A Tribute to Jim Williams

Book Chapters

(1991-2008)

A Tribute to Jim Williams

Book Chapters

(1991-2008)

4. Is Analog Circuit Design Dead?

Rumor has it that analog circuit design is dead. Indeed, it is widely reported and accepted that *rigor mortis* has set in. Precious filters, integrators, and the like seem to have been buried beneath an avalanche of microprocessors, ROMs, RAMs, and bits and bytes. As some analog people see it (peering out from behind their barricades), a digital monster has been turned loose, destroying the elegance of continuous functions with a blitzing array of flipping and flopping waveforms. The introduction of a "computerized" oscilloscope—the most analog of all instruments—with *no knobs* would seem to be the *coup de grâce*.

These events have produced some bizarre behavior. It has been kindly suggested, for instance, that the few remaining analog types be rounded up and protected as an endangered species. Colleges and universities offer few analog design courses. And some localities have defined copies of Korn and Korn publications, the *Philbrick Applications Manual,* and the *Linear Applications Handbook* as pornographic material, to be kept away from engineering students' innocent and impressionable minds. Sadly, a few well-known practitioners of the art are slipping across the border (James E. Solomon has stated, for example, that "all classical analog techniques are dead"), while more principled ones are simply leaving town.

Can all this be happening? Is it really so? Is analog dead? Or has the hysteria of the moment given rise to exaggeration and distorted judgment?

To answer these questions with any degree of intelligence and sensitivity, it is necessary to consult history. And to start this process, we must examine the patient's body.

Analog circuit design is described using such terms as subtractor, integrator, differentiator, and summing junction. These mathematical operations are performed by that pillar of analoggery, the operational amplifier. The use of an amplifier as a computing tool is not entirely obvious and was first investigated before World War II. Practical "computing amplifiers" found their first real niche inside electronic analog computers (as opposed to mechanical analog computers such as the Norden bombsight or Bush's Differential Analyzer), which were developed in the late 1940s and 1950s. These machines were, by current standards, monstrous assemblages made up of large numbers of amplifiers that could be programmed to integrate, sum, differentiate, and perform a host of mathematical operations. Individual amplifiers performed singular functions, but complex operations were performed when all the amplifiers were interconnected in any desired configuration.

The analog computer's forte was its ability to model or simulate events. Analog computers did not die out because analog simulations are no longer useful or do not approximate truth; rather, the rise of digital machines made it enticingly easy to use digital fakery to *simulate the simulations.*

Adapted from the July 22, 1991, issue of *EDN Magazine*.

Figure 4-1. Some analog types are merely leaving town.

As digital systems came on line in the late 1950s and early 1960s, a protracted and brutally partisan dispute (some recall it as more of a war) arose between the analog and digital camps. Digital methods offered high precision at the cost of circuit complexity. The analog way achieved sophisticated results at lower accuracy and with comparatively simple circuit configurations. One good op amp (eight transistors) could do the work of 100 digitally configured 2N404s. It seemed that digital circuitry was an accurate but inelegant and overcomplex albatross. Digital types insisted that analog techniques could never achieve any significant accuracy, regardless of how adept they were at modeling and simulating real systems.

This battle was not without its editorializing. One eloquent speaker was George A. Philbrick, a decided analog man, who wrote in 1963 (in *The Lightning Empiricist*, Volume II, No. 4, October, "Analogs Yesterday, Today, and Tomorrow," pp. 3–8), "In modest applications to on-line measurement and data processing, it is quite generally conceded that the advantage of continuous analog apparatus make it irresistible. This is partly owing to the simplicity and speed which its continuity makes possible, and partly to the fact that almost every input transducer is also 'analog' in character, that is to say, continuous in excursion and time."

Philbrick, however, a brilliant man, was aware enough to see that digital had at least some place in the lab: "Only the most hard-shelled of analog champions would suggest that all simulative and computational equipment be undiluted by numerical or logical adjuncts."

He continued by noting that "some analog men, perhaps overfond and defensive as regards continuous functions, really believe that analog operations are generalizations of digital ones, or that conversely digital operations are special cases of analog ones. What can be done with such people?

"While it is agreed that analog and digital techniques will increasingly cross-fertilize and interrelate," Philbrick concluded, "it is predicted that the controversy between their camps will rage on, good natured but unabated, for years to come in spite of hybrid attachments."

Although Philbrick and others were intelligent enough to prevent their analog passions from obscuring their reasoning powers, they could not possibly see what was coming in a very few years.

Figure 4-2.
Is this the fate of oscilloscopes whose innards are controlled by knobs instead of microchips?

Jack Kilby built his IC in 1958. By the middle 1960s, RTL and DTL were in common use.

While almost everyone agreed that digital approximations weren't as elegant as "the real thing," they were becoming eminently workable, increasingly inexpensive, and physically more compactable. With their computing business slipping away, the analog people pulled their amplifiers out of computers, threw the racks away, and scurried into the measurement and control business. (For a nostalgic, if not tearful, look at analog computers at the zenith of their glory, read *A Palimpsest on the Electronic Analog Art*, edited by Henry M. Paynter.)

If you have read thoughtfully to this point, it should be obvious that analog is not dead, rather just badly shaken and overshadowed in the aftermath of the war. Although measurement and control are certainly still around, the really glamorous and publicized territory has been staked out by the digital troops for some time. Hard-core guerrilla resistance to this state of affairs, while heroic, is guaranteed suicide. To stay alive, and even prosper, calls for skillful bargaining based on thorough analysis of the competition's need.

The understanding that analog is *not* dead lies in two key observations. First, to do any useful work, the digital world requires information to perform its operations upon. The information must come from something loosely referred to as "the real world." Deleting quantum mechanics, the "real world" is analog. Supermarket scales, automobile engines, blast furnaces, and the human body are all examples of systems that furnish the analog information that the silicon abacus requires to jus-

tify its existence. So long as transduction remains analog in nature, the conversion process will be required.

A further observation is that many microprocessors are being used not to replace but to enhance a fundamentally analog measurement or process. The current spate of microprocessor-controlled digital voltmeters furnishes one good example; others include digital storage oscilloscopes and smart thermometers.

If one insists on bringing ego into the arena, the digital devotee will argue that the analog content of these things is an unfortunate nuisance that must be tolerated. The analog aficionado, if permitted to speak, will counter that digital techniques exist only to aid in getting a better grip on a fundamentally analog existence. The question of who is most correct is subject to endless debate and is not really germane.

The point is that although analog is not dead, its remaining practitioners must be more systems creatures and less circuit addicts. To be sure, circuits are required to build systems, but analog technicians can only make themselves indispensable in a digital world by their recognized ability to supply what it needs to accomplish its mission.

That this is the case can be easily proven. Consider the effect on the major digital powers of a complete embargo of data converters and signal-conditioning components by the small analog nations. How can a supermarket scale compute the cost of goods it can't get weight information on? Of what use is a process controller without inputs or outputs? Think of the long lines of microprocessors waiting at the distributors for what few DIPs of analog I/O might be available! Imagine rationing of instrumentation amplifiers and V/F converters and alternate D/A and A/D days.

So it seems that analog is not so dead after all but really playing possum. By occupying this position, analoggers will stay healthy, very much alive, and need not leave town.

An uneasy but workable harmony has thus been negotiated with the dominating numerical nemesis. This compromise is not optimal, but it's certainly a more desirable and useful existence than being dead and is worthy of praise and respect by everyone.

Do all you bit pushers out there get the message?

Figure 4-3. Analoggers can stay very much alive and need not leave town.

7. Max Wien, Mr. Hewlett, and a Rainy Sunday Afternoon

One rainy Sunday afternoon, I found myself with nothing much to do. I've always treasured rainy Sundays that come supplied with spare time. With my first child on the way, I've taken a particular devotion to them lately. So I wandered off to my lab (no true home is complete without a lab).

I surveyed several breadboards in various states of inexplicable nonfunction and some newly acquired power transistors that needed putting away. Neither option offered irresistibly alluring possibilities. My attention drifted, softly coming to rest on the instrument storage area. On the left side of the third shelf sat a Hewlett-Packard series 200 oscillator. (No lab is complete without an HP series 200 oscillator, see Figure 7-1.)

The HP 200, directly descended from HP cofounder William R. Hewlett's master's degree thesis, is not simply a great instrument. Nor was it simply mighty HP's first product.[1] This machine is history. It provided a direction, methods, and standards that have been reflected in HP products to this day. There is a fundamental honesty about the thing, a sense of trustworthiness and integrity. The little box is a remarkable amalgam of elegant theoretical ideas, inspired design, careful engineering, dedicated execution, and capitalism. It answered a market need with a superior solution. The contribution was genuine, with the rewards evenly divided between Hewlett-Packard and its customers. The HP 200 is the way mother said things are supposed to be—the good guys won and nobody lost.

Digging in the lab library (no lab is complete without a library), I found my copy of William Redington Hewlett's 1939 Stanford thesis, "A New Type Resistance-Capacity Oscillator" (no lab library is complete without a copy).

Hewlett concisely stated the thesis objective (aside from graduating):

> The author has felt that there is a real need of a new type oscillator that would combine the stability of the coil-condenser type, the flexibility of operation of the beat-frequency type, and still be light and portable as well as simple in construction and adjustment.
>
> The object of this research has been development, construction, and testing of such an oscillator

Hewlett's oscillator used a resonant RC network originated by Max Wien in 1891 (see the references at the end of this chapter). Wien had no source of electronic gain

1. Also, incidentally, easily their longest-lived product. The HP 200 series was sold by Hewlett-Packard until the mid-1980s, a production lifetime of almost 50 years.

Figure 7-1. One of the original Hewlett-Packard Model 200A oscillators—the good guys won and nobody lost. (Photo courtesy of Hewlett-Packard Company.)

(DeForest hadn't even dreamed of adding a third element to Edison's Effect in 1891), so he couldn't readily get anything to oscillate. Anyway, Wien was preoccupied with other problems and developed the network for AC bridge measurements.

Hewlett saw that Wien's network, combined with suitably controlled electronic gain, offered significant potential improvements over approaches then used to make oscillators. These included dynamic tuning range, amplitude and frequency stability, low distortion, and simplicity.

Hewlett had something else besides electronic gain available; he also had the new tools of feedback theory. Harold S. Black's pioneering work, "Stabilized Feedback Amplifier," appears as the fourth reference in the thesis bibliography. Similarly, Nyquist's "Regeneration Theory," a classic describing necessary conditions for oscillation, is reference number three.

Hewlett synthesized all this nicely to show that Wien's network could be made to oscillate. Then he added a single (quite literally) crucial element. The oscillator's gain must be carefully controlled to support stable sinusoidal oscillation. If gain is too low, oscillation will not occur. Conversely, excessive gain forces limit cycles, creating a square wave oscillator. The problem is to introduce an amplitude regulation mechanism that does not generate output waveform distortion. Hewlett describes the elegant solution:

> The last requirement, an amplitude-limiting device that will not introduce distortion, is more difficult to achieve. It is well known that the gain of an amplifier with negative feedback is 1/ß, providing Aß is large compared to 1. Thus if a resistance whose value increases with the current through it is used as past of the negative feedback network, the gain of the amplifier may be made to decrease with an increase in the input voltage. If an amplifier of this type is used as part of the oscillator, it can be adjusted so that oscillations will just start. As oscillations

build up, the gain of the amplifier will be reduced, thus reducing the tendency to oscillate and causing the amplitude of oscillations to reach a stable value. If this value is low enough, the tubes will operate class A, and no serious distortion will be introduced. Furthermore, any distortion that is produced, due to the nonlinear characteristics of the tubes, will be reduced by a factor of Aß by the action of the negative feedback.

For the variable resistance, a small tungsten lamp may be used. It is a well known property of such lamps that as the current through them increases, the filament warms up, thereby increasing the lamp resistance. Figure 2 shows how the resistance of a 110 volt, 6 watt, tungsten lamp changes with the current through it. It may seem that the maximum rate of change of resistance is when the load current is less than 20 milliamperes, and so to get maximum effect, the lamp should be operated in this region. In Fig. 3 is shown a complete diagram of the oscillator. The negative feedback is applied from the plate of the output tube to the cathode of the input tube. The lamp is placed from cathode to ground, so as to increase the feedback and reduce the gain of the amplifier as the oscillation builds up.

The only requirement placed on the lamp is that it be operated at such a temperature that the time rate of change of cooling be small compared to half the period of the lowest frequency. As the radiation

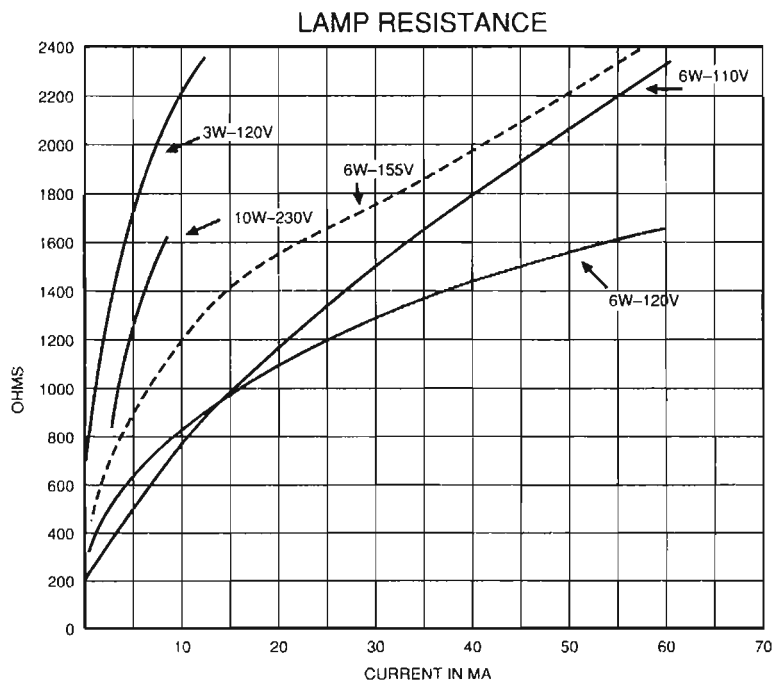


Figure 7-2. Hewlett's Figure 2 plotted lamp I-V characteristics. (Courtesy Stanford University Archives.)

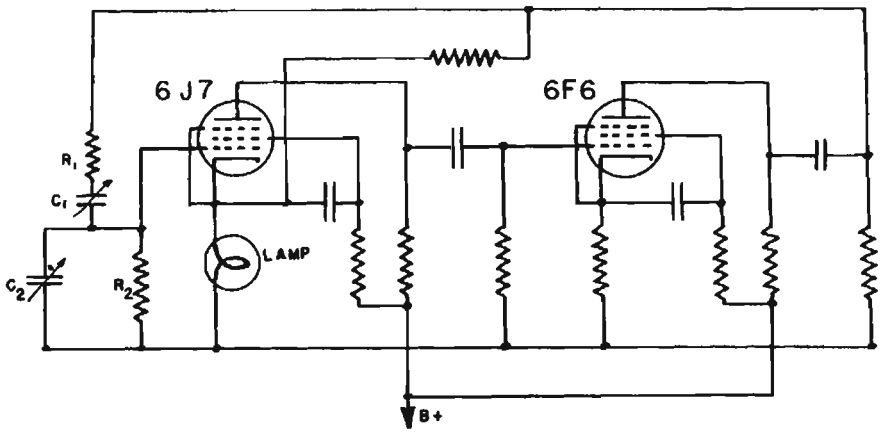


Figure 7-3. Hewlett's Figure 3 detailed the oscillator circuit. Note Wien network and lamp (Courtesy Stanford University Archives.)

is proportional to the fourth power of the absolute temperature, and as most of the energy is lost through radiation, this requirement may be easily met by not operating the lamp at too high a current. Under these conditions, the life of the lamp should be almost infinite.

Hewlett's use of the lamp is elegant because of its hardware simplicity.[2] More importantly, it is elegant because it is a beautiful example of lateral thinking. The *whole* problem was considered in an interdisciplinary spirit, not just as an electronic one. This is the signature of superior problem solving and good engineering.

The lamp solved a tricky problem, completing the requirements for a practical instrument. The design worked very well. It covered a frequency range of 20 to 20,000 cycles (it was cycles then, not Hertz) in three decade ranges with dial cali-

2. Hewlett may have adapted this technique from Meacham, who published it in 1938 as a way to stabilize a quartz crystal oscillator. Meacham's paper, "The Bridge Stabilized Oscillator," is in reference number five in Hewlett's thesis.

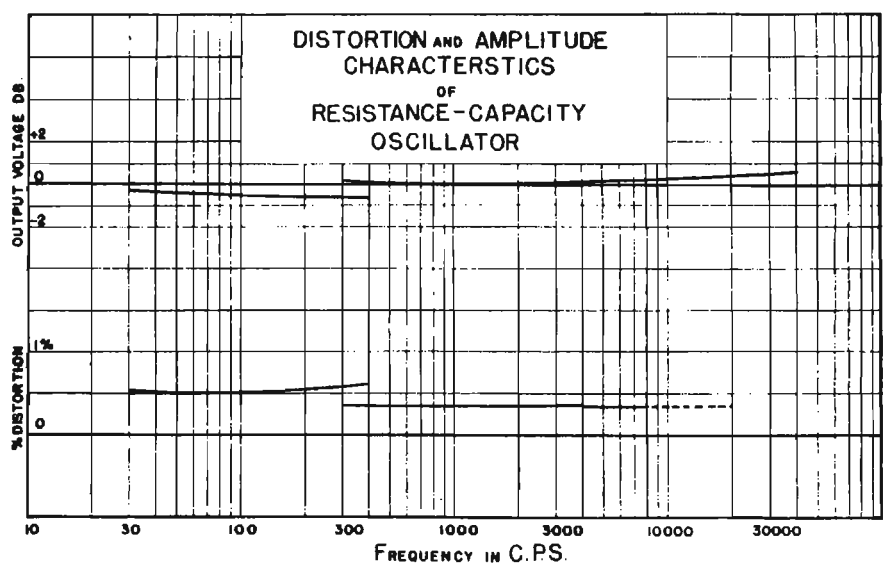


Figure 7-4. Hewlett's Figure 4 showed good distortion performance. What limited it? (Courtesy Stanford University Archives.)

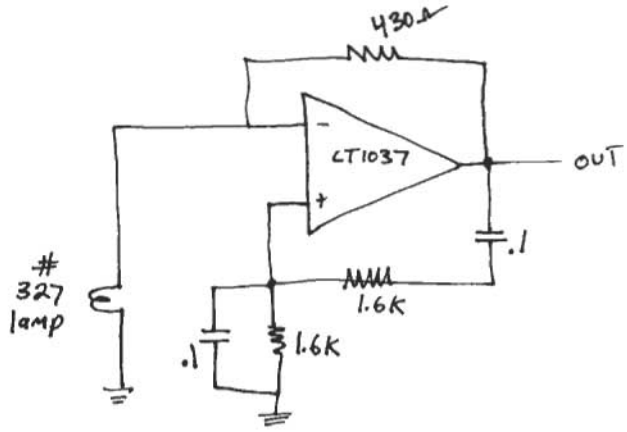


Figure 7-5. My version of Hewlett's circuit. Distortion was much better, but I was fifty years too late.

bration accuracy of 1%. The lamp maintained output amplitude stability within 0.2% at 100 cycles, varying only 1 dB from 20 to 20,000 cycles. Peering into my HP 201, I can see the light bulb, just where Hewlett, or one of his assistants, left it.

Hewlett's Figure 4 showed distortion well within 0.5% over the output range. This distortion figure caught my attention. By contemporary standards, Hewlett's 6J7/6F6-based "op amp" had major performance limitations.[3] How good, I wondered, would Hewlett's basic circuit be with a modern op amp?

And so, some fifty years after Hewlett finished, I sat down and breadboarded the oscillator to the meter of that Sunday afternoon rain. My circuit is shown in Figure 7-5.

This circuit is identical to Hewlett's, except that I have managed to replace two vacuum tubes with 94 monolithic transistors, resistors, and capacitors.[4] (I suppose this constitutes progress.) After establishing the 430 Ω value, the circuit produced a very nice sine wave. Connecting my (HP) distortion analyzer, I was pleased to measure only 0.0025% distortion (Figure 7-6). Then, I went ahead and endowed the basic circuit with multiple output ranges as shown in Figure 7-7.

This also worked out well. As Hewlett warned, distortion increases as oscillator

3. For those tender in years, the 6J7 and 6F6 are thermionically activated FETs, descended from Lee DeForest.

4. To be precise, there are 50 transistors, 40 resistors, and 4 capacitors in the device.

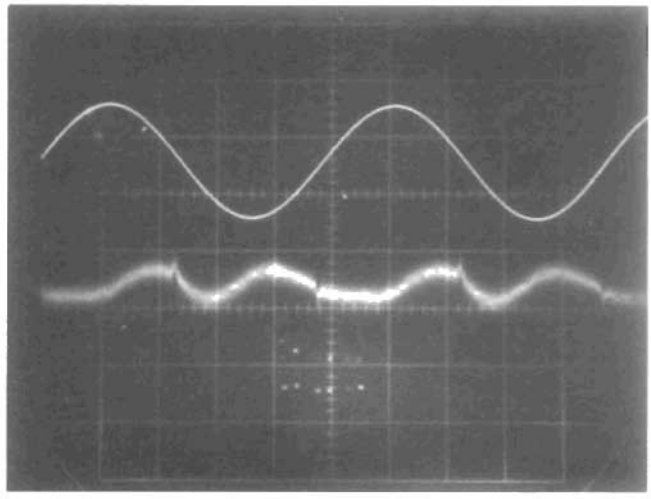


Figure 7-6. Output waveform and distortion for my first oscillator. Distortion was 0.0025%.

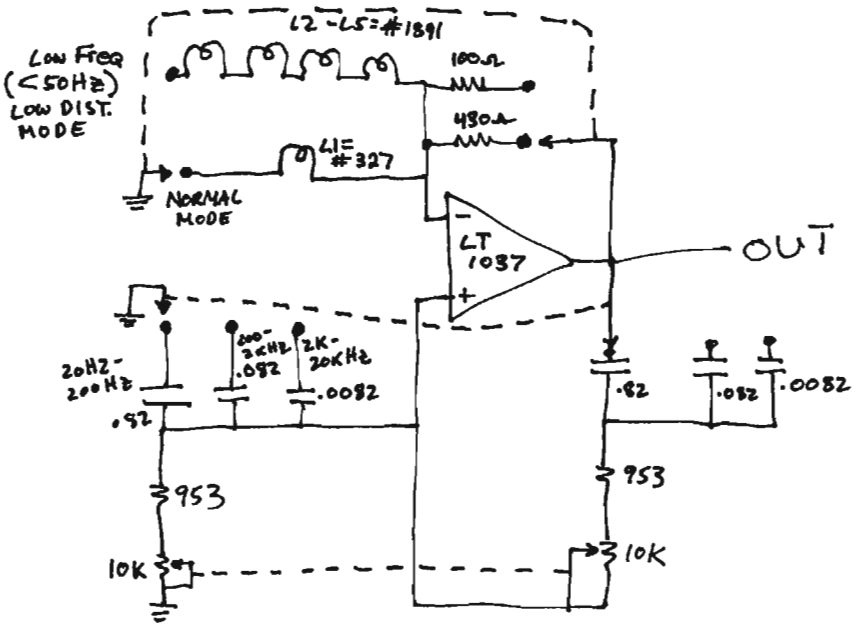


Figure 7-7. A wide range version of the basic circuit. Multiple lamps provided longer gain loop time constant, improving low frequency distortion.

frequency descends towards the lamp's thermal time constant. This effect can be attenuated by increasing the lamp's thermal time constant. The easiest way to do this is to add more and bigger lamps. This causes longer amplitude settling times, but low frequency distortion is reduced. Plotting distortion versus frequency clearly shows this (see Figure 7-8).

Looking at the plot, I wondered just how far distortion performance could be pushed using Hewlett's suppositions and conclusions as a guide. The multi-lamp experiment indicates that distortion rise at low frequencies is almost certainly due to the lamp's thermal time constant. But what causes the slight upward tilt around 15 to 20 kc? And just what limits distortion performance? Chasing all this down

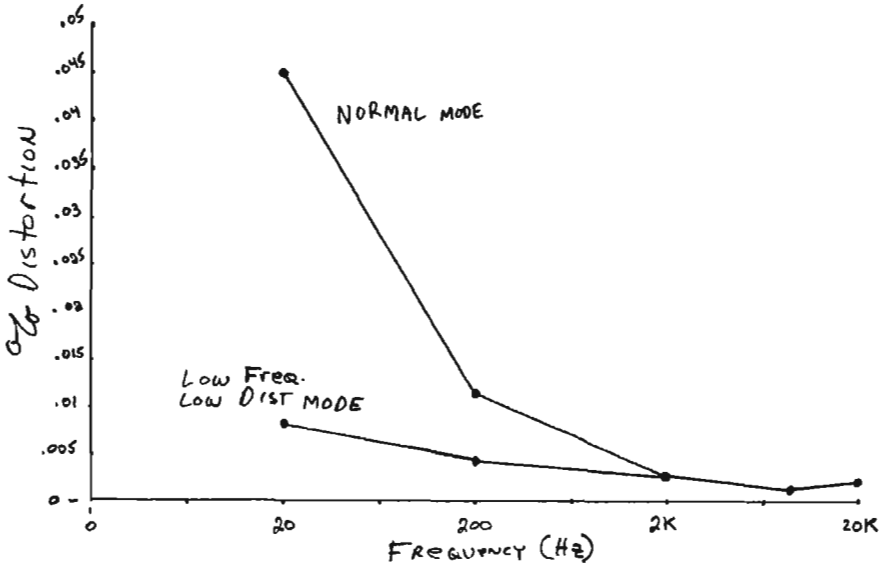


Figure 7-8. Distortion versus frequency for the wide range oscillator. The effect of the multiple lamp approach is clearly evident, but what causes increasing high frequency distortion?

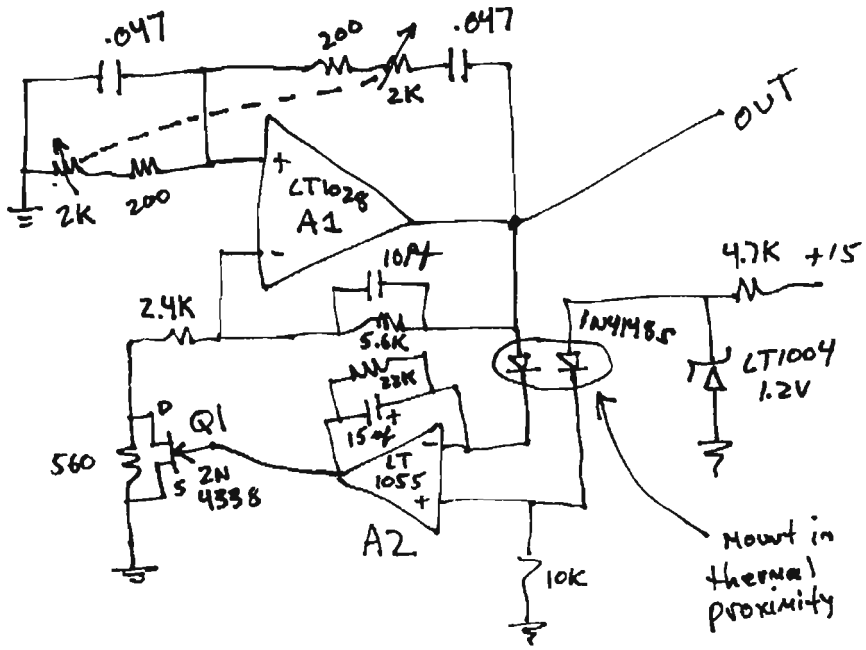


Figure 7-9. The first attempt at improving distortion. A2 and Q1 replace the lamp.

seemed an entertaining way to stay out of the rain. Of course, I couldn't ignore that I was already perilously near my analyzer's 0.0018% specification limit when interpreting results. Not to worry.

The next circuit is shown in Figure 7-9.

A1, a low noise wideband amplifier, is the oscillator. The variable resistor's decreased value maintains low noise performance by minimizing bias current induced noise. The 10 pF capacitor suppresses parasitic high frequency oscillation. A2 and associated components replace the lamp(s). A2 compares the oscillator's positive peaks with a DC reference and servo-controls Q1 to establish proper loop gain. The diode in series with the DC reference temperature compensates the rectifier diode. The large feedback capacitor sets a long time constant for A2, minimizing output ripple.

When I turned this circuit on, it oscillated, but distortion increased to a whopping 0.15%! The analyzer output showed a fierce second harmonic (twice the oscillator frequency), although A2's output seemed relatively clean (see Figure 7-10).

So, I might have gotten away with dumping the two tubes for 94 transistors, capacitors, and resistors, but replacing the lamp with a bunch of stuff was another matter! I looked apologetically at the forsaken light bulbs.

What happened? The Wien network is the same, and it's hard to believe A1 is so bad. A2's output shows some residual rectification peaking, but nothing that would unleash such a monster.

The culprit turns out to be Q1. In a FET, the channel resistance is ideally fixed by the gate-channel bias. In fact, slight modulation of channel resistance occurs as the voltage across the channel varies. Unfortunately, Q1's drain sees significant swing at the oscillator fundamental. The gate is nominally at DC, and the source grounded. This causes unwanted modulation of the amplitude stabilization loop by the oscillator's fundamental, creating distortion. The humble light bulb was beginning to look pretty good.

If you stare at this state of affairs long enough, the needed Band-Aid presents

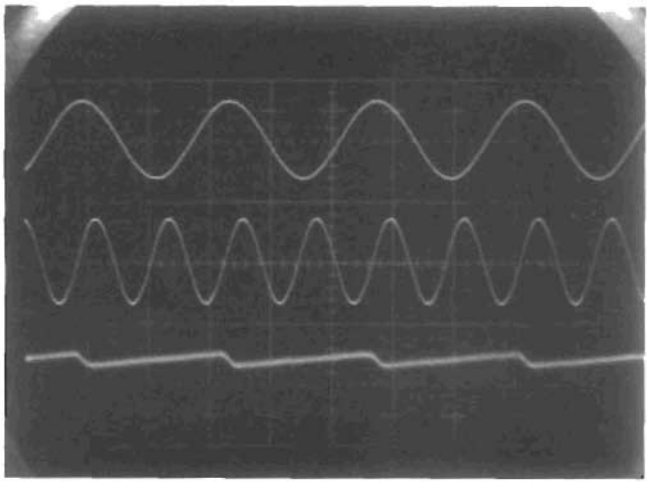


Figure 7-10. Performance for the "lampless" oscillator. Modern technology is almost 100 times worse!

Output 2V/DIV

Distortion .15%

A2 Output (AC coupled) .1V/DIV

Horiz. = 200μsec/DIV

itself and is (thank the gods) refreshingly simple. The JFET is a fairly symmetrical structure, although this circuit drives it asymmetrically from gate to source. If you arrange things so the gate is driven with a signal halfway between the drain and source, symmetry is reestablished. This symmetrical drive eliminates all even-order harmonics. Q1's new companions make things look like Figure 7-11.

With the trimmer set to provide the optimum amount of feedback, distortion dropped to just 0.0018%—the analyzer's specified limit (see Figure 7-12).

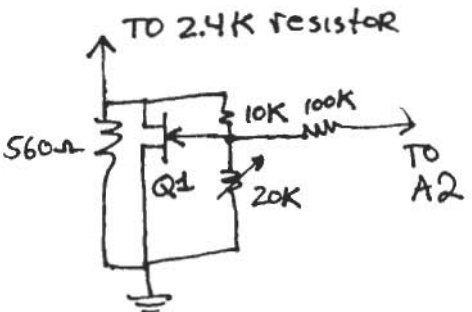


Figure 7-11. The local feedback network around Q1, intended to cure channel resistance modulation effect.

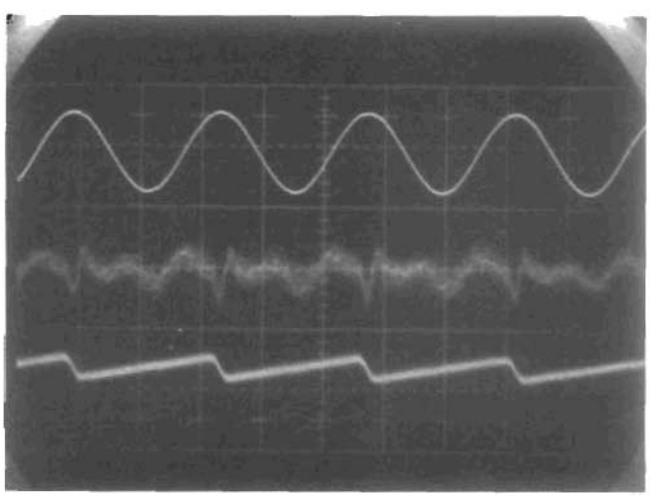


Figure 7-12. Results of Q1's local feedback fix. Distortion improves to 0.0018%....about as good as the light bulb.

Output 2V/DIV

Distortion .0018%

A2 Output (AC coupled) .1V/DIV

Horiz. = 200μsec/DIV

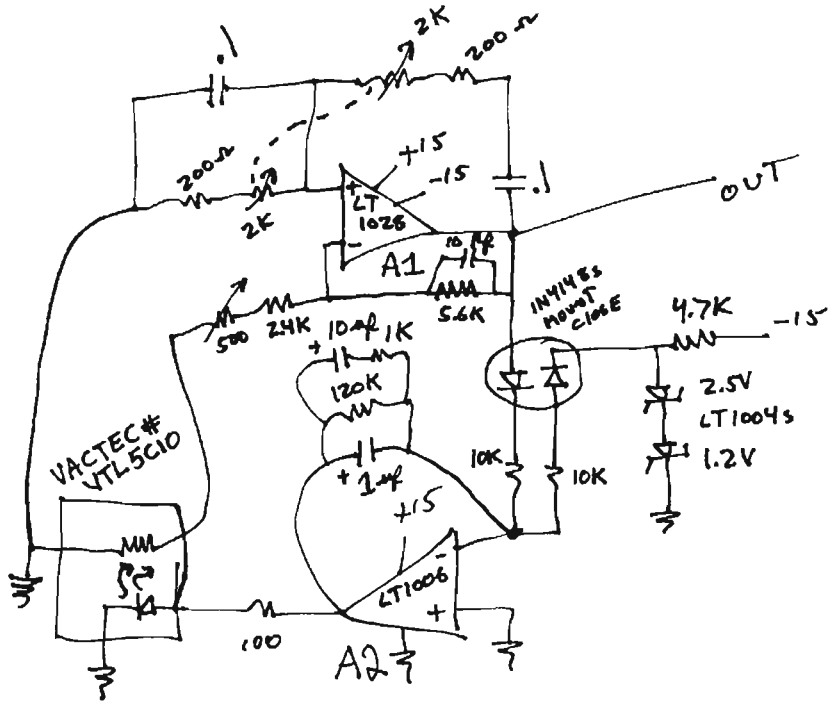


Figure 7-13. Replacing Q1 with an optically driven photocell eliminates the resistance modulation trim. A2 is now a ground-referenced integrator.

While praying that the analyzer was better than it had to be, I looked at what it was saying. Some of the first harmonic was visible, along with artifacts of the amplitude control loop's rectification peaking. No amount of fiddling with the distortion trimmer could reduce the first harmonic, although increasing A2's feedback time constant reduced rectification related content.

I didn't like the trimmer, and A2's feedback capacitor was a big dog. Also, A2 is not a true integrator and has noise gain from its positive input. This seemed more irritating than obviously relevant. Similarly annoying was the notion that if A2 ever swings positive (start-up, whatever), the electrolytic reverse biases. This ain't pertinent either but still is bad manners!

The next iteration attempted to deal with some of these issues (see Figure 7-13).

The most noticeable change is that Q1 has been replaced with an optically driven CdS photocell. These devices don't suffer from parasitic resistivity modulation, offering a way to eliminate the trim. A2, running single supply, is now a ground-sensing type configured as a true integrator. The feedback components are arranged in a weak attempt to get a long time constant with improved settling time. Lastly, the DC reference has been increased, forcing greater oscillator swing. This is a brute force play for a more favorable signal/noise ratio.

This experiment provided useful information. A2's modifications eliminated rectifier peaking artifacts from the distortion analyzer's output. The LED-driven photocell really did work, and I tossed the trimmer down to the end of the bench. The analyzer indicated 0.0015%, but I wasn't sure if I could take this "improvement" seriously. Interestingly, the second harmonic distortion product looked the same, although perhaps less noisy. It increased a bit with higher frequencies and more or less ratioed with shifts in output amplitude (facilitated by clip-leading across one of the LT1004 references). The analyzer seemed to give readings a few parts-per-million (ppm) lower for higher oscillator amplitude, suggesting signal/noise issues with the circuit, the analyzer, or both. But understanding the source of the second harmonic distortion product was clearly the key to squeezing

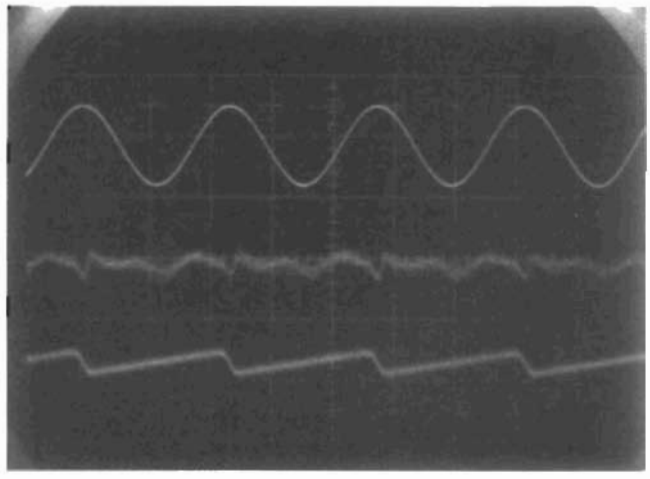


Figure 7-14. A2's increased time constant reduces rectification related distortion content.

more performance. The circuit was talking, and I was trying to listen, but I wasn't hearing (see Figure 7-14).

All this seemed to exonerate the gain control loop. That left the Wien network, the op amp, or some parasitic that wasn't on the schematic as the villain.

I considered the possible effects of voltage coefficient in the Wien network resistors and ESR or dielectric absorption in the capacitors. Sometimes when you don't know how to make things better you can learn by trying to make them worse. So I added tiny, controlled parasitic RC terms to the Wien R's and C's to test their sensitivity to component imperfections. What I found indicated that the reasonably good grades of R and C I was using were not the problem. I bolstered this conclusion by trying different R's and C's in the Wien network. Various decent grades of components all produced about the same result. That kinda left A1. Open loop gain, which degrades with frequency, could be a problem, so I decided to add a buffer to unload the amplifier. Beyond this, I couldn't do much else to increase available gain.

Now that I had license to accuse the op amp, the answer quickly seemed apparent. This circuit was in violation of a little known tenet of precision op amp circuits: Williams's Rule. Williams's Rule is simple: *always invert* (except when you can't). This rule, promulgated after countless wars with bizarre, mysterious, and stubborn effects in a variety of circuits, is designed to avoid the mercurial results of imperfect op amp common mode rejection. Common mode–induced effects are often difficult to predict and diagnose, let alone cure. A zero volt summing point is a very friendly, very reassuring place. It is (nominally) predictable, mathematically docile, and immune from the sneaky common mode dragons.

All present amplifiers have decreasing common mode rejection with frequency, and A1 is no exception. Its common mode rejection ratio (CMRR) versus frequency plot is shown in Figure 7-15.

The oscillator forces large common mode swings at A1. Since CMRR degrades with frequency, it's not surprising that I saw somewhat increased distortion at higher frequencies. This seemed at least a plausible explanation. Now I had to test the notion. Doing so required bringing the circuit into alignment with Williams's Rule. Committing A1's positive input to ground seems an enormous sacrifice in this circuit. I considered various hideous schemes to accomplish this goal. One abomination coupled the Wien network to A1's remaining input via a transformer. This approach wasn't confined to technical ugliness; in all likelihood, it would be considered obscene in some locales. I won't even sketch it, lest the publisher be hauled

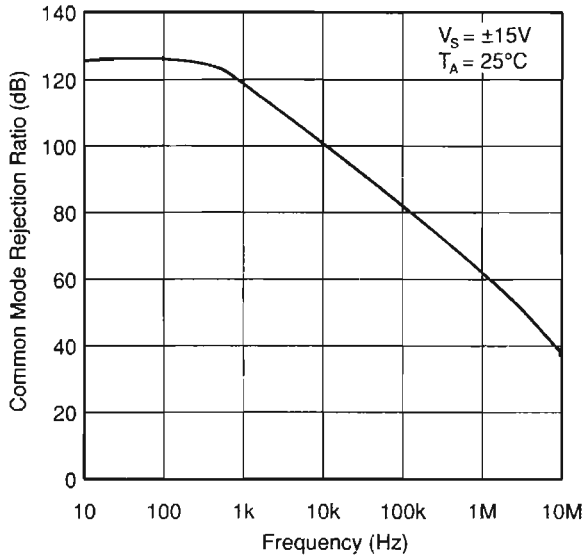


Figure 7-15. Common mode rejection ratio versus frequency for A1.

into court by some fundamentalist op amp group. Even if I could have gotten the whole perverse hulking thing to work, it just didn't feel right. I could hear Hewlett's simple, elegant little light bulb, which worked so well, laughing at me.

Somewhere in the venerable *Philbrick Applications Manual*, the writer counsels that "there is always a Way Out." The last circuit (Figure 7-16) shows what it was.

This configuration is identical to the previous one, except A3 appears along with buffer A4. A3 maintains A2's positive input at virtual ground by servocontrolling the formerly grounded nodes of the Wien network and the gain control loop. This adds a third control loop to Hewlett's basic design (this is getting to be a very busy

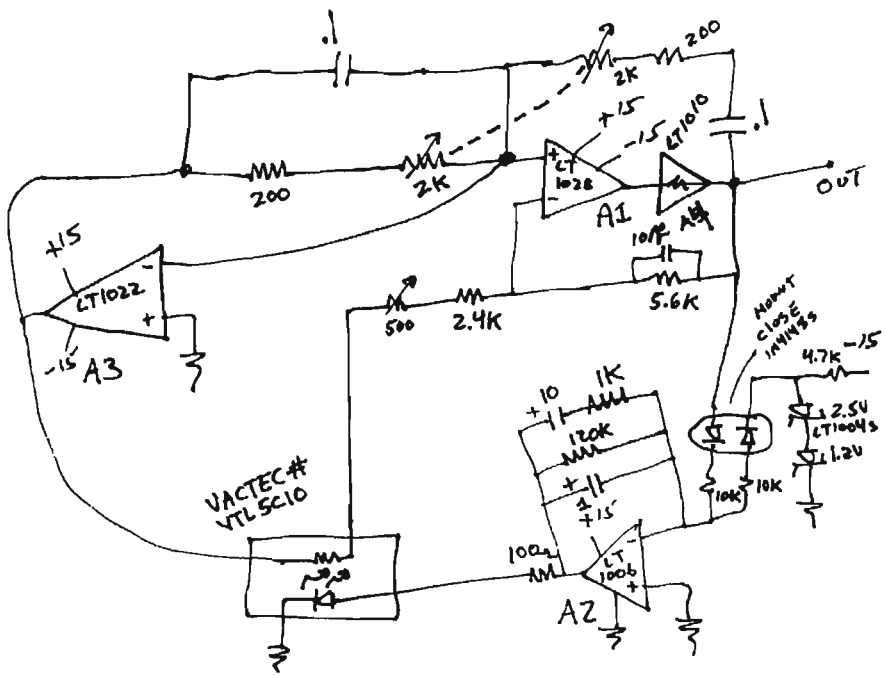


Figure 7-16. The final circuit. A3 eliminates common mode swing, allowing 0.0003% (3 ppm) distortion performance.

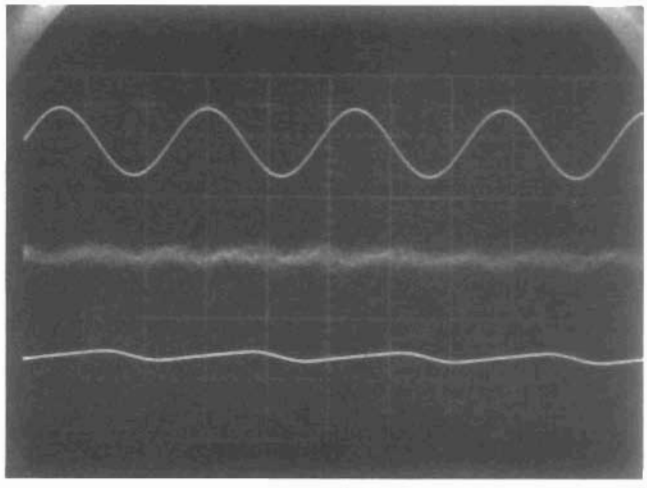


Figure 7-17. Common mode suppression runs distortion analyzer into its noise floor.

oscillator—pity poor A1, besieged by three masters) but does not adversely affect its operation. With its positive input forced to virtual ground, A1 sees no common mode swing. Williams's Rule is satisfied, and ostensibly, good things should happen.

To my utter amazement, this whole thing did not explode when I finally summoned the nerve to turn it on. Even more astonishing was the distortion analyzer's 0.0008% reading (Figure 7-17).

Its output showed only faint traces of the first harmonic outlined in noise. The analyzer was indicating more than a factor of two beyond specification, which was really asking a lot. While it's unlikely that the oscillator and analyzer have compensatory errors, it's dangerous to conclude anything. As such, I turned to some very specialized equipment to get at the truth.

The Audio Precision System One will read distortion down to 0.0003% (3 ppm). I was quite pleased to see that it couldn't find anything above this level.

After Hewlett finished *his* oscillator, he and David Packard went into their garage and built a few into boxes and then made some more kinds of instruments

Figure 7-18. Bill Hewlett and David Packard building oscillators at the Hewlett-Packard Company, located in their garage. (Photo courtesy Hewlett-Packard Company)

(Figure 7-18). After I finished *my* oscillator, I went into the kitchen and made a few hot dogs for dinner (mustard, chili sauce, no beans) and then made some other stuff. So, not only was Hewlett a lot cleverer than me, he also had somewhat different priorities. However, he did eventually get around to dinner, and I understand he ate pretty well. My hot dogs tasted pretty good.

