

THESIS

AN OPTICAL COMMUNICATIONS LINK

For the degree of M.Sc.
in Electrical Engineering

by

M. H. MOORE

Supervisor: Professor J.L.N. Besseling

The copyright of this thesis is held by the
University of Cape Town.

Reproduction of the whole or any part
may be made for study purposes only, and
not for publication.

The copyright of this thesis vests in the author. No quotation from it or information derived from it is to be published without full acknowledgement of the source. The thesis is to be used for private study or non-commercial research purposes only.

Published by the University of Cape Town (UCT) in terms of the non-exclusive license granted to UCT by the author.

Contents.

<u>Section.</u>		<u>Page.</u>
	Acknowledgments	ii
	Abstract	iii
1.0	Viability of an Optical Communications Link	1.
2.0	The Optical Channel	4.
2.1	A Comparison of Photodiode and Avalanche Photodiode Characteristics.	11.
2.2	The Receiver Frontend	14.
2.3	Lens Systems	19.
3.0	Choice of System Configuration	25.
3.1	Modulation of the Optical Carrier	27.
3.2	The FM System as Applied to the Optical Channel	31.
4.0	Development of a Practical System	39.
4.1	Major System Parameters	39.
4.1.1	Choice of Emitter and Detector	39.
4.1.2	The Power Supply	40.
4.1.3	Modulation Scheme Values	43.
4.2	The Receiver	46.
4.2.4	The Receiver Demodulator	46.
4.2.5	The Receiver Bandpass Filter	51.
4.2.6	The Frontend Circuit	55.
4.2.7	Component Values and Detailed Analysis	59.
4.2.8	The Receiver Lens	71.
4.2.9	The Receiver Audio Amplifier	73.
4.2.10	Receiver Power Saving Circuitry	75.
4.2.11	Receiver Circuitry - General	78.

4.3	The Transmitter	80.
4.3.12	The Transmitter Modulator	80.
4.3.13	Microphone Preamplifier and Level Compressor	86.
4.3.14	Transmitter Pulse Shaper	92.
4.3.15	The Output Stage	96.
4.3.16	The Transmitter Lens	101.
5.0	Predicted Performance of the FM Communicator Design	105.
6.0	Breadboard Testing	107.
7.0	Prototype Design	113.
8.0	Future Development	120.
	Bibliography	121.

Appendices:

1.	Device Data	125.
2.	Background Radiation and Optical Filtering	129.
3.	Double-tuned Bandpass Filter Circuit Analysis	131.
4.	J.f.e.t. Selection Test Jig	133.

Acknowledgments.

I would particularly like to thank my supervisor, Professor J.L.N. Besseling, for suggesting the project, and for providing many ideas and much valuable assistance throughout its development. Of the numerous others who have had a hand in the work, I must single out for special thanks Messrs J.H. van Nierop, who took over as my supervisor during the closing stages; John Greene, of Plessey, who gave of his specialised opto-electronic knowledge; Dr David Swingler, who advised on modulation theory; Malcolm Attfield, who helped with the printed circuit production; David Kenyon, who produced the prototypes with admirable speed in the face of adversity; and last but certainly not least my sister, whose typing endeavours helped the thesis to meet its deadline.

Thanks are also due to the Council for Scientific and Industrial Research, who provided financial support during the first two years of my research.

Abstract:

The thesis describes the development of a specification for, and prototypes of, an opto-electric voice communications link. The introductory sections deal with the generalised optical communications channel, the available hardware, and some of the research that has been done in the field. The information presented is used to motivate the type of system to be developed. The emphasis is placed on cost, though not as an overriding consideration. Modulation systems are examined, and frequency modulation of an m.f. subcarrier is chosen. The development of the practical system is covered in detail, in particular the receiver frontend circuitry. Considerable discrepancies between the design figures and actual measured performance are analysed, and their causes located. A practical mechanical design is presented, with suggested modifications for production. In conclusion, future developments in the field are examined.

Viability of an Optical Communications Link.

To justify the already considerable research and development effort that has been put into the field of optical communications, there must be some considerable advantages resulting from the use of optical carriers. This is particularly true in areas where large quantities of information must be transferred very rapidly, as in links between computers.¹ The world's expanding medium- and long-distance telephone systems are tending toward the use of extremely high carrier frequency systems, either microwave, using atmospheric or low-loss coaxial cable links, or optical, using fibre-optic lightpipe techniques. Thus a potential advantage of the optical channel is the very wide modulation bandwidth available.

A further advantage, of major practical importance in point-to-point communications, is the essentially directional character of nearly all optical links. A beam divergence of 1 milliradian is quite feasible with a lens of a few centimetres diameter. By comparison, a microwave dish would have to be several metres across to achieve the same effect. The combination of a very narrow beam, with no dispersion, and limited to line-of-sight range, makes for an inherently very secure communications system. To eavesdrop on such a system without revealing one's presence to the users would be extremely difficult. Thus it was that the U.S. Navy chose a short-range optical link to provide ship-to-ship communication during fuelling at sea, while denying Russian electronic eavesdroppers the chance to listen in.²

Directionality has another important effect, in that interference between two optical communications links working on the same wavelength is very unlikely. Thus the overcrowding of the electromagnetic spectrum that has occurred at radio frequencies will not be a problem in the optical domain. Licensing of optical transmitters should not be necessary, and very large numbers of systems should

/ be....

1. Refs. 21, 24, 26

2. Ref. 22

be able to share the same bandwidth.

