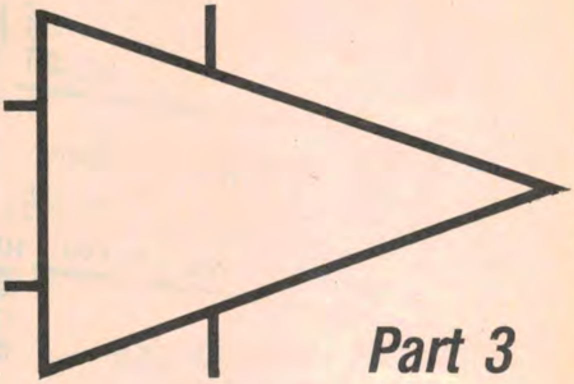
by BRYAN MAHER

OP AMPS Explained



Part 3

High gain operational amplifiers are not known for their low noise performance. In this article, we examine the causes of noise and discover what can be done about it.

The proliferation of high gain linear integrated amplifiers in the market place has done wonders in promoting circuit designs of high accuracy and linearity. Unfortunately, such progress has done little to advance the cause of low noise design and in many cases has mitigated against it. The reason for this is found buried deep in the fundamental philosophy of the operational amplifier – that the amplifier block G should have very high gain. Fig. 1, a common basic circuit, is expanded in Fig. 2 to show the action of the differential first stage.

In part one of this series we saw that at x, the inverting input of a high gain amplifier G, the feedback action and the high open loop gain combine to produce a low effective impedance to ground (the "virtual earth") and a very low signal voltage. Typical values for the case of closed loop gain = 10 are given in Table 1. The three integrated circuits named are available over the counter from a number of manufacturers.

In the application given in Table 1, all three types will give excellent results as regards DC stability, gain stability, linearity, and low sine wave distortion. Note, however, that the signal level applied to the input of the long tail pair is measured in microvolts. It is well known that every semiconductor junction generates noise, again usually of the order of some microvolts. As the input point x is applied directly to the first transistor junction (ie, to the base), it follows that if the signal voltage at x is as small as the junction noise voltage, then we will have a very noisy amplifier.

Because of the high again of G, the basic operational amplifier is fundamentally noisier than other designs. This is no slight on any manufacturer nor on any integrated circuit type. It is a statement true for all operational amplifier configurations. However, there is no need to immediately consign all our hard-won integrated circuits to the dust-bin.

Looking again at Table 1, it is clear that, for volt-size inputs, the voltage at x is large enough to swamp the transistor junction noise into insignificance. Also, the table implies that the 301A integrated circuit should be less noisy than the very high gain 725 type in the same amplifier configuration. Note that Table 1 shows only typical open loop gain values for each type. For a manufactured batch of any one type, the open loop gain G can vary over a 2:1 or even 3:1 range, while G also changes with operating temperature.

TABLE 1		INTEGRATED CIRCUIT FAMILY TYPE	OPEN LOOP GAIN	VOLTAGE AT x Vin = 1V Vout = 10V	VOLTAGE AT x Vin = 1mV Vout = 10mV	VIRTUAL EARTH IMPEDANCE
Ri	Rf					
1k	10k	301A	30000	330μV	0.3μV	0.3Ω
1k	10k	725	1000000	10μV	.01μV	.01Ω
1M	10M	545	50000	200μV	0.2μV	200Ω

The problem is, not all input signals are volt-size. Many signals are only of the order of a few millivolts or less, eg from phono cartridges, tape heads, microphones etc. To amplify these signals, we need to move away from the simple operational amplifier concept towards special low-noise designs.

It is agreed amongst the sages of this science that, in a semiconductor junction, there are three different effects, each a source of noise:

(a) Thermal noise, also called white or random noise;

(b) Shot or generation/recombination noise; and

(c) Flicker or 1/f noise.

The sum of all three is the noise we hear. (Other sounds can also interfere with our wanted signal, such as mains frequency hum and switching transients. These external noise sources are treated separately).

Thermal noise

Thermal noise is due to the random motion of charge carriers, usually electrons, in all conductors, resistors and semiconductors. Being a voltage proportional to the square root of bandwidth, we have little trouble with this effect for audio frequency work as long as we use good quality transistors and low resistance in the signal path. The high frequencies used for video signals are more of a problem. The noise voltage V(n) is actually equal to:

$V(n) = (4KTBR)^{½}$ where K is Boltzmann's constant, T is absolute temperature, B the bandwidth and R the series resistance.

Low noise designs

The basic idea of low noise design is simple: provide some gain in a very quiet amplifier stage ahead of the operational amplifier. If the operational amplifier sees a relatively high-level signal, noise will not be a problem. In order to design such a "front end" low noise amplifier stage, let's first consider the sources of noise in transistor junctions.

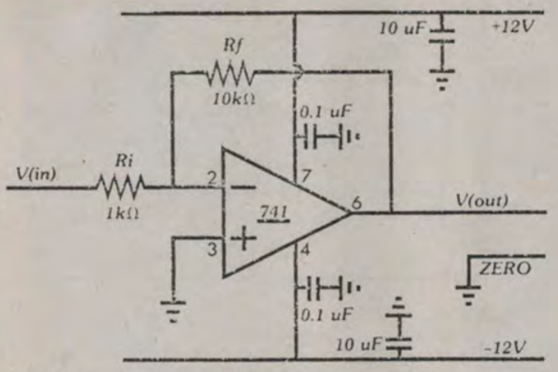


Fig. 1: basic operational amplifier circuit.