Acknowledgment

The author gratefully acknowledges William R. Hewlett's review of the manuscript and his commentary. Additionally, the cooperation of the Hewlett-Packard Company and Stanford University was beyond value.

References

Wein, Max (1891). "Measung der induction constanten mit dem 'Optischen Telephon'," *Ann. der. Phs.* **44**, 704–707.

Hewlett, William R. (1939). "A New Type Resistance-Capacity Oscillator," M.S. Thesis, Stanford University, Palo Alto, California.

Bauer. Brunton (1949). "Design Notes on the Resistance-Capacity Oscillator Circuit," Parts I and II. *Hewlett-Packard Journal*, Volume 1, Nos. 3 and 4 (November--December 1949).

Hewlett, William R. (1983). "Inventions of Opportunity: Matching Technology with Market Needs," Hewlett-Packard Company, Palo Alto, California.

Williams, Jim (1984). "Thermal Techniques in Measurement and Control Circuitry," Linear Technology Corporation Application Note 5, Linear Technology Corporation, Milpitas, California.

Williams, Jim (1990). "Bridge Circuits: Marrying Gain and Balance," Linear Technology Corporation Application Note 43, Linear Technology Corporation, Milpitas, California.

13. Should Ohm's Law Be Repealed?

When I was a kid, the Stearn family lived nearby. Dr. Stearn, his wife, and two daughters had a really nice place. The house, a big old Victorian, was grand but friendly. They had a pool, shuffleboard and tennis courts, dogs, and a horse named Fred. Inside, there was lots of spirited modern art, a terrific library with a ladder that slid around the room on a rail, and great junk food. They had a wonderful collection of old surgical instruments and some great stained glass lamps. There were also pool and billiard tables, a pinball machine, and a darkroom. One daughter, my age, had cute freckles and long, chestnut hair. Once, she even baked me chocolate chip cookies and presented them in a blue box with a ribbon. They were good. I can't be sure, but I think I missed a cue. A born engineer.

For an eight-year-old boy, it should have been a really fun place. All of the attractions were of some passing interest but really weren't even distractions. Because what Dr. Stearn had, what he *really* had, was in the basement. There, sitting on something called a "Scopemobile," next to the workbench, was a Tektronix 535. That I loved this oscilloscope is an understatement. I was beyond infatuation, long past mesmerization (see Figure 13-1).

The pure, unbounded lust I spent toward this machine probably retarded the onset of my puberty, delaying sexual nascency by at least a year.[1] I also destroyed my grade school performance. I read the mainframe manual instead of doing my homework and studied the plug-in books (they were smaller and easier to hide) in Mrs. Kemp's English class. I knew every specification and all the operating modes. I lived for that 535, and I studied it. But, best of all, I used it.

Dr. Stearn, when he wasn't doctoring or being with his family, shared his electronics hobby with me. Since no amount of pleading, scheming, bamboozling, or anything else would get my father to buy one, Dr. Stearn also shared his 535 with me. Oscillators, amplifiers, flip-flops, modulators, filters, RF stages—we circuit-hacked them all with ferocious intensity. And with that 'scope you could really *see* what was going on. You knew the excitement Leeuwenhoek felt when he looked in his microscope.

In fact, the Tektronix 535 was a sublime masterpiece. In 1956, it was so vastly superior, so far ahead of everything else, that it made a mockery of the competition. The triggered sweep worked unbelievably well, and the calibrated vertical and horizontal really were calibrated. It had an astounding 15 megacycles (it was cycles then, not Hertz) of bandwidth and something called "delayed sweep." The plug-in

---

Versions of this chapter's text have been published by Linear Technology Corporation and *EDN Magazine*.

1. Testament to the staying power of this childhood desire is the author's current ownership of copious amounts of Beaverton hardware.

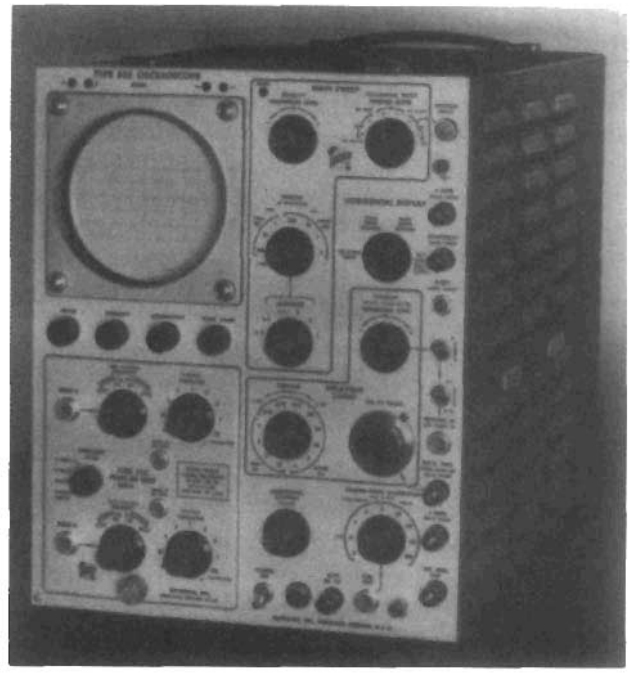


Figure 13-1. The Tektronix 535. Introduced in 1954, this vastly superior instrument made a mockery of all competition. I knew I could use it to make my breadboards work. Or so I thought. (Photo Courtesy of Tektronix, Inc.)

vertical preamplifiers greatly increased measurement capability, and I always quickly concurred when Dr. Stearn decided to buy another one.

The 535's engineering concepts and production execution were a bumpless combination of interdisciplinary technology, inspired design, attention to detail, aesthetics, and usability. It combined solid knowledge of fundamentals, unbounded thinking, and methodical discipline to produce a superior result. The thing just radiated intellectual honesty.

Using that 'scope inspired confidence bordering on arrogance. I knew I could use it to make my breadboards work. Or so I thought.

One afternoon I was having trouble getting a circuit to work. Signals looked about right, but not really, and overall performance was shaky, with odd effects. I 'scoped everything but got nowhere. Dr. Stearn came by (after all, he lived there). He listened, looked, and thought awhile. Then he moistened two fingers, and started doing a little hand dance on the circuit board. His hand moved around lightly, touching points, as he watched the 'scope. He noticed effects and, correlating them to his hand movements, iterated toward favorable results. When things looked good, he stopped his motion. He rocked his fingers gently back and forth, watching the display respond. He paused, thought, and then soldered a small capacitor between the last two points his fingers were on. To my amazement, the display looked good, and the circuit now worked. I was dumbfounded and, propelled by frustration and embarrassment, a little angry.

He explained that the circuit had a high frequency oscillation, perhaps 100 megacycles, and he suspected he'd damped it by loading the right points. His finger dance had surveyed suspect points; the capacitor was his estimate of the electrical equivalence of the finger loading.

"That's not fair," I protested. "You can't see 100 megacycles on the 'scope."

He looked right at me and spoke slowly. "The circuit doesn't care about fair, and it doesn't know what the 'scope can't see. The 'scope doesn't lie, but it doesn't always tell the truth." He then gave me a little supplementary lecture which has served me well, except when I'm foolish or frustrated enough to ignore it.

"Don't ever get too attached to a way of solving problems. Don't confuse a tool, even a very good one, with knowing something. Concentrate on understanding the problem, not applying the tool. Use any tool that will help move your thinking along, know how those tools work, and keep their limitations in mind when you use them—it's part of the responsibility of using them. If you don't do this, if you stop thinking and asking, if you simply believe what the 'scope says, you're done for. When you do that, you're not listening to the problem, and you're no longer designing the circuit. When you substitute faith in that instrument, no matter how good it is, for your judgment, you're in trouble.

"It's a tricky trap—sometimes you don't even know you're falling into it. People are very clever at fooling themselves that way. We're all human, we all very badly want things to be simple and go smoothly. But that circuit doesn't know that and it doesn't care."

That was 34 years ago. I'm still absorbing that advice, although not progressing as rapidly as I'd like. I think Doc Stearn was right. I remember him often, usually after I've been stung by me again. My interest in tools, applying them, and human tendencies continues, and hopefully I'll get better at it all.

Lately, I've been hearing quite a bit about CAD systems, computer-based workstations, and powerful software modeling techniques. At Linear Technology, where I work, we have CAD systems and they save tremendous amounts of time. They're very powerful tools, and we're learning how and when to use them efficiently. It's a tough process, but the rewards are high and well worth the effort.

Unfortunately, I see substantive and disturbing differences between what I feel these tools are and what some of them purport to be.

There is a great deal of fanfare surrounding CAD systems today (see Figure 13-2). Promotional material, admittedly always suspect, emphasizes speed, ease of use, and elimination of mundanities and odious tasks in the design process. Unbearably attractive engineers in designer clothes reside in immaculately clean and organized work areas, effortlessly "creating." Advertising text explains the ease of generating

Figure 13-2. CAD advertising assures high productivity with minimal hassle. Becoming the next Edison is only a keystroke away.

ICs, ASICs, board functions, and entire systems in weeks, even hours. Reading further, the precipitators of this nirvana are revealed: databases, expert systems, routers, models, simulators, environments, compilers, emulators, platforms, capturers, synthesizers, algorithms, virtualizers, engines, and a lot of other abstruse intellectual *frou-frou* Ohm and Kirchoff never got to. These pieces of technological manna ostensibly coalesce to eliminate messy labs, pesky nuts and bolts, and above all, those awful breadboards. Headaches vanish, fingers and the lab (if it hasn't been converted to the company health spa) are clean, the boss is thrilled, and you can go fishing. Before you leave, don't forget to trade in your subscription to *EDN* for one to *Travel and Leisure*. I can hear Edison *kvetching*: "It's not fair, I didn't have a CAD system." It's okay, Tom, you did pretty well, even if your lab was a mess.

Well, such silliness is all part of the marketing game, and not unknown wherever money may trade hands. *Caveat emptor* and all that. So maybe my acerbic musings are simply the cynicism-coated fears of a bench hacker confronting the Computer Age. Perhaps I'm just too invested in my soldering iron and moistened fingers, a cantankerous computer technopeasant deserving recuse. But I don't think so, because what I see doesn't stop at just fast-talking ad copy.

Some universities are enthusiastically emphasizing "software-based" design and "automatic" design procedures. I have spent time with a number of students and some professors who show me circuits they have designed on their computers. Some of the assumptions and simplifications the design software makes are interesting. Some of the resultant circuits are also interesting.

Such excessively spirited CAD advocacy isn't just found in ad copy or universities. Some industry trade journals have become similarly enamored of CAD methods, to the point of cavalierness. Articles alert readers to the ease of design using CAD; pristine little labeled boxes in color-coordinated figures are interconnected to form working circuits and systems. Sometimes, editorial copy is indistinguishable from advertising. An editorial titled "Electronic Design Is Now Computer Design" in the January 1988, issue of *Computer Design* informed me that,

"For the most part, the electronic details—the concerns of yesteryear about Ohm's law and Kirchoff's law, transconductance or other device parameters—have been worked out by a very select few and embedded in the software of a CAE workstation or buried deep within the functionality of an IC. Today's mainstream designers, whether they're designing a complex board-level product or an IC, don't need to fuss with electronics. They're mostly logic and system designers—computer designers—not electronics designers."

That's the road to intellectual bankruptcy; it's the kind of arrogance Doc Stearn warned about. Admittedly, this is an extreme case, but the loose climate surrounding it needs examination.

CAD is being oversold, and it shouldn't be. It shouldn't be, because it is one of the most powerful tools ever developed, with broad applicability to problem solving. If too many users are led astray by shuck and jive and become disappointed (and some already are), the rate of CAD purchase, usage, and acceptance will be slowed. In this sense, the irresponsible self-serving advisories of some CAD vendors and enthusiasts may be partially self-defeating. The associations being made between CAD tools and actual knowledge-based, idea generation, and iterative processes of design are specious, arrogant, and dangerous. They are dangerous because many of us are human. We will confuse, admittedly perhaps because our humanness begs us to, faith in the tool with the true lateral thinking and simple sweat that is design. We

will cede the judgmental, inspirational, and even accidental processes that constitute so much of what engineering is. In the rush to design efficiency, we may eliminate time and sweat at the expense of excellence. Very often the mundanities and mental grunt work aspects of problem solving provide surprises. They can force a review process that mitigates against smugness and ossification. Most of the time this doesn't occur, but when it does the effect is catalytic and intellectual left turns often follow.

In misguided hands, a group of packaged solutions or methods looking for a problem will produce nothing at worst, an amalgam of mediocrity at best.

I also said associations between CAD tools and critical elements in the design process were arrogant. They are arrogant because in their determination to streamline technology they simplify, and Mother Nature loves throwing a surprise party. Technologically driven arrogance is a dangerous brew, as any Titanic passenger will assure you.

Most good design is characterized by how the exceptions and imperfections are dealt with. In my field, linear circuits, just about everything is exceptions. A lot of the exceptions you know about, or think you do, and you're constantly learning about new exceptions. The tricky thing is that you can get things to work without even realizing that exceptions and imperfections are there, and that you could do better if only you knew. The linear circuit designers I admire are those most adept at recognizing and negotiating with the exceptions and imperfections. When they get into something they're often not sure of just what the specific issues will be, but they have a marvelous sense of balance. They know when to be wary, when to hand wave, when to finesse, when to hack, and when to use computers. These people will use CAD tools to more efficiently produce superior work. The others may be tricked, by themselves or by charlatan-hucksters, into using CAD to produce mediocrity more efficiently. (See Figure 13-3.)

The time has come to sum up. When reading, I enjoy this moment because I want to watch the author become more definitive without getting the foot in the mouth.

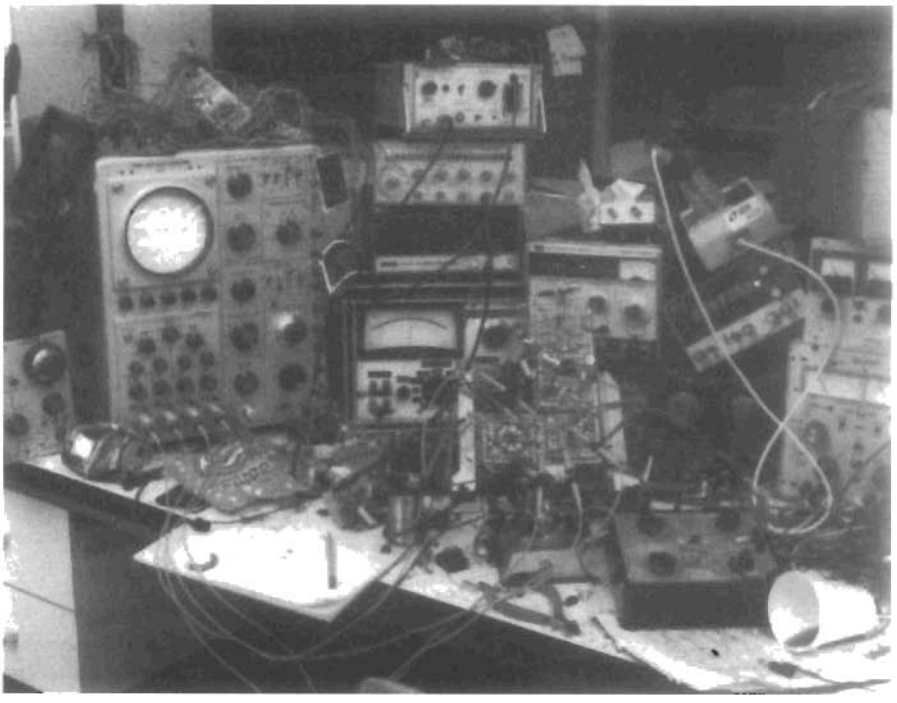


Figure 13-3. Combining other approaches with CAD yields the best circuits.

When writing, I fear this moment for the same reason. On this outing, however, I'm not so fearful. The ground seems pretty solid.

CAD-based tools and techniques, although in their infancy, will prove to be one of the most useful electrical engineering tools ever developed. In some areas, they will become effective more quickly. They have already had significant impact in digital ICs and systems, although their usefulness in linear circuit design is currently limited. As these tools emerge, the best ways to combine them with other tools will become clearer. And they will combine with other tools, not supplant them. Right now, the best simulator we have, a "virtual model" if you will, is a breadboard. In current parlance, breadboards are full parallel, infinite state machines. They have self-checking, self-generating software and heuristically generated subroutines with infinite branching capability. If you're listening, the answer, or at least the truth, is there.

I'm reasonably certain breadboardless linear circuit design is a long way off. I suspect similar sentiments apply in most engineering disciplines. The uncertainities, both known and unknown, the surprises, and the accidents require sweat and laboratories. CAD makes nail pounding easier, but it doesn't tell how to do it, or why, or when. CAD saves time and eliminates drudgery. It increases efficiency but does not eliminate the cold realities involved in making something work and selling it to someone who wants it and remains happy after the purchase.

Where I work, we eat based on our ability to ship products that work to customers that need them. We believe in CAD as a tool, and we use it. We also use decade boxes, breadboards, oscilloscopes, pulse generators, alligator clips, screwdrivers, Ohm's law, and moistened fingers. We do like Doc Stearn said back in 1956—concentrate on solving the problem, not using the tool.

23. The Zoo Circuit

History, Mistakes, and Some Monkeys Design a Circuit

This chapter is dedicated to the memory of Professor Jerrold R. Zacharias, who saved my ass.

A couple of years ago, I was asked to design a circuit for a customer. The requirements were not trivial, and the customer was having difficulty. I worked on this problem for some time and was asked to present my solution in a formal design review at the customer's location.

When I say "formal," I mean it! I came expecting to talk circuits with a few guys over a pizza. Upon arrival, I was taken to a large and very grand room, reminiscent of a movie theater. About 150 engineers were in attendence. There was every audio-visual machine known to humanity at the ready, and I was almost embarrassed to report that I had no slides, overheads, charts, or whatever (although a piece of chalk would be nice). A "senior technical management panel," positioned in a boxed-off section adjacent to the lectern, was to present a prepared list of questions. A video camera duly recorded the proceedings. The whole thing was chaired by somebody who introduced himself as "Dr. So-and-So, senior vice-president of engineering." Everybody in the place talked in whispers and nodded his head a lot. I found myself alternating between intimidation and amusement.

I gave a fairly stiff presentation, clutching my dear little piece of chalk the whole time. Things seemed to go okay, but not great, and then the panel began with their prepared list of questions. The first question went something like, "Can you explain, precisely, where the ideas for this and that piece of the circuit came from? Can you detail what design procedures, programs, and methodologies were helpful?"

I considered various acceptable answers, but decided to simply tell the truth: "Most of the ideas came from history, making mistakes, and the best source of help was some monkeys at the San Francisco Zoo."

You could have heard a pin before it dropped. There was absolute silence for a bit, and then some guy stood up and asked me to elaborate "a little." Everybody cracked up, the mood shifted, and we finally began to really *talk* about the circuit.

This customer originally came to me with a need for a "CMOS voltage-to-frequency converter." The performance requirements were as follows:

Output frequency	0–10 kHz
Input voltage	0–5 V
Linearity	0.04%
Drift	100 ppm/°C
PSRR	100 ppm/V
Temperature range	0°–55°C
Step response	< 5 cycles of output frequency
Output pulse	5 V CMOS-compatible
Power supply	Single 9 V battery (6.5–10 V)
Power consumption	200 μA maximum
Cost	< $6.00/100,000 pieces

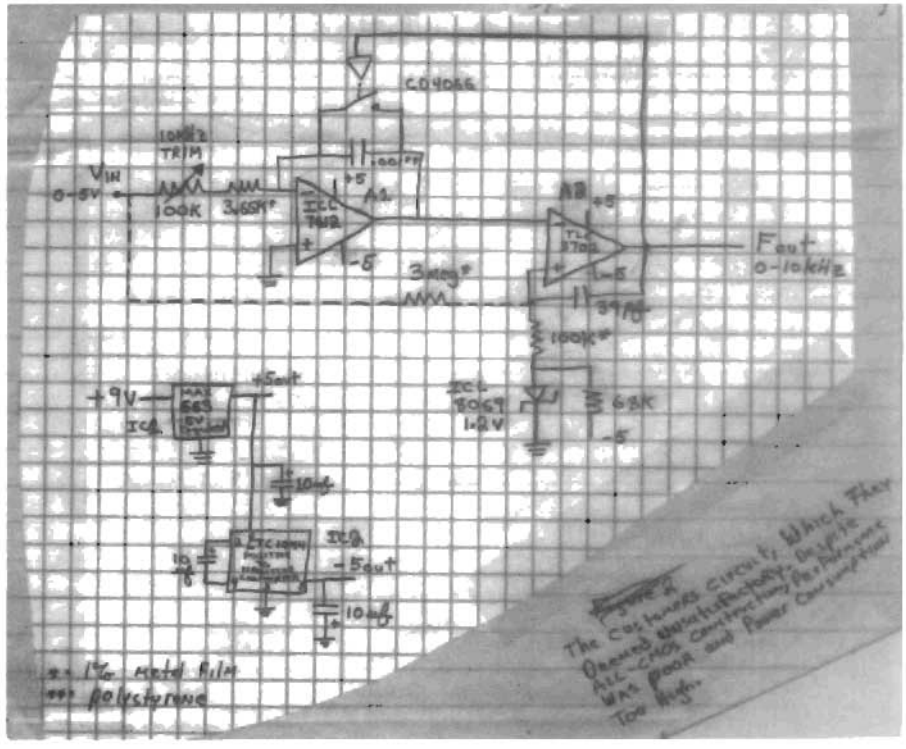


Figure 23-1. The customer's circuit, which was deemed unsatisfactory. Despite all-CMOS construction, performance was poor and power consumption too high.

These people had been working on a design for several months. It functioned, but was described as wholly unsatisfactory. I asked why they needed CMOS and was assured that "the low power requirement is nonnegotiable." Without further comment, I asked them to send me their breadboard. It arrived the next morning, and looked like Figure 23-1.

This is probably the most obvious way to design a V/F converter. The 9 V battery is regulated to 5 V by IC1 and a –5 V rail is derived by IC2. The input voltage causes current flow into A1's summing point. A1 responds by integrating negative, as shown in Figure 23-2, trace A. When A1's output goes low enough, A2 trips high (see trace B in Figure 23-2), turning on the CD4066 switch and resetting the integrator. Local positive feedback around A2 (A2's positive input is trace C) "hangs up" the reset, ensuring a complete integrator discharge. When the positive feedback decays, A1 begins to ramp again. The ramp slope, and hence the repetition frequency, depends upon the input voltage-dependent current into A1's summing point.

As soon as I saw the schematic, I knew I couldn't salvage any portion of this design. A serious drawback to this approach is A1's integrator reset time. This time, "lost" in the integration, results in significant linearity error as the operating frequency approaches it. The circuit's 6 µsec reset (see Figure 23-2, traces A and B) interval introduces a 0.6% error at 1 kHz, rising to 6% at 10 kHz. Also, variations in the reset time contribute additional errors. I added the 3 M resistor (shown in dashed lines) in a half-hearted attempt to improve these figures. This resistor causes A2's trip point to vary slightly with input, partially compensating for the integrator's "lost" reset time. This Band-Aid did improve linearity by more than an order of magnitude, to about 0.4%, but it ain't the way to go.

There are other problems. Quiescent current consumption of this entirely CMOS circuit is 190 µA, rising to a monstrous 700 µA at 10 kHz. Additionally, the polystyrene capacitor's drift alone is –120 ppm/°C, eating up the entire budget. The 1.2

A = 0.5 V/Div.
B = 10 V/Div.
C = 10 V/Div.
Horiz. = 10 μsec/Div.

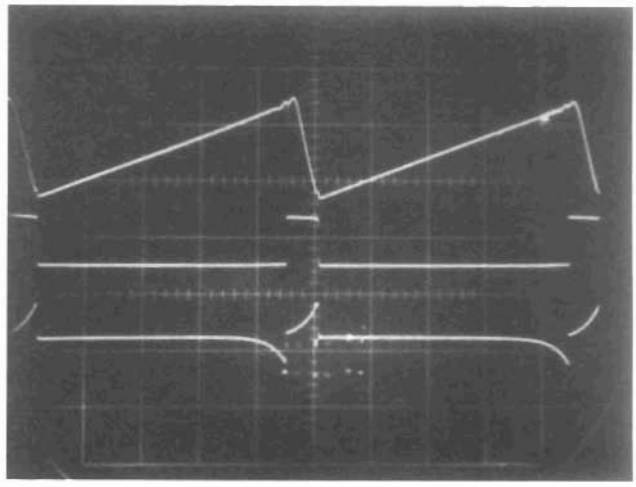


Figure 23-2. Wave forms for Figure 23-1's circuit. Finite reset time prevents good linearity performance.

V reference and the input resistor-trimmer could easily double this figure. There are a host of other problems, but what is really needed is an approach with inherently better linearity and lower power consumption.

There are many ways to convert a voltage to a frequency. The "best" approach in an application varies with desired precision, speed, response time, dynamic range, and other considerations.

Figure 23-3's concept potentially achieves high linearity by enclosing Figure 23-1's integrator in a charge-dispensing loop.

In this approach, C2 charges to $-V_{ref}$ during the integrator's ramping time. When the comparator trips, C2 is discharged into A1's summing point, forcing its output high. After C2's discharge, A1 begins to ramp and the cycle repeats. Because the loop acts to force the average summing currents to zero, the integrator time constant and reset time do not affect frequency. Gain drift terms are V_{ref}, C2, and the input resistor. This approach yields high linearity (typically 0.01%) into the megahertz range.

Figure 23-4 is conceptually similar, except that it uses feedback current instead of charge to maintain the op amp's summing point. Each time the op amp's output trips the comparator, the current sink pulls current from the summing point. Current is pulled from the summing point for the timing reference's duration, forcing the integrator positive. At the end of the current sink's period, the integrators output again heads negative. The frequency of this action is input related.

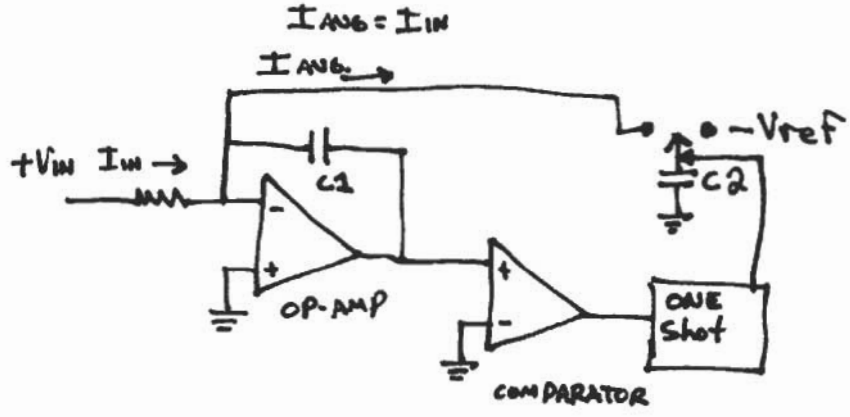


Figure 23-3. Conceptual charge-dispensing type voltage-to-frequency converter.

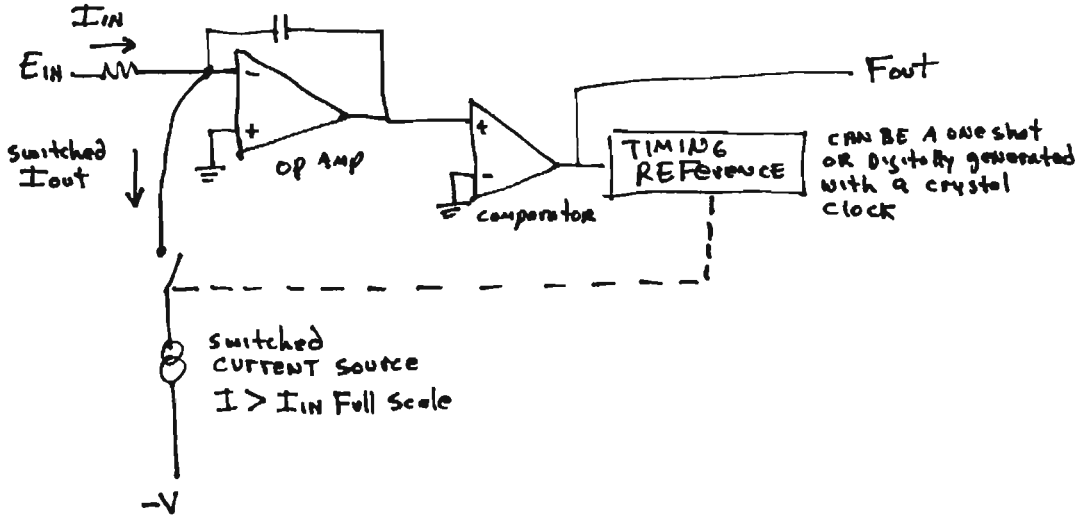


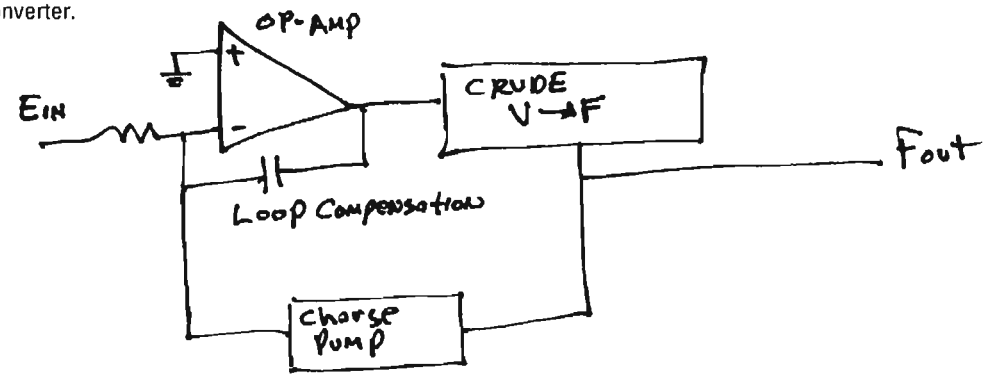
Figure 23-4. Current balance voltage-to-frequency converter.

Figure 23-5 uses DC loop correction. This arrangement offers all the advantages of charge and current balancing except that response time is slower. Additionally, it can achieve exceptionally high linearity (0.001%), output speeds exceeding 100 MHz, and very wide dynamic range (160 dB). The DC amplifier controls a relatively crude V/F converter. This V/F converter is designed for high speed and wide dynamic range at the expense of linearity and thermal stability. The circuit's output switches a charge pump whose output, integrated to DC, is compared to the input voltage.

The DC amplifier forces the V/F converter operating frequency to be a direct function of input voltage. The DC amplifier's frequency compensation capacitor, required because of loop delays, limits response time. Figure 23-6 is similar, except that the charge pump is replaced by digital counters, a quartz time base, and a DAC. Although it is not immediately obvious, this circuit's resolution is not restricted by the DAC's quantizing limitations. The loop forces the DAC's LSB to oscillate around the ideal value. These oscillations are integrated to DC in the loop compensation capacitor. Hence, the circuit will track input shifts much smaller than a DAC LSB. Typically, a 12-bit DAC (4096 steps) will yield one part on 50,000 resolution. Circuit linearity, however, is set by the DAC's specification.

Figure 23-5. Loop-charge pump voltage-to-frequency converter.

If you examine these options, Figure 23-3 looks like the winner for the customer's application. The specifications call for step response inside 5 cycles of output fre-



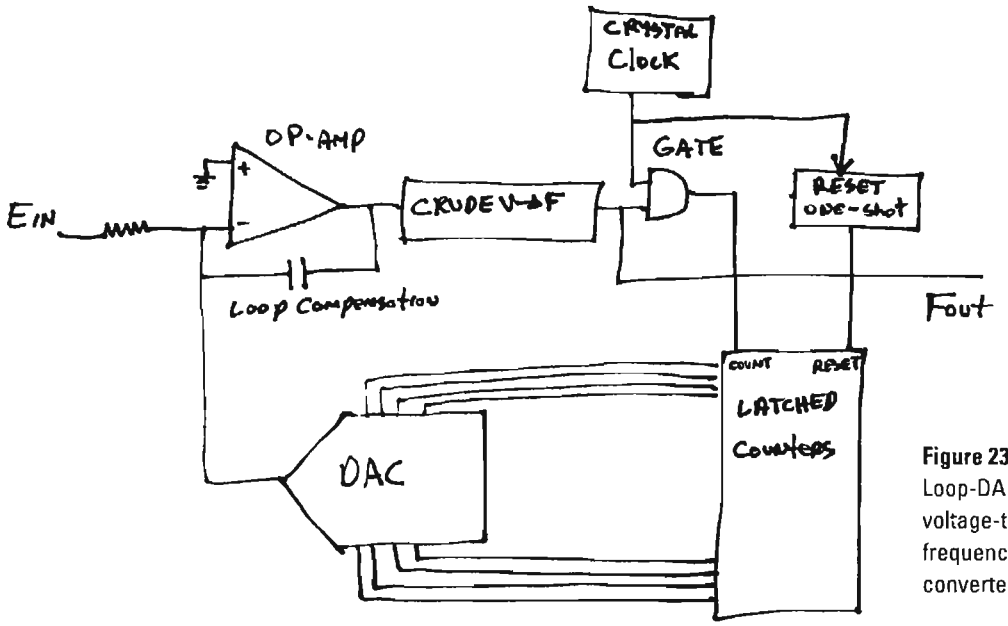


Figure 23-6. Loop-DAC voltage-to-frequency converter.

quency. This eliminates the circuits in Figures 23-4, 23-5, and 23-6 with their DC amplifiers' response time lag. Figure 23-4 requires a timing reference and a precision switched current source, implying some degree of complexity. In theory, Figure 23-3's approach can meet all the specifications without undue complexity.

This technique is not new. I first saw it back in 1964 in a copy of the *GE Transistor Manual*. T. P. Sylvan used a discrete op amp and a unijunction transistor to form the loop. Hewlett-Packard built rack-mounting V/F converters in the early 1960s which also relied on this approach. In 1972, R.A. Pease developed a commercially produced modular version (Teledyne-Philbrick Model 4701) using a single op amp which routinely achieved 0.01% linearity with commensurate drift performance. Pease's circuit is particularly relevant, and a version of it is shown in Figure 23-7.

Assume C1 sits at a small negative potential. A1's negative input is below its zero-biased positive input, and its output is high. The zener bridge clamps high (at $V_Z + V_{D4} + V_{D2}$) and C2 charges via D6, D7, and D8. The input voltage forces current through R1, and C1 begins to charge positively (trace A, Figure 23-8). When C1 crosses zero volts, A1's output (trace B) goes low and the zener bridge clamps negative, discharging C2 (C2's current is trace C) via the D5–C1 path. The resultant charge removal from C1 causes it to rapidly discharge (trace A). R2–C3 provides positive feedback to A1's positive input (trace D), reinforcing this action and hanging up A1's output long enough for a complete C2 discharge. When the R2–C3 feedback decays, A1's output returns high and the cycle repeats. The frequency of this sequence is directly proportional to the input voltage derived current through R1. Drift terms include R1, C2, and the zener, as well as residual diode mismatches. In theory, all the diode drops cancel and do not contribute toward drift. The R2–C3 "one shot" time constant is not critical, as long as it allows enough time for C2 to completely discharge. Similarly, "integrator" C1's value is unimportant as long as it averages A1's negative input to zero.

Q1 and associated components form a start-up loop. Circuit start-up or input overdrive can cause the circuit's AC-coupled feedback to latch. If this occurs, A1 goes negative and wants to stay there. R3 and C4 slowly charge negative, biasing

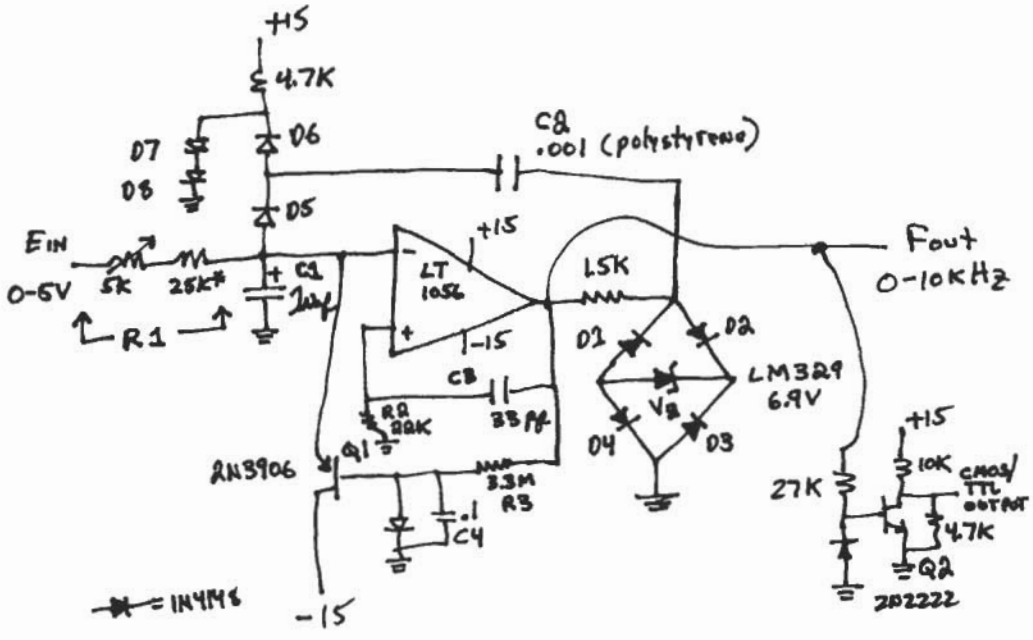


Figure 23-7. A version of Pease's elegant voltage-to-frequency converter circuit.

Q1. Q1 turns on, pulling C1 toward the –15 V rail, initiating normal circuit action. Once the circuit starts, C4 assumes a small positive potential and Q1 goes off. Q2, a simple level shifter, furnishes a logic-compatible output.

Pease's 1972 circuit is a very elegant, practical incarnation of Figure 23-3. With care, it will meet all the customer's requirements except two. It requires a split ±15 V supply, and pulls well over 10 mA. The job now boils down to dealing with these issues.

Figure 23-9 shows my first attempt at adapting Pease's circuit to my customer's needs. Operation is similar to Pease's circuit. When the input current-derived ramp (trace A, Figure 23-10) at C1A's negative input crosses zero, C1A's output (trace B) drops low, pulling charge through C1. This forces the negative input below zero. C2 provides positive feedback (trace D is the positive input), allowing a complete discharge for C1 (C1 current is trace C). When C2 decays, C1A's output goes high, clamping at the level set by D1, D2, and V_{ref}. C1 receives charge, and recycling

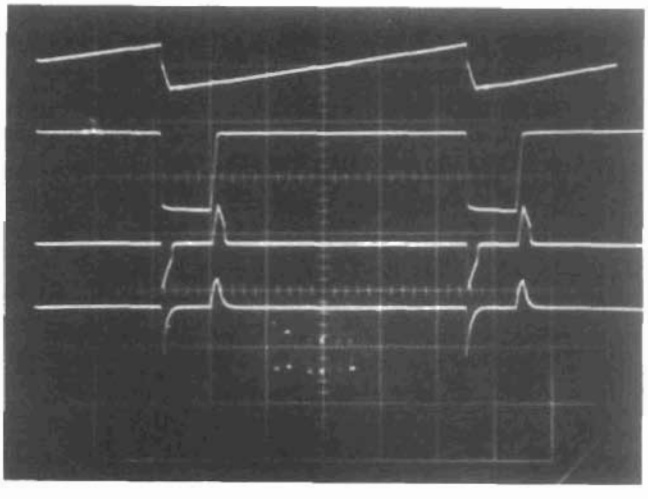


Figure 23-8. Wave forms for the Pease-type voltage-to-frequency converter.

A = 0.02 V/Div.
B = 20 V/Div.
C = 20 mA/Div.
D = 20 V/Div.
Horiz. = 20 μsec/Div.

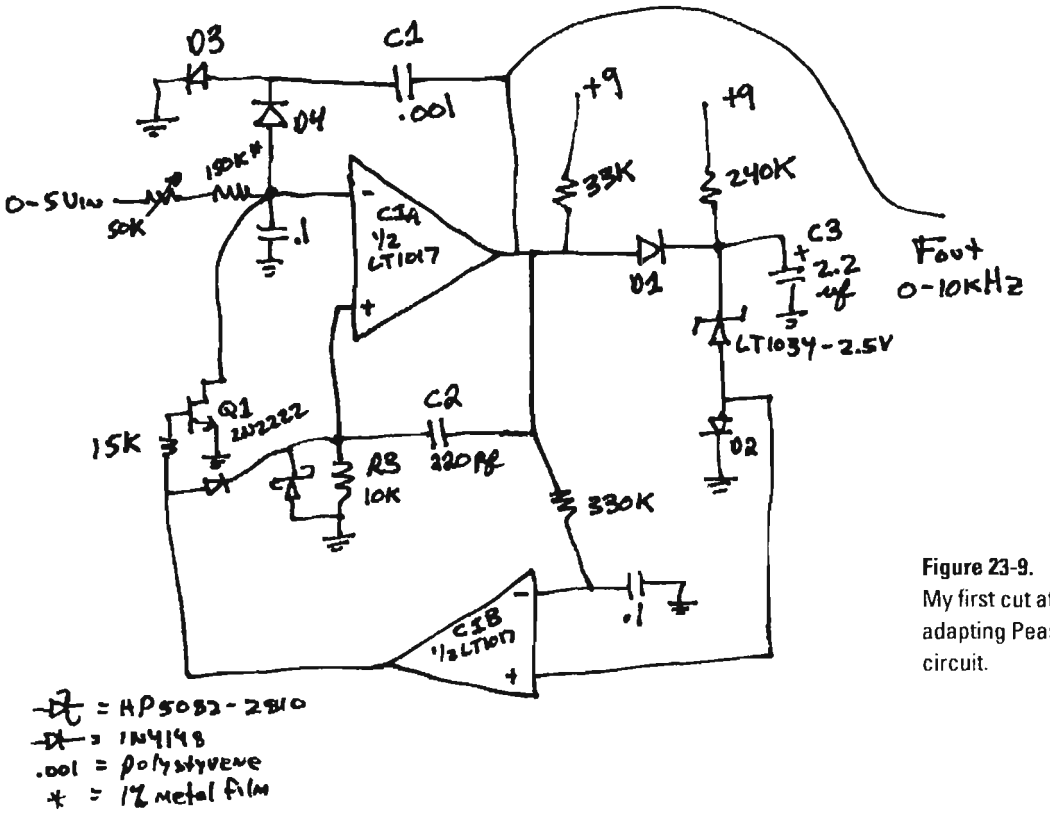


Figure 23-9. My first cut at adapting Pease's circuit.

occurs when C1A's negative input again arrives at zero. The frequency of this action is related to the input voltage. Diodes D3 and D4 provide steering and are temperature compensated by D1 and D2. C1A's sink saturation voltage is uncompensated but small. (These temperature coefficient assumptions are first order and will require more care later.) Although the LT1017 and LT1034 have low operating currents, this circuit pulls almost 400 µA. The AC current paths include C1's charge-discharge cycle, and C2's branch. The DC path through D2 and V_{ref} is particularly costly. C1's charging must occur quickly enough for 10 kHz operation, meaning the clamp seen by C1A's output must have low impedance at this frequency. C3 helps, but significant current still must come from somewhere to keep impedance low. C1A's current-limited output (≈30 µA source) cannot do the job unaided, and the resistor from the supply is required. Even if C1A could supply the necessary current, V_{ref}'s settling time would be an issue. Dropping C1's value will reduce impedance requirements proportionally and would seem to solve the problem. Unfortunately, such reduction magnifies the effects of stray capacitance at the D3–D4 junction. It also mandates increasing R_{in}'s value to keep scale factor constant. This lowers operating currents at C1A's negative input, making bias current and offset more significant error sources.

C1B, Q1, and associated components form a start-up loop which operates in similar fashion to the one in Pease's circuit (Figure 23-7).

Figure 23-11 shows an initial attempt at dealing with these issues. This scheme is similar to Figure 23-9, except that Q1 and Q2 appear. V_{ref} receives switched bias via Q1, instead of being on all the time. Q2 provides the sink path for C1. These transistors invert C1A's output, so its input pin assignments are exchanged. R1 provides a light current from the supply, improving reference settling time. This

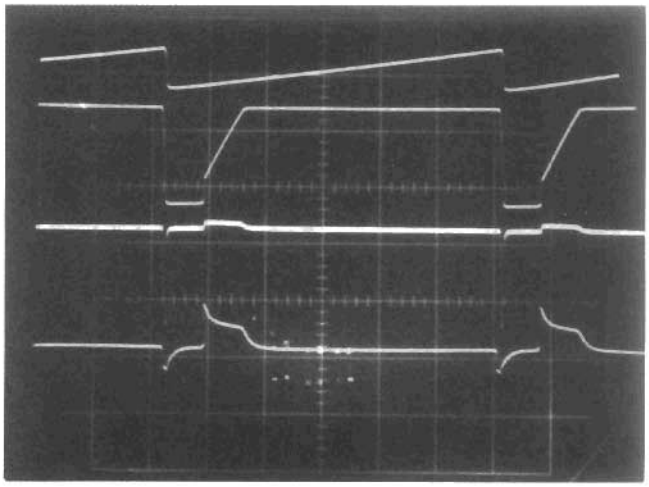


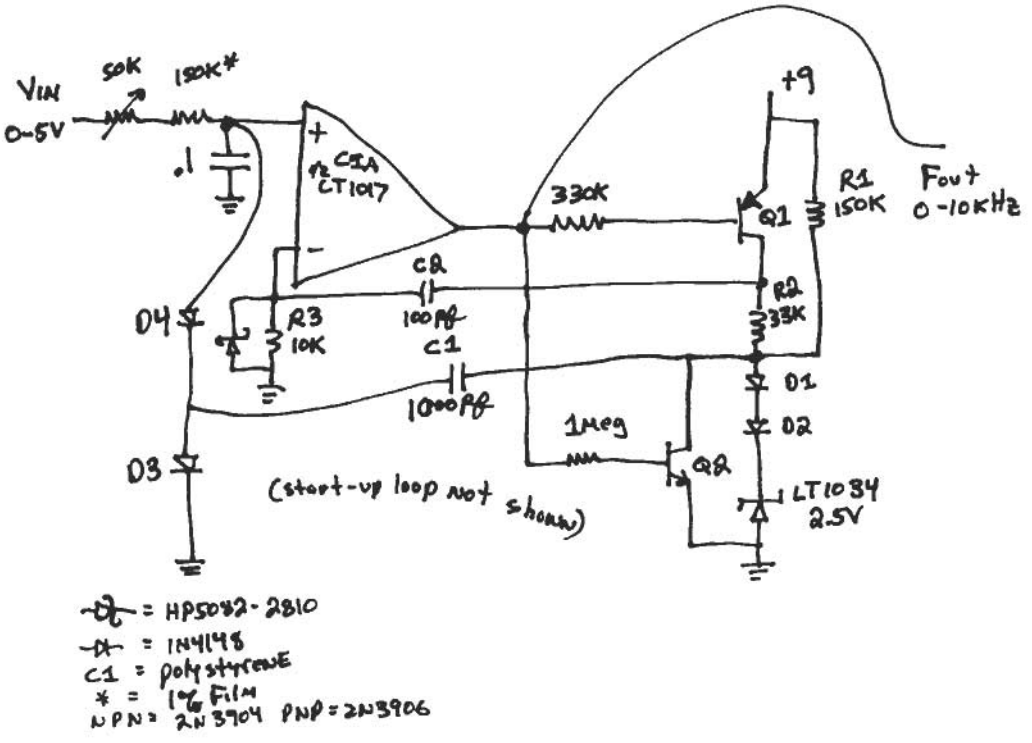
Figure 23-10. Wave forms for the circuit in Figure 23-9.

