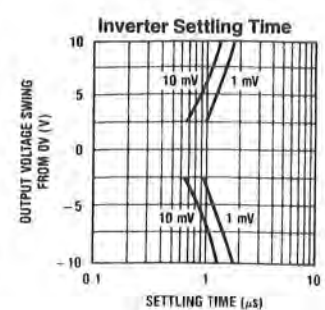
## Typical Performance Characteristics



TL/H/5655-2

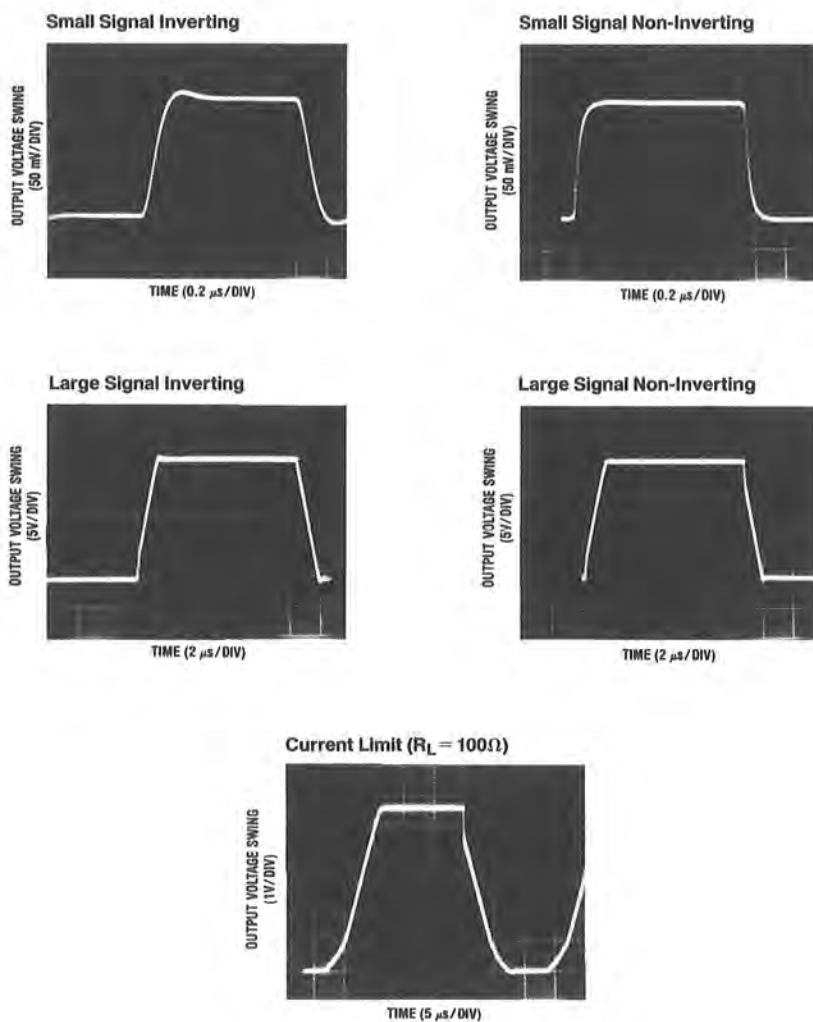


# Typical Performance Characteristics (Continued)



TL/H/5655-3

### Pulse Response $R_L = 2\text{ k}\Omega$ , $C_L = 10\text{ pF}$



TL/H/5655-4

### Application Hints

The LF411 series of internally trimmed JFET input op amps (BI-FET II™) provide very low input offset voltage and guaranteed input offset voltage drift. These JFETs have large reverse breakdown voltages from gate to source and drain eliminating the need for clamps across the inputs. Therefore, large differential input voltages can easily be accommodated without a large increase in input current. The maximum differential input voltage is independent of the supply voltages. However, neither of the input voltages should be allowed to exceed the negative supply as this will cause large currents to flow which can result in a destroyed unit.

Exceeding the negative common-mode limit on either input will force the output to a high state, potentially causing a reversal of phase to the output. Exceeding the negative common-mode limit on both inputs will force the amplifier output to a high state. In neither case does a latch occur since raising the input back within the common-mode range again puts the input stage and thus the amplifier in a normal operating mode.