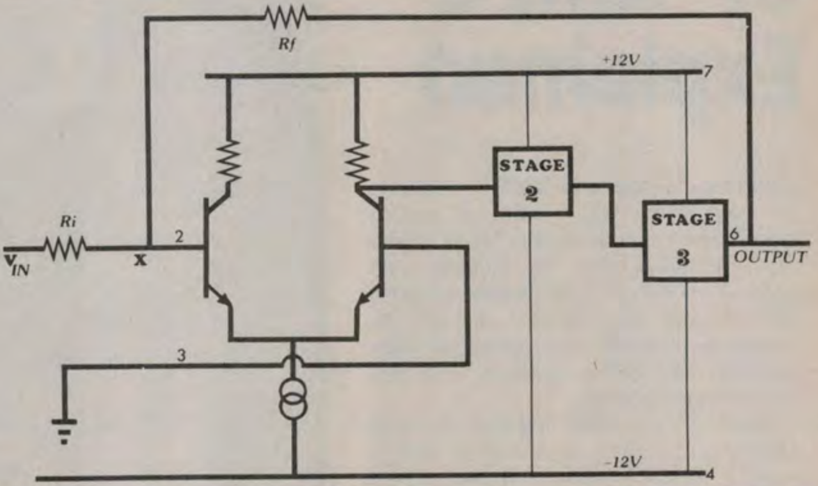
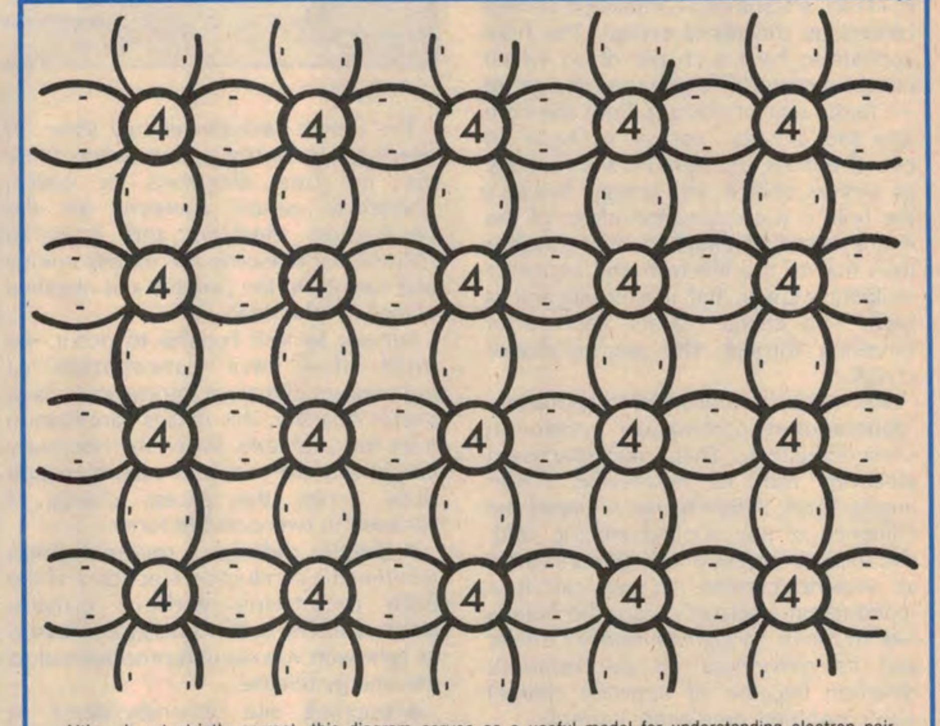


Fig. 2 (right): symbolic diagram of a common op amp IC showing the differential input stage. Ri is the input resistor, Rf the feedback resistor and 'x' the inverting input.



Although not strictly correct, this diagram serves as a useful model for understanding electron pair bonding in a crystal structure. The circles represent an atomic core having a charge of +4, while the minus signs indicate the positions of the valence electrons.

Incidentally, if resistors are used in series with the signal, they should be high-quality metal film or wirewound types. Never use carbon or composition types as these generate large amounts of noise due to their granular construction. In fact, experiments have shown that the "granular" noise generated by carbon composition resistors can be five times higher than the theoretical random noise of the same value resistor.

Shot noise

The generation of electron-hole pairs in a semiconductor, and their subsequent erratic recombination, is a second source of transistor noise, indeed the predominant one. To understand these effects, let's consider just what happens inside a semiconductor. Essentially, two separate actions occur simultaneously due to what are called the intrinsic and extrinsic properties of the semiconductor material.

The intrinsic properties of a semiconductor describe what goes on inside a pure crystal of semiconductor material, for example silicon, simply because the temperature is greater than absolute zero. In order to measure the intrinsic properties, it is imperative that the material be super-refined to no more than one impurity atom for every million million silicon atoms. Such super-pure silicon, in single-crystal, form exhibits the following true semiconductor properties:

- the conductivity is zero at absolute zero temperature because of the absence of charge carriers; and
- the conductivity increases as temperature increases due to the increase in the number of charge carriers.