A = 50 mV/Div.
B = 2 V/Div.
C = 2 mA/DIv.
D = 1 V/Div.
Horiz. = 20 μsec/Div.

arrangement decreases supply current to about 300 μA, a significant improvement. Several problems do exist, however. Q1's switched operation is really effective only at higher frequencies. In the lower ranges, C1A's output is low most of the time, biasing Q1 on and wasting power. Additionally, when C1A's output switches, Q1 and Q2 simultaneously conduct during the transition, effectively shunting R2 across the supply. Finally, the base currents of both transistors flow to ground and are lost. Figure 23-12 shows the wave form traces for this circuit. The basic temperature compensation is as before, except that Q2's saturation term replaces the comparator's. This temperature compensation scheme looks okay, but we're still hand waving.

Figure 23-11. The second try. Q1 and Q2 switch the reference, saving some power.

Figure 23-13 is better. Q1 is gone, Q2 remains, but Q3, Q4, and Q5 have been



A = 5 V/Div.
B = 5 V/Div.
C = 2 V/Div.
D = 100 µA/Div.
Horiz. = 10 µsec/Div.

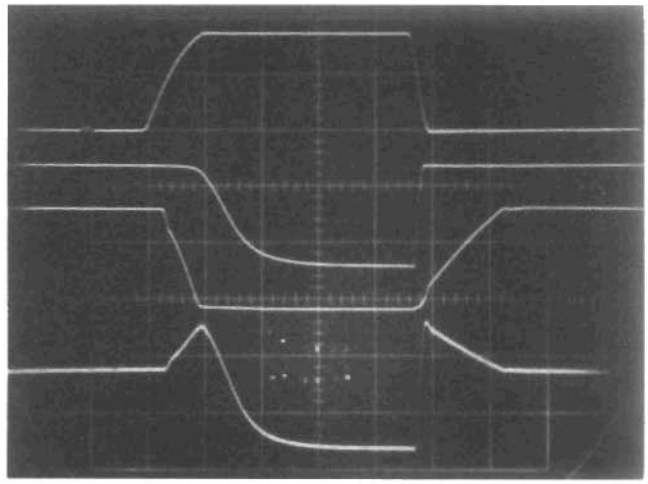


Figure 23-12. Figure 11's wave forms. Traces A, B, C, and D are C1A output, Q1 collector, Q2 collector, and R2 current, respectively. Q1–Q2 simultaneous conduction problem is evident in trace D.

added. V_{ref} and its associated diodes are biased from R1. Q3, an emitter-follower, is used to source current to C1. Q4 temperature compensates Q3's V_{be}, and Q5 switches Q3.

This method has some distinct advantages. The V_{ref} string can operate at greatly reduced current because of Q3's current gain. Also, Figure 23-11's simultaneous conduction problem is largely alleviated because Q5 and Q2 are switched at the same voltage threshold out of C1A. Q3's base and emitter currents are delivered to C1. Q5's currents are wasted, although they are much smaller than Q3's. Q2's small base current is also lost. The values for C2 and R3 have been changed. The time constant is the same, but some current reduction occurs due to R3's increase.

Operating wave forms are shown in Figure 23-14, and include C1's output (trace

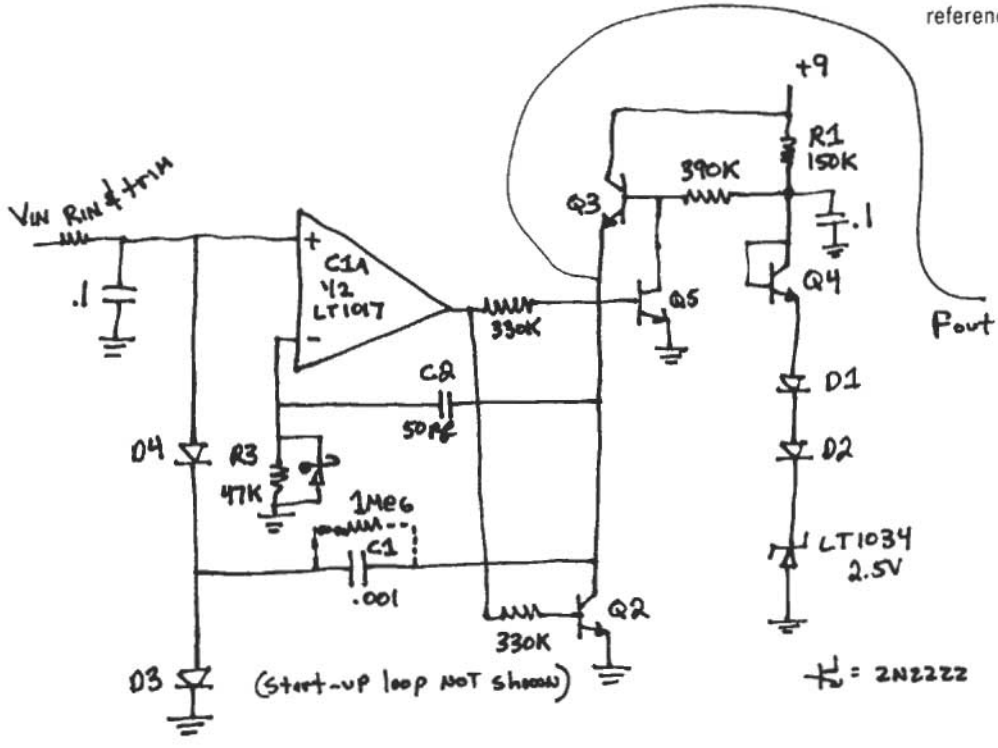


Figure 23-13. A better scheme for switching the reference.

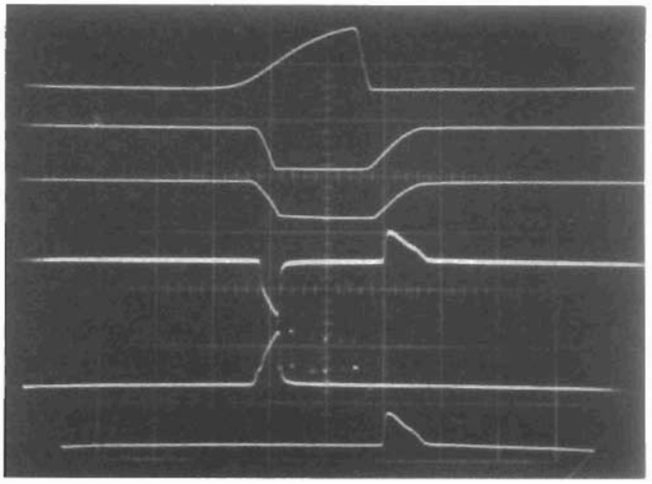


Figure 23-14. Figure 23-13's operation. Traces D, E, and F reveal no simultaneous conduction problems.

A), Q5's collector (trace B), Q2's collector (trace C), Q2's collector current (trace D), C1's current (trace E), and Q3's emitter current (trace F). Note that the current steering is clean, with no simultaneous conduction problems.

This circuit's 200 µA power consumption was low enough to make other specifications worth checking. Linearity came in at 0.05%, and dropped to 0.02% when I added a 1 M resistor (dashed lines) across C1. The D4–Q2 path cannot *fully* switch C1 because of junction drop limitations. The resistor squeezes the last little bit of charge out of C1, completing the discharge and improving linearity.

Power supply rejection ratio (PSRR) was not good enough. Supply shifts show up as current changes through R1. The LT1034 is relatively insensitive to this, but the Q4, D1, D2 trio shift value. As such, I measured 0.1%/V PSRR. R1 really needs to be a current source, or some compensation mechanism must be used.

Temperature compensation was next. Now it was time to stop hand waving and take a hard look. Q4 supposedly compensates Q3, with D1 and D2 opposing D3 and D4. Unfortunately, these devices operate under different dynamic and DC conditions, making precise cancellation difficult. In practice, R1's value should be established to source the current through Q4–D1–D2, which provides optimum circuit temperature coefficient. Assuming perfect cancellation, and no LT1034 or input resistor drift, we still must deal with Q2's V_{ce} saturation term. At 100 mV saturation, Q2 will drift about +0.3%/°C (see the Motorola 2N2222 data sheet), causing about a –300 µV/°C shift in the voltage C1 discharges toward. This works out to about –100 ppm/°C (C1 charges to 3 V) temperature coefficient, which will force a similar *positive* shift in output frequency. C1, a polystyrene type, drifts about –120 ppm/°C, contributing further overall positive temperature coefficient (as C1, or the voltage it charges to, gets smaller, the circuit must oscillate faster to keep the summing point at zero). So the best case is about 220 ppm/°C, and reality dictates that all the other junctions won't match precisely. Temperature testing confirmed all this. Initially, the breadboard showed about 275 ppm/°C, and, by varying R1, bottomed out at about 200 ppm/°C. This certainly wasn't production-worthy engineering but pointed the way toward a solution.

How could I reduce the temperature coefficient and fix the PSRR? Additionally, power consumption was still marginal, although linearity was close. Replacing R1 with a current source offered hope for PSRR, but reliable temperature compensation and lower power needed another approach. I pined for inspiration but got nothing. I was stuck.

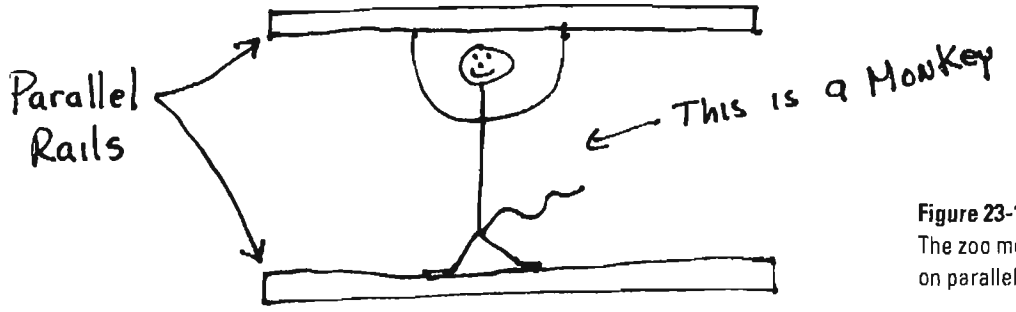


Figure 23-15. The zoo monkey on parallel rails.

Something that *had* inspired me for a couple of months was a physician I'd been seeing. We really had a good time together—a couple of playful kids. There was much dimension to this woman, and I really enjoyed just how relaxed I felt being with her. Things were going quite nicely, and I sometimes allowed myself the luxury of wondering what would become of us.

One weekday afternoon, we played hookey and went to the San Francisco Zoo. The weather was gorgeous, no crowds, and the Alfa ran great. (On our second date it threw a fan belt.) We saw bears, elephants, tigers, birds, and ate lots of junk food. The lions got fed; they were *loud* and *hungry*. Strolling around, eating cheeseburgers, and doing just fine, we came to the monkeys.

These guys are actors; they love an audience. There was the usual array of grinning, simian catcalls, cheeping, squawking, lots of jungle bar performances, wondrous feats of balance, and other such theatrics. One character particularly caught my eye. He did a little routine between two parallel rails. First, he hung by his hands as shown in figure 23-15.

Then, very quickly, he flipped over, simultaneously rotating, so he ended up inverted (see Figure 23-16).

He did this over and over at great speed; it was his act. Standing there, watching the little fellow do his inverting routine between the rails, I saw my circuit problems simply melt. I felt very lucky. I had a good lady, and a good circuit too.

If you look inside a CMOS logic inverter, the output stage looks like Figure 23-17.

The MOS output transistors connect the output terminal to the supply or ground rail. The input circuitry is arranged so only one transistor is on at a time; simultaneous conduction cannot occur. Typically, channel-on resistance is 100–200 Ω. There are no junction effects; the transistor channels are purely ohmic. The device's input pin appears almost purely capacitive, drawing only picoamperes of bias current.

Figure 23-18 shows what happens when the CMOS inverter is dropped into the gizzard of Figure 23-13's circuit. C1 is charged and discharged via the CMOS inverter's ohmic output transistors. Q3 now drives the inverter's supply pin, and Q2 goes away. Along with Q2's departure goes its 100 ppm/°C temperature coefficient

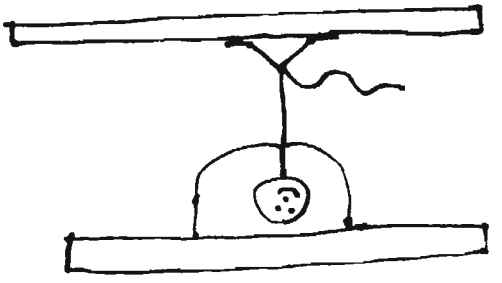


Figure 23-16. The zoo monkey on parallel rails, inverted.

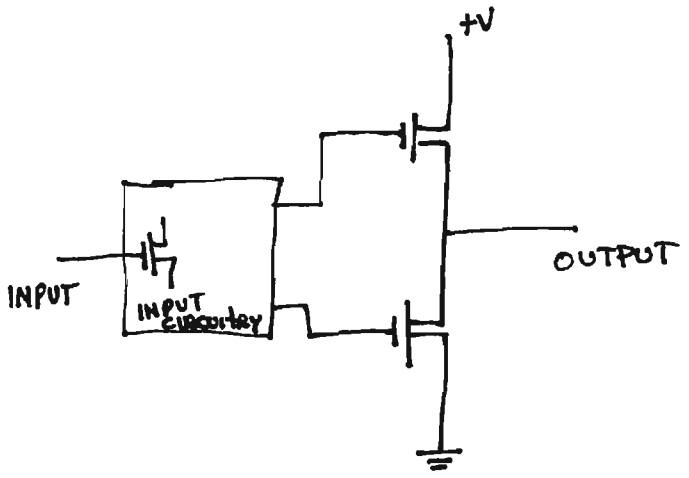


Figure 23-17. Conceptual CMOS inverter.

error. Also. Q2's base current is eliminated, along with Q5's base and collector current losses.

This scheme promises both lower temperature drift and lower power. Assuming ideal junction compensation, the remaining uncompensated drift terms are C1's −120 ppm temperature coefficient and the input resistor. Unfortunately, this configuration does nothing to fix the PSRR problem. The only realistic fix for that is to replace R1 with a current source. The current source doesn't have to be very stable but must run with only 2 V of headroom because the circuit has to work down to 6.5 V. The simplest alternative is the monolithic LM134. This three-terminal, resistor-programmable device will function with only 800 mV across it, although it does have a 0.33%/°C temperature coefficient. This temperature coefficient seemed small enough to avoid causing any trouble. The LT1034 shouldn't care, but what about D1, D2, and Q4? When I calculated the effect of current-source shift with temperature on these devices, I realized I had just inherited the world. It came out

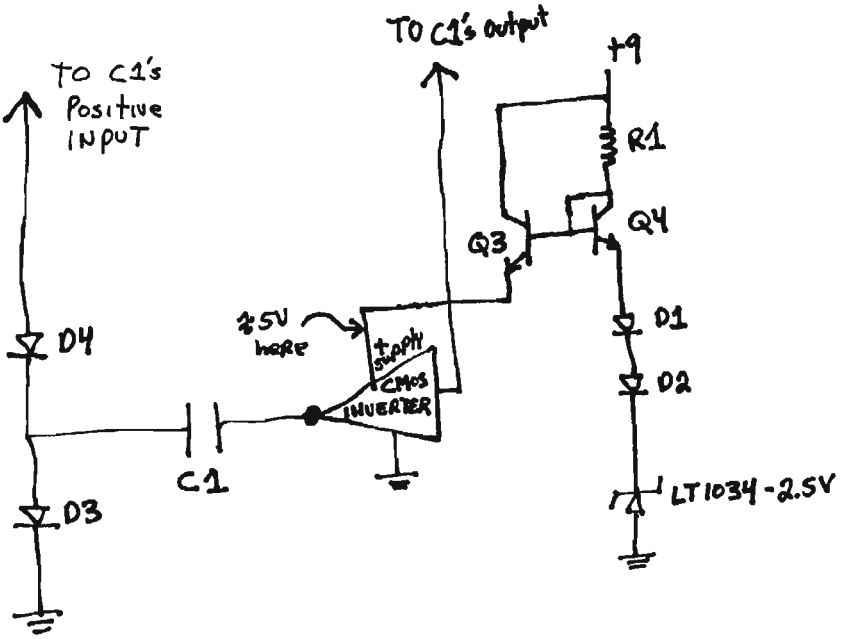


Figure 23-18. Adding the CMOS inverter to the circuit in Figure 23-13.

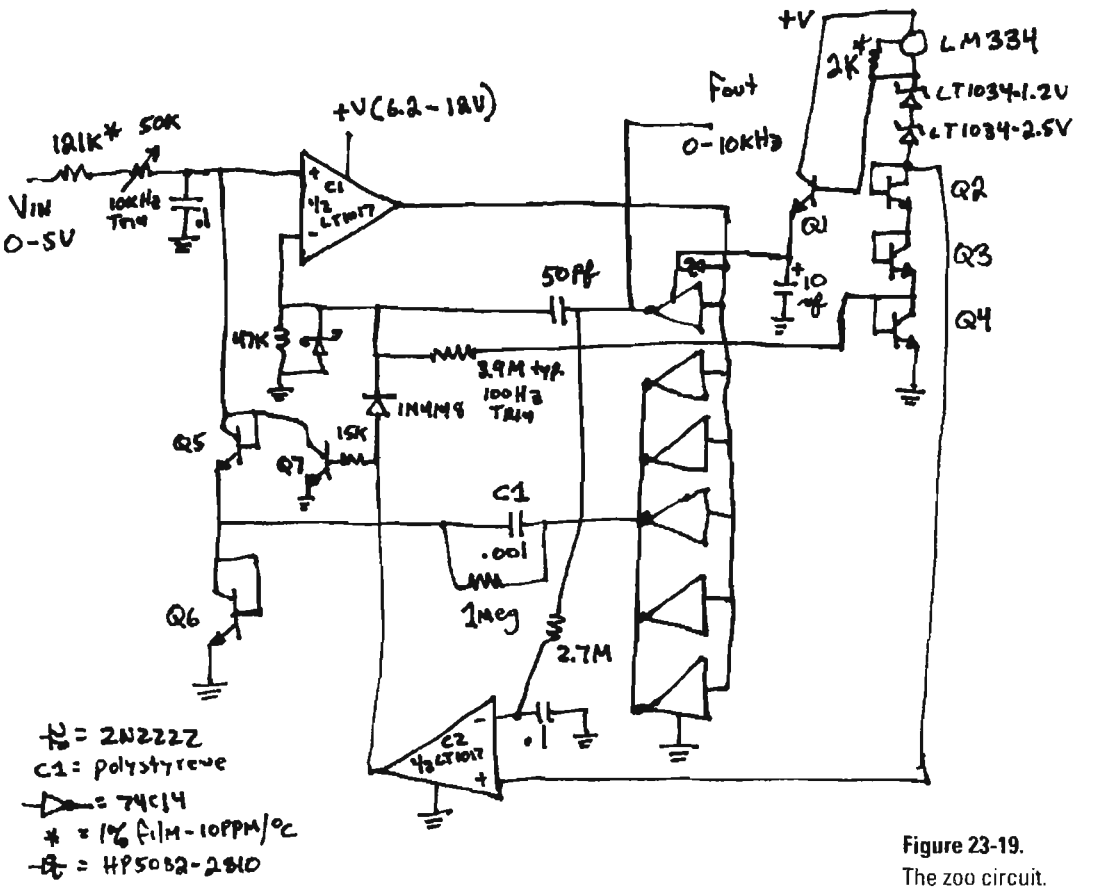


Figure 23-19.
The zoo circuit.

positive 180 ppm/°C! This tends to cancel the capacitor's –120 ppm/°C term. Additionally, increasing the LT1034's reference voltage by about 50% would pull the compensation down to +120 ppm/°C, further reducing drift. This also aids overall temperature coefficient by making the residual junction mismatches a smaller percentage of the total reference voltage. The current source's low headroom capability allows this, while maintaining operation down to $V_{supply} = 6.2$ V. The sole uncompensated term is the input resistor, which can be specified for low temperature drift.

Figure 23-19 is the final circuit. It meets or exceeds every customer specification.

A 0–5 V input produces a 0–10 kHz output, with a linearity of 0.02%. Gain drift is 40 ppm/°C, and PSRR is inside 40 ppm/V. Maximum current consumption is 145 µA, descending to 80 µA for $V_{in} = 0$. Other specifications appear in Table 2's summary. Much of this circuit should be, by now, familiar. Some changes have occurred, but nothing too drastic. The diodes have been replaced with transistors for lower leakage and more consistant matching. Also, paralleling the CMOS inverters provides lower resistance switching. The start-up loop has also been modified.

To maintain perspective, it's useful to review circuit operation. Assume C1's positive input is slightly below its negative input (C2's output is low). The input voltage causes a positive-going ramp at C1's positive input (trace A, Figure 23-20). C1's output is low, biasing the CMOS inverter outputs high. This allows current to flow from Q1's emitter, through the inverter supply pin to the 0.001 µF capacitor. The 10 µF capacitor provides high-frequency bypass, maintaining a low impedance

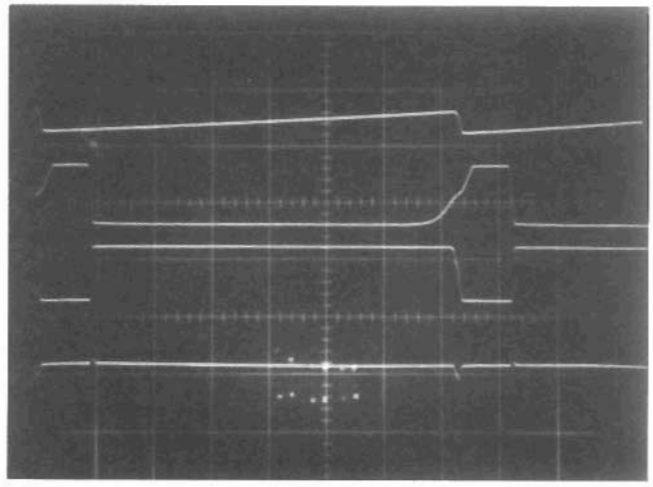


Figure 23-20. Figure 23-19's wave forms.

at Q1's emitter. Diode connected Q6 provides a path to ground. The voltage that the 0.001 µF unit charges to is a function of Q1's emitter potential and Q6's drop. When the ramp at C1's positive input goes high enough, C1's output goes high (trace B) and the inverters switch low (trace C). The Schottky clamp prevents CMOS inverter input overdrive. This action pulls current from C1's positive input capacitor via the Q5–0.001 µF route (trace D). This current removal resets C1's positive input ramp to a potential slightly below ground, forcing C1's output to go low. The 50 pF capacitor connected to the circuit output furnishes AC positive feedback, ensuring that C1's output remains positive long enough for a complete discharge of the 0.001 µF capacitor. As in Figure 23-13, the 1 MΩ resistor completes C1's discharge.

The Schottky diode prevents C1's input from being driven outside its negative common-mode limit. When the 50 pF unit's feedback decays, C1 again switches low and the entire cycle repeats. The oscillation frequency depends directly on the input voltage–derived current.

Q1's emitter voltage must be carefully controlled to get low drift. Q3 and Q4 temperature compensate Q5 and Q6 while Q2 compensates Q1's V_{be}. The two LT1034s are the actual voltage reference and the LM334 current source provides excellent supply immunity (better than 40 ppm/V PSRR) and also aids circuit temperature coefficient. It does this by utilizing the LM334's 0.3%/°C temperature coefficient to slightly temperature modulate the voltage drop in the Q2–Q4 trio. This correction's sign and magnitude directly oppose that of the –120 ppm/°C 0.001 µF polystyrene capacitor, aiding overall circuit stability.

The Q1 emitter-follower delivers charge to the 0.001 µF capacitor efficiently. Both base and collector current end up in the capacitor. The paralleled CMOS inverters provide low loss SPDT reference switching without significant drive losses. Additionally, the inverter specified is a Schmitt input type, minimizing power loss due to C1's relatively slow rising edges. The 0.001 µF capacitor, as small as accuracy permits, draws only small transient currents during its charge and discharge cycles. The 50 pF–47 K positive feedback combination draws insignificantly small switching currents. Figure 23-21, a plot of supply current versus operating frequency, reflects the low power design. At zero frequency, the LT1017's quiescent current and the 35 µA reference stack bias accounts for all current drain. There are no other paths for loss. As frequency scales up, the charge–discharge cycle of the 0.001 µF capacitor introduces the 7 µA/kHz increase shown. A smaller value

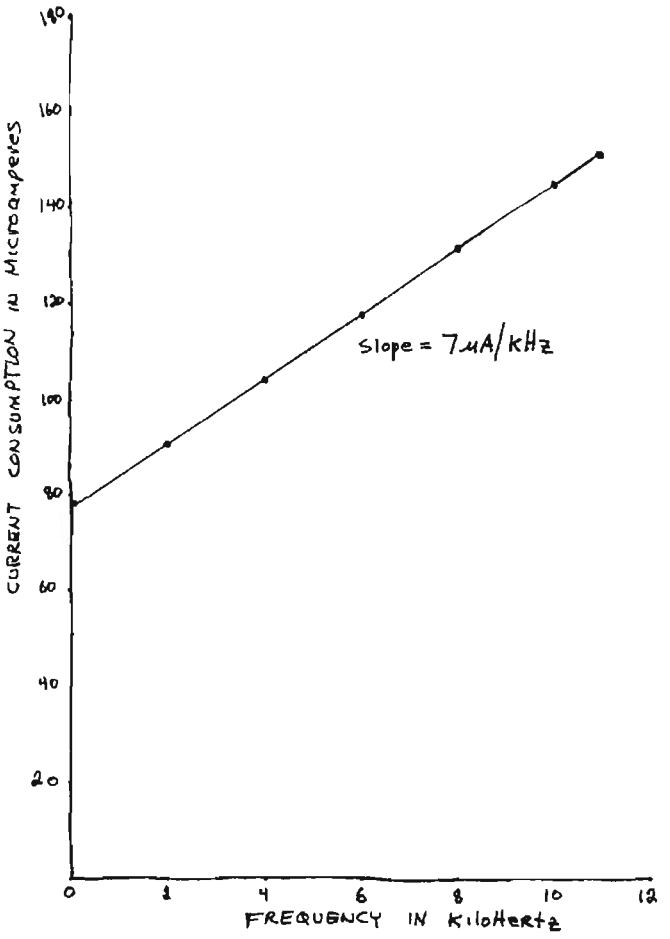


Figure 23-21. Current consumption versus frequency for Figure 23-19.

capacitor would cut power, but the effects of stray capacitance, charge imbalance in the 74C14, and LT1017 bias currents would introduce accuracy errors. For example, if C1 is reduced to 100 pf (along with other appropriate changes), the circuit consumes only 90 μA at 10 kHz, but linearity degrades to .05%.

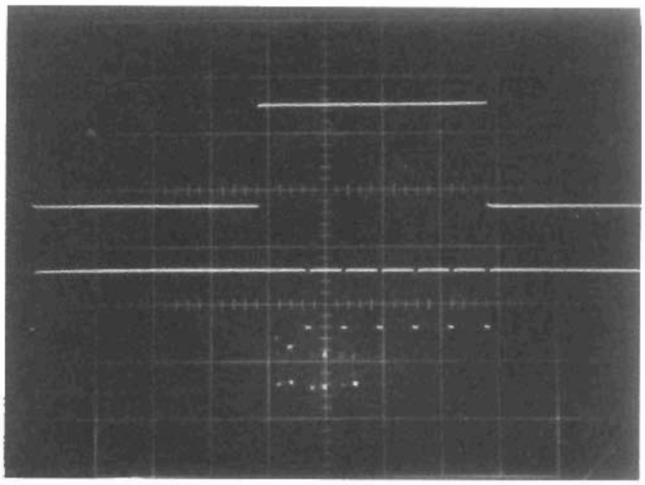
Circuit start-up or overdrive can cause the circuit's AC-coupled feedback to latch. If this occurs, C1's output goes high. C2, detecting this via the inverters and the 2.7 M–0.1 μF lag, also goes high. This lifts C1's negative input and grounds the positive input with Q7, initiating normal circuit action.

Because the charge pump is directly coupled to C1's output, response is fast. Figure 23-22 shows the output (trace B) settling within one cycle for a fast input step (trace A).

To calibrate this circuit, apply 50 mV and select the value at C1's input for a 100 Hz output. Then, apply 5 V and trim the input potentiometer for a 10 kHz output.

Here's what the customer ended up getting:

Summary: Voltage-to-Frequency Converter	
Output frequency	0–10 kHz
Input voltage	0–5 V
Linearity	0.02%
Drift	40 ppm/°C



A = 2V/DIV
B = 5V/DIV
Horiz. = 200 μsec/DIV

Figure 23-22. Figure 23-18's step response.

PSRR	40 ppm/V
Temperature range	0–70° C
Step response	1 cycle of output frequency
Output pulse	5 V CMOS-compatible
Power supply	Single 9 V battery (6.2–12 V)
Power consumption	145 μA maximum, 80 μA quiescent
Cost	< $6.00/100,000 pieces

The zoo circuit made my customer happy, even if it is almost entirely bipolar. The inverter is the only piece of CMOS in the thing. I'm fairly certain the customer wouldn't mind if I had used 12AX7s[1] as long as it met specifications. It runs well in production, and they make lots of them, which makes my boss and the stockholders happy.

This circuit has received some amount of attention in the technical community. I am aware of some spectacularly complex mathematical descriptions of it, along with some arcane explanations of its behavior. Similarly, it has been shown that the circuit could have only been arrived at with the aid of a computer. Given this undue credit, the least I could do is come clean about the circuit's humble origins.

I hope it was as much fun to read about the circuit as it was to build it.

References

1. "Voltage to Frequency Converter," *General Electric Transistor Manual*, page 346, General Electric Company, Syracuse, New York, 1964.
2. R.A. Pease, "A New Ultra-Linear Voltage-to-Frequency Converter," 1973 *NEREM Record*, Vol. I, page 167.
3. R.A. Pease, assignee to Teledyne, "Amplitude to Frequency Converter," U.S. patent 3,746,968, filed September, 1972.
4. J. Williams, "Micropower Circuits for Signal Conditioning," 10 kHz Voltage-to-Frequency Converter, pp. 10-13, Linear Technology Corporation, Application Note 23, 1987.
5. J. Williams, "Designs for High Performance Voltage-to-Frequency Converters," Linear Technology Corporation, Application Note 14, 1986.

1. For those tender of years, 12AX7s are thermionically activated FETs, descended from Lee DeForest.

1. The Importance of Fixing

Fall 1968 found me at MIT preparing courses, negotiating thesis topics with students, and getting my laboratory together. This was fairly unremarkable behavior for this locale, but for a 20 year old college dropout the circumstances were charged; the one chance at any sort of career. For reasons I'll never understand, my education, from kindergarten to college, had been a nightmare, perhaps the greatest impedance mismatch in history. I got hot. The Detroit Board of Education didn't. Leaving Wayne State University after a dismal year and a half seemed to close the casket on my circuit design dreams.

All this history conspired to give me an outlook blended of terror and excitement. But mostly terror. Here I was, back in school, but on the other side of the lectern. Worse yet, my research project, while of my own choosing, seemed open ended and unattainable. I was so scared I couldn't breathe out. The capper was my social situation. I was younger than some of my students, and my colleagues were at least 10 years past me. To call things awkward is the gentlest of verbiage.

The architect of this odd brew of affairs was Jerrold R. Zacharias, eminent physicist, Manhattan Project and Radiation Lab alumnus, and father of atomic time. It was Jerrold who waved a magic wand and got me an MIT appointment, and Jerrold who handed me carte blanche a lab and operating money. It was also Jerrold who made it quite clear that he expected results. Jerrold was not the sort to tolerate looking foolish, and to fail him promised a far worse fate than dropping out of school.

Against this background I received my laboratory budget request back from review. The utter, untrammeled freedom he permitted me was maintained. There were no quibbles. Everything I requested, even very costly items, was approved, without comment or question. The sole deviation from this I found annoying. He threw out my allocation for instrument repair and calibration. His hand written comment: "You fix everything."

It didn't make sense. Here I was, under pressure for results, scared to pieces, and I was supposed to waste time screwing around fixing lab equipment? I went to see Jerrold. I asked. I negotiated. I pleaded, I ranted, and I lost. The last thing I heard chasing me out of his office was, "You fix everything."

I couldn't know it, but this was my introduction to the next ten years. An unruly mix of airy freedom and tough intellectual discipline that

would seemingly be unremittingly pounded into me. No apprenticeship was ever more necessary, better delivered, or, years later, as appreciated.

I cooled off, and the issue seemed irrelevant, because nothing broke for a while. The first thing to finally die was a high sensitivity, differential 'scope plug-in, a Tektronix 1A7. Life would never be the same.

The problem wasn't particularly difficult to find once I took the time to understand how the thing worked. The manual's level of detail and writing tone were notable; communication was *the* priority. This seemed a significant variance from academic publications, and I was impressed. The instrument more than justified the manual's efforts. It was gorgeous. The integration of mechanicals, layout, and electronics was like nothing I had ever seen. Hours after the thing was fixed I continued to probe and puzzle through its subtleties. A common mode bootstrap scheme was particularly interesting; it had direct applicability to my lab work. Similarly, I resolved to wholesale steal the techniques used for reducing input current and noise.

Over the next month I found myself continually drifting away from my research project, taking apart test equipment to see how it worked. This was interesting in itself, but what I really wanted was to test my

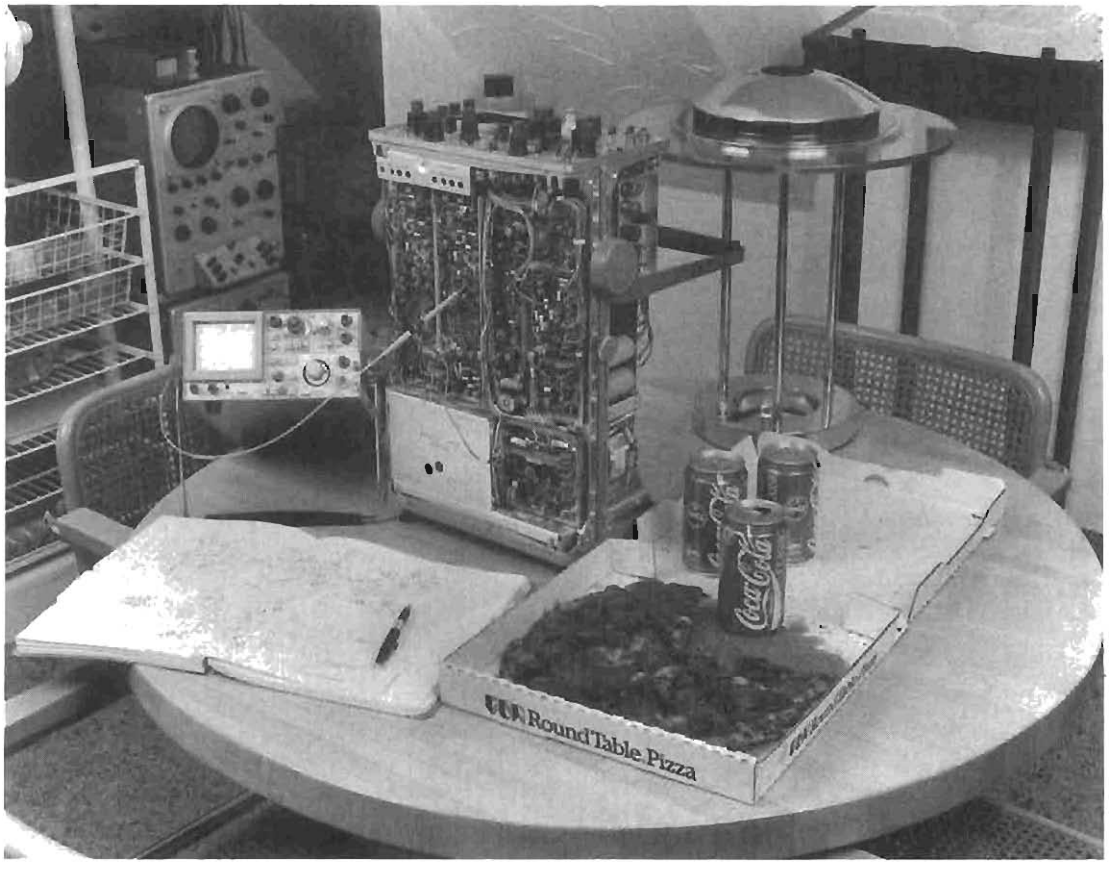


Figure 1–1. Oh boy, it's broken! Life doesn't get any better than this.

understanding by having to fix it. Unfortunately, Tektronix, Hewlett-Packard, Fluke, and the rest of that ilk had done their work well; the stuff didn't break. I offered free repair services to other labs who would bring me instruments to fix. Not too many takers. People had repair budgets . . . and were unwilling to risk their equipment to my unproven care. Finally, in desperation, I paid people (in standard MIT currency—Coke and pizza) to deliberately disable my test equipment so I could fix it. Now, their only possible risk was indigestion. This offer worked well.

A few of my students became similarly hooked and we engaged in all forms of contesting. After a while the "breakers" developed an armada of incredibly arcane diseases to visit on the instruments. The "fixers" countered with ever more sophisticated analysis capabilities. Various games took points off for every test connection made to an instrument's innards, the emphasis being on how close you could get utilizing panel controls and connectors. Fixing without a schematic was highly regarded, and a consummately macho test of analytical skill and circuit sense. Still other versions rewarded pure speed of repair, irrespective of method.[1] It really was great fun. It was also highly efficient, serious education.

The inside of a broken, but well-designed piece of test equipment is an extraordinarily effective classroom. The age or purpose of the instrument is a minor concern. Its instructive value derives from several perspectives.

It is always worthwhile to look at how the designer(s) dealt with problems, utilizing available technology, and within the constraints of cost, size, power, and other realities. Whether the instrument is three months or thirty years old has no bearing on the quality of the thinking that went into it. Good design is independent of technology and basically timeless. The clever, elegant, and often interdisciplinary approaches found in many instruments are eye-opening, and frequently directly applicable to your own design work. More importantly, they force self-examination, hopefully preventing rote approaches to problem solving, with their attendant mediocre results. The specific circuit tricks you see are certainly adaptable and useful, but not nearly as valuable as studying the thought process that produced them.

The fact that the instrument is broken provides a unique opportunity. A broken instrument (or anything else) is a capsulized mystery, a puzzle with a definite and very singular "right" answer. The one true reason why that instrument doesn't work as it was intended to is really there. You are forced to measure your performance against an absolute, non-negotiable standard; the thing either works or it doesn't when you're finished.

---

1. A more recent development is "phone fixing." This team exercise, derived by Len Sherman (the most adept fixer I know) and the author, places a telephone-equipped person at the bench with the broken instrument. The partner, somewhere else, has the schematic and a telephone. The two work together to make the fix. A surprise is that the time-to-fix seems to be less than if both parties are physically together. This may be due to dilution of ego factors. Both partners simply must speak and listen with exquisite care to get the thing fixed.

The reason all this is so valuable is that it brutally tests your thinking process. Fast judgments, glitzy explanations, and specious, hand-waving arguments cannot be costumed as "creative" activity or true understanding of the problem. After each ego-inspired lunge or jumped conclusion, you confront the uncompromising reality that the damn thing still doesn't work. The utter closedness of the intellectual system prevents you from fooling yourself. When it's finally over, and the box works, and you know why, then the real work begins. You get to try and fix you. The bad conclusions, poor technique, failed explanations, and crummy arguments all demand review. It's an embarrassing process, but quite valuable. You learn to dance with problems, instead of trying to mug them.

It's scary to wonder how much of this sort of sloppy thinking slips into your own design work. In that arena, the system is not closed. There is no arbitrarily right answer, only choices. Things can work, but not as well as they might if your thinking had been better. In the worst case, things work, but for different reasons than you think. That's a disaster, and more common than might be supposed. For me, the most dangerous point in a design comes when it "works." This ostensibly "proves" that my thinking is correct, which is certainly not necessarily true. The luxury the broken instrument's closed intellectual system provides is no longer available. In design work, results are open to interpretation and explanation and that's a very dangerous time. When a design "works" is a very delicate stage; you are psychologically ready for the kill and less inclined to continue testing your results and thinking. That's a precarious place to be, and you have to be so careful not to get into trouble. The very humanness that drives you to solve the problem can betray you near the finish line.

What all this means is that fixing things is excellent exercise for doing design work. A sort of bicycle with training wheels that prevent you from getting into too much trouble. In design work you have to mix a willingness to try anything with what you hope is critical thinking. This seemingly immiscible combination can lead you to a lot of nowheres. The broken instrument's narrow, insistent test of your thinking isn't there, and you can get in a lot deeper before you realize you blew it. The embarrassing lessons you're forced to learn when fixing instruments hopefully prevent this. This is the major reason I've been addicted to fixing since 1968. I'm fairly sure it was also Jerrold's reason for bouncing my instrument repair allocation.

There are, of course, less lofty adjunct benefits to fixing. You can often buy broken equipment at absurdly low cost. I once paid ten bucks for a dead Tektronix 454A 150MHz portable oscilloscope. It had clearly been systematically sabotaged by some weekend-bound calibration technician and tagged "Beyond Repair." This machine required thirty hours to uncover the various nasty tricks played in its bowels to ensure that it was scrapped.

This kind of devotion highlights another, secondary benefit of fixing. There is a certain satisfaction, a kind of service to a moral imperative,

that comes from restoring a high-quality instrument. This is unquestionably a gooey, hand-over-the-heart judgment, and I confess a long-term love affair with instrumentation. It just seems sacrilege to let a good piece of equipment die. Finally, fixing is simply a lot of fun. I may be the only person at an electronics flea market who will pay more for the busted stuff!

11. Tripping the Light Fantastic

Introduction

Where do good circuits come from, and what is a good circuit? Do they only arrive as lightning bolts in the minds of a privileged few? Are they synthesized, or derived after careful analysis? Do they simply evolve? What is the role of skill? Of experience? Of luck? I can't answer these weighty questions, but I do know how the best circuit I ever designed came to be.

What is a good circuit, anyway? Again, that's a fairly difficult question, but I can suggest a few guidelines. Its appearance should be fundamentally simple, although it may embody complex and powerful theoretical elements and interactions. That, to me, is the essence of elegance. The circuit should also be widely utilized. An important measure of a circuit's value is if lots of people use it, and are satisfied after they have done so. Finally, the circuit should also generate substantial revenue. The last time I checked, they still charge money at the grocery store. My employer is similarly faithful about paying me, and, in both cases, it's my obligation to hold up my end of the bargain.

So, those are my thoughts on good circuits, but I never addressed the statement at the end of the first paragraph. How did my best circuit come to be? That's a long story. Here it is.

The Postpartum Blues

Towards the end of 1991 I was in a rut. I had finished a large high-speed amplifier project in August. It had required a year of constant, intense, and sometimes ferocious effort right up to its conclusion. Then it was over, and I suddenly had nothing to do. I have found myself abruptly disconnected from an absorbing task before, and the result is always the same. I go into this funky kind of rut, and wonder if I'll ever find anything else interesting to do, and if I'm even capable of doing anything anymore.

Portions of this text have appeared in the January 6, 1994 issue of *EDN* magazine and publications of Linear Technology Corporation. They are used here with permission.

I've been dating me a long time, so this state of mind doesn't promote quite the panic and urgency it used to. The treatment is always the same. Keep busy with mundane chores at work, read, cruise electronic junk stores, fix things and, in general, look available so that some interesting problem might ask me to dance. During this time I can do some of the stuff I completely let go while I was immersed in whatever problem owned me. The treatment always seems to work, and usually takes a period of months. In this case it took exactly three.

What's a Backlight?

Around Christmas my boss, Bob Dobkin, asked me if I ever thought about the liquid crystal display (LCD) backlights used in portable computers. I had to admit I didn't know what a backlight was. He explained that LCD displays require an illumination source to make the display readable, and that this source consumed about half the power in the machine. Additionally, the light source, a form of fluorescent lamp, requires high-voltage, high-frequency AC drive. Bob was wondering how this was done, with what efficiency, and if we couldn't come up with a better way and peddle it. The thing sounded remotely interesting. I enjoy transducer work, and that's what a light bulb is. I thought it might be useful to get my hands on some computers and take a look at the backlights. Then I went off to return some phone calls, attend to other housekeeping type items, and, basically, maintain my funk.