There are several disadvantages associated with optical communications. All point-to-point links are restricted to line of sight distances, and the transmitter and receiver have to be aimed at each other. There is a direct relation between ease of aiming and range, for the broader the transmitted beam, the lower the source radiant intensity, and the greater the receiver acceptance angle, the smaller the effective detector area. Thus a system that is relatively non-directional will of necessity have a very limited range (the Holobeam-U.S.N. system), while a link of several kilometres will require extreme care in setting-up due to its directionality. ³

Atmospheric attenuation is another problem associated mainly with long-range optical links. It is essential to avoid operation at wavelengths in the various carbon dioxide and water vapour absorption bands. Failure to do so can result in almost complete attenuation of the desired signal in a few hundred metres of air path. Fortunately there is only one seriously absorptive frequency in the near infrared area of the spectrum, and that is slightly lower than those of the most commonly used optical sources. Haze, fog and rain will all attenuate the optical carrier to some extent, and high reliability atmospheric links must take this into account. ⁴ However, ranges of up to 48 km have been reported in experimental systems with clear atmospheric conditions. ⁵

An important difference between radio frequency and optical frequency systems lies in the nature of the devices which capture the signal at the receiver. The radio antenna is a low-impedance voltage source, easily matched to bipolar transistor circuitry, and exhibiting essentially no noise other than the thermal noise associated with its source resistance. The photodetector ⁶ is a current source with a very high source resistance, shunted by a small capacitance, and this makes a low-noise receiver design difficult. In addition, practical photodetectors suffer from internally-generated noise resulting from leakage currents.

3. Ref. 25

4. Ref. 27

5. Ref. 23

6. Refs. 2,3,5,28

/ The.....

The high frequency of optical carriers means that the individual photon energy is much higher than at radio frequencies. Hence the rate of photon emission for a given power output is lower, and the optical signal itself is noisy, because of the random emission of these photons. The noise current in the detector is proportional to the square root of the signal current-bandwidth-electronic charge product. Thus at low signal currents, the signal itself may limit the noise performance of the system, although in practice this is unlikely. In addition, the photodetector usually intercepts some unwanted background radiation, which also generates a noise current in the detector. This optical background noise will often be the limiting factor in the performance of any atmospheric optical link.

SUMMARY:

Optical communications brings the prospect of very wide bandwidth short-to medium-range channels, with many users sharing the same frequencies, but separated spatially so that no interference results. The optical link is directional, requiring care in aiming, but conferring a high degree of secrecy, desirable for military applications. Atmospheric conditions affect the range achievable, and quantum noise in the detector limits the minimum detectable signal.

The Optical Channel.

The optical communications channel can be considered as consisting of two transducers and a variety of noise sources. The electro-optical transducers are potential limiting factors in the performance of an optical channel, since the optical source's efficiency is usually low, and the detector's leakage generates shot noise. Further sources of noise are the quantised nature of the light detected by the receiver, and the thermal and shot noise contributions of the interfacing circuitry in the receiver. In any optical communication system, a measure of the performance will be the optical signal power required to equal the total noise power at the detector. The transmitter's optical power output per unit solid angle and the area of the receiving transducer will then determine the range of the system.

1. The detector.

Detectors currently in use fall into two categories with different performance characteristics: those with, and those without noise-free post-detection gain. The conventional p-n and p-i-n photodiodes belong to the latter group, while the former comprises the avalanche photodiodes and photomultiplier tubes. All of these devices convert optical power to electrical current. Incoming photons transfer their energy to electrons. These excited electrons are then able to pass some potential barrier, and a current flows in an external circuit. Its magnitude is $I = SP_o$, where P_o = incident optical power, and S = sensitivity (A/W), dependent on wavelength. This current is quantised, and so is noisy. The mean square noise current is $i_{ns}^2 = 2qIB$, where q = electron charge, B = noise bandwidth, and I is the total current. Generally, the signal current flows in a load resistor R which itself generates a thermal noise power $4kTB$ (T = absolute temperature). The signal power in R is $i_s^2 R$, and the shot noise power contribution, assuming a noise-free amplifier, is $2qBR(i_s + i_d + i_b)$. The shot noise has three sources: the signal current itself i_s , the dark (leakage) current of the detector i_d , /and.....

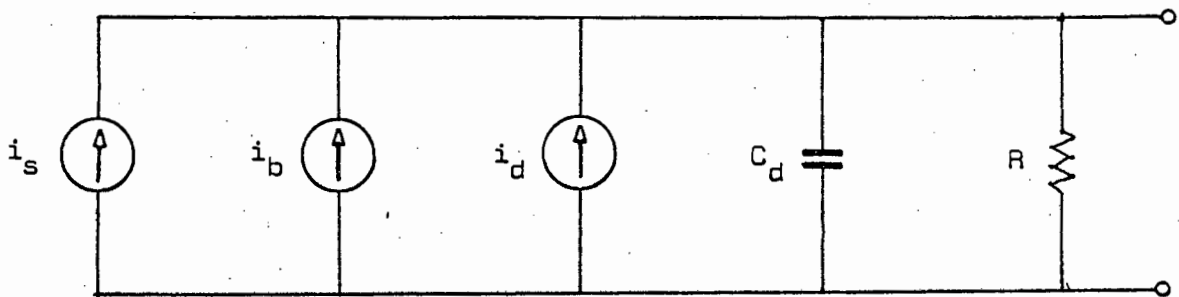
and the current due to background radiation perceived by the detector i_b . Thus the signal-to-noise power relation is:

$$\frac{S}{N} = \frac{i_s^2 R}{4kTB + 2qB(i_s + i_d + i_b)}$$

An alternative form of the above considers the mean square signal-to-noise current ratio:

$$\frac{S}{N} = \frac{i_s^2}{4kTB/R + 2qB(i_s + i_d + i_b)}$$

Practical detectors are limited in their speed of response by the capacitance appearing across their terminals. A simple generalised detector equivalent circuit is:



$$i_s + i_b = S.P_{opt} \quad R = \text{load resistor}$$

A high-frequency breakpoint occurs at $f = 1/2\pi RC_d$. This apparently inherent bandwidth limitation can be circumvented, as will be shown later. Meanwhile it is clear that if B is proportional to $1/R$, then the mean square shot noise current is independent of R . The thermal noise, however, falls with increasing R . Thus it is likely that a wideband system, with R small, will be limited by thermal noise. Conversely, very narrowband receivers with large R will be shot noise limited. If possible, it is desirable to make R large to reduce the thermal noise contribution.