The reason for this relation between the number of charge carriers and temperature is that, at absolute zero temperature, all electrons are busy forming valence bonds, the structural ties which hold the solid material together as in Fig. 3. These valence bonds are essential, as they lock atoms together into the crystal shape.

The electrons have energy levels in a range known as the "valence band of energy levels". Such electrons are improperly called "valence electrons", an expression used by everyone (including your author), though it is strictly wrong because you cannot put name tags on individual electrons. At temperatures above absolute zero only a few electrons absorb thermal energy, the rest absorbing nothing.

Those that do absorb energy find it impossible to absorb just a little bit; instead they find they have to absorb a lot of energy or nothing at all. Having absorbed this energy, such electrons now find that they are able to break away from valence bond duty and wander off through the crystal. Their energy level can be anywhere within a range of energies called the "conduction band of energies".

The reason why they have to absorb a lot of energy to get into such a condition is that between the valence band of energies and the conduction band of energies is a great void, a range of energies which are not available to any electron. This void in the energy levels is

OP AMPS Explained

called the "energy gap" and for silicon is 1.1 electron volts wide at room temperature as indicated in Fig. 4. If you are wondering why no electron can possess an energy level within this gap, the answer is that to do so its momentum would have to be a complex number, whereas in practice only real numbers are possible.

When an electron acquires thermal energy and moves to a higher energy level, if leaves behind one unoccupied energy level. This we call a "hole" which is, in fact, a result of a combination effect caused by the entire crystal. The hole appears to have a charge of +1 which complements the electron's charge of −1 (basic unit of charge). Both the hole and the electron appear to have an effective mass, momentum and mobility as well as charge and energy. Because the hole is a combination effect of the entire crystal, its effective mass is higher than that of the electron and hence its mobility is lower. Still, it is mobile and as such is a charge carrier capable of travelling through the semiconductor crystal.

Now we are in a position to appreciate "generation-recombination" noise in semiconductors. That much-discussed electron, with its new-found higher energy level, is free to move under the influence of any applied electric field, and this movement of charge constitutes an electric current. So we call it a "conduction electron". Also, the hole is free to move, though its mobility is less, and its movement (in the opposite direction because of opposite charge) also constitutes an electric current.

Leakage current

The sum of these two components is the leakage current or "saturation current" we observe in semiconductor diodes and transistors. Such leakage current is embarrassing in many aspects of circuit design and is one reason why silicon took over from germanium in the search for better semiconductor materials as it has less leakage current. The story we have just told is called "electron-hole pair" generation.

But worse is yet to come! You see, nature is "energy lazy" and always likes things to be in a state of lowest possible energy. That infamous "conduction electron", however, is in an elevated energy state, as is the hole. So the system is not in equilibrium.

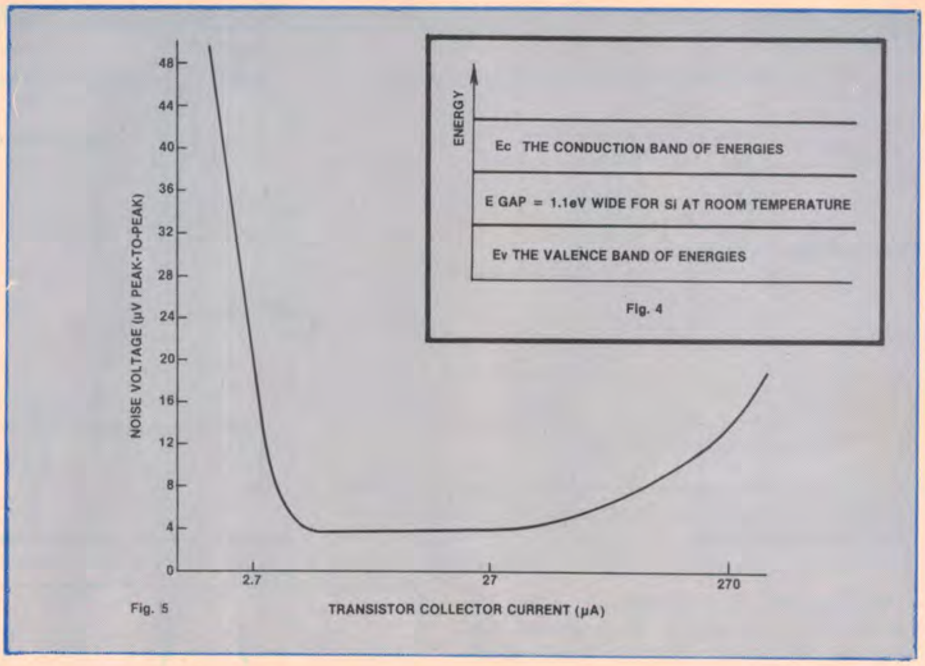


Fig. 4

Fig. 5

The lowest possible energy state for the intrinsic crystal occurs when there are no free electrons or holes. Therefore, nature contrives for the conduction electron and hole to recombine, releasing the excess energy and cancelling the positive and negative charges in the process.

But not so fast! For this to occur, we must also have cancellation of momentum. Because momentum is a vector quantity, this makes cancellation a lot less probable. When the necessary conditions are met and recombination does occur, the excess energy is released in two possible forms:

- direct radiative recombination wherein the conduction electron and the hole recombine without outside assistance. The excess energy is released as light with wavelength proportional to the energy release.
- trapping site recombinations in which the conduction electron and the hole use an intermediary energy level associated with either an impurity atom or a crystal imperfection.