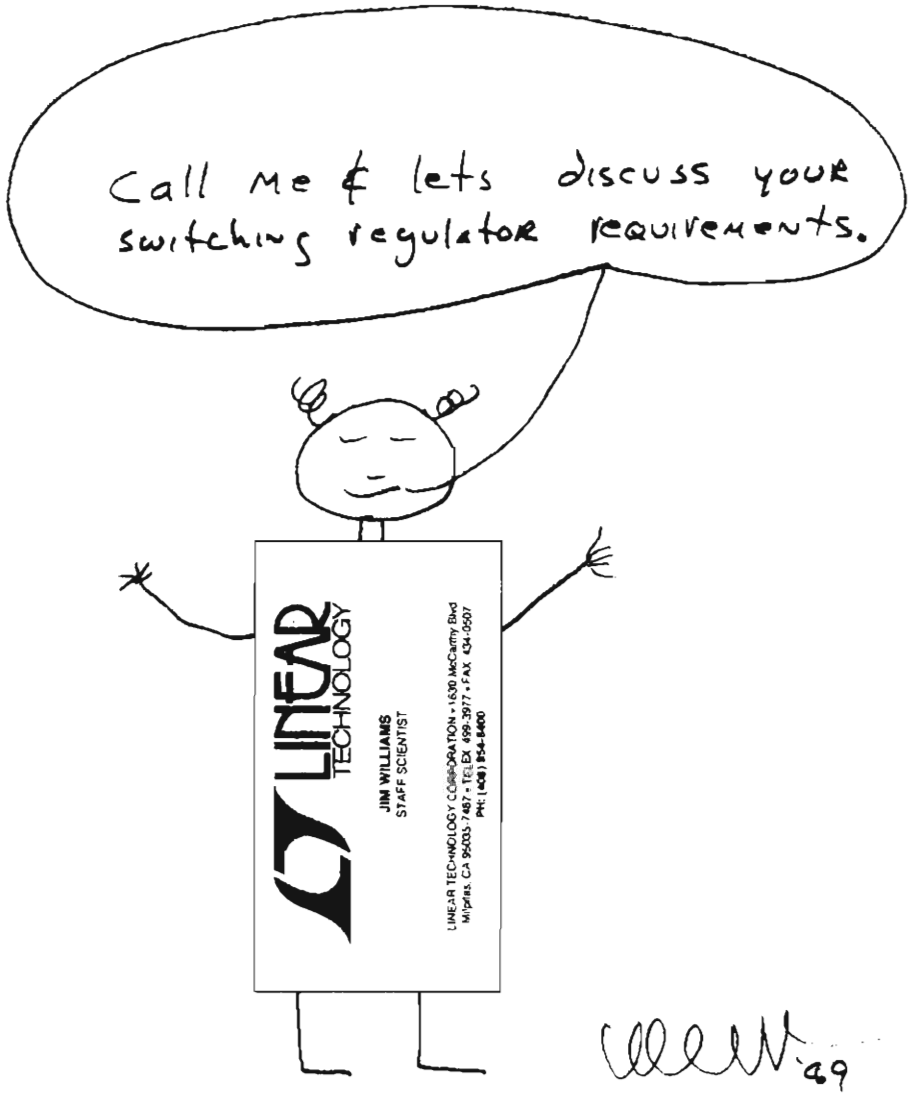
A Call from Some Guy Named Steve

Three days later the phone rang. The caller, a guy named Steve Young from Apple Computer, had seen a cartoon (Figure 11–1) I stuck on the back page of an application note in 1989. Since the cartoon invited calls, he was doing just that. Steve outlined several classes of switching power supply problems he was interested in. The application was portable computers, and a more efficient backlight circuit was a priority. Dobkin's interest in backlights suddenly sounded a lot less academic.

This guy seemed like a fairly senior type, and Apple was obviously a prominent computer company. Also, he was enthusiastic, seemed easy to work with and quite knowledgeable. This potential customer also knew what he wanted, and was willing to put a lot of front end thinking and time in to get it. It was clear he wasn't interested in a quick fix; he wanted true, "end-to-end" system oriented thinking.

What a customer! He knew what he wanted. He was open and anxious to work, had time and money, and was willing to sweat to get better solutions. On top of all that, Apple was a large and successful company with excellent engineering resources. I set up a meeting to introduce him to Dobkin and, hopefully, get something started.

Application Note 35



Linear Technology Corporation
1630 McCarthy Blvd., Milpitas, CA 95035-7487 • (408) 432-1900
FAX: (408) 434-0507 • TELEX: 499-3977

Figure 11–1.
This invitation appeared in a 1989 application note. Some guy named Steve Young from Apple Computer took me up on it. (Reproduced with permission of Linear Technology Corporation)

The meeting went well, things got defined, and I took the backlight problem. I still wasn't enthralled with backlights, but here was an almost ideal customer falling in through the roof so there really wasn't any choice.

Steve introduced me to Paul Donovan, who would become my primary Apple contact. Donovan outlined the ideal backlight. It should have the highest possible efficiency, that is, the highest possible display luminosity with the lowest possible battery drain. Lamp intensity should be smoothly and continuously variable over a wide range with no hysteresis, or "pop-on," and should not be affected by supply voltage changes. RF emissions should meet FCC and system requirements. Finally, parts count and board space should be minimal. There was a board height requirement of .25".

Getting Started—The Luddite Approach to Learning

I got started by getting a bunch of portable computers and taking them apart. I must admit that the Luddite in me enjoyed throwing away most of the computers while saving only their display sections. One thing I immediately noticed was that almost all of them utilized a purchased, board-level solution to backlight driving. Almost no one actually built the function. The circuits invariably took the form of an adjustable output step-down switching regulator driving a high voltage DC-AC inverter (Figure 11–2). The AC high-voltage output was often about 50kHz, and approximately sinusoidal. The circuits seemed to operate on the assumption that a constant voltage input to the DC-AC inverter would produce a fixed, high voltage output. This fixed output would, in turn, produce constant lamp light emission. The ballast capacitor's function was not entirely clear, but I suspected it was related to lamp characteristics. There was no form of feedback from the lamp to the drive circuitry.

Was there something magic about the 50kHz frequency? To see, I built up a variable-frequency high voltage generator (Figure 11–3) and drove the displays. I varied frequency while comparing electrical drive power

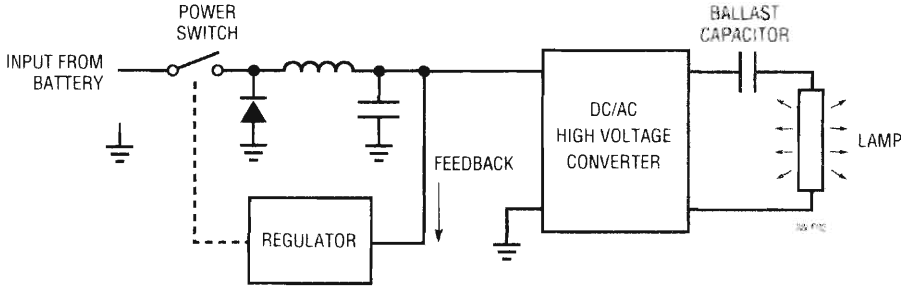


Figure 11–2. Architecture of a typical lamp driver board. There is no form of feedback from the lamp.

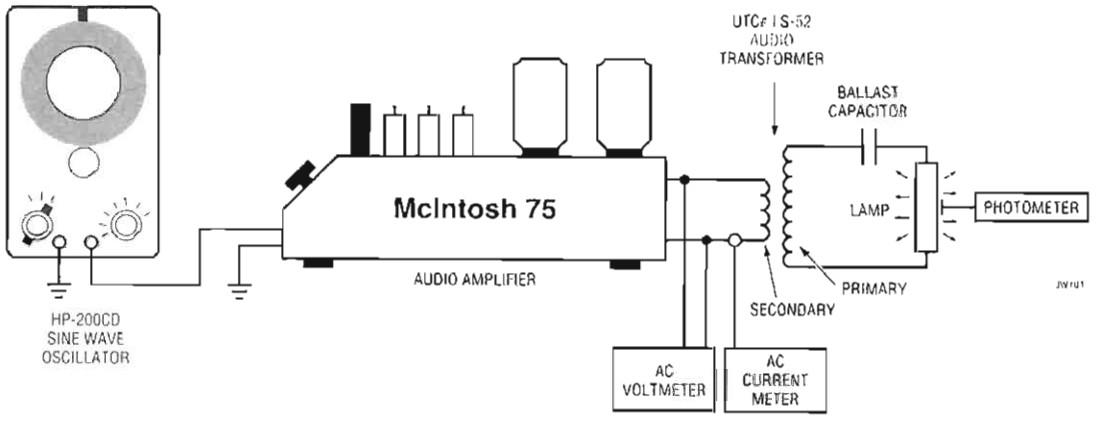


Figure 11–3. Variable frequency high-voltage test setup for evaluating lamp frequency sensitivity.

to optical emission. Lamp conversion efficiency seemed independent of frequency over a fairly wide range. I did, however, notice that higher frequencies tended to introduce losses in the wiring running to the lamp. These losses occurred at all frequencies, but became pronounced above about 100kHz or so. Deliberately introducing parasitic capacitances from the wiring or lamp to ground substantially increased the losses. The lesson was clear. The lamp wiring was an inherent and parasitic part of the circuit, and any stray capacitive path was similarly parasitic.

Armed with this information I returned to the computer displays. I modified things so that the wire length between the inverter board and display was minimized. I also removed the metal display housing in the lamp area. The result was a measurable decrease in inverter drive power for a given display intensity. In two machines the improvement approached 20%! My modifications weren't very practical from a mechanical integrity viewpoint, but that wasn't relevant. Why hadn't these computers been originally designed to take advantage of this "free" efficiency gain?

Playing around with Light Bulbs

I removed lamps from the displays. They all appeared to have been installed by the display vendor, as opposed to being selected and purchased by the computer manufacturer. Even more interesting was that I found identical backlight boards in different computers driving different types of lamps. There didn't seem to be any board changes made to accommodate the various lamps. Now, I turned my attention to the lamps.

The lamps seemed to be pretty complex and wild animals. I noticed that many of them took noticeable time to arrive at maximum intensity. Some types seemed to emit more light than others for a given input power. Still others had a wider dynamic range of intensities than the rest, although all had a seemingly narrow range of intensity control. Most striking was that every lamp's emissivity varied with ambient tempera-

ture. Experimenting with a hair dryer, a can of "cold spray" and a photometer, I found that each lamp seemed to have an optimum operating temperature range. Excursions above or below this region caused emittance to fall.

I put a lamp into a reassembled display. With the display warmed up in a 25°C environment I was able to increase light output by slightly ventilating the lamp enclosure. This increased steady-state thermal losses, allowing the lamp to run in its optimum temperature range. I also saw screen illumination shifts due to the distance between the light entry point at the display edge and the lamp. There seemed to be some optimum distance between the lamp and the entry point. Simply coupling the lamp as closely as possible did not provide the best results. Similarly, the metallic reflective foil used to concentrate the lamp's output seemed to be sensitive to placement. Additionally, there was clearly a trade-off between benefits from the foil's optical reflection and its absorption of high voltage field energy. Removing the foil decreased input energy for a given lamp emission level. I could watch input power rise as I slipped the foil back along the lamp's length. In some cases, with the foil fully replaced, I could draw sparks from it with my finger!

I also assembled lamps, displays, and inverter boards in various unoriginal combinations. In some cases I was able to increase light output, at lower input power drain, over the original "as shipped" configuration.

Grandpa Would Have Liked It

I tried a lot of similarly simple experiments and slowly developed a growing suspicion that nobody, at least in my sample of computers, was making any serious attempt at optimizing (or they did not know how to optimize) the backlight. It appeared that most people making lamps were simply filling tubes up with gas and shipping them. Display manufacturers were dropping these lamps into displays and shipping them. Computer vendors bought some "backlight power supply" board, wired it up to the display, took whatever electrical and optical efficiency they got, and shipped the computer.

If I allowed this conclusion, several things became clear. Development of an efficient backlight required an interdisciplinary approach to address a complex problem. There was worthwhile work to be done. I could contribute to the electronic portion, and perhaps the thermal design, but the optical engineering was beyond me. It was not, however, beyond Apple's resources. Apple had some very good optical types. Working together, it seemed we had a chance to build a better backlight with its attendant display quality and battery life advantages. Apple would get a more saleable product and my company would develop a valued customer. And, because the whole thing was beginning to get interesting, I could get out of my rut. The business school types would call this "synergistic" or "win-win." Other people who "do lunch" a lot on company money would

call it "strategic partnering." My grandfather would have called it "such a deal."

Goals for the backlight began to emerge. For best overall efficiency, the display enclosure, optical design, lamp, and electronics had to be simultaneously considered. My job was the electronics, although I met regularly with Paul Donovan, who was working on the other issues. In particular, I was actively involved in setting lamp specifications and evaluating lamp vendors.

The electronics should obviously be as efficient as possible. The circuit should be physically compact, have a low parts count, and assemble easily. It should have a wide, continuous dimming range with no hysteresis or "pop-on," and should meet all RF and system emission requirements. Finally, it must regulate lamp intensity against wide power supply shifts, such as when the computer's AC adapter is plugged in.

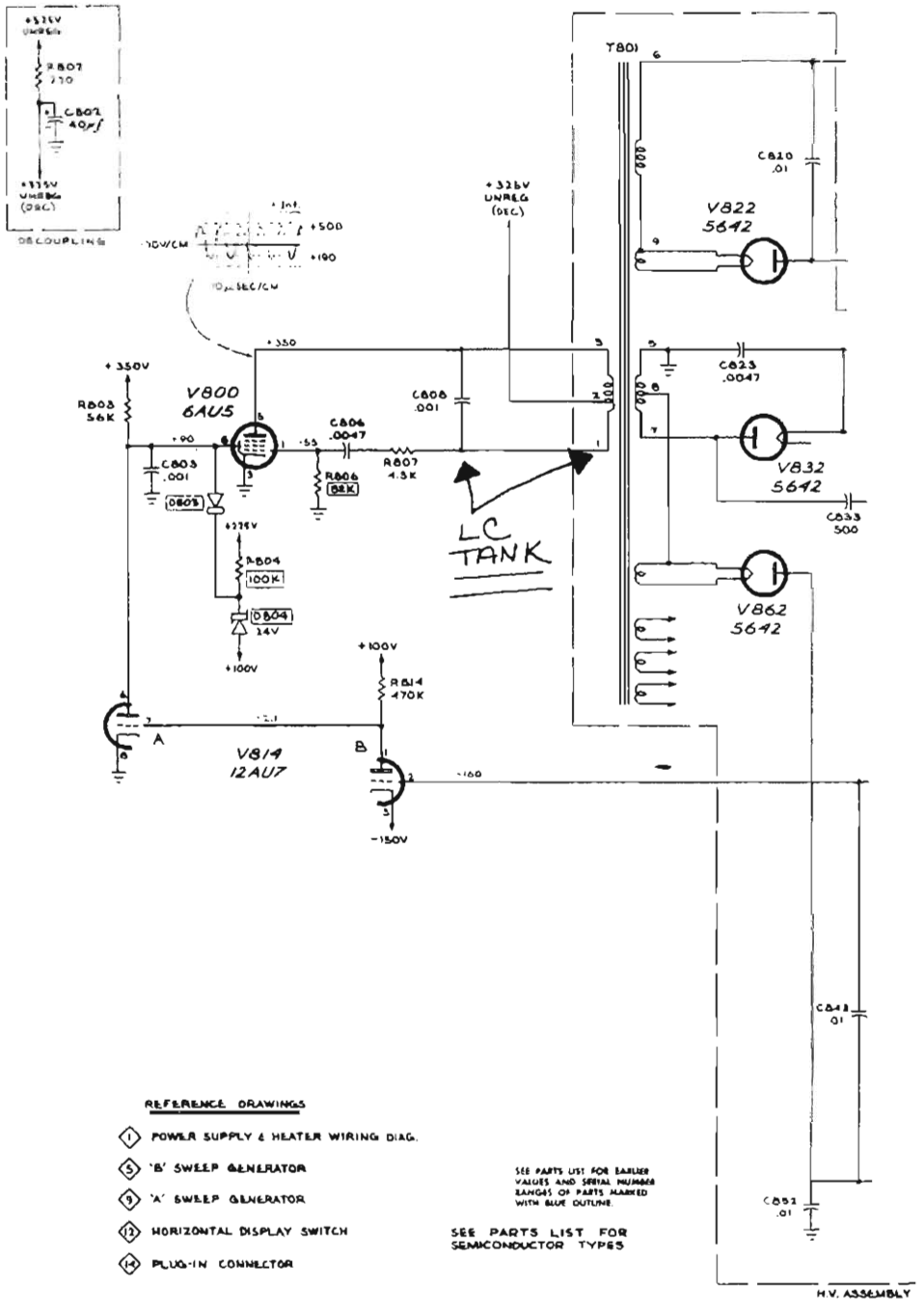
Help from Dusty Circuits

Where, I wondered, had I seen circuitry which contained any or all of these characteristics? Nowhere. But, one place to start looking was oscilloscopes. Although oscilloscope circuits do not accomplish what I needed to do, oscilloscope designers use high frequency sine wave conversion to generate the high voltage CRT supply. This technique minimizes noise and reduces transformer and capacitor size. Additionally, by doing the conversion at the CRT, long high voltage runs from the main power supply are eliminated.

I looked at the schematic of the high voltage converter in a Tektronix 547 (Figure 11–4). The manual's explanation (Figure 11–5) says the capacitor (C808) and transformer primary form a resonant tank circuit. More subtly, the "transformer primary" also includes the complex impedance reflected back from the secondary and its load. But that's a detail for this circuit and for now. A CRT is a relatively linear and benign load. The backlight's loading characteristics would have to be evaluated and matched to the circuit.

This CRT circuit could not be used to drive a fluorescent backlight tube in a laptop computer. For one reason, this circuit is not very efficient. It does not have to be. A 547 pulls over 500 watts, so efficiency in this circuit was not a big priority. Latter versions of this configuration were transistorized (Figure 11–6, Tektronix 453), but used basically the same architecture. In both circuits the resonating technique is employed, and a feedback loop enforces voltage regulation. For another reason, the CRT requires the high voltage to be rectified to DC. The backlight requires AC, eliminating the rectifier and filter. And, the CRT circuit had no feedback. Some form of feedback for the fluorescent lamp seemed desirable.

The jewel in the CRT circuit, however, was the resonating technique used to create the sine wave. The transformer does double duty. It helps create the sine wave while simultaneously generating the high voltage.



TYPE 547 OSCILLOSCOPE

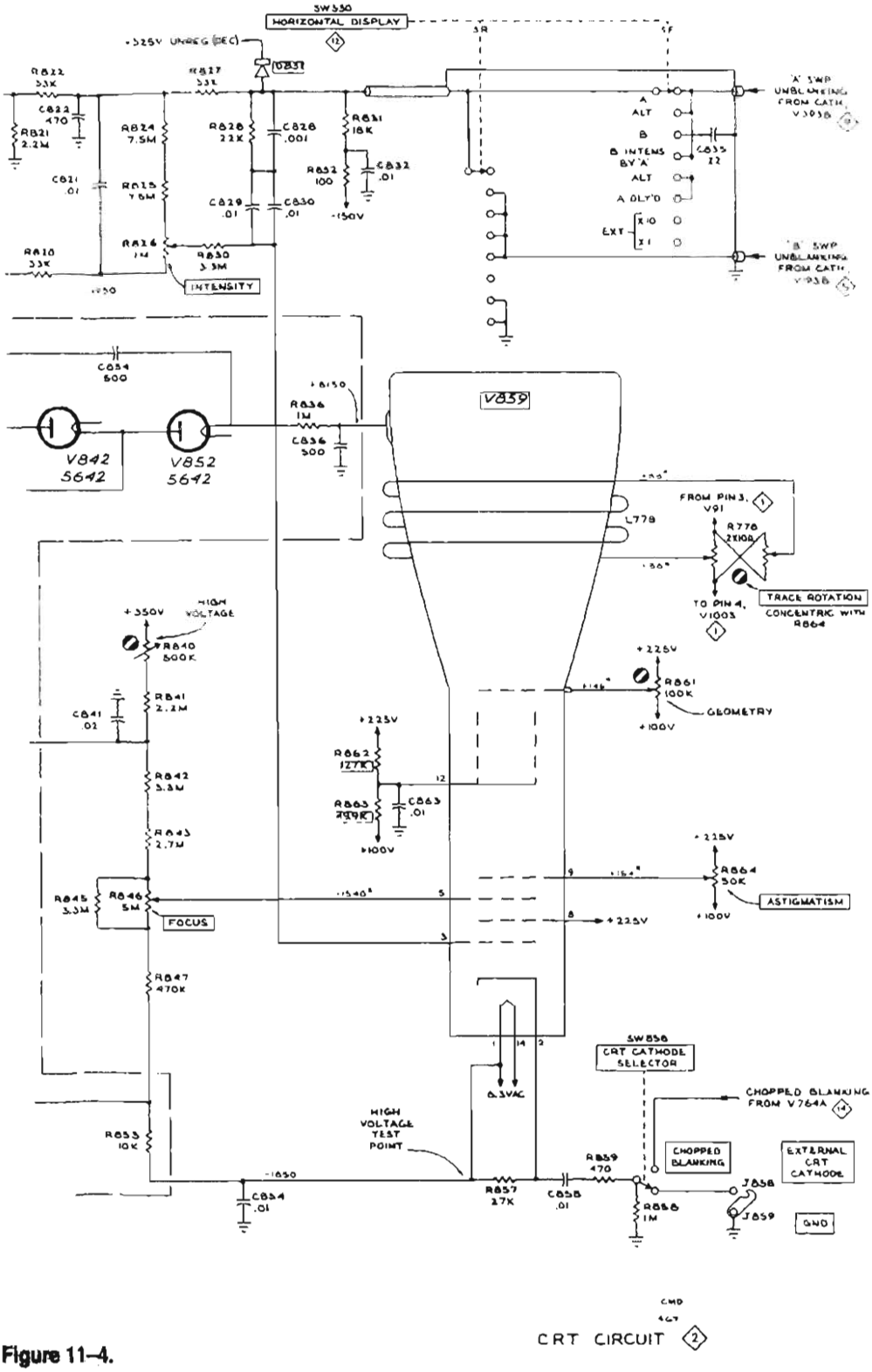


Figure 11–4.
CRT supply used in Tektronix 547. C808 resonates with transformer, creating sine wave drive. (Figure reproduced with permission of Tektronix, Inc.)

Figure 11–5. Tektronix 547 manual explains resonant operation. (Figure reproduced with permission of Tektronix, Inc.)

Crt Circuit

The crt circuit (see Crt schematic) includes the crt, the high-voltage power supply, and the controls necessary to focus and orient the display. The crt (Tektronix Type T5470-31-2) is an aluminized, 5-inch, flat-faced, glass crt with a helical post-accelerator and electrostatic focus and deflection. The crt circuit provides connections for externally modulating the crt cathode. The high-voltage power supply is composed of a dc-to-50-kc power converter, a voltage-regulator circuit, and three high-voltage outputs. Front-panel controls in the crt circuit adjust the trace rotation (screwdriver adjustment), intensity, focus, and astigmatism. Internal controls adjust the geometry and high-voltage output level.

High-Voltage Power Supply. The high-voltage power supply is a dc-to-ac converter operating at approximately 50 kc with the transformer providing three high-voltage outputs. The use of a 50-kc input to the high-voltage transformer permits the size of the transformer and filter components to be kept small. A modified Hartley oscillator converts dc from the +325-volt unregulated supply to the 50-kc input required by high-voltage transformer T801. C808 and the primary of T801 form the oscillator resonant tank circuit. No provisions are made for precise tuning of the oscillator tank since the exact frequency of oscillation is not important.

Voltage Regulation. Voltage regulation of the high-voltage outputs is accomplished by regulating the amplitude of oscillations in the Hartley oscillator. The —1850-volt output is referenced to the +350-volt regulated supply through a voltage divider composed of R841, R842, R843, R845, R846, R847, R853, and variable resistors R840 and R846. Through a tap on the voltage divider, the regulator circuit samples the —1850-volt output of the supply, amplifies any errors and uses the amplified error voltage to adjust the screen voltage of Hartley oscillator V800. If the —1850-volt output changes, the change is detected at the grid of V814B. The detected error is amplified by V814B and V814A. The error signal at the plate of V814A is direct coupled to the screen of V800 by making the plate-load resistor of V814A serve as

How could I combine this circuit's desirable resonating characteristics with other techniques to meet the backlight's requirements? One key was a simple, more efficient transformer drive. I knew just where to find it.

In December 1954 the paper "Transistors as On-Off Switches in Saturable-Core Circuits" appeared in *Electrical Manufacturing*. George H. Royer, one of the authors, described a "d-c to a-c converter" as part of this paper. Using Westinghouse 2N74 transistors, Royer reported 90% efficiency for his circuit. The operation of Royer's circuit is well described in this paper. The Royer converter was widely adopted, and used in designs from watts to kilowatts. It is still the basis for a wide variety of power conversion.

Royer's circuit is not an LC resonant type. The transformer is the sole energy storage element and the output is a square wave. Figure 11–7 is a conceptual schematic of a typical converter. The input is applied to a self-oscillating configuration composed of transistors, a transformer, and a biasing network. The transistors conduct out of phase switching (Figure 11–8: Traces A and C are Q1's collector and base, while Traces B and D are Q2's collector and base) each time the transformer saturates. Transformer saturation causes a quickly rising, high current to flow (Trace E).

This current spike, picked up by the base drive winding, switches the transistors. This phase opposed switching causes the transistors to exchange states. Current abruptly drops in the formerly conducting transistor and then slowly rises in the newly conducting transistor until saturation again forces switching. This alternating operation sets transistor duty cycle at 50%.

The photograph in Figure 11–9 is a time and amplitude expansion of Figure 11–8's Traces B and E. It clearly shows the relationship between transformer current (Trace B, Figure 11–9) and transistor collector voltage (Trace A, Figure 11–9).[1]

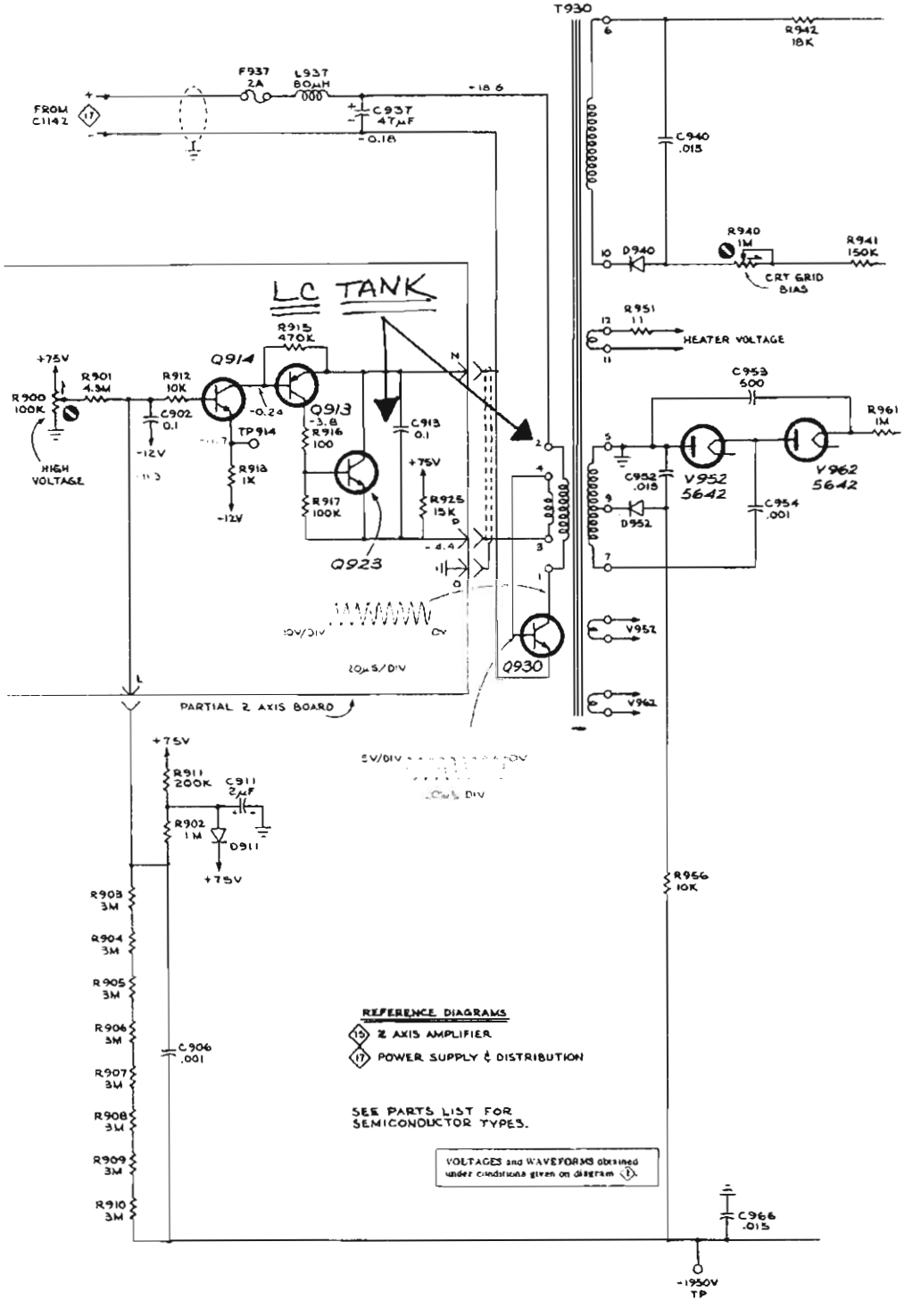
The Royer has many desirable elements which are applicable to backlight driving. Transformer size is small because core utilization is efficient. Parts count is low, the circuit self-oscillates, it is efficient, and output power may be varied over a wide range. The inherent nature of operation produces a square wave output, which is not permissible for backlight driving.

Adding a capacitor to the primary drive (Figure 11–10) should have the same resonating effect as in the Tektronix CRT circuits. The beauty of this configuration is its utter simplicity and high efficiency. As loading (e.g., lamp intensity) is varied the reflected secondary impedance changes, causing some frequency shift, but efficiency remains high.

The Royer's output power is controllable by varying the primary drive current. Figure 11–11 shows a way to investigate this. This circuit works well, except that the transistor current sink operates in its linear region, wasting power. Figure 11–12 converts the current sink to switch mode operation, maintaining high efficiency. This is obviously advantageous to the user, but also a good deal for my employer. I had spent the last six months playing with light bulbs, reminiscing over old oscilloscope circuits, taking arcane thermal measurements, and similar dalliances. All the while faithfully collecting my employer's money. Finally, I had found a place to actually sell something we made. Linear Technology (my employer) builds a switching regulator called the LT1172. Its features include a high power open collector switch, trimmed reference, low quiescent current, and shutdown capability. Additionally, it is available in an 8 pin surface-mount package, a must for board space considerations. It was also an ideal candidate for the circuit's current sink portion.

---

1. The bottom traces in both photographs are not germane and are not referenced in the discussion.



TYPE 453 OSCILLOSCOPE

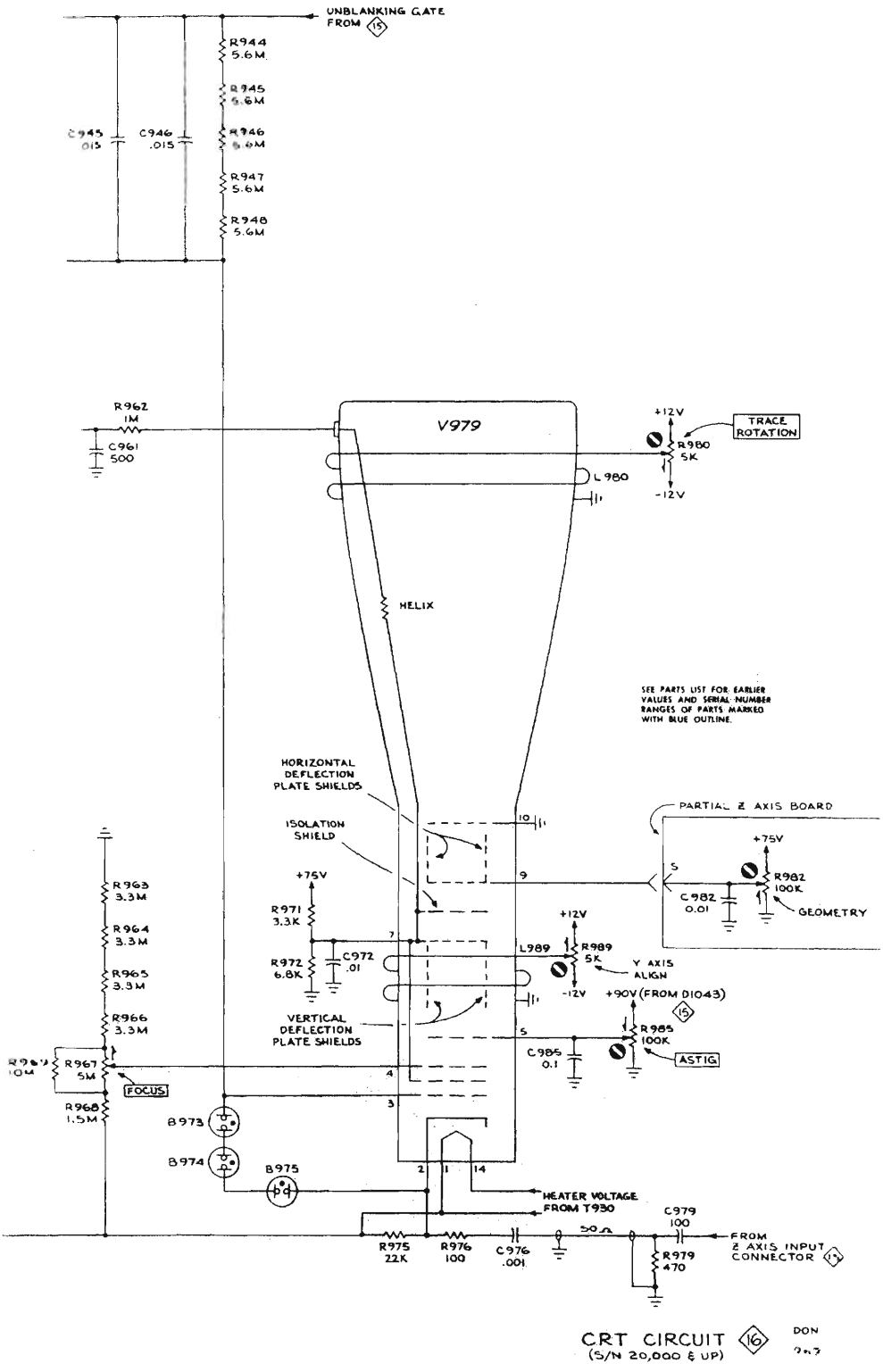


Figure 11–6.
Later model Tektronix 453 is transistorized version of 547's resonant approach. (Figure reproduced with permission of Tektronix, Inc.)

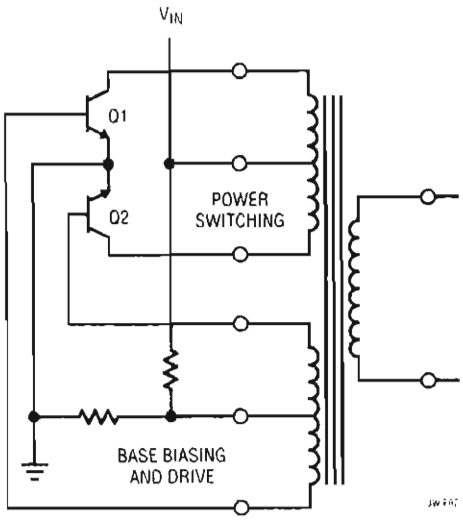


Figure 11–7. Conceptual classic Royer converter. Transformer approaching saturation causes switching.

Of Rafts and Paddles

At about this stage I sat back and stared at the wall. There comes a time in every project where you have to gamble. At some point the analytics and theorizing must stop and you have to commit to an approach and start actually doing something. This is often painful, because you never really have enough information and preparation to be confidently decisive. There are never any answers, only choices. But there comes this time when your gut tells you to put down the pencil and pick up the soldering iron.

Physicist Richard Feynman said, "If you're not confused when you start, you're not doing it right." Somebody else, I think it was an artist, said, "Inspiration comes while working." Wow, are they right. With circuits, as in life, never wait for your ship to come in. Build a raft and start paddling.

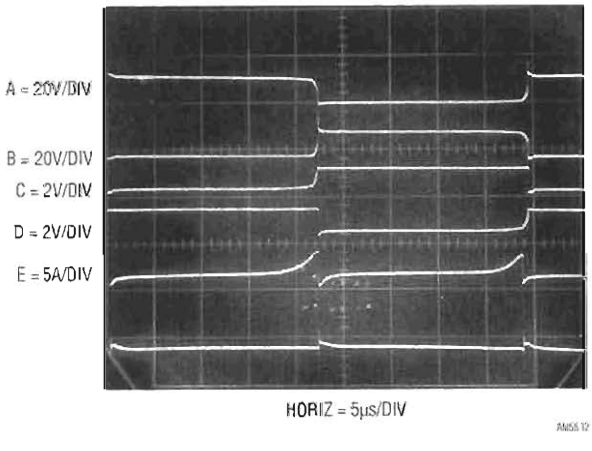


Figure 11–8. Waveforms for the classic Royer circuit.

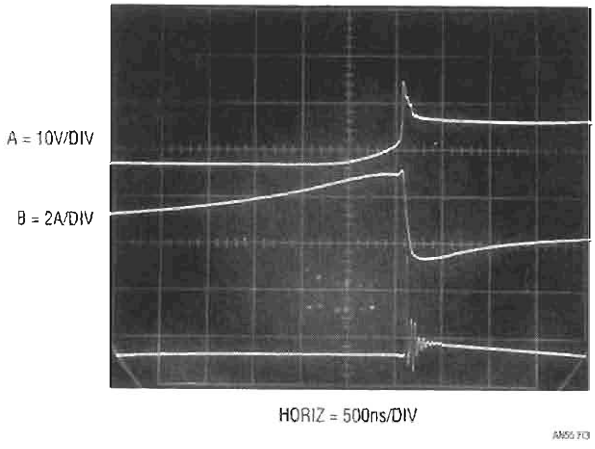


Figure 11–9. Detail of transistor switching. Turn-off (Trace A) occurs just as transformer heads into saturation (Trace B).

Everything was still pretty fuzzy, but I had learned a few things. A practical, highly efficient LCD backlight design is a classic study of compromise in a transduced electronic system. Every aspect of the design is interrelated, and the physical embodiment is an integral part of the electrical circuit. The choice and location of the lamp, wires, display housing, and other items have a major effect on electrical characteristics. The greatest care in every detail is required to achieve a practical, high efficiency LCD backlight. Getting the lamp to light is just the beginning!

A good place to start was to reconsider the lamps. These "Cold Cathode Fluorescent Lamps" (CCFL) provide the highest available efficiency for converting electrical energy to light. Unfortunately, they are optically and electrically highly nonlinear devices.

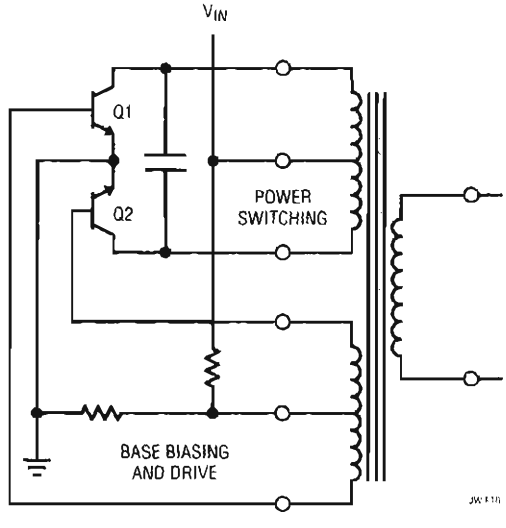


Figure 11–10. Adding the resonating capacitor to the Royer.

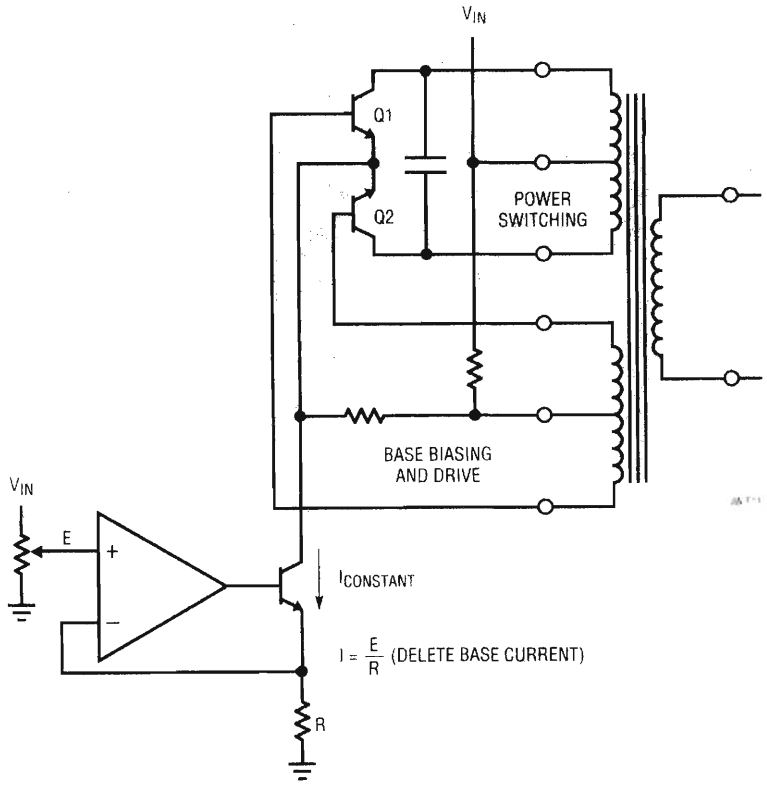


Figure 11–11. Current sink permits controlling Royer power, but is inefficient.

Cold Cathode Fluorescent Lamps (CCFLs)

Any discussion of CCFL power supplies must consider lamp characteristics. These lamps are complex transducers, with many variables affecting their ability to convert electrical current to light. Factors influencing conversion efficiency include the lamp's current, temperature, drive waveform characteristics, length, width, gas constituents, and the proximity to nearby conductors.

These and other factors are interdependent, resulting in a complex overall response. Figures 11–13 through 11–16 show some typical characteristics. A review of these curves hints at the difficulty in predicting lamp behavior as operating conditions vary. The lamp's current and temperature are clearly critical to emission, although electrical efficiency may not necessarily correspond to the best optical efficiency point. Because of this, both electrical and photometric evaluation of a circuit is often required. It is possible, for example, to construct a CCFL circuit with 94% electrical efficiency which produces less light output than an approach with 80% electrical efficiency (see Appendix C, "A Lot of Cut-off Ears and No Van Goghs—Some Not-So-Great Ideas"). Similarly, the performance of a very well matched lamp-circuit combination can be

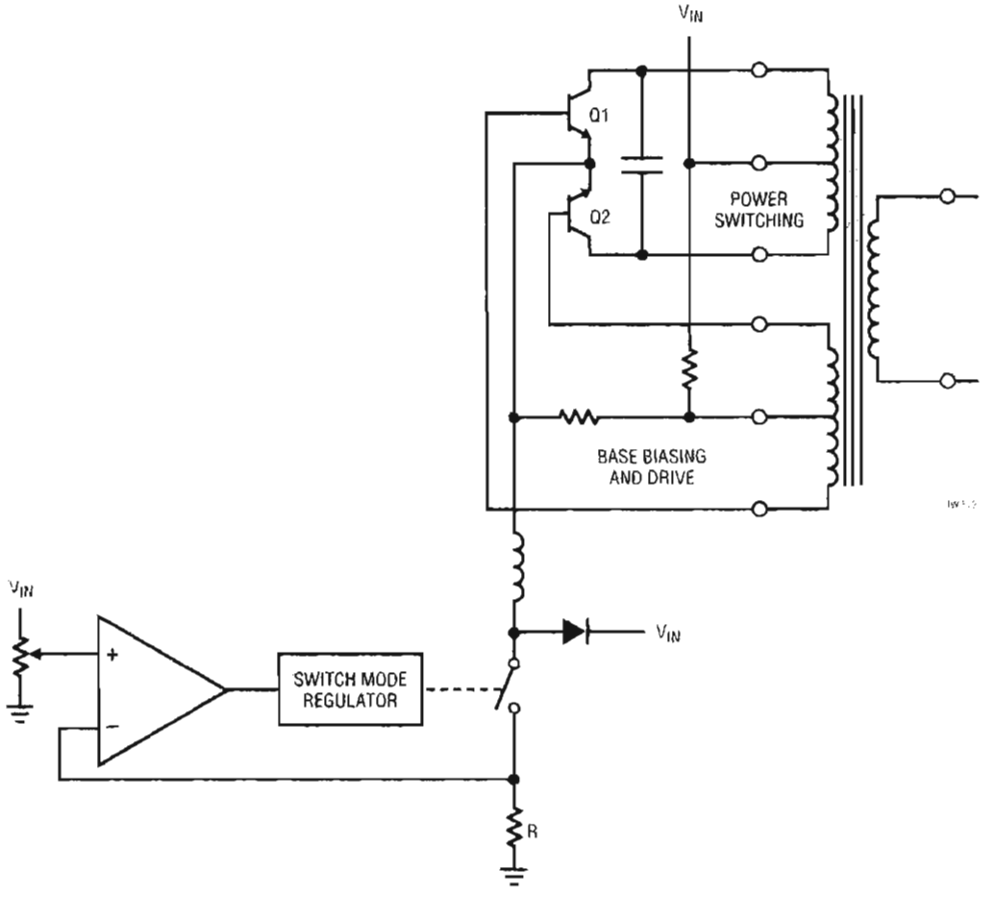


Figure 11–12. Switched mode current sink restores efficiency.

severely degraded by a lossy display enclosure or excessive high voltage wire lengths. Display enclosures with too much conducting material near the lamp have huge losses due to capacitive coupling. A poorly designed display enclosure can easily degrade efficiency by 20%. High voltage wire runs typically cause 1% loss per inch of wire.