Background radiation conditions will largely determine the shot noise contribution. In an optical channel using a light pipe, background will be

/ negligible.....

negligible, and detector leakage current or thermal noise will dominate. In a point-to-point optical link, background radiation due to reflected sunlight will swamp any other effects, and the system will be quantum limited. Under these conditions the use of a detector with noise-free gain does not confer any advantage over the cheaper photodiode detectors.

The actual amount of signal intercepted at the detector is dependent on its apparent area. A large area increases signal-to-noise ratio by increasing i_s . However, it also increases C_d in the case of junction detectors, with a consequent reduction in bandwidth. If a narrow field of view is acceptable, a lens can be used in front of a small detector to increase the apparent area very considerably. Since the source is an approximate "point source at infinity", incoming radiation is collimated (parallel), and can be focused by an ideal lens to a single spot. Thus the apparent detector area is approximately the lens area. A further advantage results from the reduced angle of acceptance, in that the incident background radiation is of the same order as that received without any lens. Signal-to-noise improvements of the order of 40 dB have been reported when using large diameter, high-quality optics. The circle of confusion of the lens used will reduce the effective area below the ideal value, if the detector diameter is less than that of the circle of confusion. Alternatively, a slight defocused spot can be used in a tradeoff between sensitivity and acceptance angle. ¹

2. The Source.

Two broad categories of sources can be distinguished: light sources which can be modulated directly, and those which require external modulation. The second category is generally less suited to any kind of portable equipment, since the external modulator is usually bulky, whether mechanical (Aldis lamp) or electrical (Kerr and Pockels effect modulators). These last offer the possibility of very high information bandwidths using continuous-wave lasers as primary sources, but for voice communication they represent an expensive, overcomplicated solution.

Sources which can be modulated directly, that is, electrically, form three subgroups : incandescent, light emitting diodes, and junction laser diodes.

Incandescent sources have been used in modulated light voice communicators (the Wehrmacht used such a system in World War II), but their simplicity brings with it several problems. Frequency response is very limited, because of thermal capacity of the filament. Communications bandwidth (3kHz) can be achieved using pre-emphasis. To improve matters, special low-mass filaments might be used, but that creates a fragility problem, undesirable in a portable unit. The limited bandwidth available means that no noise-reducing modulation schemes can be used. Also, detector design is complicated by having to operate in the $1/f$ region.

Light-emitting diodes are GaAs or GaAsP junction electroluminescent devices in which optical output power is proportional to forward current. The relatively large junction area typical of present-day emitters results in a substantial (hundreds of pF) capacitance across the diode. This capacitance limits the maximum modulating frequency. Intensity modulation at tens of megahertz can be achieved with suitable low-impedance drive circuits. Electro-optical conversion efficiency is low, of the order of a few percent. To make matters worse, many l.e.d.s emit almost uniformly into a complete hemisphere. This makes it impossible to capture a large fraction of the emitted light with a lens system. In some cases l.e.d.s are manufactured with an integral epoxy lens. This has two advantages : the emitted radiation can be more easily captured, since most of it emerges in a narrow cone, and the actual efficiency is raised. This latter effect is due to the very high refractive index of gallium arsenide ($N = 3,6$), resulting in a critical angle for the GaAs-air interface of 16° . Much radiation thus never gets out of the GaAs crystal. Addition of an epoxy dome on a l.e.d. does not automatically mean a small emitting angle, however. Some such devices still retain the broad polar pattern characteristic of hemispherical GaAs emitters.

/ Most

Most high-power l.e.d.s. are GaAs diodes, with a peak emission at about 900 nm in the near infrared. The optical bandwidth between half-power points is about 40 to 50 nm. This relatively broad peak represents a disadvantage for systems which are to operate in high background radiation conditions. It is not possible to use a really narrow-band (say 5 nm) optical interference filter at the detector to exclude the ambient radiation, as the filter would then remove most of the signal as well.

Two categories of GaAs emitters exist: diffused diodes and silicon-doped epitaxial diodes. The former group are typified by low efficiency (0,2 to 2%) and fast response time (10 ns or less). Silicon-doped devices have a longer peak wavelength (930 to 970 nm), higher efficiency (6 to 30%), but a considerably slower response time (up to 500 ns). Thus silicon-doped emitters are preferred for applications below 1 MHz, and diffused emitters for higher frequencies or short pulse times.

The third important category of directly-modulated sources are the p-n junction lasers. A p-n junction is constructed within an optical cavity, which is arranged to provide optical feedback. When the device current density reaches a threshold level, the amount of radiation fed back is sufficient to initiate laser action. Power output of a junction laser is thus low until the threshold current density is reached, after which it increases approximately linearly with current. Since the efficiency will depend on the diode current, it is convenient to consider the incremental efficiency:

$$\Delta\eta = \frac{P}{(I - I_0) V} \quad \text{where} \quad \begin{array}{l} P = \text{emitted power} \\ I_0 = \text{threshold current} \\ V = \text{bias voltage} \end{array}$$

Power is dissipated in the internal loss resistance R , so that the total power efficiency is

$$\eta_p = \frac{(I - I_0) V \Delta\eta}{IV + I^2 R}$$

/ Depending...

Depending on the type of structure employed, values of $\Delta\eta$ of up to 40% have been achieved. However, typical overall efficiencies are considerably lower, in the 5 to 10% region.