William Shockley, soon after inventing the transistor, realised that direct radiative recombination was almost insignificant compared to the more numerous trapping site recombinations, the traps being energy states due to impurities or crystal dislocations. Such energy states usually lie deep in the energy gap.

Recombinations also produce a small voltage signal in the semiconductor. Because many recombinations go on every second in a semiconductor, apparently arbitrarily, the voltage signals so generated appear as a never-ending string of pulses of random amplitude, randomly spaced in time. This we call "shot" or "popcorn" noise. The fact that it is by far the largest source of noise forces us to consider it seriously.

Transistors, as you know, are made of "doped" semiconductor material, ie silicon to which other elements such as aluminium, gallium, indium, arsenic phosphorus or antimony have been deliberately added. Such doping does marvellous things to the energy gap and makes transistor action possible, thus giving rise to the "extrinsic" properties. But still the intrinsic generation-recombination noise phenomena persists in all transistors to plague us.

Design precautions

For the integrated circuit and transistor designer, minimisation of generation-recombination noise means a reduction in the number of trap sites present. A reduction in the impurity content (eg of elements such as copper) and a reduction in the number of crystal imperfections both help in this regard. The former requires better silicon purification before doping. The latter demands greater care in manufacture and subsequent handling to minimise crystal dislocations, as well as careful design of the crystal slice shape.

In addition, it has been found that high current gain goes hand-in-hand with low noise. This is not surprising as the small base volume associated with a high gain transistor provides less opportunity for recombinations.

For the circuit designer, shot noise reduction involves choosing low-noise integrated circuits and transistors; high h_{FE}; choice of optimum collector current; and minimisation of the number of semiconductor junctions present. This latter statement outlaws Darlingtons,

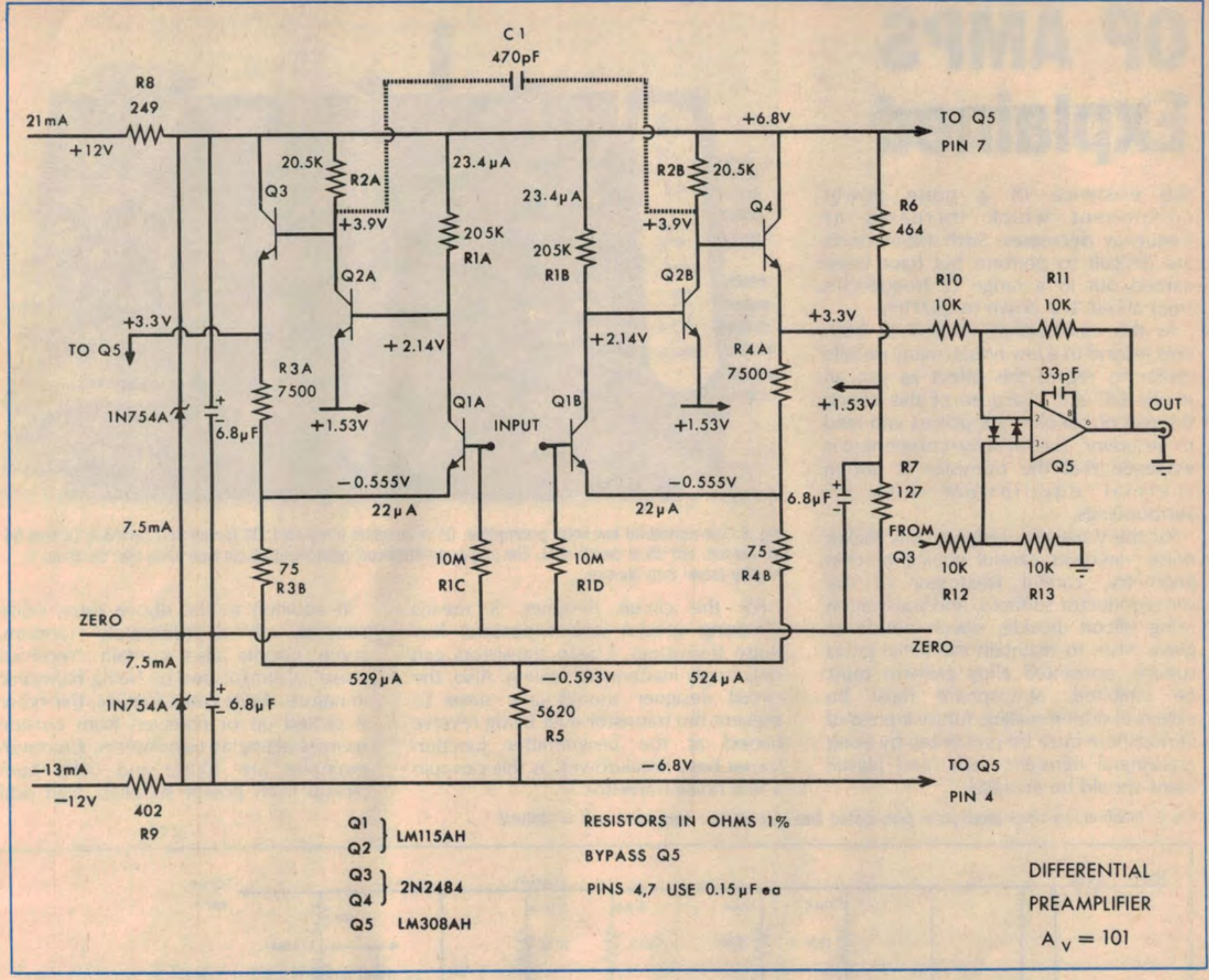


Fig. 6: Practical low noise preamplifier with closed loop gain = 101. Low noise results from moderate open loop gain (9000), avoiding the very small "voltage at x" problem.