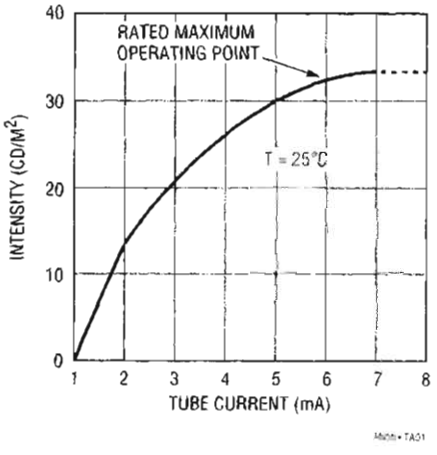


Figure 11–13. Emissivity for a typical 6mA lamp; curve flattens badly above 6mA.

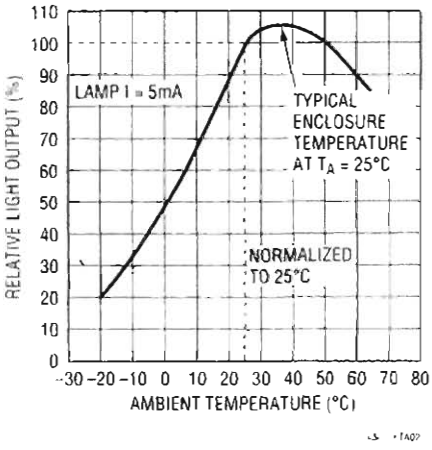


Figure 11–14. Ambient temperature effects on emissivity of a typical 5mA lamp. Lamp and enclosure must come to thermal steady state before measurements are made.

CCFL Load Characteristics

These lamps are a difficult load to drive, particularly for a switching regulator. They have a "negative resistance" characteristic; the starting voltage is significantly higher than the operating voltage. Typically, the start voltage is about 1000V, although higher and lower voltage lamps are common. Operating voltage is usually 300V to 400V, although other lamps may require different potentials. The lamps will operate from DC, but migration effects within the lamp will quickly damage it. As such, the waveform must be AC. No DC content should be present.

Figure 11–17A shows an AC driven lamp's characteristics on a curve tracer. The negative resistance induced "snapback" is apparent. In Figure 11–17B, another lamp, acting against the curve tracer's drive, produces oscillation. These tendencies, combined with the frequency compensation problems associated with switching regulators, can cause severe loop instabilities, particularly on start-up. Once the lamp is in its operating region it assumes a linear load characteristic, easing stability criteria. Lamp operating frequencies are typically 20kHz to 100kHz and a sine-

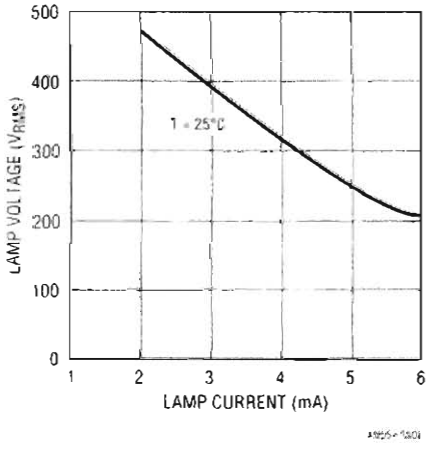


Figure 11–15. Current vs. voltage for a lamp in the operating region.

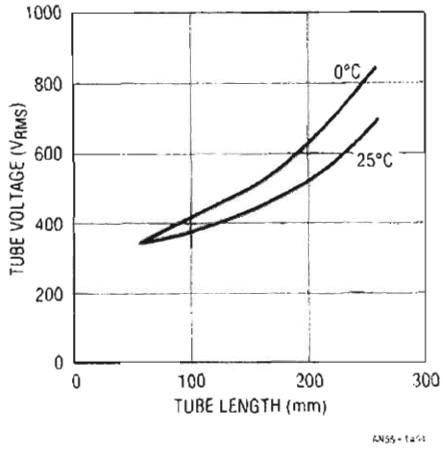


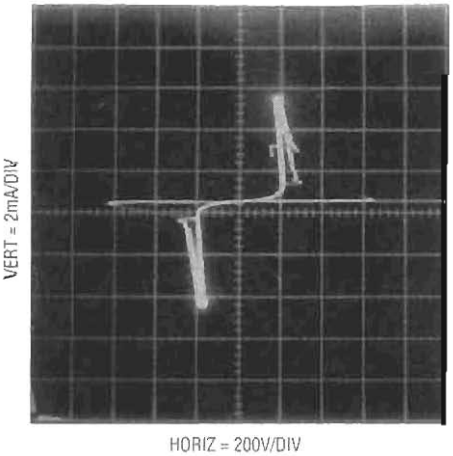
Figure 11–16. Running voltage vs. lamp length at two temperatures. Start-up voltages are usually 50% to 200% higher over temperature.

like waveform is preferred. The sine drive's low harmonic content minimizes RF emissions, which could cause interference and efficiency degradation. A further benefit of the continuous sine drive is its low crest factor and controlled risetimes, which are easily handled by the CCFL. CCFL's RMS current-to-light output efficiency is degraded by high crest factor drive waveforms.[2]

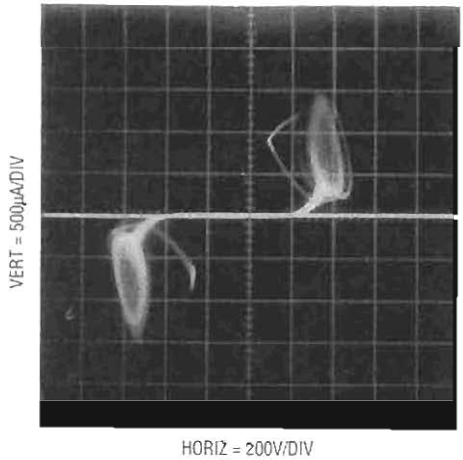
CCFL Power Supply Circuits

Figure 11–18's circuit meets CCFL drive requirements. Efficiency is 88% with an input voltage range of 4.5V to 20V. This efficiency figure will be degraded by about 3% if the LT1172 V_{IN} pin is powered from the same supply as the main circuit V_{IN} terminal. Lamp intensity is continuously and smoothly variable from zero to full intensity. When power is

Figure 11–17. Negative resistance characteristic for two CCFL lamps. "Snap-back" is readily apparent, causing oscillation in 11–17B. These characteristics complicate power supply design.

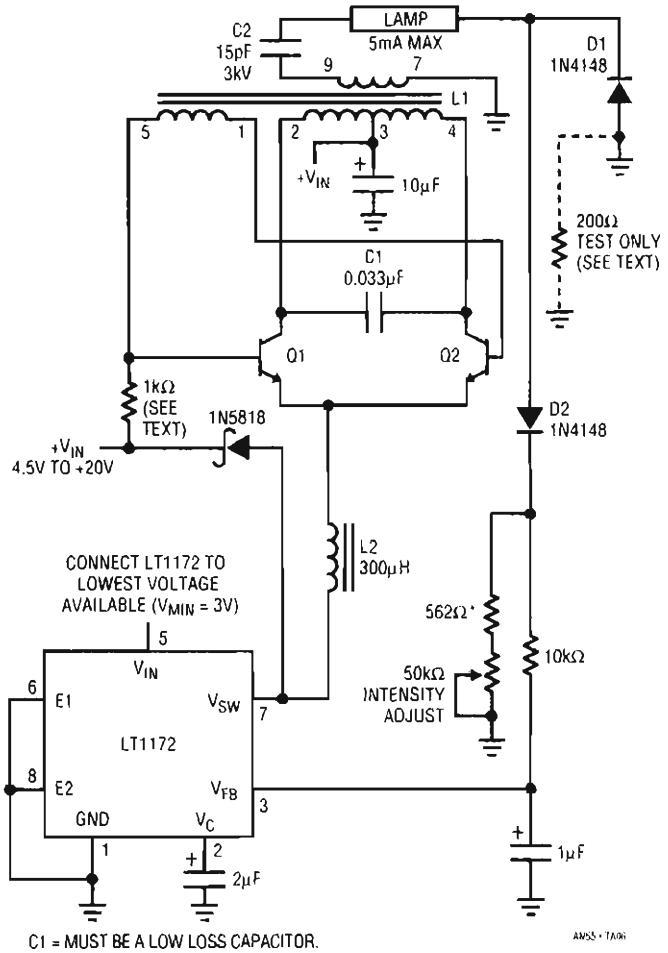


17A



17B

2. See Appendix C, "A Lot of Cut-off Ears and No Van Goghs—Some Not-So-Great Ideas."



C1 = MUST BE A LOW LOSS CAPACITOR.
METALIZED POLYCARB
WIMA FKP2 OR MKP-20 (GERMAN) RECOMMENDED
L1 = SUMIDA 6345-020 OR COILTRONICS CTX110092-1
PIN NUMBERS SHOWN FOR COILTRONICS UNIT
L2 = COILTRONICS CTX300-4
Q1, Q2 = ZETEX ZTX849 OR ROHM 2SC5001
* = 1% FILM RESISTOR
DO NOT SUBSTITUTE COMPONENTS

COILTRONICS (305) 781-8900, SUMIDA (708) 956-0666

Figure 11–18. An 88% efficiency cold cathode fluorescent lamp (CCFL) power supply.

applied the LT1172 switching regulator's feedback pin is below the device's internal 1.2V reference, causing full duty cycle modulation at the V_{SW} pin (Trace A, Figure 11–19). L2 conducts current (Trace B) which flows from L1's center tap, through the transistors, into L2. L2's current is deposited in switched fashion to ground by the regulator's action.

L1 and the transistors comprise a current driven Royer class converter which oscillates at a frequency primarily set by L1's characteristics (including its load) and the .033µF capacitor. LT1172 driven L2 sets the magnitude of the Q1-Q2 tail current, and hence L1's drive level. The 1N5818 diode maintains L2's current flow when the LT1172 is off. The LT1172's 100kHz clock rate is asynchronous with respect to the push-pull converter's (60kHz) rate, accounting for Trace B's waveform thickening.

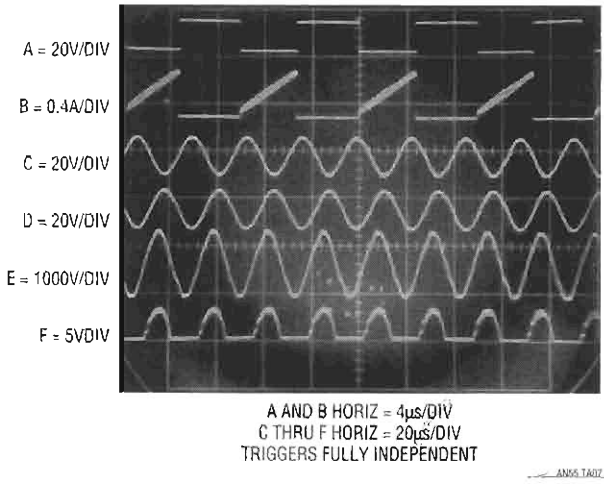


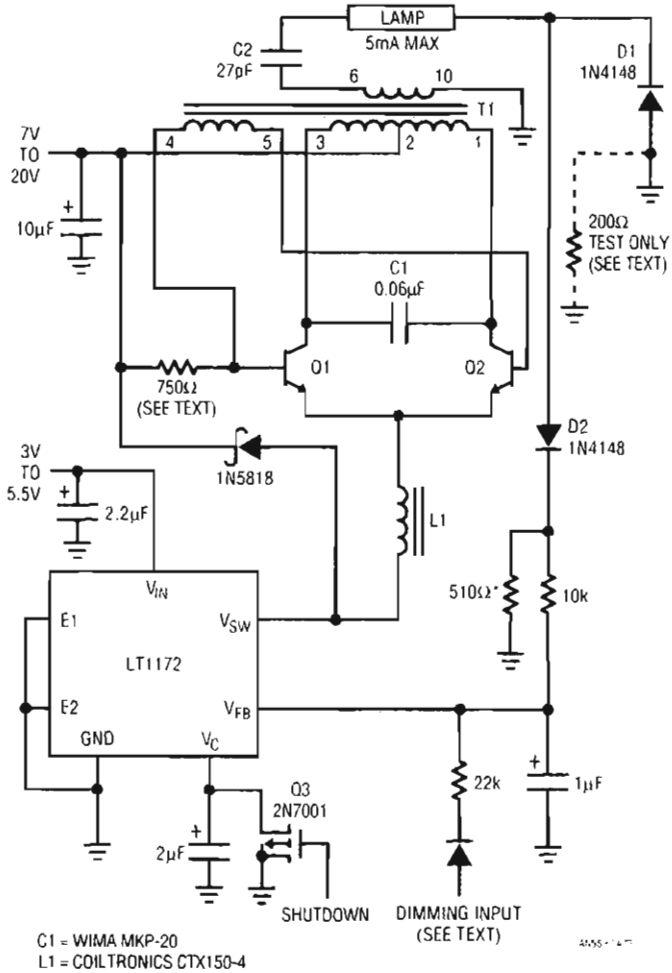
Figure 11–19. Waveforms for the cold cathode fluorescent lamp power supply. Note independent triggering on Traces A and B, and C through F.

The .033μF capacitor combines with L1's characteristics to produce sine wave voltage drive at the Q1 and Q2 collectors (Traces C and D, respectively). L1 furnishes voltage step-up, and about 1400V p-p appears at its secondary (Trace E). Current flows through the 15pF capacitor into the lamp. On negative waveform cycles the lamp's current is steered to ground via D1. Positive waveform cycles are directed, via D2, to the ground referred 562Ω–50k potentiometer chain. The positive half-sine appearing across the resistors (Trace F) represents ½ the lamp current. This signal is filtered by the 10k–1μF pair and presented to the LT1172's feedback pin. This connection closes a control loop which regulates lamp current. The 2μF capacitor at the LT1172's V_C pin provides stable loop compensation. The loop forces the LT1172 to switch-mode modulate L2's average current to whatever value is required to maintain a constant current in the lamp. The constant current's value, and hence lamp intensity, may be varied with the potentiometer. The constant current drive allows full 0%–100% intensity control with no lamp dead zones or "pop-on" at low intensities. Additionally, lamp life is enhanced because current cannot increase as the lamp ages. This constant current feedback approach contrasts with the open loop, voltage type drive used by other approaches. It greatly improves control over the lamp under all conditions.

This circuit's 0.1% line regulation is notably better than some other approaches. This tight regulation prevents lamp intensity variation when abrupt line changes occur. This typically happens when battery powered apparatus is connected to an AC powered charger. The circuit's excellent line regulation derives from the fact that L1's drive waveform never changes shape as input voltage varies. This characteristic permits the simple 10kΩ–1μF RC to produce a consistent response. The RC averaging characteristic has serious error compared to a true RMS conversion, but the error is constant and "disappears" in the 562Ω shunt's value. The base drive resistor's value (nominally 1kΩ) should be selected to provide

full V_{CE} saturation without inducing base overdrive or beta starvation. A procedure for doing this is described in the following section, "General Measurement and Optimization Considerations."

Figure 11–20's circuit is similar, but uses a transformer with lower copper and core losses to increase efficiency to 91%. The trade-off is slightly larger transformer size. Value shifts in C1, L2, and the base drive resistor reflect different transformer characteristics. This circuit also features shutdown via Q3 and a DC or pulse width controlled dimming input. Figure 11–21, directly derived from Figure 11–20, produces 10mA output to drive color LCDs at 92% efficiency. The slight efficiency improvement comes from a reduction in LT1172 "housekeeping" current as a percentage

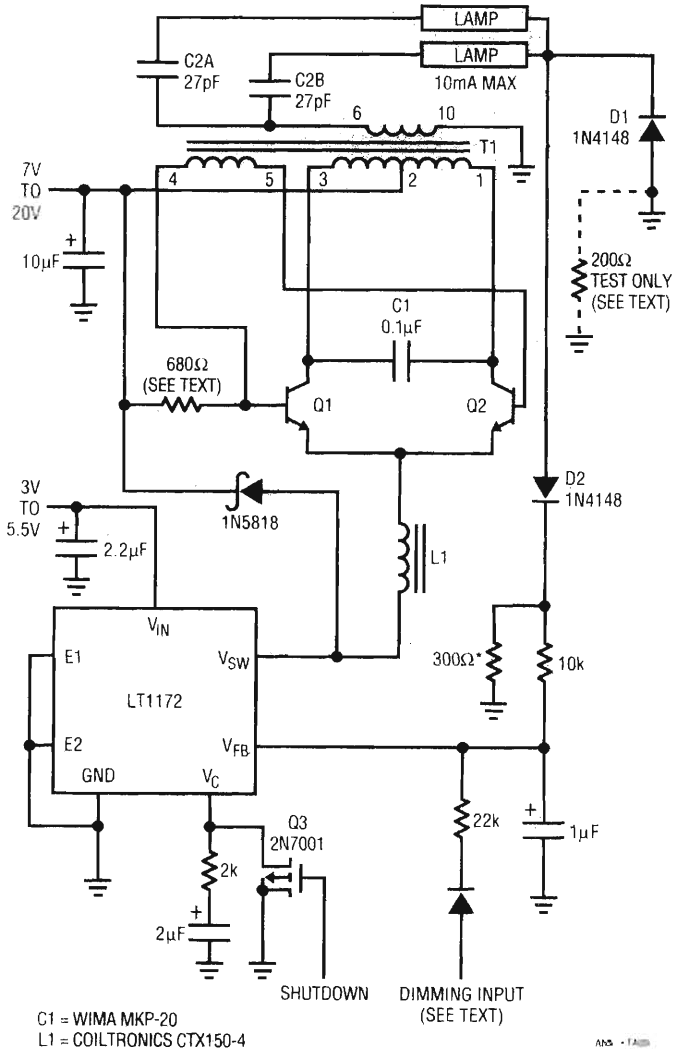


C1 = WIMA MKP-20
L1 = COILTRONICS CTX150-4
Q1, Q2 = ZETEX ZTX849 OR ROHM 2SC5001
T1 = COILTRONICS CTX110600-1 OR SUMIDA EPS-207
PIN NUMBERS SHOWN FOR COILTRONICS UNIT
* = 1% FILM RESISTOR
DO NOT SUBSTITUTE COMPONENTS

COILTRONICS (305) 781-8900, SUMIDA (708) 956-0666

Figure 11–20. A 91% efficient CCFL supply for 5mA loads features shutdown and dimming inputs.

of total current drain. Value changes in components are the result of higher power operation. The most significant change involves driving two tubes. Accommodating two lamps involves separate ballast capacitors but circuit operation is similar. Two lamp designs reflect slightly different loading back through the transformer's primary. C2 usually ends up in the 10pF to 47pF range. Note that C2A and B appear with their lamp loads in parallel across the transformer's secondary. As such, C2's value is often smaller than in a single tube circuit using the same type lamp. Ideally the transformer's secondary current splits evenly between the C2-lamp branches, with the total load current being regulated. In practice, differences between C2A and B and differences in lamps and lamp wiring layout preclude a perfect current split. Practically, these differences are small, and the



C1 = WIMA MKP-20
L1 = COILTRONICS CTX150-4
Q1, Q2 = ZETEX ZTX849 OR ROHM 2SC5001
T1 = COILTRONICS CTX110600-1 OR SUMIDA EPS-207
PIN NUMBERS SHOWN FOR COILTRONICS UNIT
* = 1% FILM RESISTOR
DO NOT SUBSTITUTE COMPONENTS

COILTRONICS (305) 781-8900, SUMIDA (708) 956-0666

Figure 11–21. A 92% efficient CCFL supply for 10mA loads features shutdown and dimming inputs. Two lamps are typical of color displays.

lamps appear to emit equal amounts of light. Layout and lamp matching can influence C2's value. Some techniques for dealing with these issues appear in the section "Layout Issues."

General Measurement and Optimization Considerations

Several points should be kept in mind when observing operation of these circuits. L1's high voltage secondary can only be monitored with a wide-band, high voltage probe fully specified for this type of measurement. *The vast majority of oscilloscope probes will break down and fail if used for this measurement.* Tektronix probe types P6007 and P6009 (acceptable) or types P6013A and P6015 (preferred) must be used to read L1's output.

Another consideration involves observing waveforms. The LT1172's switching frequency is completely asynchronous from the Q1-Q2 Royer converter's switching. As such, most oscilloscopes cannot simultaneously trigger and display all the circuit's waveforms. Figure 11–19 was obtained using a dual beam oscilloscope (Tektronix 556). LT1172 related Traces A and B are triggered on one beam, while the remaining traces are triggered on the other beam. Single beam instruments with alternate sweep and trigger switching (e.g., Tektronix 547) can also be used, but are less versatile and restricted to four traces.

Obtaining and verifying high efficiency[3] requires some amount of diligence. The optimum efficiency values given for C1 and C2 are typical, and will vary for specific types of lamps. An important realization is that the term "lamp" includes the total load seen by the transformer's secondary. This load, reflected back to the primary, sets transformer input impedance. The transformer's input impedance forms an integral part of the LC tank that produces the high voltage drive. Because of this, circuit efficiency must be optimized with the wiring, display housing and physical layout arranged exactly the same way they will be built in production. Deviations from this procedure will result in lower efficiency than might otherwise be possible. In practice, a "first cut" efficiency optimization with "best guess" lead lengths and the intended lamp in its display housing usually produces results within 5% of the achievable figure. Final values for C1 and C2 may be established when the physical layout to be used in production has been decided on. C1 sets the circuit's resonance point, which varies to some

3. The term "efficiency" as used here applies to electrical efficiency. In fact, the ultimate concern centers around the efficient conversion of power supply energy into light. Unfortunately, lamp types show considerable deviation in their current-to-light conversion efficiency. Similarly, the emitted light for a given current varies over the life and history of any particular lamp. As such, this publication treats "efficiency" on an electrical basis; the ratio of power removed from the primary supply to the power delivered to the lamp. When a lamp has been selected, the ratio of primary supply power to lamp-emitted light energy may be measured with the aid of a photometer. This is covered in Appendix B, "Photometric Measurements." See also Appendix D, "Perspectives on Efficiency."

extent with the lamp's characteristics. C2 ballasts the lamp, effectively buffering its negative resistance characteristic. Small values of C2 provide the most load isolation, but require relatively large transformer output voltage for loop closure. Large C2 values minimize transformer output voltage, but degrade load buffering. Also, C1's "best" value is somewhat dependent on the lamp type used. Both C1 and C2 must be selected for given lamp types. Some interaction occurs, but generalized guidelines are possible. Typical values for C1 are 0.01µF to .15µF. C2 usually ends up in the 10pF to 47pF range. C1 must be a low-loss capacitor and substitution of the recommended devices is not recommended. A poor quality dielectric for C1 can easily degrade efficiency by 10%. C1 and C2 are selected by trying different values for each and iterating towards best efficiency. During this procedure, ensure that loop closure is maintained by monitoring the LT1172's feedback pin, which should be at 1.23V. Several trials usually produce the optimum C1 and C2 values. Note that the highest efficiencies are not necessarily associated with the most esthetically pleasing waveshapes, particularly at Q1, Q2, and the output.

Other issues influencing efficiency include lamp wire length and energy leakage from the lamp. The high voltage side of the lamp should have the smallest practical lead length. Excessive length results in radiative losses, which can easily reach 3% for a 3 inch wire. Similarly, no metal should contact or be in close proximity to the lamp. This prevents energy leakage, which can exceed 10%.[4]

It is worth noting that a custom designed lamp affords the best possible results. A jointly tailored lamp-circuit combination permits precise optimization of circuit operation, yielding highest efficiency.

Special attention should be given to the layout of the circuit board, since high voltage is generated at the output. The output coupling capacitor must be carefully located to minimize leakage paths on the circuit board. A slot in the board will further minimize leakage. Such leakage can permit current flow outside the feedback loop, wasting power. In the worst case, long term contamination build-up can increase leakage inside the loop, resulting in starved lamp drive or destructive arcing. It is good practice for minimization of leakage to break the silk screen line which outlines transformer T1. This prevents leakage from the high voltage secondary to the primary. Another technique for minimizing leakage is to evaluate and specify the silk screen ink for its ability to withstand high voltages.

4. A very simple experiment quite nicely demonstrates the effects of energy leakage. Grasping the lamp at its low-voltage end (low field intensity) with thumb and forefinger produces almost no change in circuit input current. Sliding the thumb-forefinger combination towards the high-voltage (higher field intensity) lamp end produces progressively greater input currents. Don't touch the high-voltage lead or you may receive an electrical shock. Repeat: Do not touch the high-voltage lead or you may receive an electrical shock.

Efficiency Measurement

Once these procedures have been followed efficiency can be measured. Efficiency may be measured by determining lamp current and voltage. Measuring current involves measuring RMS voltage across a temporarily inserted 200Ω .1% resistor in the ground lead of the negative current steering diode. The lamp current is

$$\text{Ilamp} = \frac{\text{ERMS}}{200} \times 2$$

The ×2 factor is necessitated because the diode steering dumps the current to ground on negative cycles. The 200Ω value allows the RMS meter to read with a scale factor numerically identical to the total current. Once this measurement is complete, the 200Ω resistor may be deleted and the negative current steering diode again returned directly to ground. Lamp RMS voltage is measured at the lamp with a properly compensated high voltage probe. Multiplying these two results gives power in watts, which may be compared to the DC input supply E × I product. In practice, the lamp's current and voltage contain small out of phase components but their error contribution is negligible.

Both the current and voltage measurements require a wideband true RMS voltmeter. The meter must employ a thermal type RMS converter—the more common logarithmic computing type based instruments are inappropriate because their bandwidth is too low.

The previously recommended high voltage probes are designed to see a 1MΩ–10pF–22pF oscilloscope input. The RMS voltmeters have a 10 meg Ω input. This difference necessitates an impedance matching network between the probe and the voltmeter. Details on this and other efficiency measurement issues appear in Appendix A, "Achieving Meaningful Efficiency Measurements."

Layout

The physical layout of the lamp, its leads, the display housing, and other high voltage components, is an integral part of the circuit. Poor layout can easily degrade efficiency by 25%, and higher layout induced losses have been observed. Producing an optimal layout requires attention to how losses occur. Figure 11–22 begins our study by examining potential parasitic paths between the transformer's output and the lamp. Parasitic capacitance to AC ground from any point between the transformer output and the lamp creates a path for undesired current flow. Similarly, stray coupling from any point along the lamp's length to AC ground induces parasitic current flow. All parasitic current flow is wasted, causing the circuit to produce more energy to maintain the desired current flow in D1 and D2. The high-voltage path from the transformer to the display housing should be as short as possible to minimize losses. A good rule of thumb is

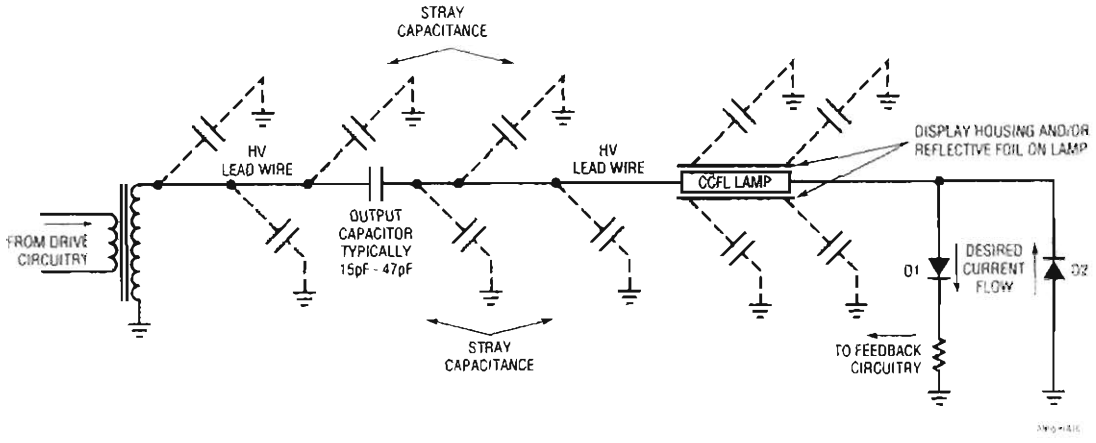


Figure 11–22. Loss paths due to stray capacitance in a practical LCD installation. Minimizing these paths is essential for good efficiency.

to assume 1% efficiency loss per inch of high voltage lead. Any PC board ground or power planes should be relieved by at least ¼" in the high voltage area. This not only prevents losses, but eliminates arcing paths.

Parasitic losses associated with lamp placement within the display housing require attention. High voltage wire length within the housing must be minimized, particularly for displays using metal construction. Ensure that the high voltage is applied to the shortest wire(s) in the display. This may require disassembling the display to verify wire length and layout. Another loss source is the reflective foil commonly used around lamps to direct light into the actual LCD. Some foil materials absorb considerably more field energy than others, creating loss. Finally, displays supplied in metal enclosures tend to be lossy. The metal absorbs significant energy and an AC path to ground is unavoidable. Direct grounding of a metal enclosed display further increases losses. Some display manufacturers have addressed this issue by relieving the metal in the lamp area with other materials.

The highest efficiency "in system" backlights have been produced by careful attention to these issues. In some cases the entire display enclosure was re-engineered for lowest losses.

Layout Considerations for Two-Lamp Designs

Systems using two lamps have some unique layout problems. Almost all two lamp displays are color units. The lower light transmission characteristics of color displays necessitate more light. Therefore, display manufacturers use two tubes to produce more light. The wiring layout of these two tube color displays affects efficiency and illumination balance in the lamps. Figure 11–23 shows an "x-ray" view of a typical display. This symmetrical arrangement presents equal parasitic losses. If C1 and C2 and the lamps are matched, the circuit's current output splits evenly and equal illumination occurs.

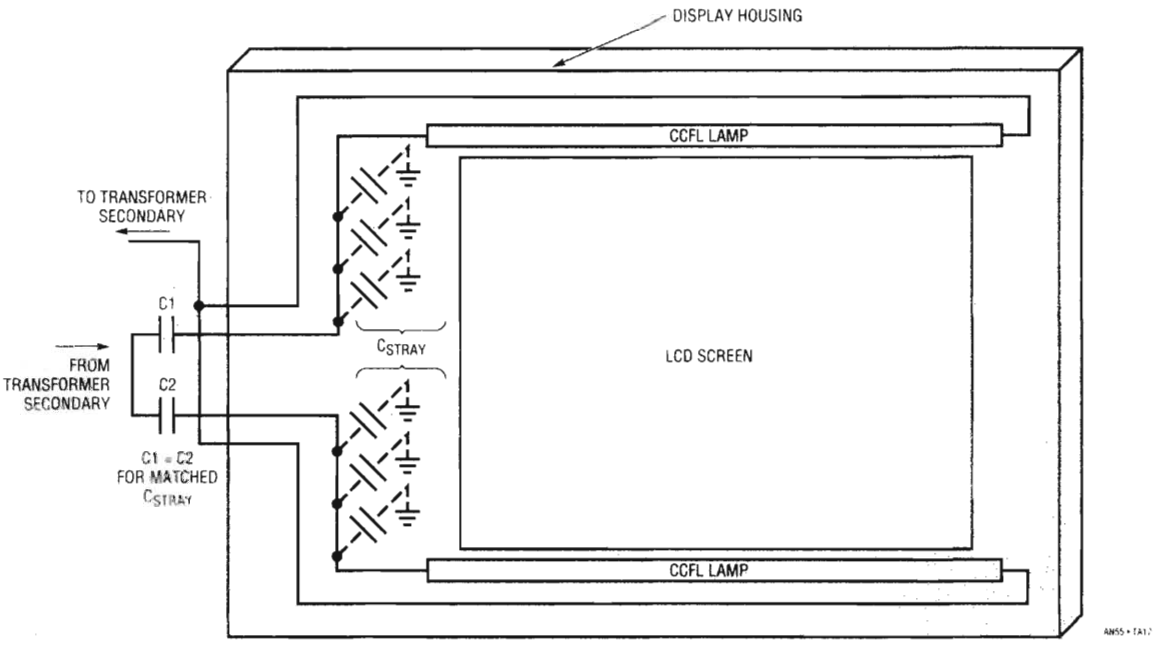


Figure 11–23. Loss paths for a "best case" dual lamp display. Symmetry promotes balanced illumination.

Figure 11–24's display arrangement is less friendly. The asymmetrical wiring forces unequal losses, and the lamps receive imbalanced current. Even with identical lamps, illumination may not be balanced. This condition is correctable by skewing C1's and C2's values. C1, because it drives greater parasitic capacitance, should be larger than C2. This tends to equalize the currents, promoting equal lamp drive. It is important to realize that this compensation does nothing to recapture the lost energy—efficiency is still compromised. There is no substitute for minimizing loss paths.

In general, imbalanced illumination causes fewer problems than might be supposed. The effect is very difficult for the eye to detect at high intensity levels. Unequal illumination is much more noticeable at lower levels. In the worst case, the dimmer lamp may only partially illuminate. This phenomenon is discussed in detail in the section "Thermometering."

Feedback Loop Stability Issues

The circuits shown to this point rely on closed loop feedback to maintain the operating point. All linear closed loop systems require some form of frequency compensation to achieve dynamic stability. Circuits operating with relatively low power lamps may be frequency compensated simply by overdamping the loop. Figures 11–18 and 11–20 use this approach. The higher power operation associated with color displays requires more attention to loop response. The transformer produces much higher output

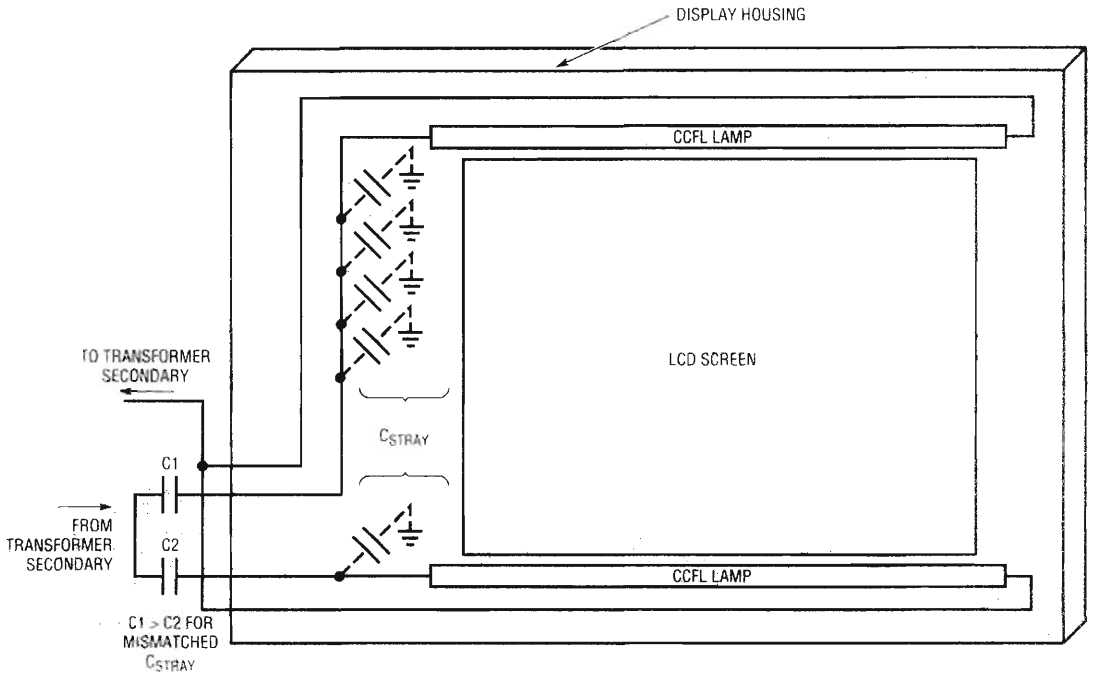


Figure 11–24. Symmetric losses in a dual lamp display. Skewing C1 and C2 values compensates imbalanced loss paths, but not wasted energy.

voltages, particularly at start-up. Poor loop damping can allow transformer voltage ratings to be exceeded, causing arcing and failure. As such, higher power designs may require optimization of transient response characteristics.

Figure 11–25 shows the significant contributors to loop transmission in these circuits. The resonant Royer converter delivers information at

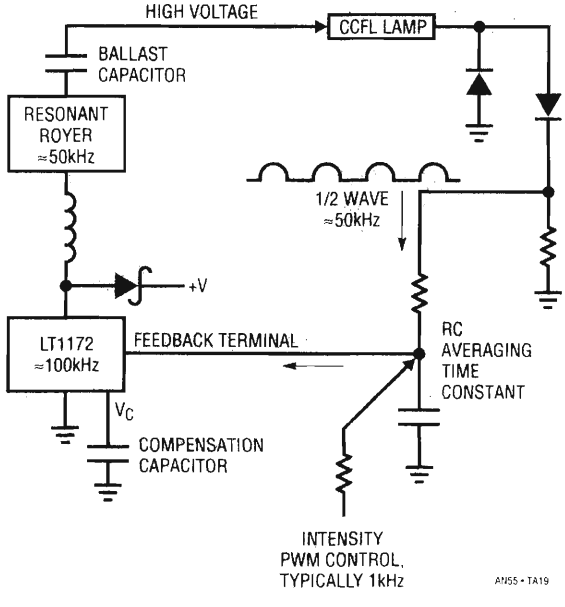


Figure 11–25. Delay terms in the feedback path. The RC time constant dominates loop transmission delay and must be compensated for stable operation.

about 50kHz to the lamp. This information is smoothed by the RC averaging time constant and delivered to the LT1172's feedback terminal as DC. The LT1172 controls the Royer converter at a 100kHz rate, closing the control loop. The capacitor at the LT1172 rolls off gain, nominally stabilizing the loop. This compensation capacitor must roll off the gain bandwidth at a low enough value to prevent the various loop delays from causing oscillation.

Which of these delays is the most significant? From a stability viewpoint, the LT1172's output repetition rate and the Royer's oscillation frequency are sampled data systems. Their information delivery rate is far above the RC averaging time constant's delay and is not significant. The RC time constant is the major contributor to loop delay. This time constant must be large enough to turn the half wave rectified waveform into DC. It also must be large enough to average any intensity control PWM signal to DC. Typically, these PWM intensity control signals come in at a 1kHz rate. The RC's resultant delay dominates loop transmission. It must be compensated by the capacitor at the LT1172. A large enough value for this capacitor rolls off loop gain at low enough frequency to provide stability. The loop simply does not have enough gain to oscillate at a frequency commensurate with the RC delay.

This form of compensation is simple and effective. It ensures stability over a wide range of operating conditions. It does, however, have poorly damped response at system turn-on. At turn-on, the RC lag delays feedback, allowing output excursions well above the normal operating point. When the RC acquires the feedback value, the loop stabilizes properly. This turn-on overshoot is not a concern if it is well within transformer breakdown ratings. Color displays, running at higher power, usually require large initial voltages. If loop damping is poor, the overshoot may be dangerously high. Figure 11–26 shows such a loop responding to turn-on. In this case the RC values are 10kΩ and 4.7μf, with a 2μf compensation capacitor. Turn-on overshoot exceeds 3500 volts for over 10

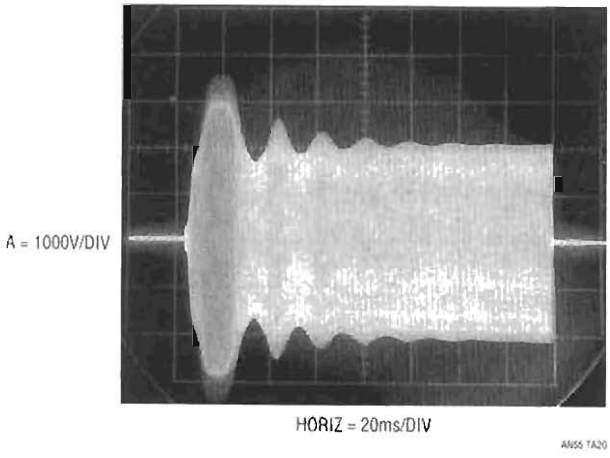


Figure 11–26. Destructive high voltage overshoot and ring-off due to poor loop compensation. Transformer failure and field recall are nearly certain. Job loss may also occur.

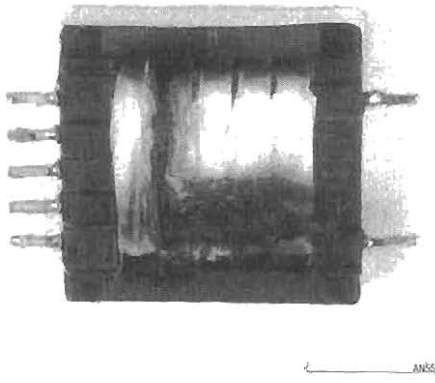


Figure 11–27. Poor loop compensation caused this transformer failure. Arc occurred in high voltage secondary (lower right). Resultant shorted turns caused overheating.

milliseconds! Ring-off takes over 100 milliseconds before settling occurs. Additionally, an inadequate (too small) ballast capacitor and excessively lossy layout force a 2000 volt output once loop settling occurs. This photo was taken with a transformer rated well below this figure. The resultant arcing caused transformer destruction, resulting in field failures. A typical destroyed transformer appears in Figure 11–27.

Figure 11–28 shows the same circuit, with the RC values reduced to 10kΩ and 1μf. The ballast capacitor and layout have also been optimized. Figure 11–28 shows peak voltage reduced to 2.2 kilovolts with duration down to about 2 milliseconds. Ring-off is also much quicker, with lower amplitude excursion. Increased ballast capacitor value and wiring layout optimization reduce running voltage to 1300 volts. Figure 11–29's results are even better. Changing the compensation capacitor to a 3kΩ–2μf network introduces a leading response into the loop, allowing faster acquisition. Now, turn-on excursion is slightly lower, but greatly reduced in duration. The running voltage remains the same.

The photos show that changes in compensation, ballast value, and layout result in dramatic reductions in overshoot amplitude and duration. Figure 11–26's performance almost guarantees field failures, while Figures 11–28 and 11–29 do not overstress the transformer. Even with

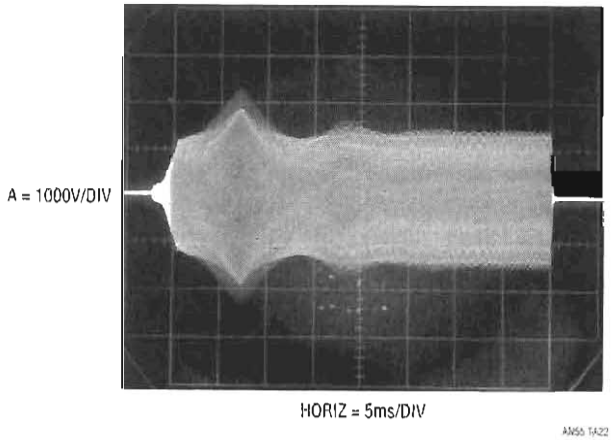


Figure 11–28. Reducing RC time constant improves transient response, although peaking, ring-off, and run voltage are still excessive.

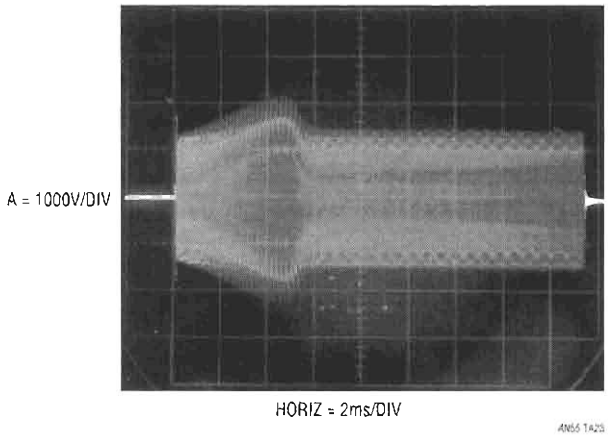


Figure 11–29. Additional optimization of RC time constant and compensation capacitor reduces turn-on transient. Run voltage is large, indicating possible lossy layout and display.

the improvements, more margin is possible if display losses can be controlled. Figures 11–26–11–29 were taken with an exceptionally lossy display. The metal enclosure was very close to the foil wrapped lamps, causing large losses with subsequent high turn-on and running voltages. If the display is selected for lower losses, performance can be greatly improved.

Figure 11–30 shows a low loss display responding to turn-on with a 2μf compensation capacitor and 10kΩ-1μf RC values. Trace A is the transformer's output while Traces B and C are the LT1172's Vcompensation and feedback pins, respectively. The output overshoots and rings badly, peaking to about 3000 volts. This activity is reflected by overshoots at the Vcompensation pin (the LT1172's error amplifier output) and the feedback pin. In Figure 11–31, the RC is reduced to 10kΩ–.1μf. This substantially reduces loop delay. Overshoot goes down to only 800 volts—a reduction of almost a factor of four. Duration is also much shorter. The Vcompensation and feedback pins reflect this tighter control. Damping is much better, with slight overshoot induced at turn-on. Further reduction of the RC to 10kΩ–.01μf (Figure 11–32) results in even faster loop capture, but a new problem appears. In Trace A, lamp turn on is so fast that the overshoot does not register in the photo. The

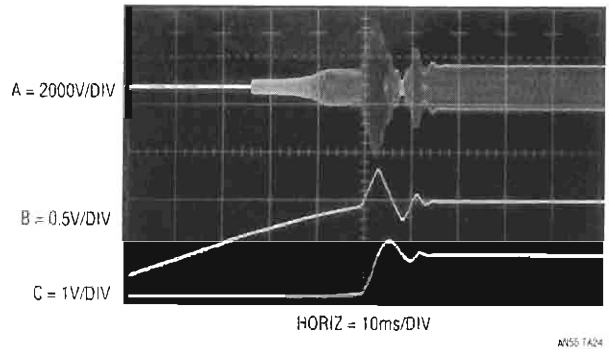


Figure 11–30. Waveforms for a lower loss layout and display. High voltage overshoot (Trace A) is reflected at compensation node (Trace B) and feedback pin (Trace C).

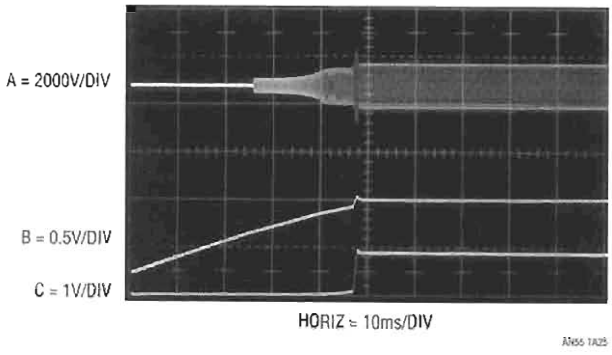


Figure 11–31. Reducing RC time constant produces quick, clean loop behavior. Low loss layout and display result in 650 VRMS running voltage.