Unlike conventional GaAs emitters, laser diodes emit in the plane of the junction. Since high current densities are required to achieve lasing action, the emitting area is small. This results in a source with very high radiance, of the order of $10^6 \text{ W/cm}^2 \text{ sr}$, compared with l.e.d. radiances of the order of at most $10^2 \text{ W/cm}^2 \text{ sr}$. Commercially available laser diodes are capable of relatively high peak output powers. Single diodes with 10 watt peak output are readily available, and arrays with a total output of several hundred watts have been produced. However, diode heating caused by the very high current levels required results in a severe restriction on maximum pulse duty cycle. At 300° K , typical laser diodes are limited to less than 0.1% duty cycle.

The effect of this low duty cycle is twofold. Average output power is quite low, of the order of milliwatts. More seriously, if high repetition rates are to be maintained, pulse duration must be reduced to the order of nanoseconds. Generation of a 10 A, 10 ns current pulse requires special drive circuitry and extreme care in parts layout. A further problem associated with frequency operation is the delay between application of the current pulse and emission of light. This delay can amount to several nanoseconds, and thus represents an upper limit on pulse rates. If extremely high data rates are essential, it is possible, at great sacrifice in efficiency, to use a prebias technique. A bias current of just less than the threshold value will reduce the turn-on delay to negligible levels. In this way modulation frequencies of 10 to 100 MHz at 0.1% duty cycle are theoretically possible.

The polar characteristics of junction lasers are due to the small size of the emitting area, resulting in diffraction limiting. The fan-shaped radiation pattern is not realised in practice, because of lack of phase coherence

/ across.....

across the whole length of the junction. Even so, the beam angle between half-intensity points is of the order of 30° to 40° , considerably less than lensless l.e.d.s.

The spectral characteristics of laser diodes vary with current density. In the area of interest, with current appreciably greater than the threshold value, nearly all of the output power is confined to a number of discrete optical modes within a spectral range of several nanometers (less than 5 nm). This narrow-band spectral characteristic permits the use at the receiving detector of interference filtering to remove most of the noise-generating background radiation. Account must be taken of the shift in peak emitted wavelength with diode temperature. This shift is caused by variation in band gap energy and in refractive index of the semiconductor, and amounts to about $+0,3 \text{ nm}/^\circ\text{K}$. Since the junction temperature can change appreciably during a current pulse, any interference filter used will have to be wider than the instantaneous bandwidth of the laser.

Laser diodes capable of continuous operation at room temperature have been developed ¹, but they are not yet available commercially except at very high prices (R200 and upwards). Typical output powers and efficiencies are comparable with conventional l.e.d.s., but the emitting area is much smaller, so that very small beam angles can be achieved by collimation of the output.

A Comparison of L.e.d. and Laser Diode Characteristics.

The characteristics presented below are typical figures, and do not relate to specific devices:

	<u>L.e.d.</u>	<u>Laser</u>	
Peak power	2×10^{-2}	10	Watts
Average power	10^{-3}	10^{-2}	Watts
Maximum duty cycle	100	0,1	%
Maximum pulse length	unlimited	200	ns
1. Ref. 31		/Maximum	

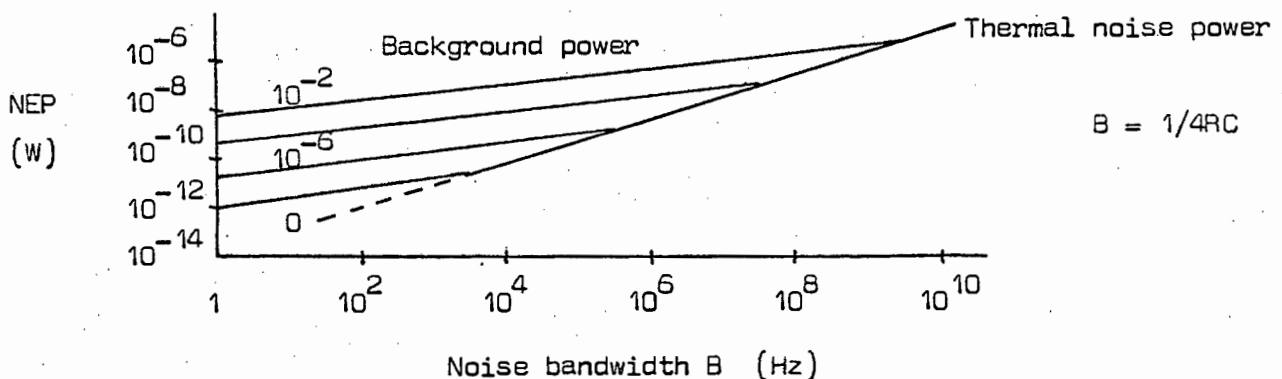
Maximum repetition rate	5×10^7	10^5	Hz
Minimum useful drive current	5×10^{-2}	5	Amps
Radiance	10^{-2} to 10^2	10^6	W/cm ² sr
Operating life	10^5	10^3	hours
Unit cost	1 to 5	10 to 50	Rands

In general, the l.e.d. is more suited to a portable voice link because of its low cost, simpler drive requirements and greater ease of modulation. If maximum range were the only criterion, a laser-based system using a pulse modulation scheme would offer definite advantages. System cost and power consumption would however be high.

A Comparison of Photodiode and Avalanche Photodiode Characteristics.

P-i-n Photodiode:

Consider a p-i-n photodiode with $S = 0,3 \text{ A/W}$, $i_d = 2 \times 10^{-9} \text{ A}$, $C = 7,5 \text{ pF}$ (in fact the H-P 5082-4207). One can plot total NEP in watts against bandwidth, with background radiation power as parameter.



$$\text{Thermal noise power} = 4kTB \quad (1,6 \times 10^{-20} B \text{ at room temperature})$$

$$\text{Shot noise power} = i_{ns}^2 R = 2qIBR \quad (I = i_s + i_d + i_b ; i_s = 0)$$

$$\text{Total noise power, } P_n = 2qIBR + 4kTB \text{ watts}$$

Consider the incident power P_o required to give a S/N ratio of unity.