super-beta transistors and active "tail" circuits. As field effect transistors are in general noisier than bipolar junction transistors, FET's should be avoided if at all possible. To prevent extra noise being generated because of a drop in h_{FE} at high frequencies, the circuit designer should start with transistors capable of high h_{FE} at frequencies much higher than the intended passband.

Flicker noise

Sometimes also referred to as "excess noise", flicker or 1/f noise is a surface phenomenon caused by the random trapping of charge carriers in fast surface energy states. Recall what we said about crystal imperfections or "faults" as trapping sites. Well, imperfections are also created in the surface of the crystal simply by using a saw to cut the crystal into slices. Experiments have established

Low-noise preamplifier PCB design

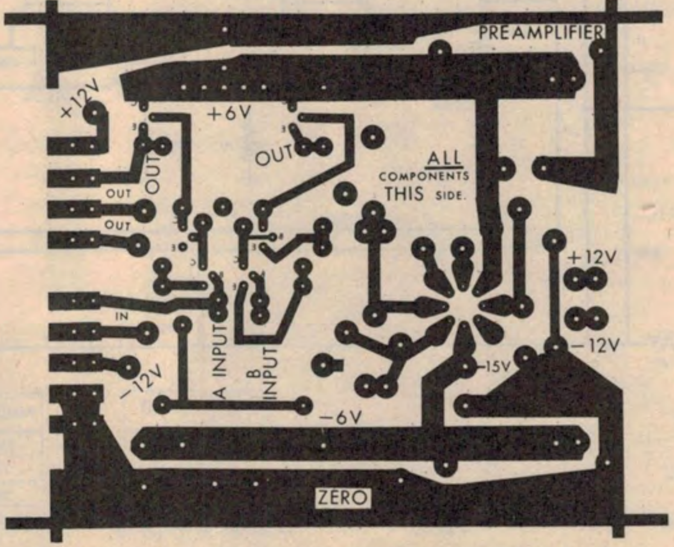


Fig. 7: actual size PCB design for the low noise preamplifier shown in Fig. 6. Note that, in this instance, the components are mounted on the copper side of the board. Component positions can be gleaned from Fig. 8.

OP AMPS Explained

the existence of a noise power component which increases as frequency decreases. Such experiments are difficult to perform but have been carried out in a range of frequencies from about 1Hz down to .0001Hz.

As this corresponds to periods from one second to a few hours, many people prefer to regard this effect as just an erratic DC drift. Because of this effect, the use of coupling capacitors can lead to problems. This 1/f noise component is evidence that the transistor is not in thermal equilibrium with its surroundings.

For the transistor designer, low flicker noise involves careful choice of chip geometry, careful treatment of the semiconductor surfaces, and passivation using silicon dioxide, silicon nitride or glass. Also, to maintain the initial good surface, unwanted alloy growths must be inhibited; atmosphere must be excluded during sealing, future ingress of atmosphere must be prevented by using glass-metal hermetic seals, and plastic cases should be avoided.