Vcompensation (Trace B) and feedback nodes (Trace C) reflect this with exceptionally fast response. Unfortunately, the RC's light filtering causes ripple to appear when the feedback node settles. As such, Figure 11–31's RC values are probably more realistic for this situation.

The lesson from this exercise is clear. The higher voltages involved in color displays mandate attention to transformer outputs. Under running conditions, layout and display losses can cause higher loop compliance voltages, degrading efficiency and stressing the transformer. At turn-on, improper compensation causes huge overshoots, resulting in possible transformer destruction. Isn't a day of loop and layout optimization worth a field recall?

Extending Illumination Range

Lamps operating at relatively low currents may display the "thermometer effect," that is, light intensity may be nonuniformly distributed along lamp length. Figure 11–33 shows that although lamp current density is uniform, the associated field is imbalanced. The field's low intensity, combined with its imbalance, means that there is not enough energy to maintain uniform phosphor glow beyond some point. Lamps displaying the thermometer effect emit most of their light near the positive electrode, with rapid emission fall-off as distance from the electrode increases.

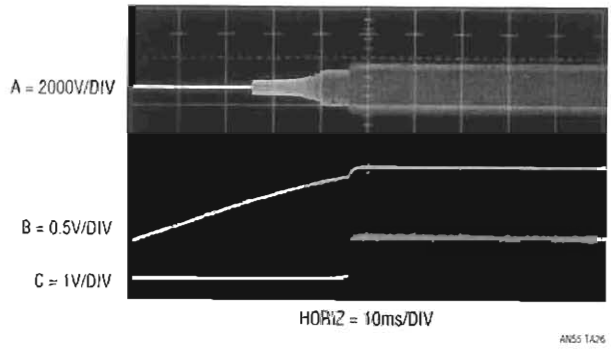


Figure 11–32. Very low RC value provides even faster response, but ripple at feedback pin (Trace C) is too high. Figure 11–31 is the best compromise.

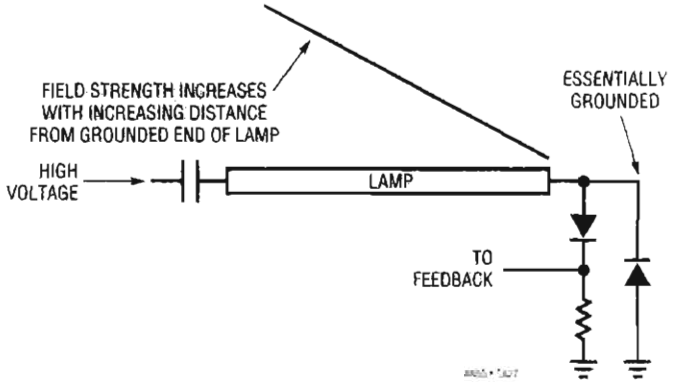
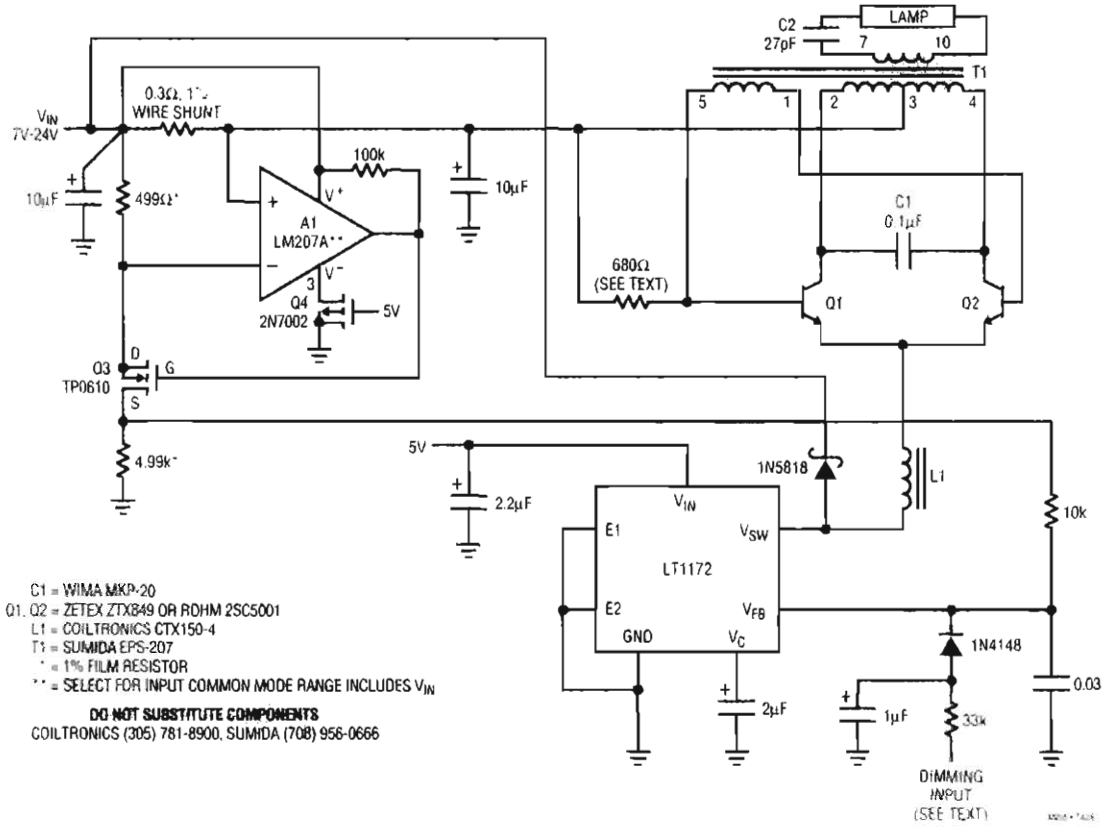


Figure 11–33. Field strength vs. distance for a ground referred lamp. Field imbalance promotes uneven illumination at low drive levels.

Placing a conductor along the lamp's length largely alleviates "thermometering." The trade-off is decreased efficiency due to energy leakage (see Note 4 and associated text). It is worth noting that various lamp types have different degrees of susceptibility to the thermometer effect.

Some displays require an extended illumination range. "Thermometering" usually limits the lowest practical illumination level. One acceptable way to minimize "thermometering" is to eliminate the large



field imbalance. Figure 11–34's circuit does this. This circuit's most significant aspect is that the lamp is fully floating—there is no galvanic connection to ground as in the previous designs. This allows T1 to deliver symmetric, differential drive to the lamp. Such balanced drive eliminates field imbalance, reducing thermometering at low lamp currents. This approach precludes any feedback connection to the now floating output. Maintaining closed loop control necessitates deriving a feedback signal from some other point. In theory, lamp current proportions to T1's or L1's drive level, and some form of sensing this can be used to provide feedback. In practice, parasitics make a practical implementation difficult.[5]

Figure 11–34 derives the feedback signal by measuring Royer converter current and feeding this information back to the LT1172. The Royer's drive requirement closely proportions to lamp current under all conditions. A1 senses this current across the .3Ω shunt and biases Q3, closing a local feedback loop. Q3's drain voltage presents an amplified, single ended version of the shunt voltage to the feedback point, closing the main loop. The lamp current is not as tightly controlled as before, but .5% regulation over wide supply ranges is possible. The dimming in this circuit is controlled by a 1kHz PWM signal. Note the heavy filtering (33kΩ–2μf) outside the feedback loop. This allows a fast time constant, minimizing turn-on overshoot.[6]

In all other respects, operation is similar to the previous circuits. This circuit typically permits the lamp to operate over a 40:1 intensity range without "thermometering." The normal feedback connection is usually limited to a 10:1 range.

The losses introduced by the current shunt and A1 degrade overall efficiency by about 2%. As such, circuit efficiency is limited to about 90%. Most of the loss can be recovered at moderate cost in complexity. Figure 11–35's modifications reduce shunt and A1 losses. A1, a precision micropower type, cuts power drain and permits a smaller shunt value without performance degradation. Unfortunately, A1 does not function when its inputs reside at the V+ rail. Because the circuit's operation requires this, some accommodation must be made.[7]

At circuit start-up, A1's input is pulled to its supply pin potential (actually, slightly above it). Under these conditions, A1's input stage is shut off. Normally, A1's output state would be indeterminate but, for the amplifier specified, it will always be high. This turns off Q3, permitting the LT1172 to drive the Royer stage. The Royer's operation causes Q1's collector swing to exceed the supply rail. This turns on the 1N4148, the BAT-85 goes off, and A1's supply pin rises above the supply rail. This "bootstrapping" action results in A1's inputs being biased within the am-

5. See Appendix C, "A Lot of Cut-Off-Ears and No Van Goghs—Some Not-So-Great Ideas," for details.
6. See section "Feedback Loop Stability Issues."
7. In other words, we need a hack.

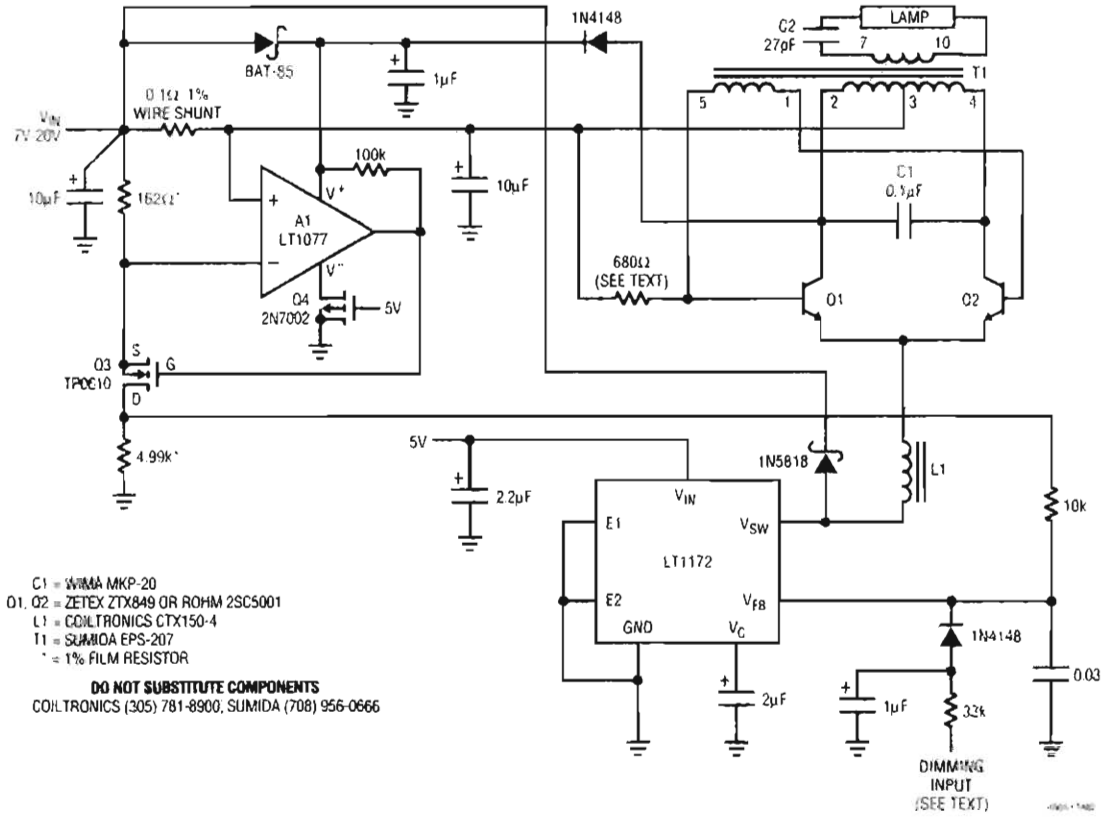


Figure 11–35. The "low thermometer" circuit using a micropower, precision topside sensing amplifier. Supply bootstrapping eliminates input common mode requirement, permitting a 1.6% efficiency gain.

plifier's common mode range, and normal circuit operation commences. The result of all this is a 1.6% efficiency gain, permitting an overall circuit efficiency of just below 92%.

Epilogue

Our understanding with Apple Computer gave them six months sole use of everything I learned while working with them. After that, we were free to disclose the circuit and most attendant details to anyone else, which we did. It found immediate use in other computers and applications, ranging from medical equipment to automobiles, gas pumps, retail terminals and anywhere else LCD displays are used. The development work consumed about 20 months, ending in August, 1993. Upon its completion I immediately fell into a rut, certain I would never do anything worthwhile again.

References

1. Blake, James W. *The Sidewalks of New York.* (1894).
2. Bright, Pittman, and Royer. "Transistors As On-Off Switches in Saturable Core Circuits." *Electrical Manufacturing* (December 1954): Available from Technomic Publishing, Lancaster, PA.
3. Sharp Corporation. *Flat Panel Displays.* (1991).
4. Kitchen, C., and L. Counts. *RMS-to-DC Conversion Guide.* Analog Devices, Inc. (1986).
5. Williams, Jim. "A Monolithic IC for 100MHz RMS-DC Conversion." Linear Technology Corporation, *Application Note 22* (September 1987).
6. Hewlett-Packard. "1968 Instrumentation. Electronic Analytical-Medical." *AC Voltage Measurement* (1968): 197–198.
7. Hewlett-Packard. *Model 3400RMS Voltmeter Operating and Service Manual.* (1965).
8. Hewlett-Packard. *Model 3403C True RMS Voltmeter Operating and Service Manual.* (1973).
9. Ott, W.E. "A New Technique of Thermal RMS Measurement." *IEEE Journal of Solid State Circuits* (December 1974).
10. Williams, J.M., and T.L. Longman. "A 25MHz Thermally Based RMS-DC Converter." *IEEE ISSCC Digest of Technical Papers* (1986).
11. O'Neill, P.M. "A Monolithic Thermal Converter." *H.P. Journal* (May 1980).
12. Williams, J. "Thermal Technique in Measurement and Control Circuitry," "50MHz Thermal RMS-DC Converter." Linear Technology Corporation, *Application Note 5* (December 1984).
13. Williams, J., and B. Huffman. "Some Thoughts on DC-DC Converters": Appendix A, "The +5 to 10 ±15V Converter—A Special Case." Linear Technology Corporation, *Application Note 29* (October 1988).
14. Baxendall, P.J. "Transistor Sine-Wave LC Oscillators." *British Journal of IEEE* (February 1960): Paper No. 2978E.
15. Williams, J. "Temperature Controlling to Microdegrees." Massachusetts Institute of Technology, Education Research Center (1971): out of print.
16. Fulton, S.P. "The Thermal Enzyme Probe." Thesis, Massachusetts Institute of Technology (1975).
17. Williams, J. "Designer's Guide to Temperature Measurement." *EDN* part II (May 20, 1977).
18. Williams. J. "Illumination Circuitry for Liquid Crystal Displays." Linear Technology Corporation, *Application Note 49* (August 1992).
19. Olsen, J.V. "A High Stability Temperature Controlled Oven." Thesis, Massachusetts Institute of Technology (1974).
20. MIT Reports on Research. *The Ultimate Oven.* (March 1972).
21. McDermott, James. "Test System at MIT Controls Temperature of Microdegrees." *Electronic Design* (January 6, 1972).
22. Williams, Jim. "Techniques for 92% Efficient LCD Illumination." Linear Technology Corporation, *Application Note 55* (August 1993).

Appendix A

Achieving Meaningful Efficiency Measurements

Obtaining reliable efficiency data for the CCFL circuits presents a high order difficulty measurement problem. Establishing and maintaining accurate AC measurements is a textbook example of attention to measurement technique. The combination of high frequency, harmonic laden waveforms and high voltage makes meaningful results difficult to obtain. The choice, understanding, and use of test instrumentation is crucial. Clear thinking is needed to avoid unpleasant surprises![1]

Probes

The probes employed must faithfully respond over a variety of conditions. Measuring across the resistor in series with the CCFL is the most favorable circumstance. This low voltage, low impedance measurement allows use of a standard 1X probe. The probe's relatively high input capacitance does not introduce significant error. A 10X probe may also be used, but frequency compensation issues (discussion to follow) must be attended to.

The high voltage measurement across the lamp is considerably more demanding on the probe. The waveform fundamental is at 20kHz to 100kHz, with harmonics into the MHz region. This activity occurs at peak voltages in the kilovolt range. The probe must have a high fidelity response under these conditions. Additionally, the probe should have low input capacitance to avoid loading effects which would corrupt the measurement. The design and construction of such a probe requires significant attention. Figure 11–A1 lists some recommended probes along with their characteristics. As stated in the text, almost all standard oscilloscope probes will fail[2] if used for this measurement. Attempting to circumvent the probe requirement by resistively dividing the lamp voltage also creates problems. Large value resistors often have significant voltage coefficients and their shunt capacitance is high and uncertain. As such, simple voltage dividing is not recommended. Similarly, common high voltage probes intended for DC measurement will have large errors because of AC effects. The P6013A and P6015 are the favored probes; their 100MΩ input and small capacitance introduces low loading error. The penalty for their 1000X attenuation is reduced output, but the recommended voltmeters (discussion to follow) can accommodate this.

All of the recommended probes are designed to work into an oscilloscope input. Such inputs are almost always 1MΩ paralleled by (typically)

1. It is worth considering that various constructors of Figure 11–18 have reported efficiencies ranging from 8% to 115%.
2. That's twice I've warned you nicely.

10pF–22pF. The recommended voltmeters, which will be discussed, have significantly different input characteristics. Figure 11–A2's table shows higher input resistances and a range of capacitances. Because of this the probe must be compensated for the voltmeter's input characteristics. Normally, the optimum compensation point is easily determined and adjusted by observing probe output on an oscilloscope. A known-amplitude square wave is fed in (usually from the oscilloscope calibrator) and the probe adjusted for correct response. Using the probe with the voltmeter presents an unknown impedance mismatch and raises the problem of determining when compensation is correct.

The impedance mismatch occurs at low and high frequency. The low frequency term is corrected by placing an appropriate value resistor in shunt with the probe's output. For a 10MΩ voltmeter input, a 1.1MΩ resistor is suitable. This resistor should be built into the smallest possible BNC equipped enclosure to maintain a coaxial environment. No cable connections should be employed; the enclosure should be placed directly between the probe output and the voltmeter input to minimize stray capacitance. This arrangement compensates the low frequency impedance mismatch. Figure 11–A4 shows the impedance-matching box attached to the high voltage probe.

Correcting the high frequency mismatch term is more involved. The wide range of voltmeter input capacitances combined with the added shunt resistor's effects presents problems. How is the experimenter to know where to set the high frequency probe compensation adjustment? One solution is to feed a known value RMS signal to the probe-voltmeter combination and adjust compensation for a proper reading. Figure 11–A3 shows a way to generate a known RMS voltage. This scheme is simply a standard backlight circuit reconfigured for a constant voltage output. The op amp permits low RC loading of the 5.6K feedback termination without introducing bias current error. The 5.6kΩ value may be series or parallel trimmed for a 300V output. Stray parasitic capacitance in the feedback network affects output voltage. Because of this, all feedback associated nodes and components should be rigidly fixed and the entire circuit built into a small metal box. This prevents any significant change in the parasitic terms. The result is a known $300V_{RMS}$ output.

Now, the probe's compensation is adjusted for a 300V voltmeter indication, using the shortest possible connection (e.g., BNC-to-probe adapter) to the calibrator box. This procedure, combined with the added resistor, completes the probe-to-voltmeter impedance match. If the probe compensation is altered (e.g., for proper response on an oscilloscope) the voltmeter's reading will be erroneous.[3] It is good practice to verify the

3. The translation of this statement is to hide the probe when you are not using it. If anyone wants to borrow it, look straight at them, shrug your shoulders, and say you don't know where it is. This is decidedly dishonest, but eminently practical. Those finding this morally questionable may wish to reexamine their attitude after producing a day's worth of worthless data with a probe that was unknowingly readjusted.

Figure 11–A1. Characteristics of some wideband high voltage probes. Output impedances are designed for oscilloscope inputs.

TEKTRONIX PROBE TYPE	ATTENUATION FACTOR	ACCURACY	INPUT RESISTANCE	INPUT CAPACITANCE	RISE TIME	BAND-WIDTH	MAXIMUM VOLTAGE	DERATED ABOVE	DERATED TO AT FREQUENCY	COMPENSATION RANGE	ASSUMED TERMINATION RESISTANCE
P6007	100X	3%	10MΩ	2.2pF	14ns	25MHz	1.5kV	200kHz	$700V_{RMS}$ at 10MHz	15-55pF	1M
P6009	100X	3%	10MΩ	2.5pF	2.9ns	120MHz	1.5kV	200kHz	$450V_{RMS}$ at 40MHz	15-47pF	1M
P6013A	1000X	Adjustable	100MΩ	3pF	7ns	50MHz	12kV	100kHz	$800V_{RMS}$ at 20MHz	12-60pF	1M
P6015	1000X	Adjustable	100MΩ	3pF	1.4ns	250MHz	20kV	100kHz	$2000V_{RMS}$ at 20MHz	12-47pF	1M

Figure 11–A2. Pertinent characteristics of some thermally based RMS voltmeters. Input impedances necessitate matching network and compensation for high voltage probes.

MANUFACTURER AND MODEL	FULL SCALE RANGES	ACCURACY AT 1MHz	ACCURACY AT 100kHz	INPUT RESISTANCE AND CAPACITANCE	MAXIMUM BANDWIDTH	CREST FACTOR
Hewlett-Packard 3400 Meter Display	1mV to 300V, 12 Ranges	1%	1%	0.001V to 0.3V Range = 10M and < 50pF, 1V to 300V Range = 10M and < 20pF	10MHz	10:1 At Full Scale, 100:1 At 0.1 Scale
Hewlett-Packard 3403C Digital Display	10mV to 1000V, 6 Ranges	0.5%	0.2%	10mV and 100mV Range = 20M and 20pF ±10%, 1V to 1000V Range = 10M and 24pF ±10%	100MHz	10:1 At Full Scale, 100:1 At 0.1 Scale
Fluke 8920A Digital Display	2mV to 700V, 7 Ranges	0.7%	0.5%	10M and < 30pF	20MHz	7:1 At Full Scale 70:1 At 0.1 Scale

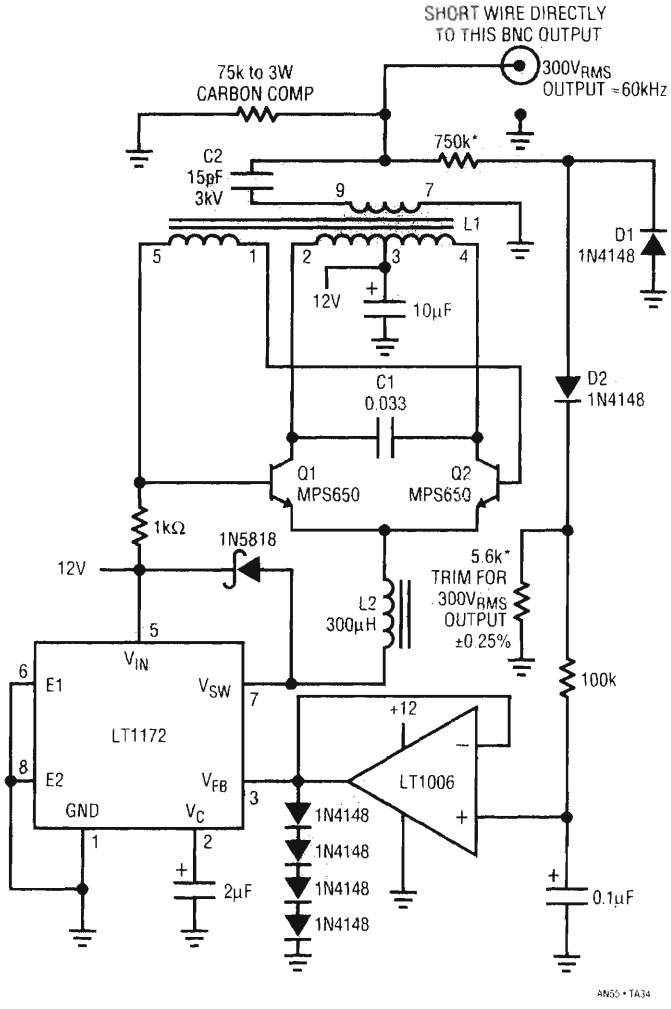


Figure 11–A3. High voltage RMS calibrator is voltage output version of CCFL circuit.

C1 = MUST BE A LOW LOSS CAPACITOR.
METALIZED POLYCARB
WIMA FKP2 OR MKP-20 (GERMAN) RECOMMENDED
L1 = SUMIDA 6345-020 OR COILTRONICS CTX110092-1
PIN NUMBERS SHOWN FOR COILTRONICS UNIT
L2 = COILTRONICS CTX300-4
Q1, Q2 = AS SHOWN OR BCP 56 (PHILLIPS SO PACKAGE)
* = 1% FILM RESISTOR (10kΩ TO 75kΩ RESISTORS IN SERIES)
DO NOT SUBSTITUTE COMPONENTS

COILTRONICS (305) 781-8900, SUMIDA (708) 956-0666

calibrator box output before and after every set of efficiency measurements. This is done by directly connecting, via BNC adapters, the calibrator box to the RMS voltmeter on the 1000V range.

RMS Voltmeters

The efficiency measurements require an RMS responding voltmeter. This instrument must respond accurately at high frequency to irregular and harmonically loaded waveforms. These considerations eliminate almost all AC voltmeters, including DVMs with AC ranges.

Figure 11–A4. The impedance matching box (extreme left) mated to the high voltage probe. Note direct connection. No cable is used.

There are a number of ways to measure RMS AC voltage. Three of the most common include average, logarithmic, and thermally responding. Averaging instruments are calibrated to respond to the average value of the input waveform, which is almost always assumed to be a sine wave. Deviation from an ideal sine wave input produces errors. Logarithmically based voltmeters attempt to overcome this limitation by continuously computing the input's true RMS value. Although these instruments are "real time" analog computers, their 1% error bandwidth is well below 300kHz and crest factor capability is limited. Almost all general purpose DVMs use such a logarithmically based approach and, as such, are not suitable for CCFL efficiency measurements. Thermally based RMS voltmeters are direct acting thermo-electronic analog computers. They respond to the input's RMS heating value. This technique is explicit, relying on the very definition of RMS (e.g., the heating power of the waveform). By turning the input into heat, thermally based instruments achieve vastly higher bandwidth than other techniques.[4] Additionally, they are insensitive to waveform shape and easily accommodate large crest factors. These characteristics are necessary for the CCFL efficiency measurements.

Figure 11–A5 shows a conceptual thermal RMS-DC converter. The input waveform warms a heater, resulting in increased output from its associated temperature sensor. A DC amplifier forces a second, identical, heater-sensor pair to the same thermal conditions as the input driven pair. This differentially sensed, feedback enforced loop makes ambient temperature shifts a common mode term, eliminating their effect. Also, although the voltage and thermal interaction is non-linear, the input-output RMS voltage relationship is linear with unity gain.

The ability of this arrangement to reject ambient temperature shifts depends on the heater-sensor pairs being isothermal. This is achievable by thermally insulating them with a time constant well below that of ambient shifts. If the time constants to the heater-sensor pairs are matched, ambient temperature terms will affect the pairs equally in phase and amplitude.

4. Those finding these descriptions intolerably brief are commended to references 4, 5, and 6.

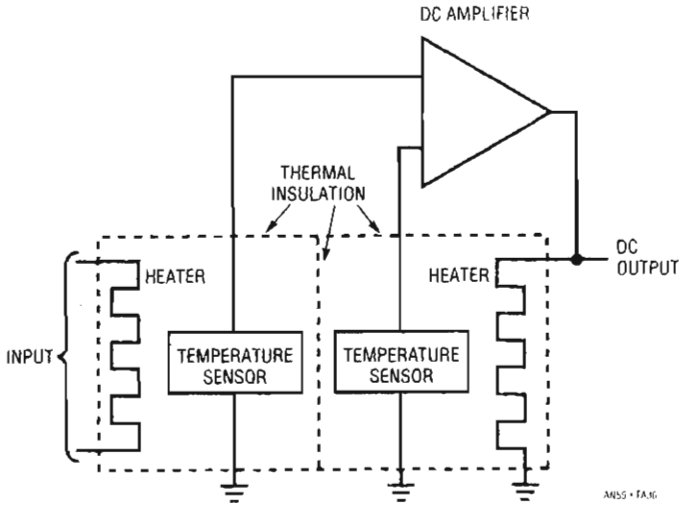


Figure 11–A5. Conceptual thermal RMS-DC converter.

The DC amplifier rejects this common mode term. Note that, although the pairs are isothermal, they are insulated from each other. Any thermal interaction between the pairs reduces the system's thermally based gain terms. This would cause unfavorable signal-to-noise performance, limiting dynamic operating range.

Figure 11–A5's output is linear because the matched thermal pair's nonlinear voltage-temperature relationships cancel each other.

The advantages of this approach have made its use popular in thermally based RMS-DC measurements.

The instruments listed in Figure 11–A2, while considerably more expensive than other options, are typical of what is required for meaningful results. The HP3400A and the Fluke 8920A are currently available from their manufacturers. The HP3403C, an exotic and highly desirable instrument, is no longer produced but readily available on the secondary market.

Figure 11–A6 shows equipment in a typical efficiency test setup. The RMS voltmeters (photo center and left) read output voltage and current via high voltage (left) and standard 1X probes (lower left). Input voltage is read on a DVM (upper right). A low loss clip-on ammeter (lower right) determines input current. The CCFL circuit and LCD display are in the foreground. Efficiency, the ratio of input to output power, is computed with a hand held calculator (lower right).

Calorimetric Correlation of Electrical Efficiency Measurements

Careful measurement technique permits a high degree of confidence in the accuracy of the efficiency measurements. It is, however, a good idea to check the method's integrity by measuring in a completely different domain. Figure 11–A7 does this by calorimetric techniques. This arrangement, identical to the thermal RMS voltmeter's operation (Figure 11–A5),

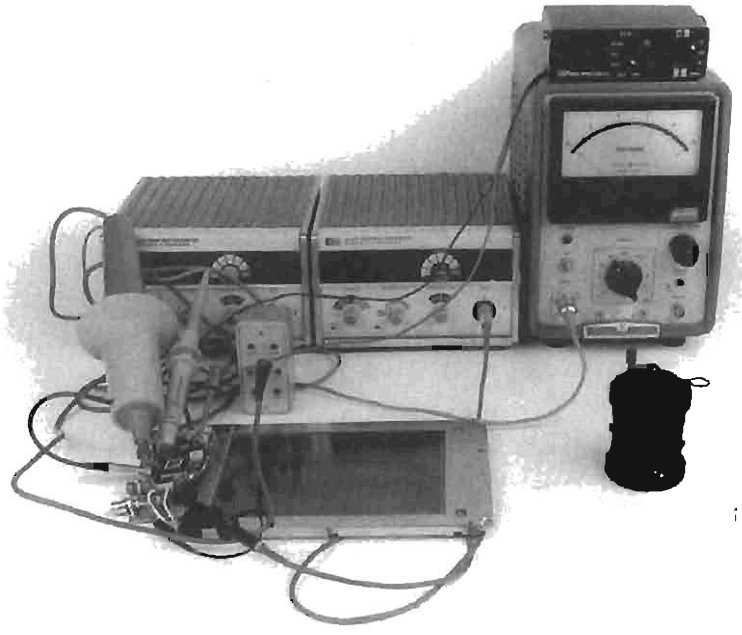


Figure 11–A6. Typical efficiency measurement instrumentation. RMS voltmeters (center left) measure output voltage and current via appropriate probes. Clip-on ammeter (right) gives low loss input current readings. DVM (upper right) measures input voltage. Hand calculator (lower right) is used to compute efficiency.

determines power delivered by the CCFL circuit by measuring its load temperature rise. As in the thermal RMS voltmeter, a differential approach eliminates ambient temperature as an error term. The differential amplifier's output, assuming a high degree of matching in the two thermal enclosures, proportions to load power. The ratio of the two cells' $E \times I$ products yields efficiency information. In a 100% efficient system, the amplifier's output energy would equal the power supplies' output. Practically it is always less, as the CCFL circuit has losses. This term represents the desired efficiency information.

Figure 11–A8 is similar except that the CCFL circuit board is placed within the calorimeter. This arrangement nominally yields the same information, but is a much more demanding measurement because far less heat is generated. The signal-to-noise (heat rise above ambient) ratio is unfavorable, requiring almost fanatical attention to thermal and instru-

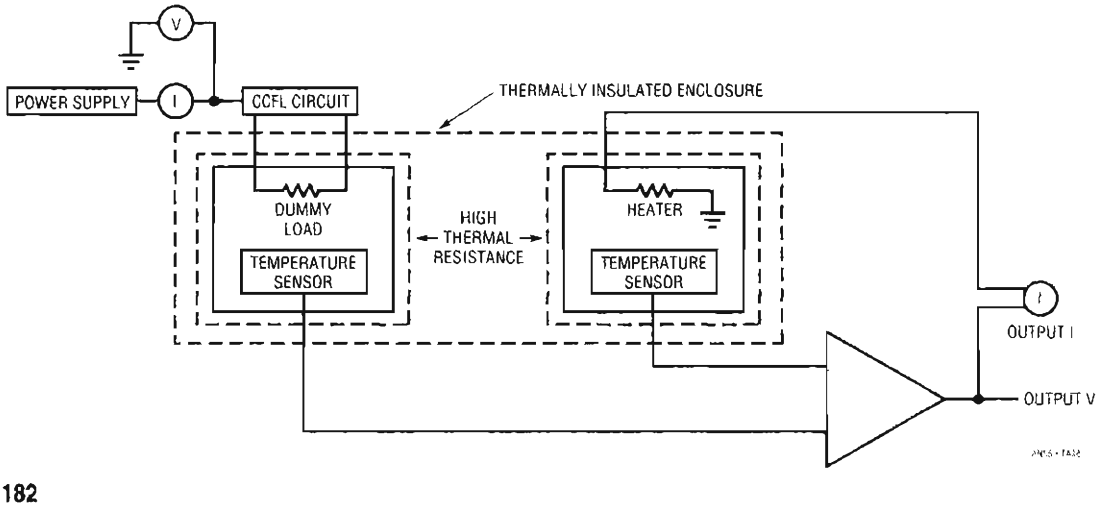


Figure 11–A7. Efficiency determination via calorimetric measurement. Ratio of power supply to output energy gives efficiency information.

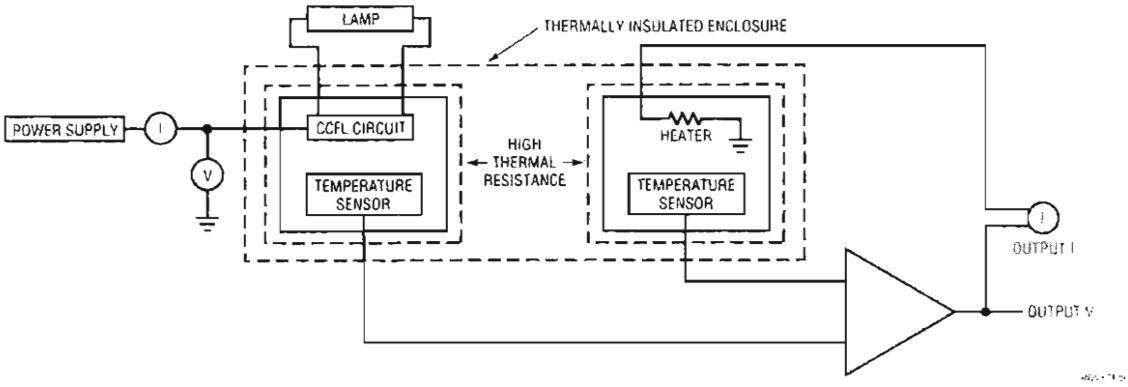


Figure 11–A8. The calorimeter measures efficiency by determining circuit heating losses.

mentation considerations.[5] It is significant that the total uncertainty between electrical and both calorimetric efficiency determinations was 3.3%. The two thermal approaches differed by about 2%. Figure 11–A9 shows the calorimeter and its electronic instrumentation. Descriptions of this instrumentation and thermal measurements can be found in the References section following the main text.

5. Calorimetric measurements are not recommended for readers who are short on time or sanity.

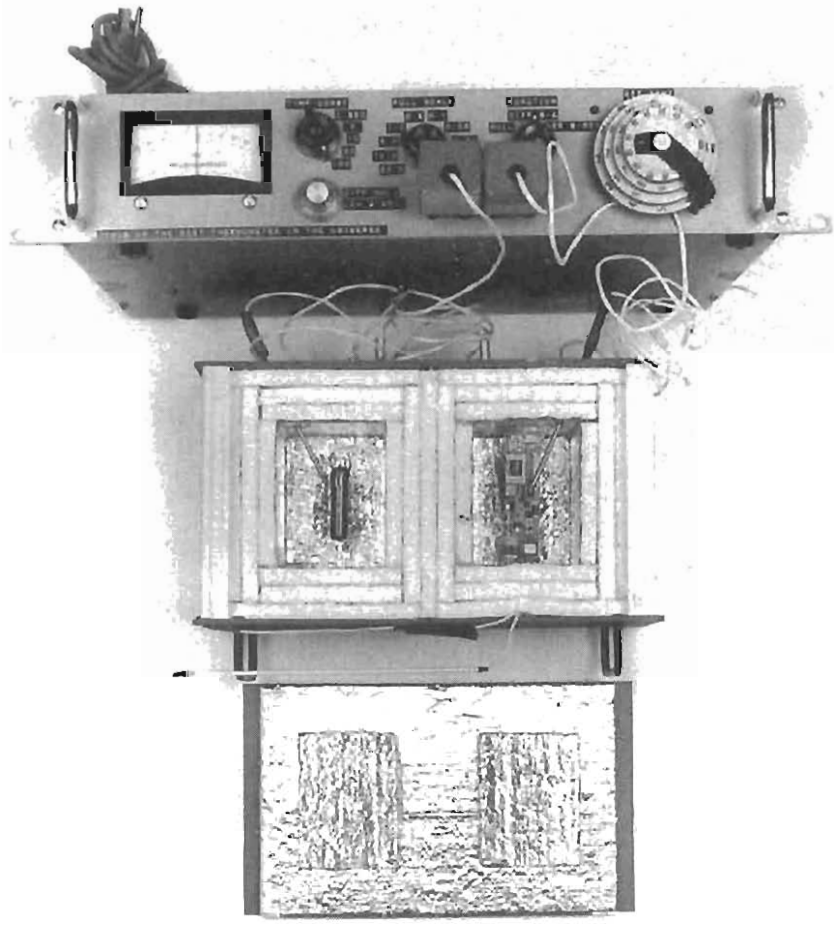


Figure 11–A9. The calorimeter (center) and its instrumentation (top). Calorimeter's high degree of thermal symmetry combined with sensitive servo instrumentation produces accurate efficiency measurements. Lower portion of photo is calorimeter's top cover.

Appendix B

Photometric Measurements

In the final analysis the ultimate concern centers around the efficient conversion of power supply energy to light. Emitted light varies monotonically with power supply energy,[1] but certainly not linearly. In particular, bulb luminosity may be highly nonlinear, particularly at high power, vs. drive power. There are complex trade-offs involving the amount of emitted light vs. power consumption and battery life. Evaluating these trade-offs requires some form of photometer. The relative luminosity of lamps may be evaluated by placing the lamp in a light tight tube and sampling its output with photodiodes. The photodiodes are placed along the lamp's length and their outputs electrically summed. This sampling technique is an uncalibrated measurement, providing relative data only. It is, however, quite useful in determining relative bulb emittance under various drive conditions. Figure 11–B1 shows this "glometer," with its uncalibrated output appropriately scaled in "brights." The switches allow various sampling diodes along the lamp's length to be disabled. The photodiode signal conditioning electronics are mounted behind the switch panel.

Calibrated light measurements call for a true photometer. The Tektronix J-17/J1803 photometer is such an instrument. It has been found

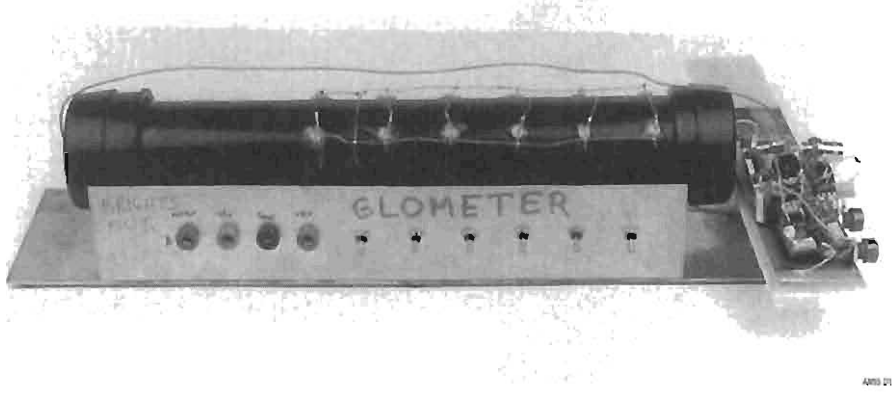


Figure 11–B1.
The "glometer" measures relative lamp emissivity. CCFL circuit mounts to the right. Lamp is inside cylindrical housing. Photodiodes (center) convert light to electrical output (lower left) via amplifiers (not visible in photo).

1. But not always! It is possible to build highly electrically efficient circuits that emit less light than "less efficient" designs. See Appendix C. "A Lot of Cut-Off Ears and No Van Goghs—Some Not-So-Great Ideas."

particularly useful in evaluating display (as opposed to simply the lamp) luminosity under various drive conditions. The calibrated output permits reliable correlation with customer results.[2] The light tight measuring head allows evaluation of emittance evenness at various display locations. This capability is invaluable when optimizing lamp location and/or ballast capacitor values in dual lamp displays.

Figure 11–B2 shows the photometer in use evaluating a display.

2. It is unlikely that customers would be enthusiastic about correlating the "brights" units produced by the aforementioned glometer.

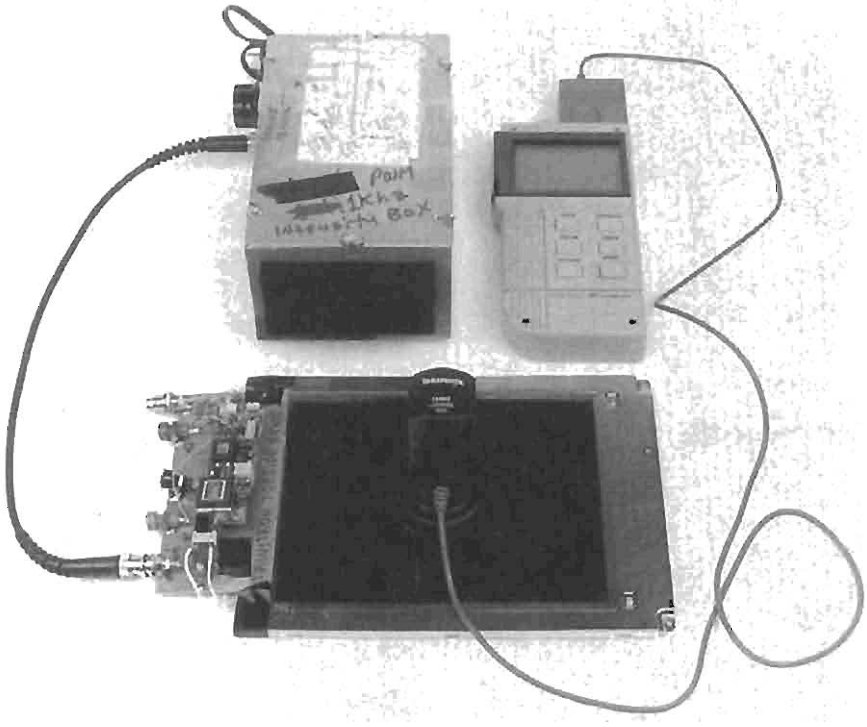


Figure 11–B2.
Apparatus for calibrated photometric display evaluation. Photometer (upper right) indicates display luminosity via sensing head (center). CCFL circuit (left) intensity is controlled by a calibrated pulse width generator (upper left).

Appendix C

A Lot of Cut-Off Ears and No Van Goghs—Some Not-So-Great Ideas

The hunt for a practical CCFL power supply covered (and is still covering) a lot of territory. The wide range of conflicting requirements combined with ill-defined lamp characteristics produces plenty of unpleasant surprises. This section presents a selection of ideas that turned into disappointing breadboards. Backlight circuits are one of the deadliest places for theoretically interesting circuits the author has ever encountered.

Not-So-Great Backlight Circuits

Figure 11–C1 seeks to boost efficiency by eliminating the LT1172's saturation loss. Comparator C1 controls a free running loop around the Royer by on-off modulation of the transistor base drive. The circuit delivers bursts of high voltage sine drive to the lamp to maintain the feedback

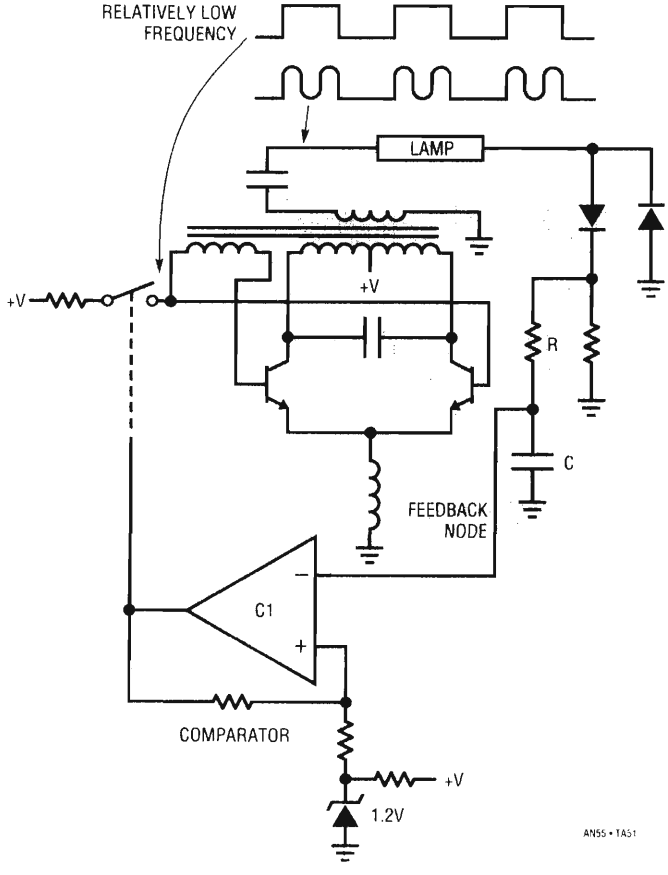


Figure 11–C1. A first attempt at improving the basic circuit. Irregular Royer drive promotes losses and poor regulation.

node. The scheme worked, but had poor line rejection, due to the varying waveform vs. supply seen by the RC averaging pair. Also, the "burst" modulation forces the loop to constantly re-start the bulb at the burst rate, wasting energy. Finally, bulb power is delivered by a high crest factor waveform, causing inefficient current-to-light conversion in the bulb.