P_o is a measure of the total detector system noise, the NEP or noise equivalent power.

/Hence.....

$$P_o = i_n^2 / S$$

where i_n = total noise current

$$\text{Hence } P_o = \frac{1}{S} \left(\frac{P_n}{R} \right)^{\frac{1}{2}}$$

S = sensitivity

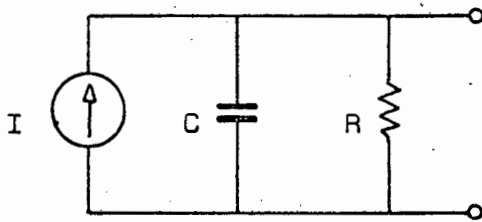
$$P_n = i_n^2 R$$

$$P_o = \frac{B^{\frac{1}{2}}}{S} \left(2qI + \frac{4kT}{R} \right)^{\frac{1}{2}}$$

$$P_o = \frac{1}{S} \left((2qIB)^{\frac{1}{2}} + 4B(kTC)^{\frac{1}{2}} \right) \quad \text{since } B = 1/4RC$$

Thus the shot noise component of P_o is proportional to $B^{\frac{1}{2}}$, while the thermal component is directly proportional to B . Hence one expects thermal noise to dominate P_o at large B .

Consider the advantage to be gained from use of a transresistance amplifier configuration (current-to-voltage converter). Assuming that the amplifier used is noiseless, one finds that the thermal noise equivalent power for a given bandwidth is reduced by a factor $1/A_V^{\frac{1}{2}}$.



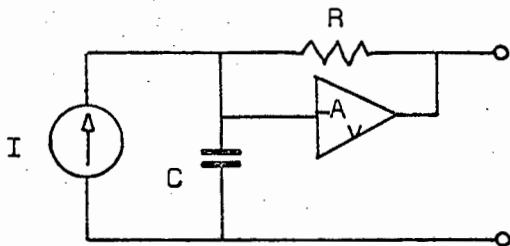
$$3 \text{ dB bandwidth} = 1/2\pi RC$$

$$\text{Noise bandwidth } B = 1/4RC$$

$$\text{Thermal noise power} = 4kTB = kT/RC$$

$$\text{Thermal N.E.P} = \frac{4B(kTC)^{\frac{1}{2}}}{S}$$

Conventional photodiode model.



$$3 \text{ dB bandwidth} = A_V / 2\pi RC$$

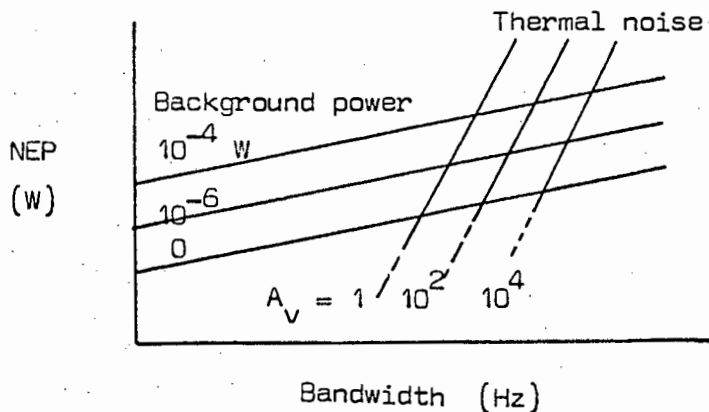
$$\text{Noise bandwidth } B = A_V / 4RC$$

$$\text{Thermal noise power} = 4kTB = kTA_V / RC$$

$$\text{Thermal NEP} = \frac{4B(kTC)^{\frac{1}{2}}}{S} \frac{1}{A_V^{\frac{1}{2}}}$$

Transresistance model

Thus use of an amplifier with a gain of 1000 would increase B by a decade and a half, for the same thermal NEP.



The effect of this feedback technique is to extend the range of bandwidths and background powers for which the detector will be shot noise limited. In cases where the background radiation is appreciable, use of a transresistance amplifier may well make the p-i-n photodiode competitive with more costly avalanche detectors.

Avalanche Photodiode:

Avalanche photodiodes belong to the class of detectors with noise-free gain. Incident photons cause avalanche breakdown of the reverse-biased diode, resulting in multiplication of the photocurrent. The multiplication factor m can be of the order of 100. Noise current is also multiplied, by a factor $m' = m^x$, where x for silicon is approximately 1,15. The expression for NEP for an avalanche photodiode is :

$$P_o = \frac{B^{\frac{1}{2}}}{S} \left(2qI \left(\frac{m'}{m} \right)^2 + \frac{4kT}{m^2 R} \right)^{\frac{1}{2}}$$

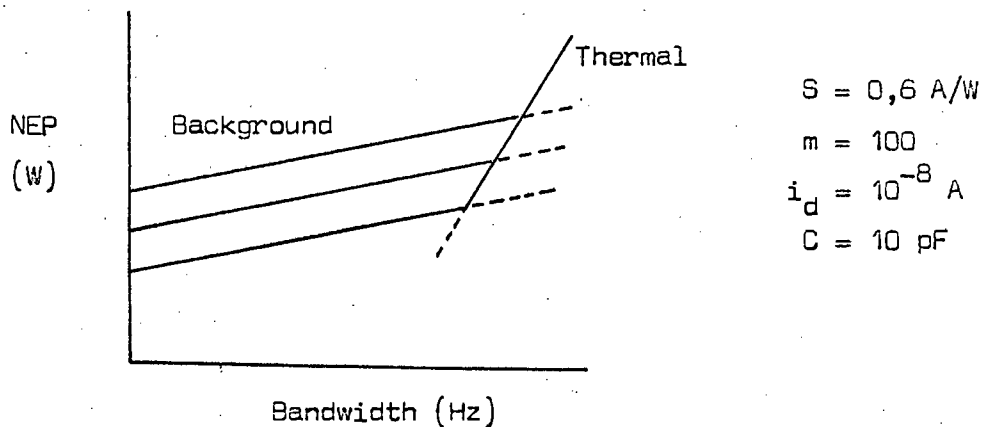
For $m = 100$, this becomes:

$$P_o = \frac{B^{\frac{1}{2}}}{S} \left(8qI + \frac{4kT}{10^4 R} \right)^{\frac{1}{2}}$$