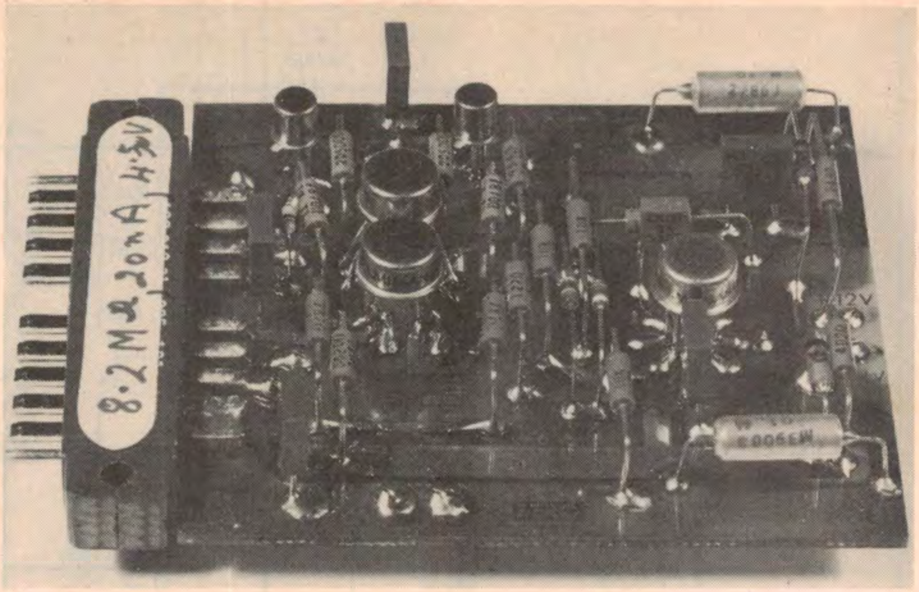
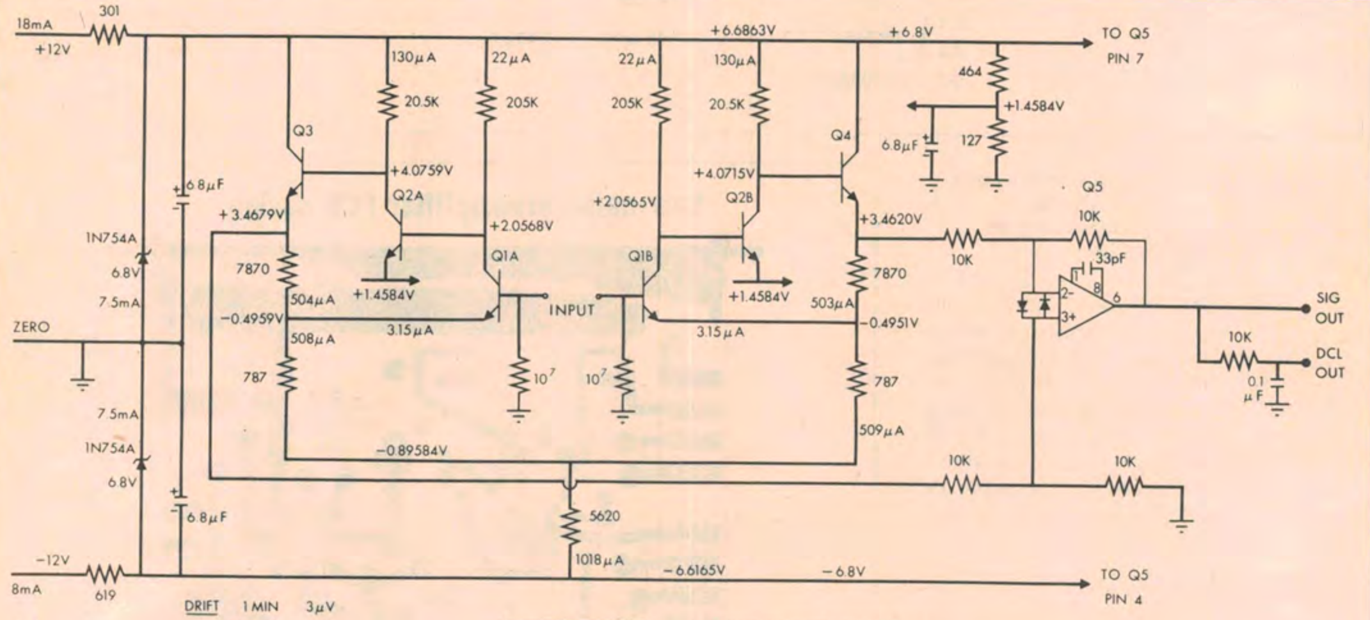


Fig. 8: The assembled low noise preamplifier. Q1 is at centre foreground, Q2 immediately behind it, Q3 and Q4 at rear left, and Q5 at centre right. The plug is an Amphenol gold plated 10-pin type. Note that the photo is slightly larger than life-size.

For the circuit designer, it means choosing special (and expensive) low noise transistors. Cheap transistors can only imply inadequate sealing. Also, the circuit designer should take steps to prevent the transistor ever being reverse biased at the base-emitter junction (zener base breakdown), as this can ruin a low noise transistor.

In addition to the above three noise sources, all approximately random, some circuits also contain "received noise", distinguished by being harmonic in nature. As its name implies, this noise is picked up or received from outside sources acting as transmitters. Common examples are 100Hz and 50Hz hum pickup from power supplies; high odd

Fig 9: Practical low noise preamplifier with closed loop gain = 11. Maximum input = 400mV



Q1, Q2 2N3117

Q3, Q4 2N2484

Q5 LM308AH

DRIFT 1 MIN 3μV

30 MIN 9μV

SLEW RATE LIMIT 0.125V/μS

(TO 460 mV INPUT)

INPUT LIMIT +400mV

−400mV

OUTPUT LIMIT +4.4V

−4.4V

R_{in}	8.18MΩ
i_{in}	20nA
T_R	1μS

RESISTORS IN OHMS 1%

BYPASS Q5

PINS 4, 7 USE 0.15μF

BANDWIDTH	NOISE μV p−p
1MHz	60
300K	36
100K	16
30K	9
10K	4.5
3K	2.7

PREAMPLIFIER

STAGE

AV = 11

ACCENT	LOW NOISE FET DC STABILITY
R_{in}	$\doteqdot 5.10^{10}\Omega$ (WITHOUT R1C, R1D)
i_{in}	5.10^{-13} A
T_R	1µS