Figure 11–C2 attempts to deal with some of these issues. It converts the previous circuit to an amplifier-controlled current mode regulator. Also, the Royer base drive is controlled by a clocked, high frequency pulse width modulator. This arrangement provides a more regular waveform to the averaging RC, improving line rejection. Unfortunately the improvement was not adequate. 1% line rejection is required to avoid annoying flicker when the line moves abruptly, such as when a charger is activated. Another difficulty is that, although reduced by the higher frequency PWM, crest factor is still non-optimal. Finally, the lamp is still forced to restart at each PWM cycle, wasting power.

Figure 11–C3 adds a "keep alive" function to prevent the Royer from turning off. This aspect worked well. When the PWM goes low, the Royer is kept running, maintaining low level lamp conduction. This eliminates the continuous lamp restarting, saving power. The "supply correc-

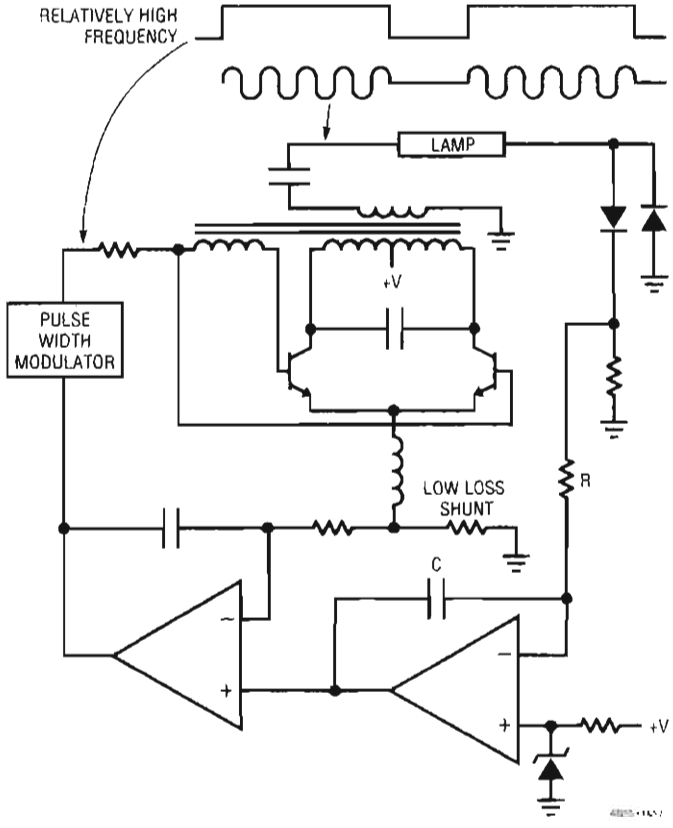


Figure 11–C2. A more sophisticated failure still has losses and poor line regulation.

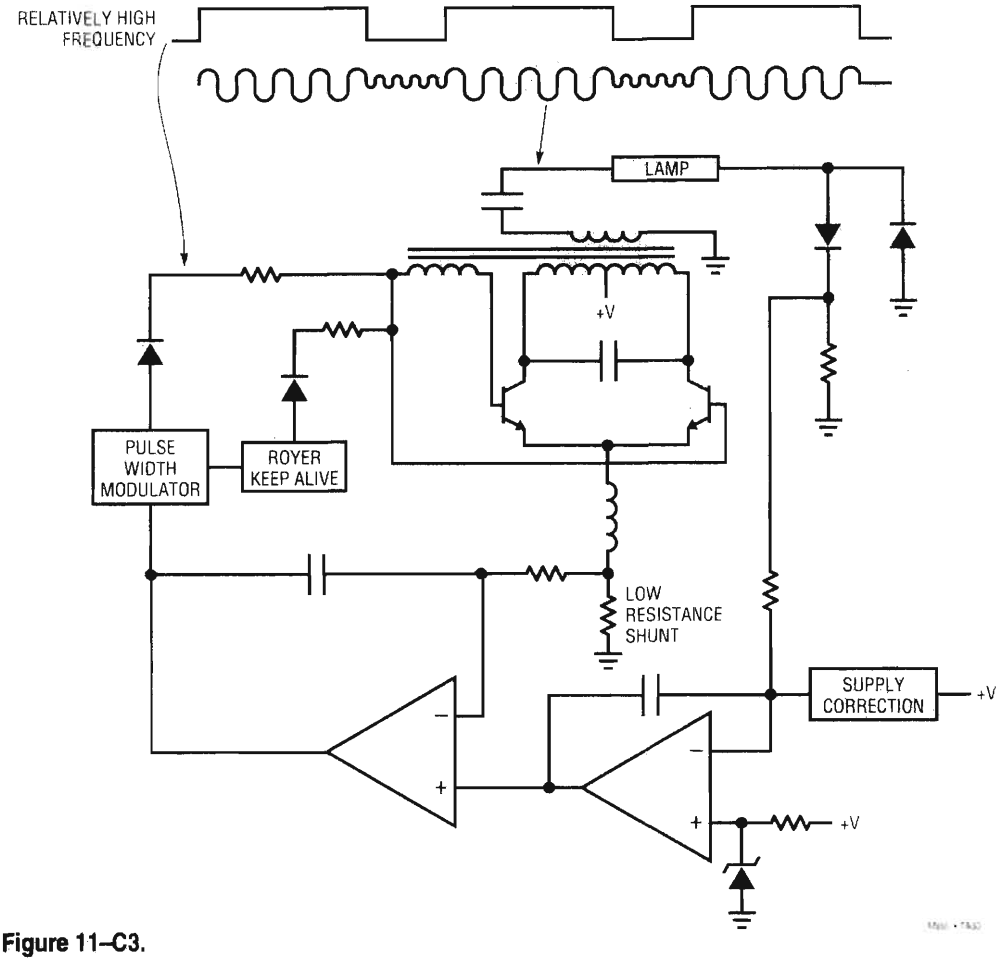


Figure 11–C3. "Keep alive" circuit eliminates turn-on losses and has 94% efficiency. Light emission is lower than "less efficient" circuits.

tion" block feeds a portion of the supply into the RC averager, improving line rejection to acceptable levels.

This circuit, after considerable fiddling, achieved almost 94% efficiency but produced less output light than a "less efficient" version of Figure 11–18! The villain is lamp waveform crest factor. The keep alive circuit helps, but the lamp still cannot handle even moderate crest factors.

Figure 11–C4 is a very different approach. This circuit is a driven square wave converter. The resonating capacitor is eliminated. The base drive generator shapes the edges, minimizing harmonics for low noise operation. This circuit works well, but relatively low operating frequencies are required to get good efficiency. This is so because the sloped drive must be a small percentage of the fundamental to maintain low losses. This mandates relatively large magnetics—a crucial disadvantage. Also, square waves have a different crest factor and rise time than sines, forcing inefficient lamp transduction.

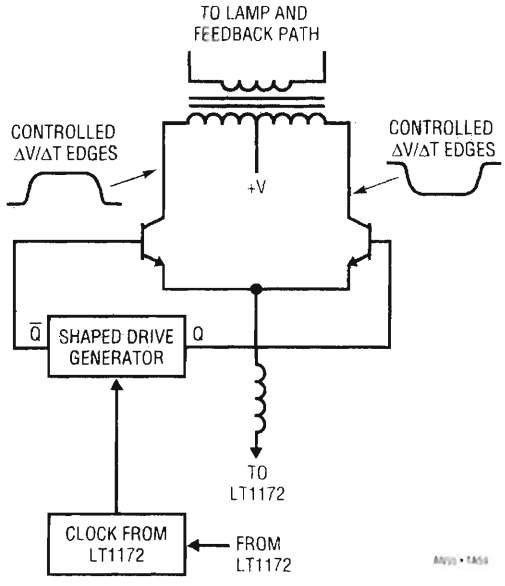


Figure 11–C4. A non-resonant approach. Slew retarded edges minimize harmonics, but transformer size goes up. Output waveform is also non-optimal, causing lamp losses.

Not-So-Great Primary Side Sensing Ideas

Figures 11–34 and 11–35 use primary side current sensing to control bulb intensity. This permits the bulb to fully float, extending its dynamic operating range. A number of primary side sensing approaches were tried before the "topside sense" won the contest.

Figure 11–C5's ground referred current sensing is the most obvious way to detect Royer current. It offers the advantage of simple signal conditioning—there is no common mode voltage. The assumption that essentially all Royer current derives from the LT1172 emitter pin path is true. Also true, however, is that the waveshape of this path's current

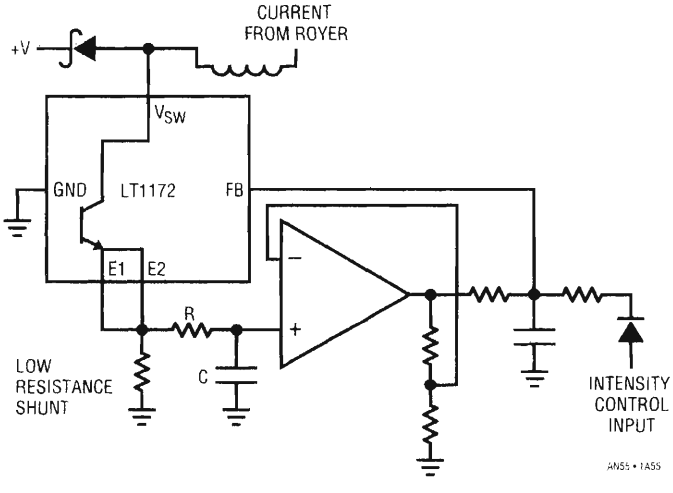


Figure 11–C5. "Bottom side" current sensing has poor line regulation due to RC averaging characteristics.

varies widely with input voltage and lamp operating current. The RMS voltage across the shunt (e.g., the Royer current) is unaffected by this, but the simple RC averager produces different outputs for the various waveforms. This causes this approach to have very poor line rejection, rendering it impractical. Figure 11–C6 senses inductor flux, which should correlate with Royer current. This approach promises attractive simplicity. It gives better line regulation but still has some trouble giving reliable feedback as waveshape changes. It also, in keeping with most flux sampling schemes, regulates poorly under low current conditions.

Figure 11–C7 senses flux in the transformer. This takes advantage of the transformer's more regular waveform. Line regulation is reasonably good because of this, but low current regulation is still poor. Figure 11–C8 samples Royer collector voltage capacitively, but the feedback signal does not accurately represent start-up, transient, and low current conditions.

Figure 11–C9 uses optical feedback to eliminate all feedback integrity problems. The photodiode-amplifier combination provides a DC feedback signal which is a function of actual lamp emission. It forces the lamp to constant emissivity, regardless of environmental or aging factors.

This approach works quite nicely, but introduces some evil problems. The lamp comes up to constant emission immediately at turn-on. There is no warm-up time required because the loop forces emission, instead of current. Unfortunately, it does this by driving huge overcurrents through the lamp, stressing it and shortening life. Typically, 2 to 5 times rated current flows for many seconds before lamp temperature rises, allowing the loop to back down drive. A subtle result of this effect occurs with lamp aging. When lamp emissivity begins to fall off, the loop increases current to correct the condition. This increase in current accelerates lamp aging, causing further emissivity degradation. The resultant downward spiral continues, resulting in dramatically shortened lamp life.

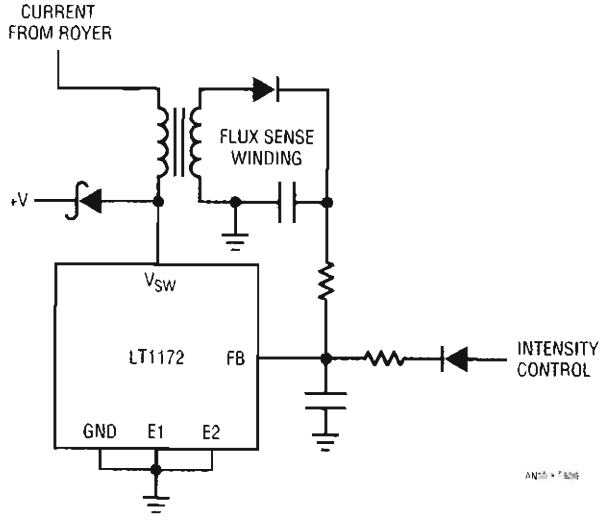


Figure 11–C6. Flux sensing has irregular outputs, particularly at low currents.

Other problems involve increased component count, photodiode mounting, and the requirement for photodiodes with predictable response or some form of trim.

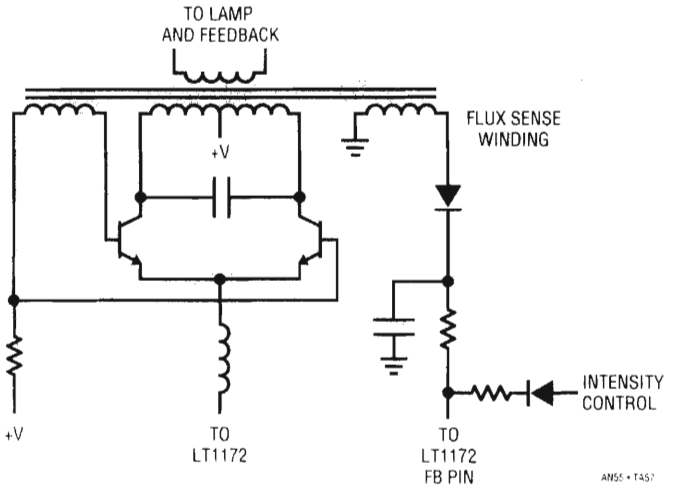


Figure 11–C7. Transformer flux sensing gives more regular feedback, but not at low currents.

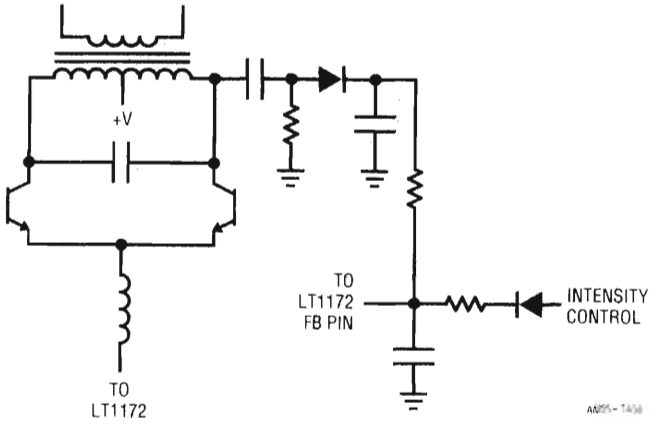


Figure 11–C8. AC couples drive waveform feedback is not reliable at low currents.

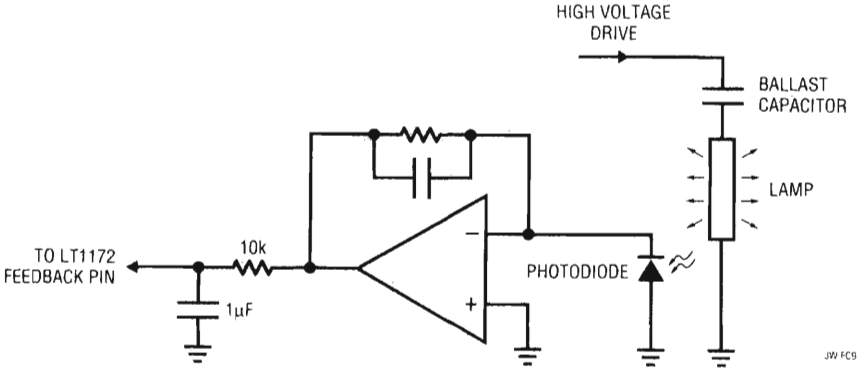


Figure 11–C9. Optically sensed feedback eliminates feedback irregularities, but introduces other problems.

Appendix D

Perspectives on Efficiency

The LCD displays currently available require two power sources, a backlight supply and a contrast supply. The display backlight is the single largest power consumer in a typical portable apparatus, accounting for almost 50% of the battery drain when the display intensity control is at maximum. Therefore, every effort must be expended to maximize backlight efficiency.

The backlight presents a cascaded energy attenuator to the battery (Figure 11–D1). Battery energy is lost in the electrical-to-electrical conversion to high voltage AC to drive the cold cathode fluorescent lamp (CCFL). This section of the energy attenuator is the most efficient; conversion efficiencies exceeding 90% are possible. The CCFL, although the most efficient electrical-to-light converter available today, has losses exceeding 80%. Additionally, the light transmission efficiency of present displays is about 10% for monochrome, with color types even lower. Clearly, overall backlight efficiency improvements must come from bulb and display improvements.

Higher CCFL circuit efficiency does, however, directly translate into increased operating time. For comparison purposes Figure 11–20's circuit was installed in a computer running 5mA lamp current. The result was a 19 minute increase in operating time.

Relatively small reductions in backlight intensity can greatly extend battery life. A 20% reduction in screen intensity results in nearly 30 minutes of additional running time. This assumes that efficiency remains reasonably flat as power is reduced. Figure 11–D2 shows that the circuits presented do reasonably well in this regard, as opposed to other approaches.

The contrast supply, operating at greatly reduced power, is not a major source of loss.

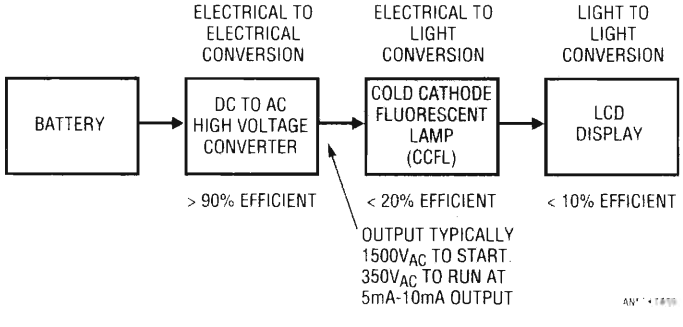


Figure 11–D1. The backlit LCD display presents a cascaded energy attenuator to the battery. DC to AC conversion is significantly more efficient than energy conversions in lamp and display.

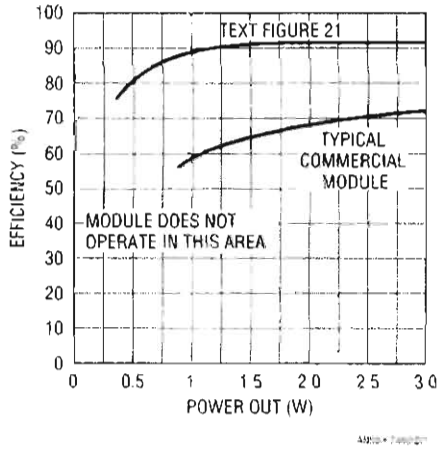


Figure 11–D2. Efficiency comparison between Figure 11–21 and a typical modular converter.

17. There's No Place Like Home

What's your choice for the single best aid to an interesting and productive circuit design career? A PhD? An IQ of 250? A CAD workstation? Getting a paper into the Solid State Circuit Conference? Befriending the boss? I suppose all of these are of some value, but none even comes close to something else. In fact, their combined benefit isn't worth a fraction of the something else. This something else even has potential economic rewards. What is this wondrous thing that outshines all the other candidates? It is, simply, a laboratory in your home. The enormous productivity advantage provided by a home lab is unmatched by anything I am familiar with. As for economic benefits, no stock tip, no real estate deal, no raise, no nothing can match the long term investment yield a home lab can produce. The laboratory is, after all, an investment in yourself. It is an almost unfair advantage.

The magic of a home lab is that it effectively creates time. Over the last 20 years I estimate that about 90% of my work output has occurred in a home lab. The ability to grab a few hours here and there combined with occasional marathon 5–20 hour sessions produces a huge accumulated time benefit. Perhaps more importantly, the time generated is highly leveraged. An hour in the lab at home is worth a day at work.

A lot of work time is spent on unplanned and parasitic activities. Phone calls, interruptions, meetings, and just plain gossiping eat up obscene amounts of time. While these events may ultimately contribute towards good circuits, they do so in a very oblique way. Worse yet, they rob psychological momentum, breaking up design time into chunks instead of allowing continuous periods of concentration. When I'm at work I do my job. When I'm at home in the lab is where the boss and stockholders get what they paid for. It sounds absurd, but I have sat in meetings praying for 6 o'clock to come so I can go home and get to work. The uninterrupted time in a home lab permits persistence, one of the most powerful tools a designer has.

I favor long, uninterrupted lab sessions of at least 5 to 10 hours, but family time won't always allow this. However, I can almost always get in two to four hours per day. Few things can match the convenience and efficiency of getting an idea while washing dishes or putting my son to sleep and being able to breadboard it *now*. The easy and instant availability of lab time makes even small amounts of time practical. Because no

Figure 17–1.
Everything is undisturbed and just as you left it. You can get right to work.

one else uses your lab, everything is undisturbed and just as you left it after the last session. Nothing is missing or broken,[1] and all test equipment is familiar. You can get right to work.

Measured over months, these small sessions produce spectacular gains in work output. The less frequent but more lengthy sessions contribute still more.

Analog circuits have some peculiar and highly desirable characteristics which are in concert with all this. They are small in scale. An analog design is almost always easily and quickly built on a small piece of copper-clad board. This board is readily shuttled between home and work, permitting continuous design activity at both locations.[2] A second useful characteristic is that most analog circuit development does not require the most sophisticated or modern test equipment. This, combined with test equipment's extremely rapid depreciation rate, has broad implications for home lab financing. The ready availability of high-quality used test equipment is the key to an affordable home lab. Clearly, serious circuit design requires high performance instrumentation. The saving grace is that this equipment can be five, twenty, or even thirty years old and still easily meet measurement requirements. The fundamental measurement perfor-

1. It is illuminating to consider that the average lifetime of an oscilloscope probe in a corporate lab is about a year. The company money and time lost due to this is incalculable. In 20 years of maintaining a home lab I have never broken a probe or lost its accessories. When personal money and time are at risk, things just seem to last longer.
2. An extreme variant related to this is reported by Steve Pietkiewicz of Linear Technology Corporation. Faced with a one-week business trip, he packed a complete portable lab and built and debugged a 15-bit A-D converter in hotel rooms.

mance of test equipment has not really changed much. Modern equipment simplifies the measurement process, offers computational capability, lower parts count, smaller size, and cost advantages (for new purchases). It is still vastly more expensive than used instrumentation. A Tektronix 454 150MHz portable oscilloscope is freely available on the surplus market for about $150.00. A new oscilloscope of equivalent capability costs at least ten times this price.

Older equipment offers another subtle economic advantage. It is far easier to repair than modern instruments. Discrete circuitry and standard-product ICs ease servicing and parts replacement problems. Contemporary processor-driven instruments are difficult to fix because their software control is "invisible," often convoluted, and almost impervious to standard troubleshooting techniques. Accurate diagnosis based on symptoms is extremely difficult. Special test equipment and fixtures are usually required. Additionally, the widespread usage of custom ICs presents a formidable barrier to home repair. Manufacturers will, of course, service their products, but costs are too high for home lab budgets. Modern, computationally based equipment using custom ICs makes perfect sense in a corporate setting where economic realities are very different. The time and dollar costs associated with using and maintaining older equipment in an industrial setting are prohibitive. This is diametrically opposed to home lab economics, and a prime reason why test equipment depreciates so rapidly.

The particular requirements of analog design combined with this set of anomalies sets guidelines for home lab purchases.[3] In general, instruments designed between about 1965 and 1980 meet most of the discussed criteria. Everybody has their own opinions and prejudices about instruments. Here are some of mine.

Oscilloscopes

The oscilloscope is probably the most important instrument in the analog laboratory. Tektronix oscilloscopes manufactured between 1964 and 1969 are my favorites. Brilliantly conceived and stunning in their execution, they define excellence. These instruments were designed and manufactured under unique circumstances. It is unlikely that test equipment will ever again be built to such uncompromising standards. Types 547 and 556 are magnificent machines, built to last forever, easily maintained, and almost a privilege to own. The widely available plug-in vertical amplifiers provide broad measurement capability. The 1A4 four-trace and 1A5 and 1A7A differential plug-ins are particularly useful. A 547 equipped with a

3. An excellent publication for instrument shopping is "Nuts and Volts," headquartered in Corona, California. Telephone 800/783-4624.

1A4 plug-in provides extensive triggering and display capability. The dual beam 556, equipped with two vertical plug-ins, is an oscilloscope driver's dream. These instruments can be purchased for less than the price of a dinner for two in San Francisco.[4] Their primary disadvantages are size and 50MHz bandwidth, although sampling plug-ins go out to 1GHz.

The Tektronix 453 and 454 portables extend bandwidth to 150MHz while cutting size down. The trade-off is lack of plug-in capability. The later (1972) Tektronix 485 portable has 350MHz bandwidth but uses custom ICs, is not nearly as ruggedly built, and is very difficult to repair. Similarly, Tektronix 7000 series plug-in instruments (1970s and 80s) feature very high performance but have custom ICs and are not as well constructed as earlier types. They are also harder to fix. The price-risk-performance ratio is, however, becoming almost irresistible. A 500MHz 7904 with plug-in amplifiers brings only $1000.00 today, and the price will continue to drop.

Sampling 'scopes and plug-ins attain bandwidths into the GHz range at low cost. The Tektronix 661, equipped with a 4S2 plug-in, has 3.9GHz bandwidth, but costs under $100.00. The high bandwidths, sensitivity, and overload immunity of sampling instruments are attractive, but their wideband sections are tricky to maintain.

Other 'scopes worthy of mention include the Hewlett-Packard 180 series, featuring small size, plug-in capability and 250MHz bandwidth. HP also built the portable 1725A, a 275MHz instrument with many good attributes. Both of these instruments utilize custom ICs and hybrids, raising the maintenance cost risk factor.

Related to oscilloscopes are curve tracers. No analog lab is complete without one of these. The Tektronix 575 is an excellent choice. It is the same size as older Tektronix lab 'scopes and is indispensable for device characterization. The more modern 576 is fully solid state, and has extended capabilities and more features. A 576 is still reasonably expensive (»$1500.00). I winced when I finally bought one, but the pain fades quickly with use. A 575 is adequate; the 576 is the one you really want.

Oscilloscopes require probes. There are so many kinds of probes and they are all so wonderful! I am a hopeless probe freak. It's too embarrassing to print how many probes I own. A good guideline is to purchase only high quality, name brand probes. There are a lot of subtleties involved in probe design and construction, particularly at high frequencies. Many off-brand types give very poor results. You will need a variety of 1× and 10× passive probes, as well as differential, high voltage, and other types. 50Ω systems utilize special probes, which give exceptionally clean results at very high frequency.

---

4. It is highly likely that Tektronix instruments manufactured between 1964 and 1969 would have appreciated at the same rate as, say, the Mercedes-Benz 300 SL . . . if oscilloscopes were cars. They meet every criterion for collectible status except one; there is no market. As such, for the few aberrants interested, they are surely the world's greatest bargain.

Active probes are also a necessity. This category includes FET probes and current probes. FET probes provide low-capacitive loading at high frequency. The 230MHz Tektronix P-6045 is noteworthy because it is easy to repair compared to other FET probes. A special type of FET probe is the differential probe. These devices are basically two matched FET probes contained within a common probe housing. This probe literally brings the advantages of a differential input oscilloscope to the circuit board. The Tektronix P6046 is excellent, and usually quite cheap because nobody knows what it is. Make sure it works when you buy it, because these probes are extraordinarily tricky to trim up for CMRR after repair. Finally, there are clip-on current probes. These are really a must, and the one to have is the DC-50MHz Tektronix P-6042. They are not difficult to fix, but the Hall effect-based sensor in the head is expensive. AC only clip-on probes are not as versatile, but are still useful. Tektronix has several versions, and the type 131 and 134 amplifiers extend probe capability and eliminate scale factor calculations. The Hewlett-Packard 428, essentially a DC only clip-on probe, features high accuracy over a 50mA to 10 amp range.

Power Supplies

There are never enough power supplies. For analog work, supplies should be metered, linear regulators with fully adjustable voltage output and current limiting. The HP 6216 is small and serves well. At higher currents (i.e., to 10 amps) the Lambda LK series are excellent. These SCR pre-regulated linear regulators are reasonably compact, very rugged, and handle any load I have ever seen without introducing odd dynamics. The SCR pre-regulator permits high power over a wide output voltage range with the low noise characteristics of a linear regulator.

Signal Sources

A lab needs a variety of signal sources. The Hewlett-Packard 200 series sine wave oscillators are excellent, cheap, and easily repaired. The later versions are solid state, and quite small. At high frequencies the HP 8601A sweep generator is a superb instrument, with fully settable and leveled output to 100MHz. The small size, high performance, and versatility make this a very desirable instrument. It does, however, have a couple of custom hybrid circuits, raising the cost-to-repair risk factor.

Function generators are sometimes useful, and the old Wavetek 100 series are easily found and repaired. Pulse generators are a must; the Datapulse 101 is my favorite. It is compact, fast, and has a full complement of features. It has fully discrete construction and is easy to maintain. For high power output the HP214A is excellent, although not small.

Voltmeters

DVMs are an area where I'm willing to risk on processor-driven equipment. The reason is that the cost is so low. The Fluke handheld DVMs are so cheap and work so well they are irresistible. There are some exceptionally good values in older DVMs too. The 5½ digit Fluke 8800A is an excellent choice, although lacking current ranges. The 4½ digit HP3465 is also quite good, and has current ranges. Another older DVM worthy of mention is the Data Precision 245–248 series. These full featured 4½ digit meters are very small, and usually sell for next to nothing. Their construction is acceptable, although their compactness sometimes makes repair challenging.

AC wideband true RMS voltmeters utilize thermal converters. These are special purpose instruments, but when you must measure RMS they are indispensable. The metered Hewlett-Packard 3400A has been made for years, and is easy to get. This instrument gives good accuracy to 10MHz. All 3400s look the same, but the design has been periodically updated. If possible, avoid the photochopped version in favor of the later models. The HP3403C goes out to 100MHz, has higher accuracy, and an autoranging digital display. This is an exotic, highly desirable machine. It is also harder to find, more expensive, and not trivial to repair.

Miscellaneous Instruments

There are literally dozens of other instruments I have found useful and practical to own. Tektronix plug-in spectrum analyzers make sense once you commit to a 'scope mainframe. Types 1L5, 1L10, and 1L20 cover a wide frequency range, but are harder to use than modern instruments. Distortion analyzers are also useful. The HP334A is very good, and has about a .01% distortion floor. The HP339A goes down to about .002%, and has a built in low distortion oscillator. It is also considerably more expensive. Both are "auto-nulling" types, which saves much knob twiddling. Frequency counters are sometimes required, and the little HP5300 series are very good general purpose units. The old 5245L is larger, but the extensive line of plug-ins makes this a very versatile instrument. Occasionally, a chart recorder makes sense, and the HP7000A (XY) and HP680 (strip) are excellent. The 7000A has particularly well thought out input amplifiers and sweep capabilities. Other instruments finding occasional use are a variable voltage reference (the Fluke 332 is huge, but there is no substitute when you need it) and a picoammeter. Kiethley picoammeters (e.g., type 610) are relatively hard to find, but read into the femtoampere range. "Diddle boxes" for both resistance and capacitance are very useful. These break down into precision and non-precision types. General Radio and ESI built excellent precision types (e.g., G.R. 1400 Series), but many have been abused . . . look (and smell) in-

side before you buy. Non-precision types by EICO and Heathkit are everywhere, and cost essentially nothing. The precision variable air capacitors built by General Radio (types 722D and the later 1422D) are particularly applicable for transducer simulation. They are also worth buying just to look at; it is hard to believe human beings could build anything so beautiful.

Oscilloscope cameras are needed to document displays. Modern data recording techniques are relegating 'scope cameras to almost antique status, which has happily depressed their price. My work involves a significant amount of waveform documentation, so I have quite a bit of specialized camera equipment. The Tektronix C-30 is a good general purpose camera which fits, via adapters, a wide variety of oscilloscopes. It is probably the best choice for occasional work. The Tektronix C27 and C12 are larger cameras, designed for plug-in 'scopes. Their size is somewhat compensated by their ease of use. However, I do not recommend them unless you do a lot of photographic documentation, or require highly repeatable results.

Finally, cables, connectors, and adapters are a must have. You need a wide variety of BNC, banana jack, and other terminator, connector, adapter, and cable hardware. This stuff is not cheap; in fact it is outrageously expensive, but there is no choice. You can't work without it and the people who make it know it.

No discussion of a home laboratory is complete without comment on its location. You will spend many hours in this lab; it should be as comfortable and pleasant a place as possible. The use of space, lighting, and furnishings should be quite carefully considered. My lab is in a large

Figure 17–2. You will spend many hours in this lab. It should be a comfortable and pleasant place.

Figure 17–3. Maintaining lab organization is painful, but increases time efficiency.

room on the second floor, overlooking a very quiet park. It is a bright, colorful room. Some of my favorite pictures and art are on the walls, and I try to keep the place fairly clean. In short, I do what I can to promote an environment conducive to working.

Over the last 20 years I have found a home lab the best career friend imaginable. It provides a time efficiency advantage that is almost unfair. More importantly, it has insured that my vocation and hobby remain happily and completely mixed. That room on the second floor maintains my enthusiasm. Engineering looks like as good a career choice at 45 as it did at 8 years old. To get that from a room full of old equipment has got to be the world's best bargain.

Figure 17–4.
It's convenient to be able to write up lab results as they occur.

CHAPTER 18

The Zoo Circuit

Jim Williams

Here is a classic example of how to adapt an old circuit to new requirements. Jim took one of my old (power-wasteful) circuits and adapted it and reengineered it to run at very low power. Nice engineering! With no degradation of accuracy. /rap

This chapter is dedicated to the memory of Professor Jerrold R. Zacharias, who saved my ass.

History, Mistakes, and Some Monkeys Design a Circuit

A couple of years ago I was asked to design a circuit for a customer. The requirements were not trivial, and the customer was having difficulty. I worked on this problem for some time and was asked to present my solution in a formal design review at the customer's location.

When I say "formal," I mean it! I came expecting to talk circuits with a few guys over a pizza. Upon arrival, I was taken to a large and very grand room, reminiscent of a movie theater. About 150 engineers were in attendance. There was every audiovisual machine known to humanity at the ready, and I was almost embarrassed to report that I had no slides, overheads, charts, or whatever (although the piece of chalk I grasped was nice). A "senior technical management panel," positioned in a boxed-off section adjacent to the lectern, was to present a prepared list of questions. A video camera duly recorded the proceedings. The whole thing was chaired by somebody who introduced himself as "Dr. So-and-So, senior vice president of engineering." Everybody in the place talked in whispers and nodded his head a lot. I found myself alternating between intimidation and amusement.

I gave a fairly stiff presentation, clutching my dear little piece of chalk the whole time. Things seemed to go okay but not great, and then the panel began with their prepared list

of questions. The first question went something like, "Can you explain, precisely, where the ideas for this and that piece of the circuit came from? Can you detail what design procedures, programs, and methodologies were helpful?"

I considered various acceptable answers but decided to simply tell the truth: "Most of the ideas came from history and making mistakes, and the best source of help was some monkeys at the San Francisco Zoo."

You could have heard a pin before it dropped. There was absolute silence for a bit, and then some guy stood up and asked me to elaborate "a little." Everybody cracked up, the mood shifted, and we finally began to really *talk* about the circuit.

This customer originally came to me with a need for a "CMOS voltage-to-frequency converter." The performance requirements were as follows:

Output frequency	0 to 10 kHz
Input voltage	0 to 5 V
Linearity	0.04%
Drift	100 ppm/°C
PSRR	100 ppm/V
Temperature range	0° to 55°C
Step response	<5 cycles of output frequency
Output pulse	5 V CMOS-compatible
Power supply	Single 9 V battery (6.5 to 10 V)
Power consumption	200 µA maximum
Cost	< $6.00/100,000 pieces

These people had been working on a design for several months. It functioned but was described as wholly unsatisfactory. I asked why they needed CMOS and was assured that "the low-power requirement is nonnegotiable." Without further comment, I asked them to send me their breadboard. It arrived the next morning, and it looked like **Figure 18-1**.

This is probably the most obvious way to design a V/F converter. The 9 V battery is regulated to 5 V by *IC1* and a −5 V rail is derived by *IC2*. The input voltage causes current flow into *A1*'s summing point. *A1* responds by integrating negative, as shown in **Figure 18-2**, trace *A*. When *A1*'s output goes low enough, *A2* trips high (see trace *B* in Figure 18-2), turning on the CD4066 switch and resetting the integrator. Local positive feedback around *A2* (*A2*'s positive input is trace *C*) "hangs up" the reset, ensuring a complete integrator discharge. When the positive feedback decays, *A1* begins to ramp again. The ramp slope, and hence the repetition frequency, depends on the input voltage-dependent current into *A1*'s summing point.

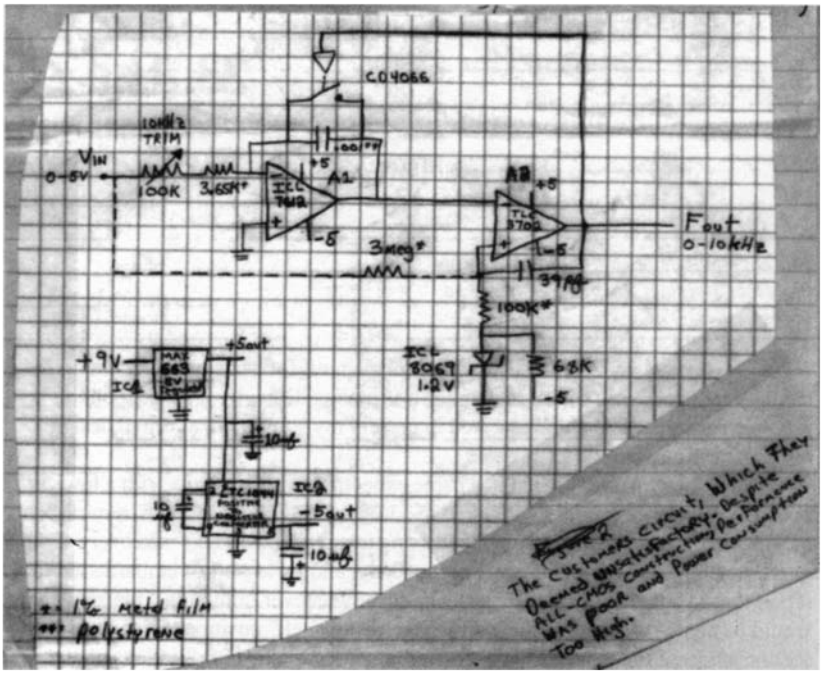


Figure 18-1: The customer's circuit, which they deemed unsatisfactory. Despite all-CMOS construction, performance was poor and power consumption too high.

As soon as I saw the schematic, I knew I couldn't salvage any portion of this design. A serious drawback to this approach is *A1*'s integrator reset time. This time, "lost" in the integration, results in significant linearity error as the operating frequency approaches it. The circuit's 6 µsec reset (see Figure 18-2, traces *A* and *B*) interval introduces a 0.6% error at 1 kHz, rising to 6% at 10 kHz.Also, variations in the reset time contribute additional errors. I added the 3 M resistor (shown in dashed lines) in a half-hearted attempt to improve these figures. This resistor causes *A2*'s trip point to vary slightly with input, partially compensating for the integrator's "lost" reset time. This Band-Aid did improve linearity by more than an order of magnitude, to about 0.4%, but it ain't the way to go.

There are other problems. Quiescent current consumption of this entirely CMOS circuit is 190 µA, rising to a monstrous 700 µA at 10 kHz. Additionally, the polystyrene capacitor's drift alone is –120 ppm/°C, eating up the entire budget. The 1.2 V reference and the input resistor-trimmer could easily double this figure. There are a host of other problems, but what is really needed is an approach with inherently better linearity and lower power consumption.

There are many ways to convert a voltage to a frequency. The "best" approach in an application varies with desired precision, speed, response time, dynamic range, and other considerations.

A = 0.5 V/Div.
B = 10 V/Div.
C = 10 V/Div.
Horiz. = 10 μsec/Div.

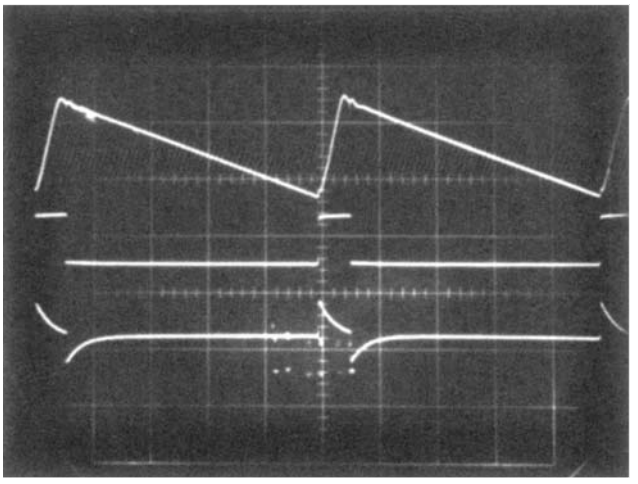


Figure 18-2: Waveforms for Fig. 18-1's circuit. Finite reset time prevents good linearity performance.

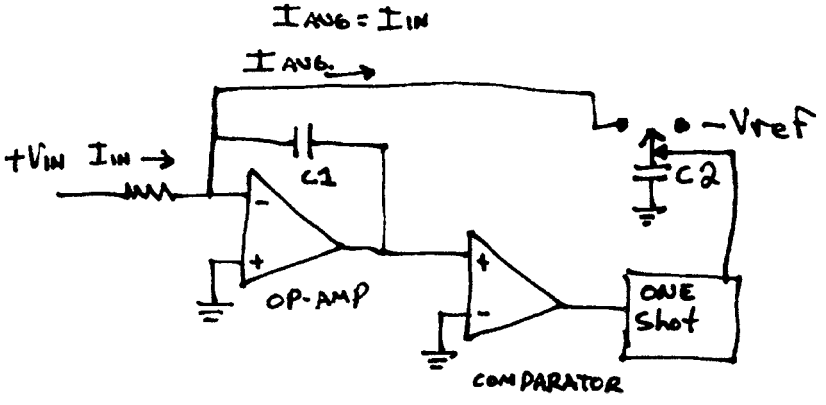


Figure 18-3: Conceptual charge-dispensing-type voltage-to-frequency converter.

Figure 18-3's concept potentially achieves high linearity by enclosing Figure 18-1's integrator in a charge-dispensing loop.

In this approach, *C2* charges to $-V_{ref}$ during the integrator's ramping time. When the comparator trips, *C2* is discharged into *A1*'s summing point, forcing its output high. After *C2*'s discharge, *A1* begins to ramp and the cycle repeats. Because the loop acts to force the average summing currents to zero, the integrator time constant and reset time do not affect frequency. Gain-drift terms are V_{ref}, *C2*, and the input resistor. This approach yields high linearity (typically 0.01%) into the megahertz range.

Figure 18-4 is conceptually similar except that it uses feedback current instead of charge to maintain the op-amp's summing point. Each time the op-amp's output trips the

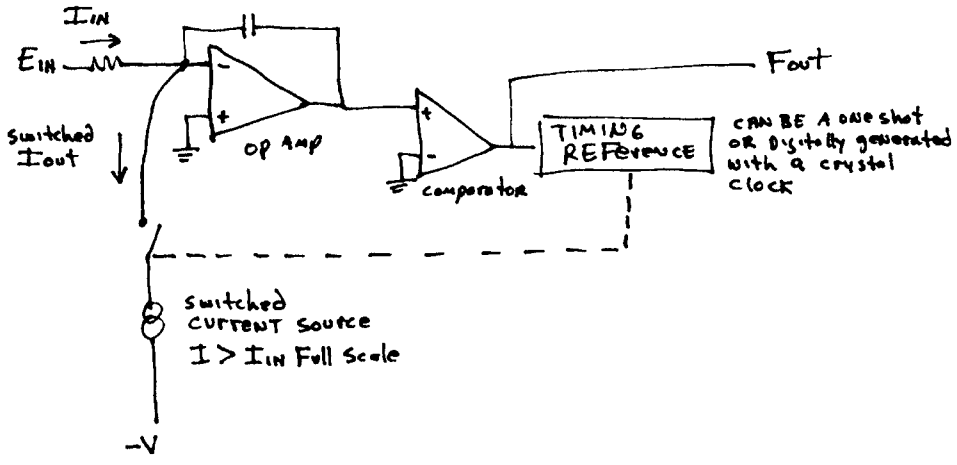


Figure 18-4: Current balance voltage-to-frequency converter.

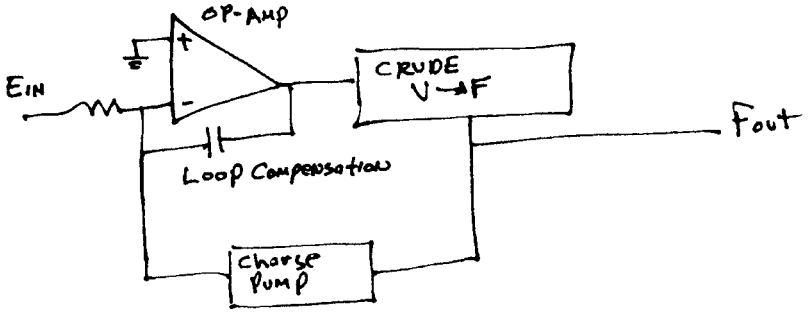


Figure 18-5: Loop-charge pump voltage-to-frequency converter.

comparator, the current sink pulls current from the summing point. Current is pulled from the summing point for the timing reference's duration, forcing the integrator positive. At the end of the current sink's period, the integrator's output again heads negative. The frequency of this action is input related.