Exceeding the positive common-mode limit on a single input will not change the phase of the output; however, if both inputs exceed the limit, the output of the amplifier may be forced to a high state.

### Application Hints (Continued)

The amplifier will operate with a common-mode input voltage equal to the positive supply; however, the gain bandwidth and slew rate may be decreased in this condition. When the negative common-mode voltage swings to within 3V of the negative supply, an increase in input offset voltage may occur.

The LF411 is biased by a zener reference which allows normal circuit operation on  $\pm 4.5\text{V}$  power supplies. Supply voltages less than these may result in lower gain bandwidth and slew rate.

The LF411 will drive a  $2\text{ k}\Omega$  load resistance to  $\pm 10\text{V}$  over the full temperature range. If the amplifier is forced to drive heavier load currents, however, an increase in input offset voltage may occur on the negative voltage swing and finally reach an active current limit on both positive and negative swings.

Precautions should be taken to ensure that the power supply for the integrated circuit never becomes reversed in polarity or that the unit is not inadvertently installed backwards in a socket as an unlimited current surge through the resulting forward diode within the IC could cause fusing of the internal conductors and result in a destroyed unit.

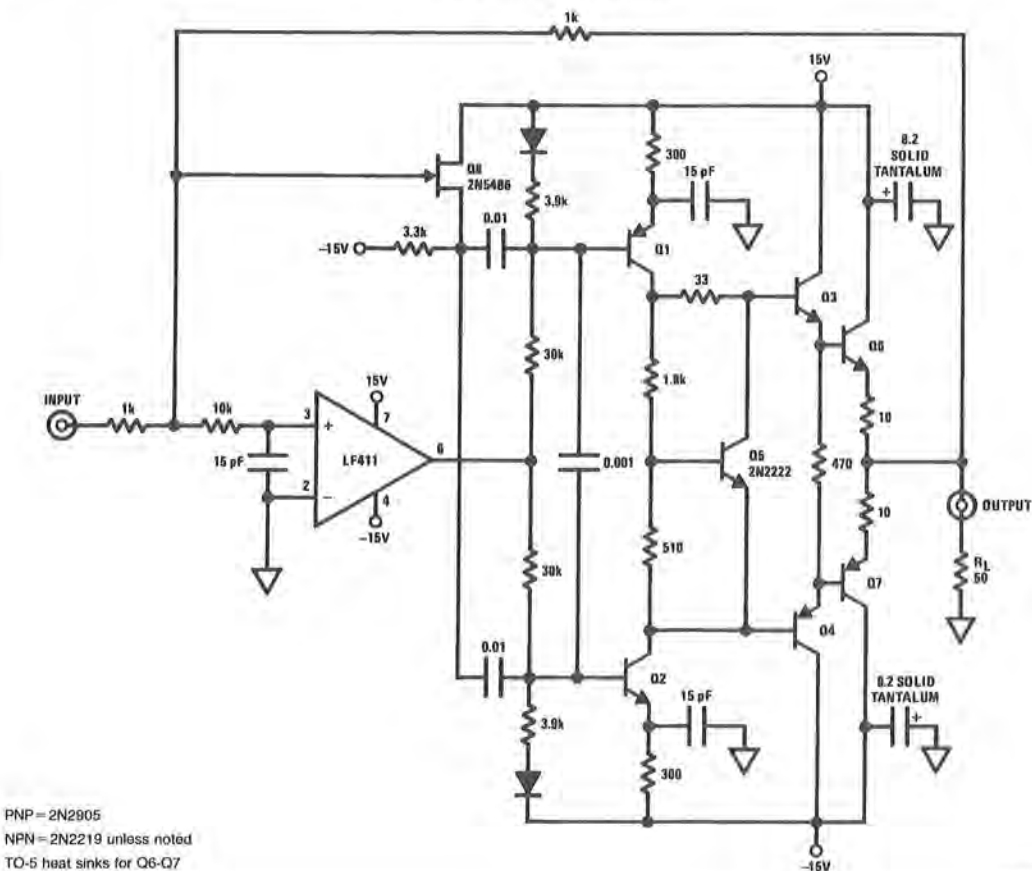
Because these amplifiers are JFET rather than MOSFET input op amps they do not require special handling.