BANDWIDTH	NOISE VOLTAGE (MICROVOLTS PEAK TO PEAK)
DC—1MHz	50
300KHz	30
100KHz	20
30KHz	15
10KHz	6
3KHz	4

TIME PERIOD	DC DRIFT
1 MIN	±1µV
30 MIN	+5µV
3 HRS	+15µV

$A_0 = 2040$

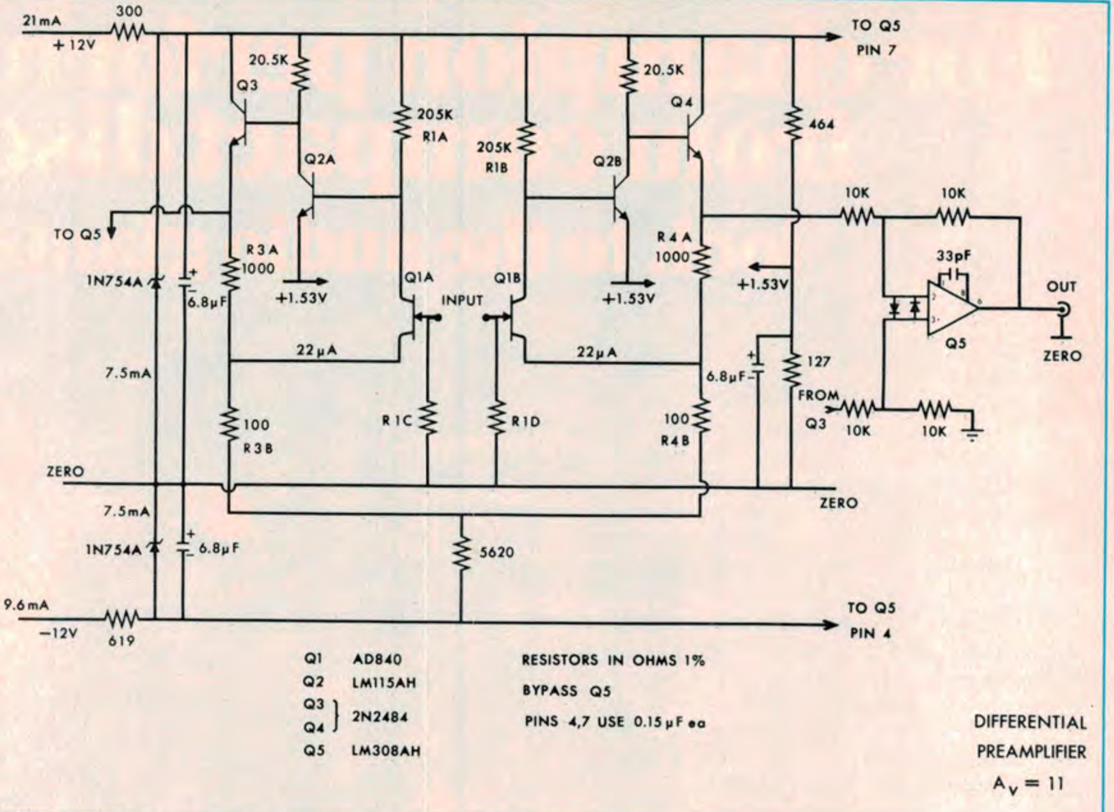


Fig. 10: FET low noise preamplifier with closed loop gain = 11. R_{in} = 50,000MΩ, i_{in} = 0.5pA.

OP AMPS Explained

harmonics picked up from 50Hz TV vertical deflection circuits; and bursts of noise from switch openings and closings.

The cure for these noise effects is obvious: better power supply design, decoupling, shielding and separation. Attempts to remove 50Hz interference by means of notch filters often fail as the radiated interference also contains a collection of higher harmonics.

Under some conditions much of the received hum, especially the 50Hz component, can be removed by the use of differential amplifiers having high common mode rejection. This is a subject which we will treat in a future part of this series. For the moment, we observe that all our integrated circuits are actually differential amplifiers, but we commonly implement them into single sided (non-differential) closed loop amplifiers.

The low noise design

From the foregoing, we see that our low noise amplifier should be designed along the following lines:

- All efforts should be concentrated on the first stage which will have a closed loop gain of about 100. After such gain, noise will not be a problem in most applications.
- The front end should be a feedback amplifier for linearity, but should only have moderate open loop gain so that the infamous "voltage at x" will not be too small.
- We construct our own amplifier for the first stage rather than use an integrated circuit.
- The front end should use bipolar junction transistors having wide bandwidth, high h_{FE} and the lowest possible noise figure; wire wound or metal film resistors; decoupled power supplies; and no coupling capacitors.
- As every junction contributes noise, the number of transistors used before gain should be kept to a minimum. Therefore we forbid active load and active tail circuits, super beta transistors and Darlington pairs.
- Tests should establish the optimum collector current.
- The input impedance should be adequate for normal requirements.

Suitable transistors

From data books, we observe that, for single NPN transistors, the 2N3117 has a very small noise figure (1dB under some conditions), with the 2N2484 and SE4010 running a close second. Amongst the dual NPN transistors, the LM114AH, LM115AH, 2N2920, LM194 and LM394 are some of the quietest. Tests were conducted to establish the collector current range for the lowest noise. The results, plotted in Fig. 5, show a low noise plateau in the curve for collector currents in the range 3 to 30 microamps. Other characteristics are also satisfactory in this range. For example, the LM114AH has h_{FE} = 500 minimum and f_T = 22MHz at 10 microamps.

A practical low-noise amplifier design is shown in Fig. 6 and uses LM115AH transistor input stages operating at 23 microamps collector current. The circuit can be broken down into two sections, the first consisting of a six transistor differential negative feedback amplifier (Q1A, Q1B, Q2A, Q2B, Q3 and Q4) with an open loop gain of 9000, a closed gain of 101 and differential output. This is followed by Q5, a differential unity gain operational amplifier.