Figure 18-5 uses DC loop correction. This arrangement offers all the advantages of charge and current balancing except that response time is slower. Additionally, it can achieve exceptionally high linearity (0.001%), output speeds exceeding 100 MHz, and very wide dynamic range (160 dB). The DC amplifier controls a relatively crude V/F converter. This V/F converter is designed for high speed and wide dynamic range at the expense of linearity and thermal stability. The circuit's output switches a charge pump whose output, integrated to DC, is compared to the input voltage.

The DC amplifier forces the V/F converter operating frequency to be a direct function of input voltage. The DC amplifier's frequency compensation capacitor, required because

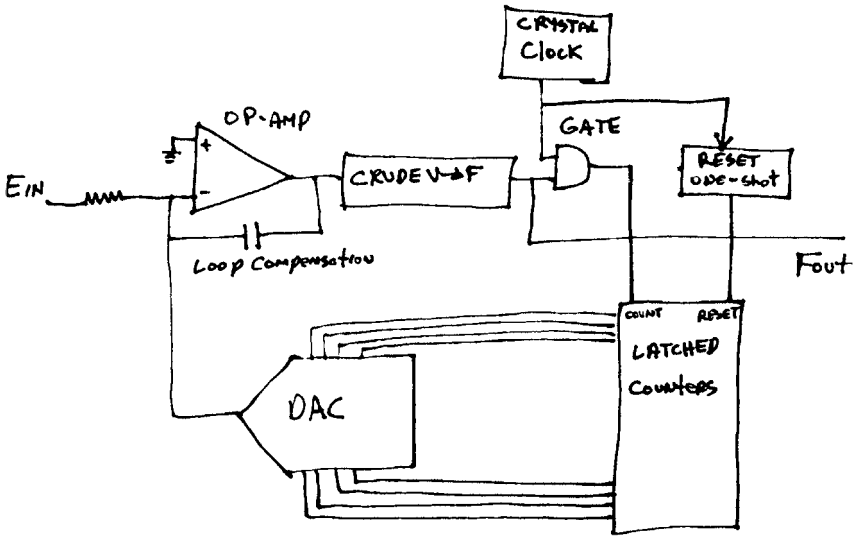


Figure 18-6: Loop-DAC voltage-to-frequency converter.

of loop delays, limits response time. **Figure 18-6** is similar except that the charge pump is replaced by digital counters, a quartz time base, and a DAC. Although it is not immediately obvious, this circuit's resolution is not restricted by the DAC's quantizing limitations. The loop forces the DAC's LSB to oscillate around the ideal value. These oscillations are integrated to DC in the loop compensation capacitor. Hence, the circuit will track input shifts much smaller than a DAC LSB. Typically, a 12-bit DAC (4096 steps) will yield one part on 50,000 resolution. Circuit linearity, however, is set by the DAC's specification.

If you examine these options, Figure 18-3 looks like the winner for the customer's application. The specifications call for step response inside five cycles of output frequency. This eliminates the circuits in Figures 18-4, 18-5, and 18-6 with their DC amplifiers' response time lag. Figure 18-4 requires a timing reference and a precision switched current source, implying some degree of complexity. In theory, Figure 18-3's approach can meet all the specifications without undue complexity.

This technique is not new. I first saw it back in 1964 in a copy of the *GE Transistor Manual*. T. P. Sylvan used a discrete op-amp and a unijunction transistor to form the loop. Hewlett-Packard built rack-mounted V/F converters in the early 1960s that also relied on this approach. In 1972, R. A. Pease developed a commercially produced modular version (Teledyne-Philbrick Model 4701) using a single op-amp that routinely achieved 0.01% linearity with commensurate drift performance. Pease's circuit is particularly relevant, and a version of it is shown in **Figure 18-7**.

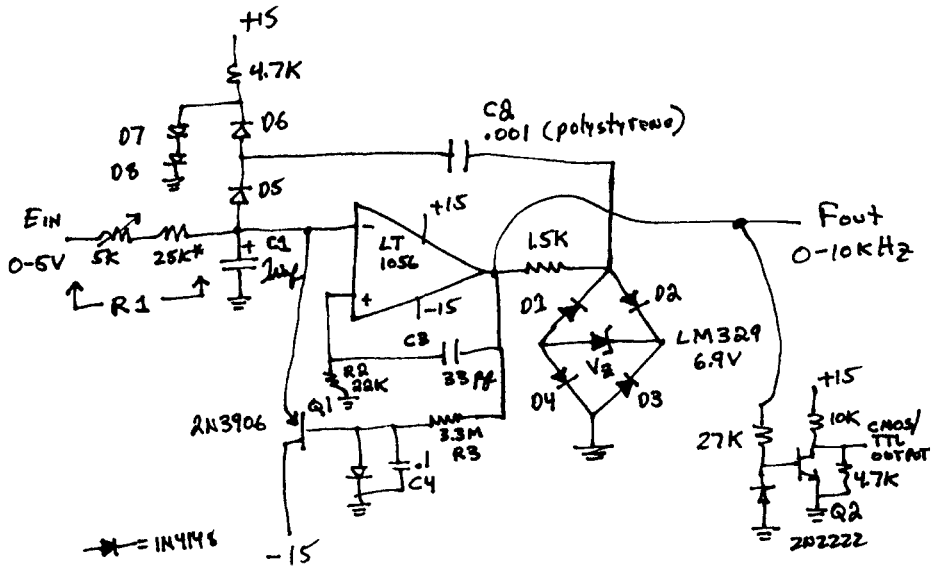
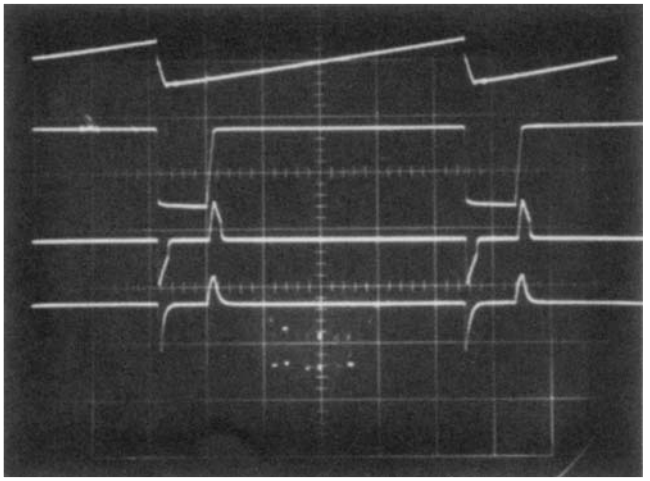


Figure 18-7: A version of Pease's elegant voltage-to-frequency converter circuit.



A = 0.02 V/Div.
B = 20 V/Div.
C = 20 mA/Div.
D = 20 V/Div.
Horiz. = 20 μsec/Div.

Figure 18-8: Waveforms for the Pease-type voltage-to-frequency converter.

Assume *C1* sits at a small negative potential. *A1*'s negative input is below its zero-biased positive input, and its output is high. The zener bridge clamps high (at $V_Z + V_{D4} + V_{D2}$) and *C2* charges via *D6*, *D7*, and *D8*. The input voltage forces current through *R1*, and *C1* begins to charge positively (trace *A*, **Figure 18-8**). When *C1* crosses zero volts, *A1*'s output (trace *B*) goes low and the zener bridge clamps negative, discharging *C2* (*C2*'s

current is trace *C*) via the *D5–C1* path. The resultant charge removal from *C1* causes it to rapidly discharge (trace *A*). *R2–C3* provides positive feedback to *A1*'s positive input (trace *D*), reinforcing this action and hanging up *A1*'s output long enough for a complete *C2* discharge. When the *R2–C3* feedback decays, *A1*'s output returns high and the cycle repeats. The frequency of this sequence is directly proportional to the input voltage derived current through *R1*. Drift terms include *R1*, *C2*, and the zener, as well as residual diode mismatches. In theory, all the diode drops cancel and do not contribute toward drift. The *R2–C3* "one-shot" time constant is not critical, as long as it allows enough time for *C2* to completely discharge. Similarly, "integrator"*C1*'s value is unimportant as long as it averages *A1*'s negative input to zero.

Q1 and associated components form a startup loop. Circuit startup or input overdrive can cause the circuit's AC-coupled feedback to latch. If this occurs, *A1* goes negative and wants to stay there. *R3* and *C4* slowly charge negative, biasing *Q1*. *Q1* turns on, pulling *C1* toward the –15 V rail, initiating normal circuit action. Once the circuit starts, *C4* assumes a small positive potential and *Q1* goes off. *Q2*, a simple level shifter, furnishes a logic-compatible output.

Pease's 1972 circuit is a very elegant, practical incarnation of Figure 18-3. With care, it will meet all the customer's requirements except two. It requires a split ±15 V supply and pulls well over 10 mA.The job now boils down to dealing with these issues.

Figure 18-9 shows my first attempt at adapting Pease's circuit to my customer's needs. Operation is similar to Pease's circuit. When the input current-derived ramp (trace *A*, **Figure 18-10**) at *C1A*'s negative input crosses zero, *C1A*'s output (trace *B*) drops low, pulling charge through *C1*. This forces the negative input below zero. *C2* provides positive feedback (trace *D* is the positive input), allowing a complete discharge for *C1* (*C1* current is trace *C*). When *C2* decays, *C1A*'s output goes high, clamping at the level set by *D1*, *D2*, and V_{ref}. *C1* receives charge, and recycling occurs when *C1A*'s negative input again arrives at zero. The frequency of this action is related to the input voltage. Diodes *D3* and *D4* provide steering and are temperature compensated by *D1* and *D2*. *C1A*'s sink saturation voltage is uncompensated but small. (These temperature coefficient assumptions are first order and will require more care later.) Although the LT1017 and LT1034 have low operating currents, this circuit pulls almost 400 µA. The AC current paths include *C1*'s charge-discharge cycle and *C2*'s branch. The DC path through *D2* and V_{ref} is particularly costly. *C1*'s charging must occur quickly enough for 10 kHz operation, meaning the clamp seen by *C1A*'s output must have low impedance at this frequency. *C3* helps, but significant current still must come from somewhere to keep impedance low. *C1A*'s current-limited output (≈30 µA source) cannot do the job unaided, and the resistor from the supply is required. Even if *C1A* could supply the necessary current, V_{ref}'s settling time would be an issue. Dropping *C1*'s value will reduce impedance requirements

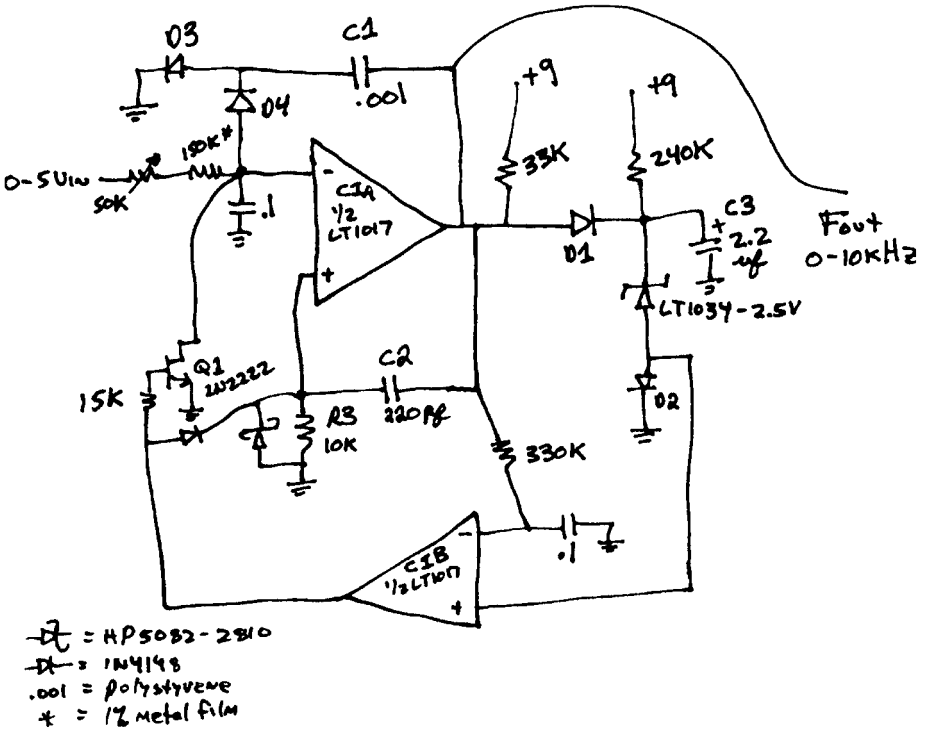
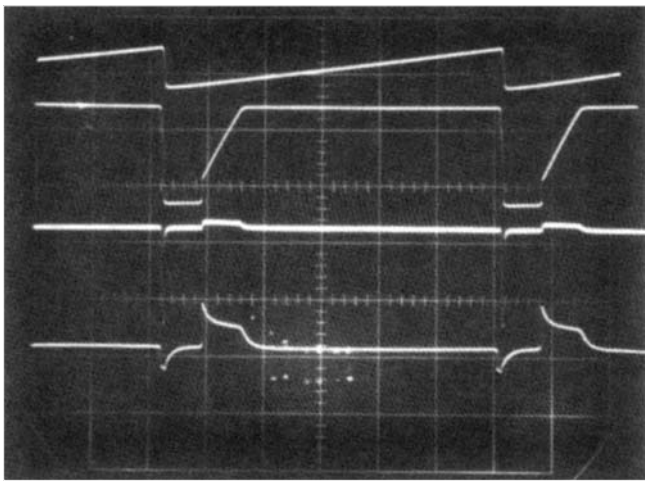


Figure 18-9: My first cut at adapting Pease's circuit.



A = 50 mV/Div.
B = 2 V/Div.
C = 2 mA/Div.
D = 1 V/Div.
Horiz. = 20 μsec/Div.

Figure 18-10: Waveforms for the circuit in Fig. 18-9.

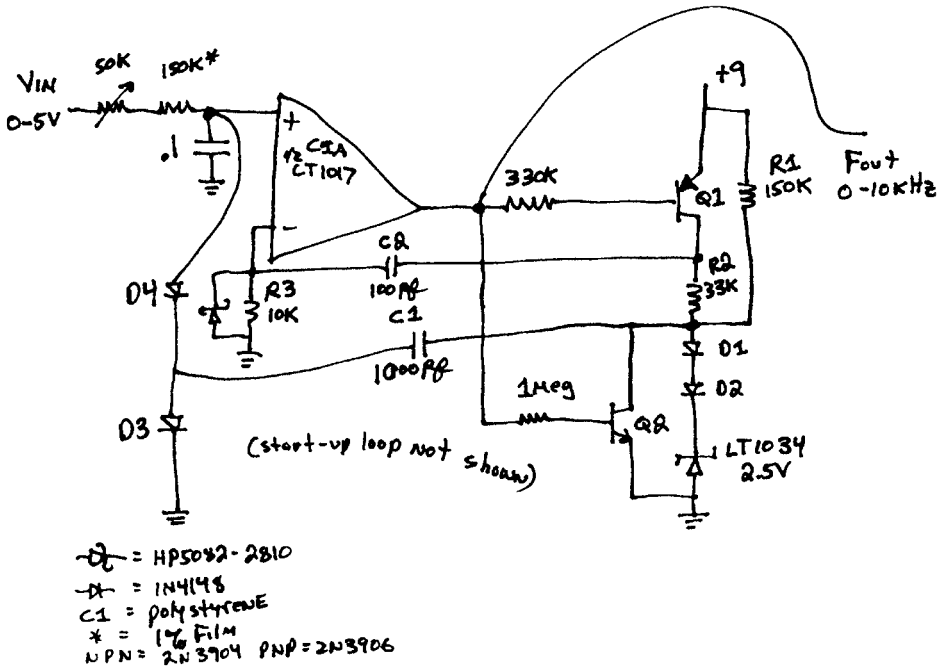


Figure 18-11: The second try. *Q1* and *Q2* switch the reference, saving some power.

proportionally and would seem to solve the problem. Unfortunately, such reduction magnifies the effects of stray capacitance at the *D3* to *D4* junction. It also mandates increasing R_{in}'s value to keep scale factor constant. This lowers operating currents at *C1A*'s negative input, making bias current and offset more significant error sources.

C1B, Q1, and associated components form a startup loop that operates in similar fashion to the one in Pease's circuit (Figure 18-7).

Figure 18-11 shows an initial attempt at dealing with these issues. This scheme is similar to Figure 18-9, except that *Q1* and *Q2* appear. V_{ref} (the top of D1) receives switched bias via *Q1* instead of being on all the time. *Q2* provides the sink path for *C1*. These transistors invert *C1A*'s output, so its input pin assignments are exchanged. *R1* provides a light current from the supply, improving reference settling time. This arrangement decreases supply current to about 300 µA, a significant improvement. Several problems do exist, however. *Q1*'s switched operation is really effective only at higher frequencies. In the lower ranges, *C1A*'s output is low most of the time, biasing *Q1* on and wasting power. Additionally, when *C1A*'s output switches, *Q1* and *Q2* simultaneously conduct during the transition, effectively shunting *R2* across the supply. Finally, the base currents of both transistors flow to ground and are lost. **Figure 18-12** shows the waveform traces for this circuit. The basic temperature

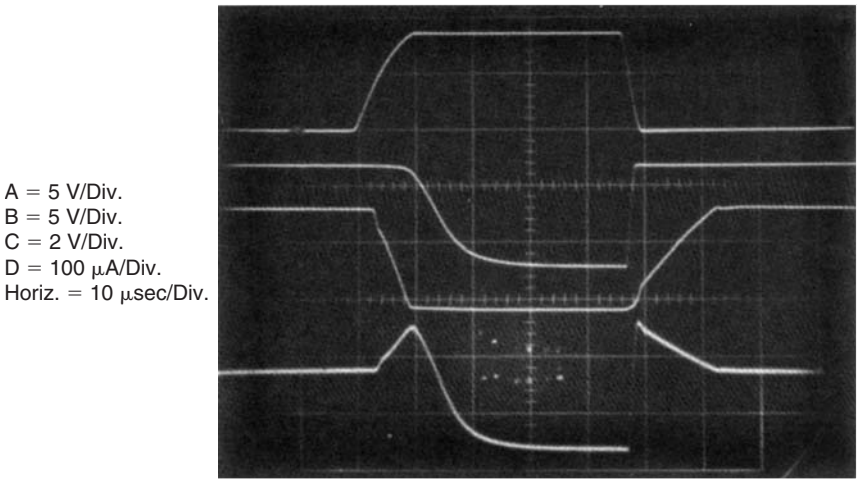


Figure 18-12: Fig. 18-11's waveforms. Traces *A, B, C,* and *D* are *C1A* output, *Q1* collector, *Q2* collector, and *R2* current, respectively. *Q1* to *Q2* simultaneous conduction problem is evident in trace *D*.

compensation is as before except that *Q2*'s saturation term replaces the comparator's. This temperature compensation scheme looks okay, but we're still hand waving.

Figure 18-13 is better. *Q1* is gone, *Q2* remains, but *Q3, Q4,* and *Q5* have been added. V_{ref} and its associated diodes are biased from *R1*. *Q3*, an emitter-follower, is used to source current to *C1*. *Q4* temperature compensates *Q3*'s V_{be}, and *Q5* switches *Q3*.

This method has some distinct advantages. The V_{ref} string can operate at greatly reduced current because of *Q3*'s current gain. Also, Figure 18-11's simultaneous conduction problem is largely alleviated because *Q5* and *Q2* are switched at the same voltage threshold out of *C1A*. *Q3*'s base and emitter currents are delivered to *C1*. *Q5*'s currents are wasted, although they are much smaller than *Q3*'s. *Q2*'s small base current is also lost. The values for *C2* and *R3* have been changed. The time constant is the same, but some current reduction occurs due to *R3*'s increase.

Operating wave forms are shown in **Figure 18-14**, and include *C1*'s output (trace *A*), *Q5*'s collector (trace *B*), *Q2*'s collector (trace *C*), *Q2*'s collector current (trace *D*), *C1*'s current (trace *E*), and *Q3*'s emitter current (trace *F*). Note that the current steering is clean, with no simultaneous conduction problems.

This circuit's 200 μA power consumption was low enough to make other specifications worth checking. Linearity came in at 0.05% and dropped to 0.02% when I added a 1 M resistor (dashed lines) across *C1*. The *D4–Q2* path cannot *fully* switch *C1* because of junction drop limitations. The resistor squeezes the last little bit of charge out of *C1*, completing the discharge and improving linearity.

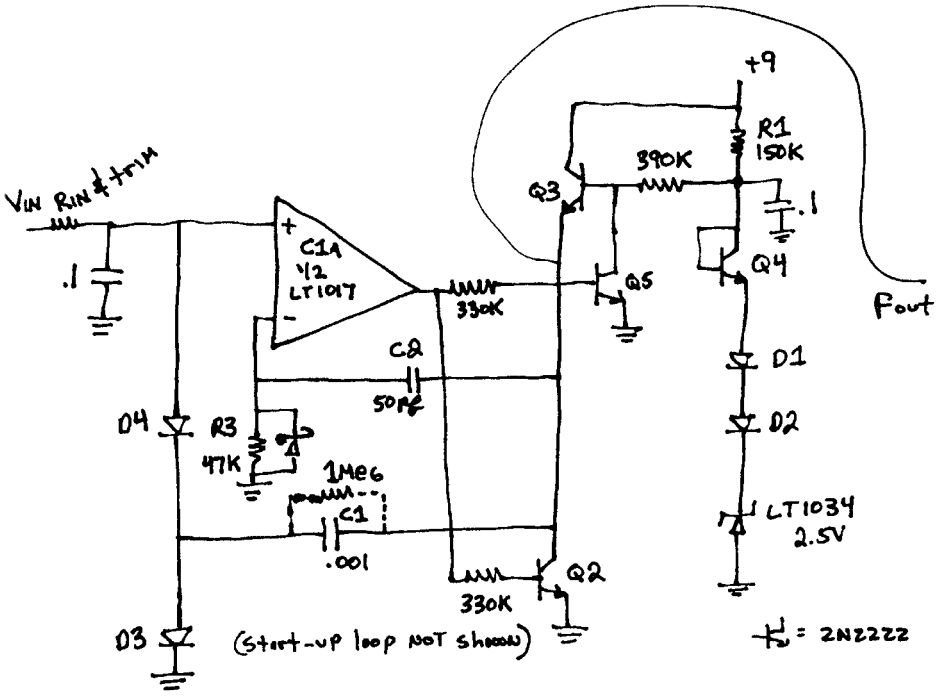
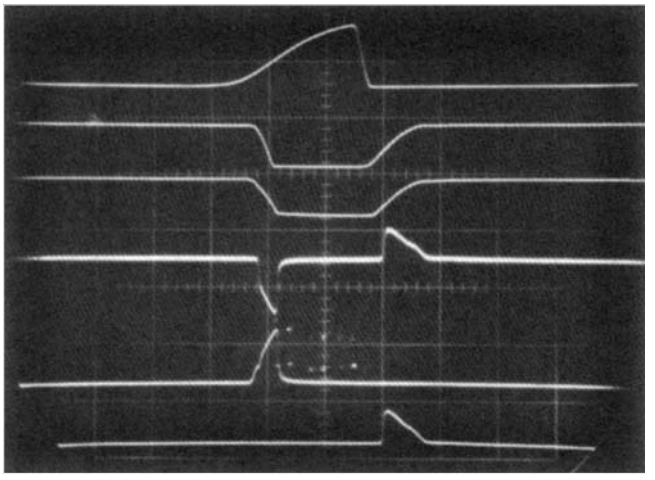


Figure 18-13: A better scheme for switching the reference.



A = 5 V/Div.
B = 5 V/Div.
C = 5 V/Div.
D = 1 mA/Div.
E = 1 mA/Div.
F = 1 mA/Div.
Horiz. = 10 μsec/Div.

Figure 18-14: Fig. 18-13's operation. Traces *D, E,* and *F* reveal no simultaneous conduction problems.

Power supply rejection ratio (PSRR) was not good enough. Supply shifts show up as current changes through *R1*. The LT1034 is relatively insensitive to this, but the *Q4*, *D1*, *D2* trio shift value. As such, I measured 0.1%/V PSRR. *R1* really needs to be a current source, or some compensation mechanism must be used.

Temperature compensation was next. Now it was time to stop hand waving and take a hard look. *Q4* supposedly compensates *Q3*, with *D1* and *D2* opposing *D3* and *D4*. Unfortunately, these devices operate under different dynamic and DC conditions, making precise cancellation difficult. In practice, *R1*'s value should be established to source the current through *Q4–D1–D2*, which provides optimum circuit temperature coefficient. Assuming perfect cancellation and no LT1034 or input resistor drift, we still must deal with *Q2*'s V_{ce} saturation term. At 100 mV saturation, *Q2* will drift about +0.3%/°C (see the Motorola 2N2222 datasheet), causing about a −300 µV/°C shift in the voltage *C1* discharges toward. This works out to about −100 ppm/°C (*C1* charges to 3 V) temperature coefficient, which will force a similar *positive* shift in output frequency. *C1*, a polystyrene type, drifts about −120 ppm/°C, contributing further overall positive temperature coefficient (as *C1*, or the voltage it charges to, gets smaller, the circuit must oscillate faster to keep the summing point at zero). So the best case is about 220 ppm/°C, and reality dictates that all the other junctions won't match precisely. Temperature testing confirmed all this. Initially, the breadboard showed about 275 ppm/°C and, by varying *R1*, bottomed out at about 200 ppm/°C. This certainly wasn't production-worthy engineering but pointed the way toward a solution.

How could I reduce the temperature coefficient and fix the PSRR? Additionally, power consumption was still marginal, although linearity was close. Replacing *R1* with a current source offered hope for PSRR, but reliable temperature compensation and lower power needed another approach. I pined for inspiration but got nothing. I was stuck.

Something that *had* inspired me for a couple of months was a physician I'd been dating. We really had a good time together—a couple of playful kids. There was much dimension to this woman, and I really enjoyed just how relaxed I felt being with her. Things were going quite nicely, and I sometimes allowed myself the luxury of wondering what would become of us.

One weekday afternoon, we played hooky and went to the San Francisco Zoo. The weather was gorgeous, no crowds, and the Alfa ran great. (On our second date it had thrown a fan belt.) We saw bears, elephants, tigers, birds, and ate lots of junk food. The lions got fed; they were *loud* and *hungry*. Strolling around, eating cheeseburgers, and doing just fine, we came to the monkeys.

These guys are actors; they love an audience. There was the usual array of grinning, simian catcalls, cheeping, squawking, lots of jungle bar performances, wondrous feats of

balance, and other such theatrics. One character particularly caught my eye. He did a little routine between two parallel rails. First, he hung by his hands as shown in **Figure 18-15**.

Then, very quickly, he flipped over, simultaneously rotating, so he ended up inverted (see **Figure 18-16**).

He did this over and over at great speed; it was his act. Standing there, watching the little fellow do his inverting routine between the rails, I saw my circuit problems simply melt. I felt very lucky. I had a good lady, and a good circuit too.

If you look inside a CMOS logic inverter, the output stage looks like **Figure 18-17**.

The MOS output transistors connect the output terminal to the supply or ground rail. The input circuitry is arranged so only one transistor is on at a time; simultaneous conduction cannot occur. Typically, channel-on resistance is 100 to 200 Ω. There are no junction effects; the transistor channels are purely ohmic. The device's input pin appears almost purely capacitive, drawing only picoamperes of bias current.

Figure 18-18 shows what happens when the CMOS inverter is dropped into the gizzard of Figure 18-13's circuit. *C1* is charged and discharged via the CMOS inverter's ohmic output transistors. *Q3* now drives the inverter's supply pin, and *Q2* goes away. Along with

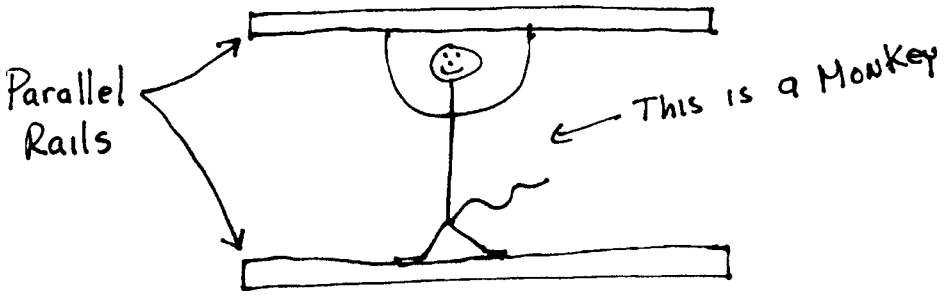


Figure 18-15: The zoo monkey on parallel rails.

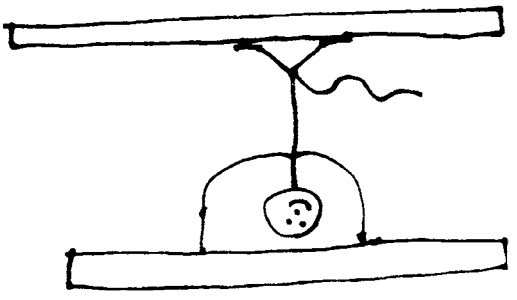


Figure 18-16: The zoo monkey on parallel rails, inverted.

Q2's departure goes its 100 ppm/°C temperature coefficient error. Also, *Q2*'s base current is eliminated, along with *Q5*'s base and collector current losses.

This scheme promises both lower temperature drift and lower power. Assuming ideal junction compensation, the remaining uncompensated drift terms are *C1*'s −120 ppm temperature coefficient and the input resistor. Unfortunately, this configuration does nothing to fix the PSRR problem. The only realistic fix for that is to replace *R1* with a current source. The current source doesn't have to be very stable but must run with only 2 V of headroom because the circuit has to work down to 6.5 V. The simplest alternative is the monolithic LM134. This three-terminal, resistor-programmable device will function with only 800 mV across it, although it does have a 0.33%/°C temperature coefficient.

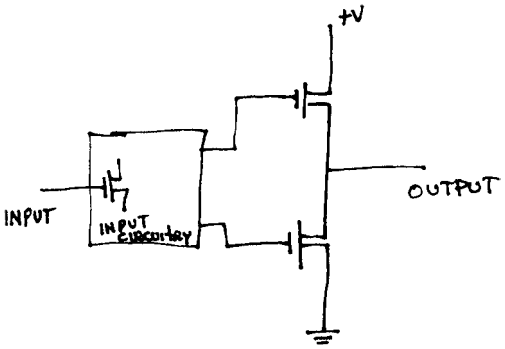


Figure 18-17: Conceptual CMOS inverter.

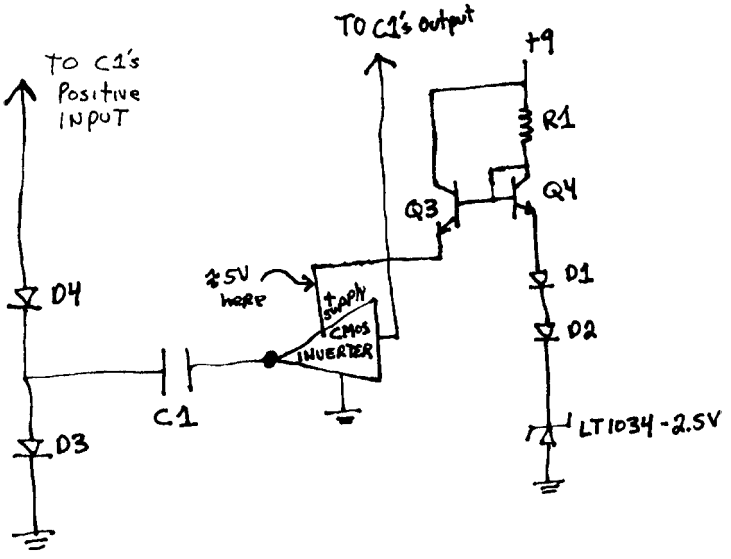


Figure 18-18: Adding the CMOS inverter to the circuit in Figure 18-13.

This temperature coefficient seemed small enough to avoid causing any trouble. The LT1034 shouldn't care, but what about *D1*, *D2*, and *Q4*?

When I calculated the effect of current-source shift with temperature on these devices, I realized I had just inherited the world. It came out *positive* 180 ppm/°C! This tends to cancel the capacitor's −120 ppm/°C term. Additionally, increasing the LT1034s reference voltage by about 50% would pull the compensation down to +120 ppm/°C, further reducing drift. This also aids overall temperature coefficient by making the residual junction mismatches a smaller percentage of the total reference voltage. The current source's low headroom capability allows this, while maintaining operation down to V_{supply} = 6.2 V. The sole uncompensated term is the input resistor, which can be specified for low temperature drift.

Figure 18-19 is the final circuit. It meets or exceeds every customer specification.

A 0 to 5 V input produces a 0 to 10 kHz output with a linearity of 0.02%. Gain drift is 40 ppm/°C, and PSRR is inside 40 ppm/V. Maximum current consumption is 145 µA,

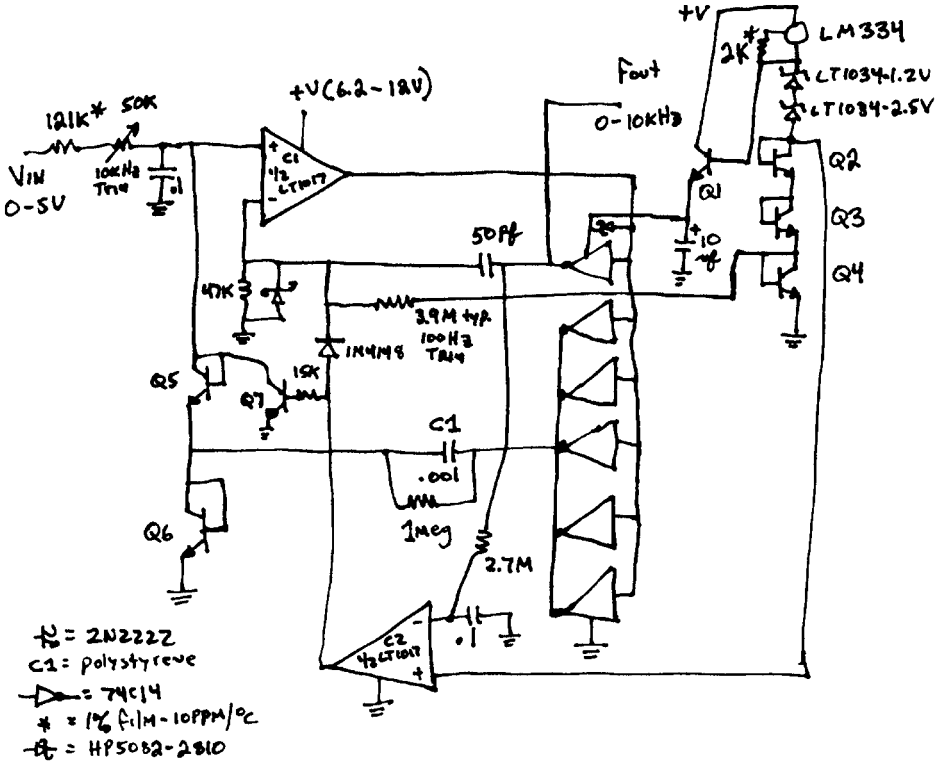


Figure 18-19: The zoo circuit.

descending to 80 µA for $V_{in} = 0$. Other specifications appear in Table 18-1. Much of this circuit should be familiar by now. Some changes have occurred, but nothing too drastic. The diodes have been replaced with transistors for lower leakage and more consistent matching. Also, paralleling the CMOS inverters provides lower-resistance switching. The startup loop has also been modified.

To maintain perspective, it's useful to review circuit operation. Assume *C1*'s positive input is slightly below its negative input (*C2*'s output is low). The input voltage causes a positive-going ramp at *C1*'s positive input (trace *A*, **Figure 18-20**). *C1*'s output is low, biasing the CMOS inverter outputs high. This allows current to flow from *Q1*'s emitter through the inverter supply pin to the 0.001 µF capacitor. The 10 µF capacitor provides high-frequency bypass, maintaining a low impedance at *Q1*'s emitter. Diode connected *Q6* provides a path to ground. The voltage that the 0.001 µF unit charges to is a function of *Q1*'s emitter potential and *Q6*'s drop. When the ramp at *C1*'s positive input goes high enough, *C1*'s output goes high (trace *B*) and the inverters switch low (trace *C*). The Schottky clamp prevents CMOS inverter input overdrive. This action pulls current from *C1*'s positive input capacitor via the *Q5*–0.001 µF route (trace *D*). This current removal resets *C1*'s positive input ramp to a potential slightly below ground, forcing *C1*'s output to go low. The 50 pF capacitor connected to the circuit output furnishes AC positive feedback, ensuring that *C1*'s output remains positive long enough for a complete discharge of the 0.001 µF capacitor. As in Figure 18-13, the 1 MΩ resistor completes *C1*'s discharge.

The Schottky diode prevents *C1*'s input from being driven outside its negative common-mode limit. When the 50 pF unit's feedback decays, *C1* again switches low and the entire cycle repeats. The oscillation frequency depends directly on the input voltage–derived current.

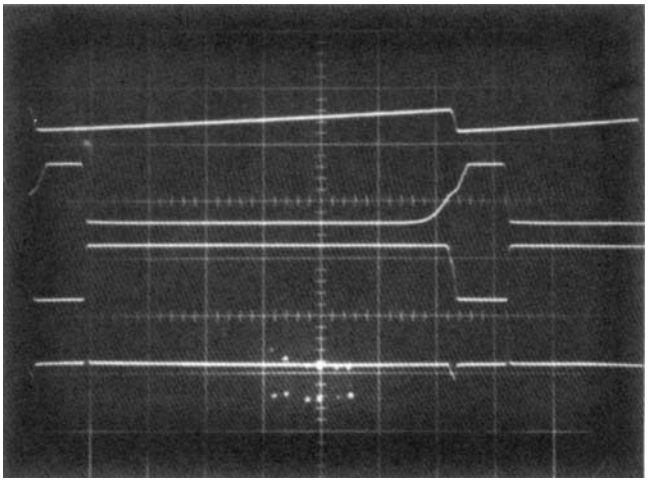


Figure 18-20: Fig. 18-19's waveforms.

Q1's emitter voltage must be carefully controlled to get low drift. *Q3* and *Q4* temperature compensate *Q5* and *Q6* while *Q2* compensates *Q1*'s V_{be}. The two LT1034s are the actual voltage reference, and the LM334 current source provides excellent supply immunity (better than 40 ppm/V PSRR) and aids circuit temperature coefficient. It does this by utilizing the LM334's 0.3%/°C temperature coefficient to slightly temperature modulate the voltage drop in the *Q2* to *Q4* trio. This correction's sign and magnitude directly oppose that of the −120 ppm/°C 0.001 µF polystyrene capacitor, aiding overall circuit stability.

The *Q1* emitter-follower delivers charge to the 0.001 µF capacitor efficiently. Both base and collector current end up in the capacitor. The paralleled CMOS inverters provide low-loss SPDT reference switching without significant drive losses. Additionally, the inverter specified is a Schmitt input type, minimizing power loss due to *C1*'s relatively slow rising edges. The 0.001 µF capacitor, as small as accuracy permits, draws only small transient currents during its charge and discharge cycles. The 50 pF−47 K positive feedback combination draws insignificantly small switching currents. **Figure 18-21**, a plot of supply current versus operating frequency, reflects the low-power design. At zero frequency, the LT1017's quiescent current and the 35 µA reference stack bias accounts for all current drain. There are no other paths for loss. As frequency scales up, the charge–discharge cycle of the 0.001 µF capacitor introduces the 7 µA/kHz increase shown. A smaller value capacitor would cut power, but the effects of stray capacitance, charge imbalance in the 74C14, and LT1017 bias currents would introduce accuracy errors. For example, if *C1* is reduced to 100 pf (along with other appropriate changes), the circuit consumes only 90 µA at 10 kHz, but linearity degrades to .05%.

Circuit startup or overdrive can cause the circuit's AC-coupled feedback to latch. If this occurs, *C1*'s output goes high. *C2*, detecting this via the inverters and the 2.7 M−0.1 µF lag, also goes high. This lifts *C1*'s negative input and grounds the positive input with *Q7*, initiating normal circuit action.

Because the charge pump is directly coupled to *C1*'s output, response is fast. **Figure 18-22** shows the output (trace *B*) settling within one cycle for a fast input step (trace *A*).

To calibrate this circuit, apply 50 mV and select the value at *C1*'s input for a 100 Hz output. Then apply 5 V and trim the input potentiometer for a 10 kHz output.

Table 18-1 shows what the customer ended up getting.

The zoo circuit made my customer happy, even if it is almost entirely bipolar. The inverter is the only piece of CMOS in the thing. I'm fairly certain the customer wouldn't mind if I had used 12AX7s[1] as long as it met specifications. It runs well in production, and they make lots of them, which makes my boss and the stockholders happy.

[1] For those tender of years, 12AX7s are thermionically activated FETs, descended from the work of Lee De Forest.

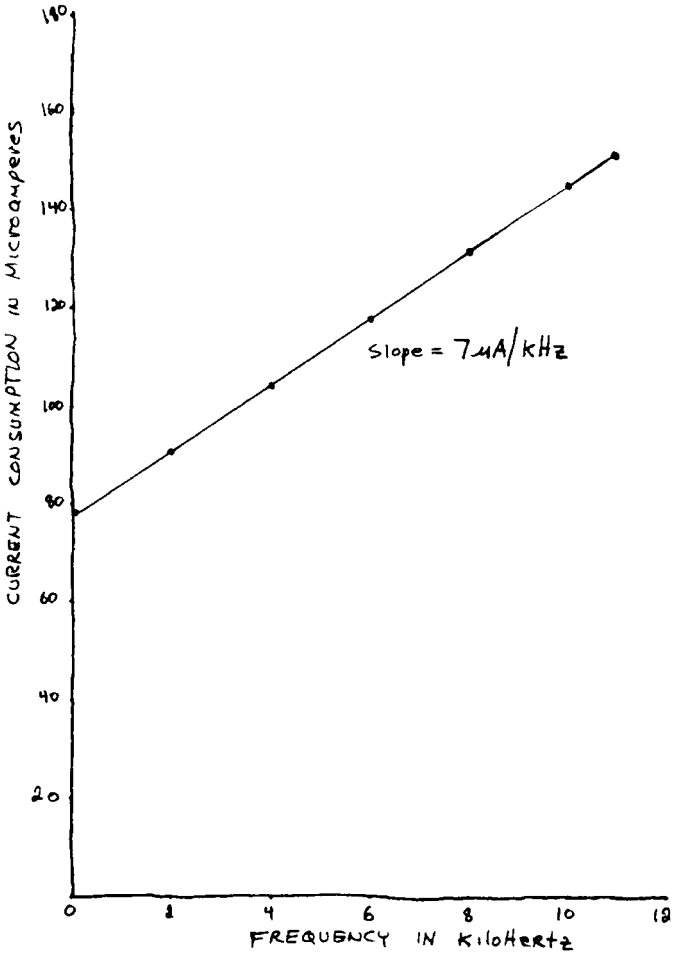


Figure 18-21: Current consumption versus frequency for Fig. 18-19.

This circuit has received some amount of attention in the technical community. I am aware of some spectacularly complex mathematical descriptions of it, along with some arcane explanations of its behavior. Similarly, it has been shown that the circuit could have only been arrived at with the aid of a computer. Given this undue credit, the least I could do is come clean about the circuit's humble origins.

I hope it was as much fun to read about the circuit as it was to build it.

Updated versions of this circuit, which draw as little as 8 μA, can be found in J. Williams' "Circuitry for Signal Conditioning and Power Conversion" (LTC Application Note 75, March 1999, pp. 1–4).

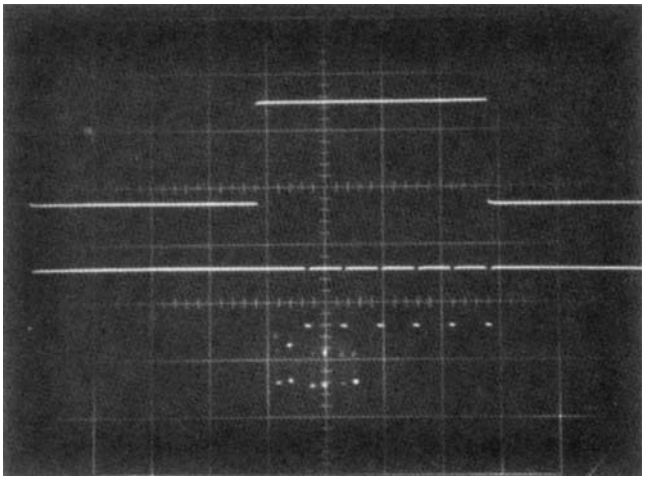


Figure 18-22: Figure 18-18's step response.

Table 18-1: Summary: voltage-to-frequency converter

Output frequency	0 to 10 kHz
Input voltage	0 to 5 V
Linearity	0.02%
Drift	40 ppm/°C
PSRR	40 ppm/V
Temperature range	0 to 70°C
Step response	1 cycle of output frequency
Output pulse	5 V CMOS-compatible
Power supply	Single 9 V battery (6.2 to 12 V)
Power consumption	145 μA maximum, 80 μA quiescent
Cost	< $6.00/100,000 pieces

References

General Electric Co, "Voltage-to-frequency converter," *General Electric Transistor Manual*, GEC, Syracuse, New York, 1964, p. 346.

Pease, R.A., "A new ultra-linear voltage-to-frequency converter," *NEREM Record*, vol. I, 1973, p. 167.

Pease, R.A., assignee to Teledyne, "Amplitude to frequency converter," U.S. patent 3,746,968, filed September 1972.

Williams, J., "Circuitry for signal conditioning and power conversion," LTC Application Note 75, March 1999, pp. 1–4.

———, "Designs for high-performance voltage-to-frequency converters," Linear Technology Corp., Application Note 14, 1986.

———, "Micropower circuits for signal conditioning,"*10 kHz Voltage-to-Frequency Converter,* Linear Technology Corp., Application Note 23, 1987, pp. 10–13.