As with most amplifiers, care should be taken with lead dress, component placement and supply decoupling in order to ensure stability. For example, resistors from the output to an input should be placed with the body close to the input to minimize "pick-up" and maximize the frequency of the feedback pole by minimizing the capacitance from the input to ground.

A feedback pole is created when the feedback around any amplifier is resistive. The parallel resistance and capacitance from the input of the device (usually the inverting input) to AC ground set the frequency of the pole. In many instances the frequency of this pole is much greater than the expected 3 dB frequency of the closed loop gain and consequently there is negligible effect on stability margin. However, if the feedback pole is less than approximately 6 times the expected 3 dB frequency, a lead capacitor should be placed from the output to the input of the op amp. The value of the added capacitor should be such that the RC time constant of this capacitor and the resistance it parallels is greater than or equal to the original feedback pole time constant.

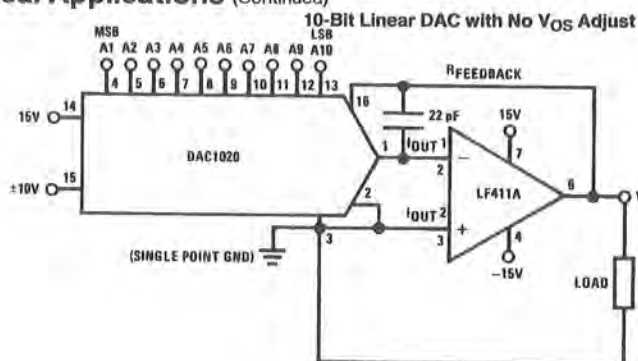
### Typical Applications

#### Ultra High Speed Current Booster



TL/H/5655-9

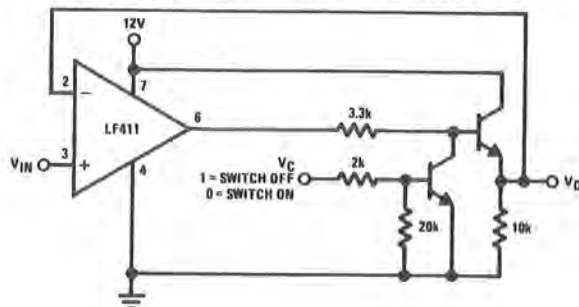
### Typical Applications (Continued)



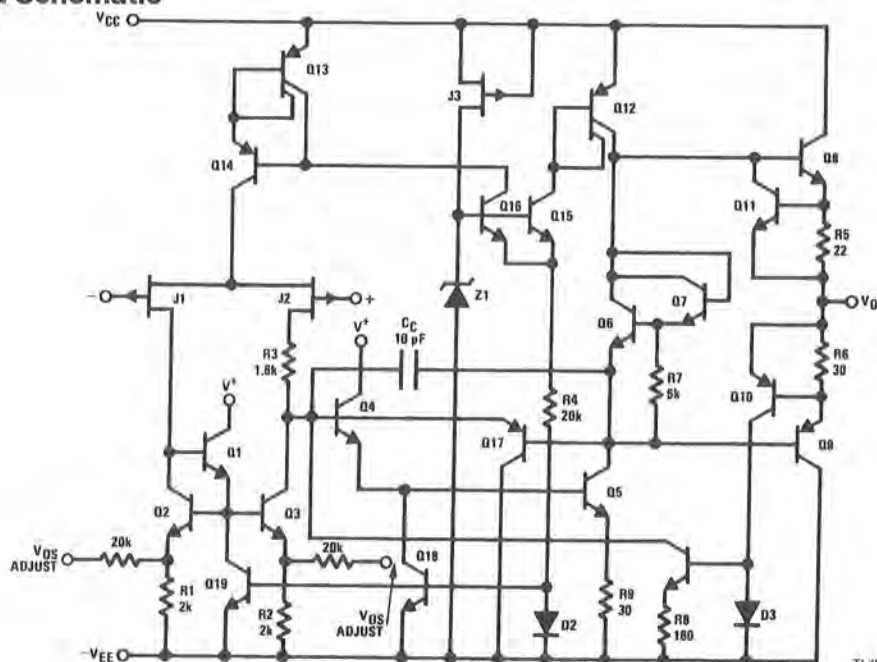
$$V_{OUT} = -V_{REF} \left( \frac{A_1}{2} + \frac{A_2}{4} + \frac{A_3}{8} + \dots + \frac{A_{10}}{1024} \right)$$