Input is applied to the Q1A and Q1B bases, the input impedance and current being 8.2MΩ and 20 nanoamps respectively. Q1A and Q1B form a long tail pair with R5 the tail resistor. R3A, R3B, R4A and R4B form voltage dividers from the emitters of Q3 and Q4 to provide differential negative feedback to the emitters of Q1A and Q1B. The closed loop gain of the six transistor first stage is equal to 1 + (R3A/R3B) = 101. This emitter feedback raises the input impedance at the bases of Q1A and Q1B and widens the bandwidth of the six transistor stage to 500kHz, which can be reduced if you like by the addition of optional C1.

Q5 resolves the differential outputs from Q3 and Q4 into a single output. Because the signal is raised by a factor of 101 before being presented to Q5, the noise of this operational amplifier is much smaller by comparison, and hence unnoticeable.

The preamplifier responds linearly to input signals up to 43 millivolts, and has excellent DC stability and low drift. The latter is a result of a symmetrical circuit arrangement, dual transistors Q1 and Q2, small transistor heating (23 microamps collector current) and the use of 1% metal film resistors throughout. The back-to-back diodes across the input to Q5 prevent heating of the latter's input transistors in case of accidental signal imbalance, such as occurs in overdrive or power switch-on surges.

Stray capacitance

To ensure small stray capacitances, the amplifier should be constructed on a small PC board as shown in Fig. 7. If a plug and socket are used, gold contact surfaces are absolutely essential.

There is, of course, no reason why the operational amplifier stage Q5 could not have gain higher than unity. This would simply involve changing R11 and R13 from 10kΩ to some higher value, with a corresponding reduction in the 33pF compensation capacitor on pins 1 and 8. This increase in gain should not be taken too far, otherwise instability may result. If your application demands a circuit suitable for higher input voltages, a change may be made in the feedback ratio R3A/R3B, R4A/R4B. For example 7870 ohms/787 ohms will result in a first stage closed loop gain of 11, as in Fig. 9, with the ability to respond linearly to inputs up to 400 millivolts.

Measured noise in both cases is 700 nanovolts RMS, ie 4 microvolts peak to peak, from DC to 10kHz, referred to the input. If the 8MΩ imput impedance is not high enough or the 20 nanoamp input current is not low enough for some applications, two courses of action may be taken. First, Q1's collector current may be reduced to 4 microamps by raising R1A and R1B to 1MΩ. If R1C and R1D are now removed, the preamplifier will have an input impedance of 100MΩ and an input current of 4.0 nanoamps, with no noticeable change in noise.

The alternative is to substitute a low noise dual field effect transistor for Q1 with minor circuit changes as in Fig. 10. The dual FET type AD840 was chosen for this role. With R1C and R1D removed, the input impedance measured 50,000MΩ with an input current of 0.5 picoamps. Noise is about double the previous example, or 1.2 microvolts RMS (6.0 microvolts peak to peak), from DC to 10kHz.

The dual transistors and FETs specified are excellent units and well worth their higher price. But your poor impoverished author thought he would define a new unit to be called the "noise microvolt-dollar product", which should be kept to a minimum. Accordingly, the circuit Fig. 6 was rebuilt using six SE4010 transistors, while R1C and R1D were changed to 1MΩ. All other components remained the same, except that no plug-socket was used.

The result was a preamplifier of much lower cost, but with noise almost the same: 700 nanovolts RMS or 4 microvolts peak to peak (referred to input) from DC to 10kHz. Input impedance with R1C and R1D measured approximately 1MΩ and the input current 0.1 microamps. Because the transistors are not pairs, DC drift was much higher but this was reduced by sliding tygon tubing over pairs of transistors and pumping the tubing full of Dow-Corning DC4 silicone grease for thermal coupling. Readers who are still awake will find a summary of these preamplifier variations in Table 2.

Next month, we will consider the implications of DC coupling, high gain and drift.

PREAMPLIFIER	Fig.	NOISE (0-10kHz)		★ Rin Ω	iin A	Cin pF	† BANDWIDTH Hz	† OPEN LOOP GAIN Ao	CLOSED LOOP GAIN Av	Vin MAX ±mV	Vout MAX ±V	DRIFT RATE µV/hr
		P-P µV	RMS µV									
BASIC Q1 = 2N2920 (ic = 4µA, R1A,B = 1M)	9	4	0.7	10^8	4.10^{-9}	20	100000	6000	11	300	3.3	20
BASIC Q1 = 2x2N3117	6	3.8	0.7	8.10^7	2.10^{-8}	20	500000	9000	100	38	3.8	300
BASIC Q1 = 2xSE4010	6	4	0.7	10^6	10^{-7}	20	300000	4000	100	38	3.8	90
BASIC Q1 = 2N2920, LM114AH	6	4	0.7	8.10^7	2.10^{-8}	20	500000	9000	100	38	3.8	6
FET Q1 = AD840	10	6	1.2	5.10^{10}	5.10^{-13}	20	500000	2040	11	120	1.32	6
FET Q1 = AD840 (16V VERSION)	10	6	1.2	5.10^{10}	4.10^{-13}	20	500000	4000	10.1	500	5	6

TABLE 2

★ R1C,R1D DELETED

† SIX TRANSISTOR SECTION, FROM LOW IMPEDANCE

Table 2: summary of preamplifier performance variations for different transistors.