Compared with normal photodiodes, the background shot noise equivalent power is doubled for given I , while the thermal NEP is reduced by a factor $1/m$, equal to 0,01 for a typical device with $m = 100$. The avalanche photodiode is superior to a p-i-n photodiode with similar current, sensitivity, and capacitance at high bandwidths or at very low background radiation levels. Thus for a daylight

/ point-

point-to-point voice communications system the p-i-n detector is to be preferred.



Note that in the above equations for the avalanche diode, $I = SP_b + i_d/m$.

The Receiver Frontend.

The advantage of a detector with low NEP can be nullified by the use of unsuitable amplification circuitry, contributing significant noise to the system. Design of a low-noise frontend is complicated by the fact that the detector is a current source. The use of bipolar transistors in the first stage of the amplifier is undesirable, when other noise sources are small, because there will always be a shot noise current component caused by the transistor base current. Reduction of the device collector current to minimise shot noise results in bandwidth degradation. A reasonable compromise can be achieved with careful design, but a better solution is to use a junction f.e.t. as the first stage. Since the j.f.e.t. has an input current which is essentially zero, its input noise current is very small. There is also a noise contribution from thermally-generated noise in the j.f.e.t.'s conducting channel, but this is a second-order effect when one is dealing with high resistance sources.

A problem common to both bipolar and field effect transistors is that of excess or $1/f$ noise. This results in a sharp increase in noise per unit bandwidth at low frequencies. Below some noise corner frequency, typically in the range 300

/ to.....

to 3000 Hertz, the spectral density of noise is inversely proportional to frequency (hence $1/f$). It is desirable, if possible, to operate outside this $1/f$ region.

When it is operating at frequencies above several kilohertz, a j.f.e.t. generates noise in two ways. Since the conducting channel of the device is an ohmic resistance (for small signals), thermal noise is generated. This can be represented by a noise voltage generator in series with the input. For a device operating in the saturation region, this voltage has the value:

$$e_n = (2,67kTB/g_m)^{\frac{1}{2}}$$

$$= \frac{1,05 \times 10^{-10}}{(g_m)^{\frac{1}{2}}} \quad V/(Hz)^{\frac{1}{2}} \text{ at } 25^\circ \text{ C.}$$

Since g_m will typically be 1 mA/V or more, a maximum value for e_n is approximately 3,3 nV.

The other major noise source in the j.f.e.t. is the gate leakage current shot noise.

$$i_n = (2qBI_g)^{\frac{1}{2}}$$

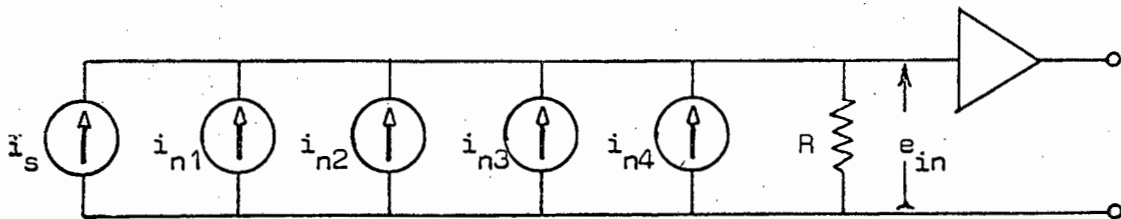
$$= \frac{1,05 \times 10^{-10}}{g_m^{\frac{1}{2}}} \quad \text{where } q = 1,6 \times 10^{-19} \text{ Coulomb}$$

$$V/(Hz)^{\frac{1}{2}} \text{ at } 25^\circ \text{ C.}$$

Gate leakage increases exponentially with temperature, so that for the ultimate low-noise performance, cooling to cryogenic temperatures would be necessary. In practice, choice of a suitable device will ensure that the gate leakage shot noise remains negligible over the commercial (0° to 70°C.) temperature range.

A Noise Model For The Receiver Frontend.

Since the signal appears as a current source at the input of the frontend amplifier, it is convenient to model the noise sources as current generators in parallel with the signal source. The noise performance is, theoretically at least not affected by the actual configuration of the frontend. A load resistor plus voltage amplifier or operational amplifier current to voltage conversion technique may be used. The latter has a bandwidth advantage, for any given resistor value.



i_s = signal photocurrent

i_{n1} = photodiode noise current = $(2qB(i_s + i_d + i_b))^{1/2}$

i_{n2} = thermal noise in load or feedback resistor = $(4kTB/R)^{1/2}$

i_{n3} = j.f.e.t. gate leakage noise = $(2qBi_g)^{1/2}$

i_{n4} = j.f.e.t. channel noise = $e_n/R = (2,67kTB/g_m)^{1/2}/R$

i_{n5} = second stage noise = $e_{n2}/A_{v1}R$, where e_{n2} is 2nd stage i/p noise.