$-10V \leq V_{REF} \leq 10V$   
 $0 \leq V_{OUT} \leq \frac{1023}{1024} V_{REF}$   
 where  $A_N = 1$  if the  $A_N$  digital input is high  
 $A_N = 0$  if the  $A_N$  digital input is low

### Single Supply Analog Switch with Buffered Output



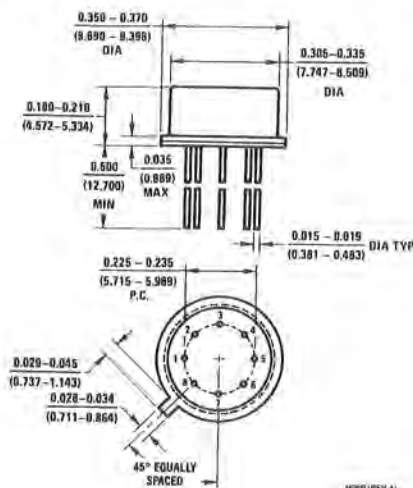
### Detailed Schematic



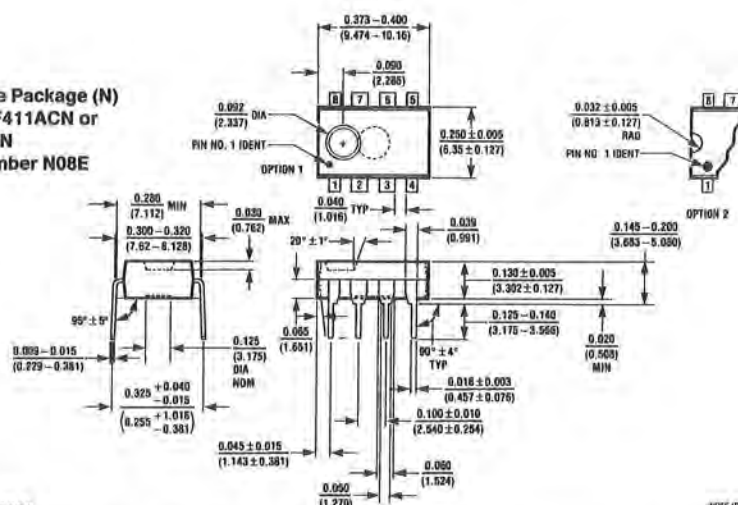
TL/H/5655-10

# Physical Dimensions inches (millimeters)

**Metal Can Package (H)**  
Order Number LF411AMH, LF411MH,  
LF411ACH or LF411CH  
NS Package Number H08B



**Molded Dual-In-Line Package (N)**  
Order Number LF411ACN or  
LF411CN  
NS Package Number N08E



## LIFE SUPPORT POLICY

NATIONAL'S PRODUCTS ARE NOT AUTHORIZED FOR USE AS CRITICAL COMPONENTS IN LIFE SUPPORT DEVICES OR SYSTEMS WITHOUT THE EXPRESS WRITTEN APPROVAL OF THE PRESIDENT OF NATIONAL SEMICONDUCTOR CORPORATION. As used herein:

1. Life support devices or systems are devices or systems which, (a) are intended for surgical implant into the body, or (b) support or sustain life, and whose failure to perform, when properly used in accordance with instructions for use provided in the labeling, can be reasonably expected to result in a significant injury to the user.
2. A critical component is any component of a life support device or system whose failure to perform can be reasonably expected to cause the failure of the life support device or system, or to affect its safety or effectiveness.



National Semiconductor  
Corporation  
2600 Semiconductor Drive  
P.O. Box 58060  
Santa Clara, CA 95052-8060  
Tel: (408) 721-5000  
TWX: (610) 326-9240

National Semiconductor  
GmbH  
Westendstrasse 193-195  
D-6000 Munchen 21  
West Germany  
Tel: (089) 5 70 05 01  
Telex: 522772

NS Japan Ltd.  
Sanso Bldg. 5F  
4-15 Nishi Shinjuku  
Shinjuku-ku,  
Tokyo 160, Japan  
Tel: 3-299-7001  
FAX: 3-299-7000

National Semiconductor  
Hong Kong Ltd.  
Southeast Asia Marketing  
Austin Tower, 4th Floor  
22-26A Austin Avenue  
Tsimshatsui, Kowloon, H.K.  
Tel: 3-7221290, 3-7243545  
Cable: NSSEAMKTG  
Telex: 52906 NSSEA HX

National Semicondutores  
Do Brasil Ltda.  
Av. Eng. Faria Lima, 650  
8 Andar  
01452 Sao Paulo, SP, Brasil  
Tel: (55/11) 212-5086  
Telex: 391-1131931 NSBR BR

National Semiconductor  
(Australia) PTY, Ltd.  
21/3 High Street  
Bayswater, Victoria 3150  
Australia  
Tel: (03) 728-6333  
Telex: AAS2098





## Voltage Regulators

### LM117/LM217/LM317 3-terminal adjustable regulator

#### General Description

The LM117/LM217/LM317 are adjustable 3-terminal positive voltage regulators capable of supplying in excess of 1.5A over a 1.2V to 37V output range. They are exceptionally easy to use and require only two external resistors to set the output voltage. Further, both line and load regulation are better than standard fixed regulators. Also, the LM117 is packaged in standard transistor packages which are easily mounted and handled.

In addition to higher performance than fixed regulators, the LM117 series offers full overload protection available only in IC's. Included on the chip are current limit, thermal overload protection and safe area protection. All overload protection circuitry remains fully functional even if the adjustment terminal is disconnected.

#### Features

- Adjustable output down to 1.2V
- Guaranteed 1.5A output current
- Line regulation typically 0.01%/V
- Load regulation typically 0.1%
- Current limit constant with temperature
- 100% electrical burn-in
- Eliminates the need to stock many voltages
- Standard 3-lead transistor package
- 80 dB ripple rejection

Normally, no capacitors are needed unless the device is situated far from the input filter capacitors in which case an input bypass is needed. An optional output capacitor can be added to improve transient response. The adjustment terminal can be bypassed to achieve very high ripple rejection ratios which are difficult to achieve with standard 3-terminal regulators.

Besides replacing fixed regulators, the LM117 is useful in a wide variety of other applications. Since the regulator is "floating" and sees only the input-to-output differential voltage, supplies of several hundred volts can be regulated as long as the maximum input to output differential is not exceeded.

Also, it makes an especially simple adjustable switching regulator, a programmable output regulator, or by connecting a fixed resistor between the adjustment and output, the LM117 can be used as a precision current regulator. Supplies with electronic shutdown can be achieved by clamping the adjustment terminal to ground which programs the output to 1.2V where most loads draw little current.

The LM117K, LM217K and LM317K are packaged in standard TO-3 transistor packages while the LM117H, LM217H and LM317H are packaged in a solid Kovar base TO-5 transistor package. The LM117 is rated for operation from  $-55^{\circ}\text{C}$  to  $+150^{\circ}\text{C}$ , the LM217 from  $-25^{\circ}\text{C}$  to  $+150^{\circ}\text{C}$  and the LM317 from  $0^{\circ}\text{C}$  to  $+125^{\circ}\text{C}$ . The LM317T and LM317MP, rated for operation over a  $0^{\circ}\text{C}$  to  $+125^{\circ}\text{C}$  range, are available in a TO-220 plastic package and a TO-202 package, respectively.

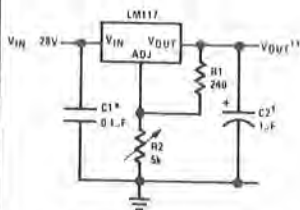
For applications requiring greater output current in excess of 3A and 5A, see LM150 series and LM138 series data sheets, respectively. For the negative complement, see LM137 series data sheet.