Total noise current is:

$$i_{nt} = (i_{n1}^2 + i_{n2}^2 + i_{n3}^2 + i_{n4}^2 + i_{n5}^2)^{1/2}$$

As a figure of merit for the frontend, one can calculate the amplifier noise figure:

$$F = \frac{i_{n1}^2 + i_{n2}^2 + i_{n3}^2 + i_{n4}^2 + i_{n5}^2}{i_{n1}^2 + i_{n2}^2}$$

Ideally, F should be unity. Although this definition of a noise figure is somewhat / different.....

different from the normal resistive source definition, it shares the property of dependence on the source characteristics. Thus any variation in the source noise will affect the value of F , that is, the amount by which the amplifier degrades the incoming signal-to-noise ratio. In general, comparisons of values of F are not valid. But, given well-defined source and load resistor conditions, F provides a valuable means not only of comparing different frontend designs, but also of evaluating the 'goodness' of a particular design as related to the given conditions.

The above noise model applies only at frequencies above the $1/f$ noise region, and below the device high-frequency limits. It can be extended to operation in the $1/f$ region by modification of the equations for j.f.e.t. noise. However, this will have to be done on an empirical basis, as the data necessary for an accurate calculation of the excess noise characteristics, using the method of Sah², is not readily available. At high frequencies, noise again increases above the values given by the simple model. Van der Ziel³ and Klaasen⁴ have arrived separately at very similar values for this additional noise current, which results from the noise voltage in the channel. This noise voltage causes a reactive current, flowing in the total gate-source capacitance. Although the models do not specifically take account of external impedances, the nature of the noise source indicates that including the external gate-source capacitances in the formula is necessary for a correct evaluation of f.e.t. amplifier noise at high frequency.

The noise current is:

$$i_{n6}^2 = 4kTB \frac{(\omega C_{in})^2}{g_m} F_1 \left(\frac{W_{gs}}{W_o}, \frac{W_{gd}}{W_o} \right)$$

where F_1 is typically 0,3 under saturation conditions. This source must be included in the expressions for total noise current and noise figure. Substitution for F_1 in the above formula gives:

$$i_{n6}^2 = \frac{1,2kTB(\omega C_{in})^2}{g_m}$$

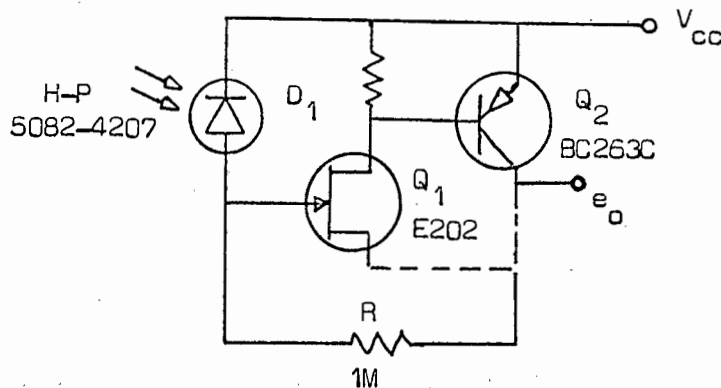
2. Ref. 35

3. Refs. 33 and 34

4. Ref. 43

A Worked Example.

Consider the frontend as used during the latter stages of development:



The noise bandwidth is 35 kHz, centred around 450 kHz. Taking the worst case (for noise figure) of operation in darkness ($i_b = 0$), with small signal level ($i_s \neq 0$), one derives the following results:

Data: $B = 3,5 \times 10^4$ Hz	$4kT = 1,6 \times 10^{-20}$ W/Hz	$i_d = 2 \times 10^{-9}$ A
$R = 10^6$ Ohm	$i_g = 10^{-10}$ A	$g_{m1} = 10^{-3}$ S (mho)
$A_{V1} = 0,68$	$C_{in} = 12$ pF	$e_{n2} = 1,6 \times 10^{-9}$ V/(Hz) $^{\frac{1}{2}}$

Photocurrent shot noise	$i_{n1} = 4,73 \times 10^{-12}$ A
Load resistor thermal noise	$i_{n2} = 2,37 \times 10^{-11}$ A
Midband j.f.e.t. shot noise	$i_{n3} = 1,06 \times 10^{-12}$ A
Midband j.f.e.t. thermal noise	$i_{n4} = 6,11 \times 10^{-13}$ A
Referred second stage noise	$i_{n5} = 4,40 \times 10^{-13}$ A
High frequency j.f.e.t. noise	$i_{n6} = 1,39 \times 10^{-11}$ A

The resulting amplifier noise figure is $F = 1,33 = 1,25$ dB. Thus although the frontend is operating at a frequency above its noise corner frequency ($i_{n6} > i_{n3}$), the degradation in signal-to-noise ratio is small, and would be smaller, but for the photodiode junction capacitance, which contributes significantly to C_{in} .

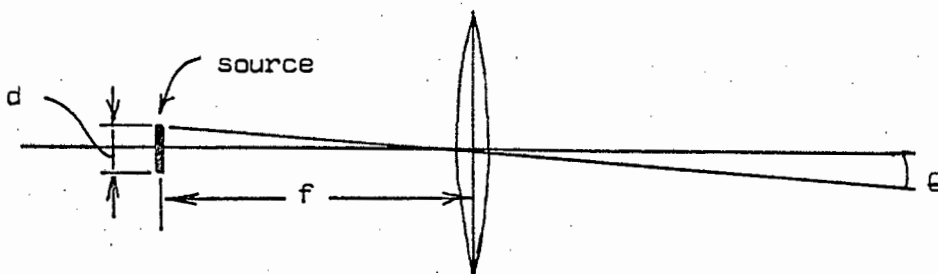
Summary:

It has been shown that the combination of operational amplifier techniques and a j.f.e.t. as the input device can produce a broadband optical receiver frontend which closely approximates the ideal noiseless amplifier. It is / desirable....

desirable for best low-noise performance to avoid operation at very low or high frequencies. Noise performance at high frequencies can be predicted, but the $1/f$ noise contribution, if not specified by the device manufacturer, must be measured.