LM117 Series Packages and Power Capability

| DEVICE | PACKAGE | RATED POWER DISSIPATION | DESIGN LOAD CURRENT |
|--------|---------|-------------------------|---------------------|
| LM117  | TO-3    | 20W                     | 1.5A                |
| LM217  | TO-39   | 2W                      | 0.5A                |
| LM317  | TO-220  | 15W                     | 1.5A                |
| LM317T | TO-202  | 7.5W                    | 0.5A                |

#### Typical Applications

1.2V-25V Adjustable Regulator

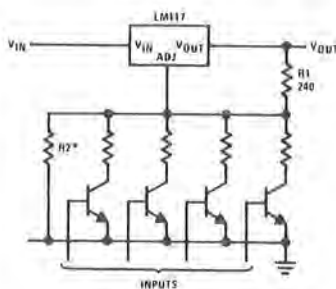


<sup>†</sup>Optional—improves transient response

<sup>\*</sup>Needed if device is far from filter capacitors

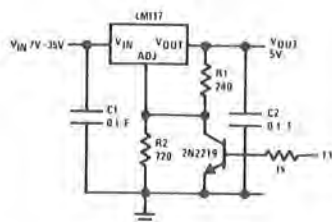
$$V_{OUT} = 1.25V \left( 1 + \frac{R_2}{R_1} \right)$$

Digitally Selected Outputs



<sup>\*</sup>Sets maximum  $V_{OUT}$

5V Logic Regulator with Electronic Shutdown\*



<sup>\*</sup>Min output  $\approx 1.2V$

### absolute maximum ratings

|                                          |                    |
|------------------------------------------|--------------------|
| Power Dissipation                        | Internally limited |
| Input–Output Voltage Differential        | 40V                |
| Operating Junction Temperature Range     |                    |
| LM117                                    | –55°C to +150°C    |
| LM217                                    | –25°C to +150°C    |
| LM317                                    | 0°C to +125°C      |
| Storage Temperature                      | –65°C to +150°C    |
| Lead Temperature (Soldering, 10 seconds) | 300°C              |

### electrical characteristics (Note 1)

| PARAMETER                            | CONDITIONS                                                                                                                                       | LM117/217 |       |      | LM317 |       |      | UNITS              |
|--------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------|-----------|-------|------|-------|-------|------|--------------------|
|                                      |                                                                                                                                                  | MIN       | TYP   | MAX  | MIN   | TYP   | MAX  |                    |
| Line Regulation                      | $T_A = 25^\circ\text{C}$ , $3\text{V} \leq V_{IN} - V_{OUT} \leq 40\text{V}$<br>(Note 2)                                                         |           | 0.01  | 0.02 |       | 0.01  | 0.04 | %/V                |
| Load Regulation                      | $T_A = 25^\circ\text{C}$ , $10\text{mA} \leq I_{OUT} \leq I_{MAX}$<br>$V_{OUT} \leq 5\text{V}$ , (Note 2)<br>$V_{OUT} \geq 5\text{V}$ , (Note 2) |           | 5     | 15   |       | 5     | 25   | mV                 |
|                                      |                                                                                                                                                  |           | 0.1   | 0.3  |       | 0.1   | 0.5  | %                  |
| Adjustment Pin Current               |                                                                                                                                                  |           | 50    | 100  |       | 50    | 100  | $\mu\text{A}$      |
| Adjustment Pin Current Change        | $10\text{mA} \leq I_L \leq I_{MAX}$<br>$2.5\text{V} \leq (V_{IN} - V_{OUT}) \leq 40\text{V}$                                                     |           | 0.2   | 5    |       | 0.2   | 5    | $\mu\text{A}$      |
| Reference Voltage                    | $3 \leq (V_{IN} - V_{OUT}) \leq 40\text{V}$ , (Note 3)<br>$10\text{mA} \leq I_{OUT} \leq I_{MAX}$ , $P \leq P_{MAX}$                             | 1.20      | 1.25  | 1.30 | 1.20  | 1.25  | 1.30 | V                  |
| Line Regulation                      | $3\text{V} \leq V_{IN} - V_{OUT} \leq 40\text{V}$ , (Note 2)                                                                                     |           | 0.02  | 0.05 |       | 0.02  | 0.07 | %/V                |
| Load Regulation                      | $10\text{mA} \leq I_{OUT} \leq I_{MAX}$ , (Note 2)<br>$V_{OUT} \leq 5\text{V}$<br>$V_{OUT} \geq 5\text{V}$                                       |           | 20    | 50   |       | 20    | 70   | mV                 |
|                                      |                                                                                                                                                  |           | 0.3   | 1    |       | 0.3   | 1.5  | %                  |
| Temperature Stability                | $T_{MIN} \leq T_J \leq T_{MAX}$                                                                                                                  |           | 1     |      |       | 1     |      | %                  |
| Minimum Load Current                 | $V_{IN} - V_{OUT} = 40\text{V}$                                                                                                                  |           | 3.5   | 5    |       | 3.5   | 10   | mA                 |
| Current Limit                        | $V_{IN} - V_{OUT} \leq 15\text{V}$<br>K and T Package                                                                                            | 1.5       | 2.2   |      | 1.5   | 2.2   |      | A                  |
|                                      | H and P Package                                                                                                                                  | 0.5       | 0.8   |      | 0.5   | 0.8   |      | A                  |
|                                      | $V_{IN} - V_{OUT} = 40\text{V}$<br>K and T Package                                                                                               |           | 0.4   |      |       | 0.4   |      | A                  |
|                                      | H and P Package                                                                                                                                  |           | 0.07  |      |       | 0.07  |      | A                  |
| RMS Output Noise, % of $V_{OUT}$     | $T_A = 25^\circ\text{C}$ , $10\text{Hz} \leq f \leq 10\text{kHz}$                                                                                |           | 0.003 |      |       | 0.003 |      | %                  |
| Ripple Rejection Ratio               | $V_{OUT} = 10\text{V}$ , $f = 120\text{Hz}$                                                                                                      |           | 65    |      |       | 65    |      | dB                 |
|                                      | $C_{ADJ} = 10\mu\text{F}$                                                                                                                        | 66        | 80    |      | 66    | 80    |      | dB                 |
| Long-Term Stability                  | $T_A = 125^\circ\text{C}$                                                                                                                        |           | 0.3   | 1    |       | 0.3   | 1    | %                  |
| Thermal Resistance, Junction to Case | H Package                                                                                                                                        |           | 12    | 15   |       | 12    | 15   | $^\circ\text{C/W}$ |
|                                      | K Package                                                                                                                                        |           | 2.3   | 3    |       | 2.3   | 3    | $^\circ\text{C/W}$ |
|                                      | T Package                                                                                                                                        |           |       |      |       | 5     |      | $^\circ\text{C/W}$ |
|                                      | P Package                                                                                                                                        |           |       |      |       | 12    |      | $^\circ\text{C/W}$ |

Note 1: Unless other wise specified, these specifications apply  $-55^\circ\text{C} \leq T_J \leq +150^\circ\text{C}$  for the LM117,  $-25^\circ\text{C} \leq T_J \leq +150^\circ\text{C}$  for the LM217 and  $0^\circ\text{C} \leq T_J \leq +125^\circ\text{C}$  for the LM317;  $V_{IN} - V_{OUT} = 5\text{V}$  and  $I_{OUT} = 0.1\text{A}$  for the TO-5 package and  $I_{OUT} = 0.5\text{A}$  for the TO-3 package and TO-220 package. Although power dissipation is internally limited, these specifications are applicable for power dissipations of 2W for the TO-5 and 20W for the TO-3 and TO-220.  $I_{MAX}$  is 1.5A for the TO-3 and TO-220 package and 0.5A for the TO-5 package.

Note 2: Regulation is measured at constant junction temperature. Changes in output voltage due to heating effects must be taken into account separately. Pulse testing with low duty cycle is used.

Note 3: Selected devices with tightend tolerance reference voltage available.