The Transmitter Lens System.

Since no commercially available sources suitable for hand-portable use produce a really narrow beam, a lens is necessary to collimate the radiation, so as to maximise the axial radiant intensity, and consequently, the range of the system. Choice of lens focal length is determined by the dimensions of the source and the desired beam angle.



The beam half-angle $\theta = \arctan (d/2f) \approx d/2f$ for small θ . Defocusing can increase the beam angle of a given lens-source combination by a factor 2, while preserving a fairly uniform intensity across the beam. Excessive defocusing (source well inside the focal distance) results in the creation of a beam with a bright ring around a dim central region, and is not desirable in this application.⁵

In order that the radiant intensity in the beam be as large as possible, the lens must collect as much of the source's radiated power as possible. A wide-aperture lens is thus preferred, with an upper limit on size being set by the intended physical size of the system. The problem of capturing the radiated power is particularly severe for sources with wide emitting angles. Given a source with polar intensity distribution $I(\theta)$, the fraction of the emitted power contained in a cone of half-angle θ is ⁶ :

5. Ref. 7

6. Ref. 2

$$F(\theta) = \frac{\int_0^\theta I(\theta) \sin\theta \, d\theta}{\int_0^{\pi/2} I(\theta) \sin\theta \, d\theta}$$

For an isotropic source, emitting uniformly into a hemisphere, $I(\theta)$ is a constant, and therefore :

$$F(\theta) = 1 - \cos\theta$$

For a Lambertian source such as a flat GaAs crystal, $I(\theta) = I_0 \cos\theta$, and

$$F(\theta) = 1 - (\cos\theta)^2$$

The inefficiency of these types of source pattern is demonstrated by calculating the value of $F(\theta)$ for the case of a lens with $F/1$ aperture (an expensive lens). For a Lambertian source, 20 % of the radiated power is collected, while for an isotropic source the figure is only 10 %. For an $F/2$ lens, the respective figures are 5,9 % and 3 % .

Sources with integral lenses do not fall into either of the above categories. $F(\theta)$ for these sources can be evaluated numerically if the device polar pattern is known. In some cases, the manufacturer quotes the axial radiant intensity J , in mW/sr , and provides a plot of $J(\theta)$. Then the power emitted into a cone of half-angle θ is given by :

$$P(\theta) = 2\pi \int_0^\theta J(\theta) \sin\theta \, d\theta$$

The advantage of a source with integral lens over isotropic or Lambertian sources can be seen by considering a practical example. The Monsanto ME60 is a low-power GaAs emitter intended for use in punched-card and paper-tape readers. Its beam is a cone of 17° half-angle between half-intensity points. An $F/1$ lens ($\theta = 26,6^\circ$) will capture over 40 % of the total radiated power, and an $F/2$ lens, about 20 % . Thus, using an $F/2$ lens, an ME60 will produce a collimated beam about six times as powerful as an isotropic emitter of the same output power.

A design problem encountered when using integral-lens sources is that of

/ predicting....

predicting the beam width when an external lens system is added. Usually neither the emitting area nor the integral lens focal length is specified. It is thus necessary to measure these parameters, or, if this is not possible, to find the correct external lens focal length by empirical means. Adjustment of the source-lens separation will allow some spread in beam angles to be obtained with any given lens, so that with luck a ready-made lens can be used, at some reduction in cost.

Summary:

At the transmitter of an optical communications link, a lens is required to capture as much of the emitted power as possible, and to produce a collimated beam. If the source has an isotropic or Lambertian characteristic, a large-aperture lens is needed. L.e.d.s with integral lenses allow a greater proportion of the optical power to be collected by a relatively small-aperture lens. In general, laser diodes fall between Lambertian and lensed sources in their polar characteristics, and thus still need large-aperture lenses for efficient light capture.

The Receiver Lens System.

As outlined earlier, the use of a lens at the receiver gives markedly greater sensitivity at the expense of sharply decreased field of view. As a result of spherical aberration, a simple lens produces an image of a point source at infinity that is a blurred spot.⁷ The diameter of this spot is approximately :

$$d_b = 0,07 \frac{f_r}{(F.no.)^3}$$

where f_r is the focal length, and F.no. is the focal number.

Now in a practical communications system, the receiving lens is in the far field of the transmitter, and is subject to uniform irradiance H . Thus the total power intercepted by a lens of area A_r is $H.A_r$. This radiation is focused to a blurred spot of diameter d_b , at a distance f_r from the lens. The irradiance

7. Refs. 8 and 10 / in the plane.....

in the plane of the image is :

$$H_i = \frac{4HA_r}{\pi d_b^2} T \quad \text{where } T \text{ is the lens transmission coefficient}$$

If the receiving diode has area A_d , and if A_d is less than the area of the blurred image, then the power received by the diode is :

$$P_r = H_i A_d = \frac{4A_d A_r HT}{\pi d_b^2}$$

The effective gain resulting from the use of a lens at the receiver can be evaluated by considering the lens capture area. For the case of the small-area detector, the blurred spot area is larger than the detector area A_d , so that the fraction of the power transmitted through the lens that is incident on the detector is the ratio of A_d to blurred spot area. The capture area of the lens is the product of the actual lens area and the fraction of the radiation incident on it which actually reaches the detector. Then one can write the simple relation $P_r = H.A_c$, where H is the incident irradiance and A_c is the capture area. Using the previous expression for P_r , one can derive a formula for the capture area :

$$\begin{aligned} A_c &= \frac{4A_d A_r T}{\pi d_b^2} \\ &= \left(\frac{d_r^2 T}{d_b^2} \right) A_d \quad \text{where } d_r \text{ is the receiving lens diameter.} \end{aligned}$$

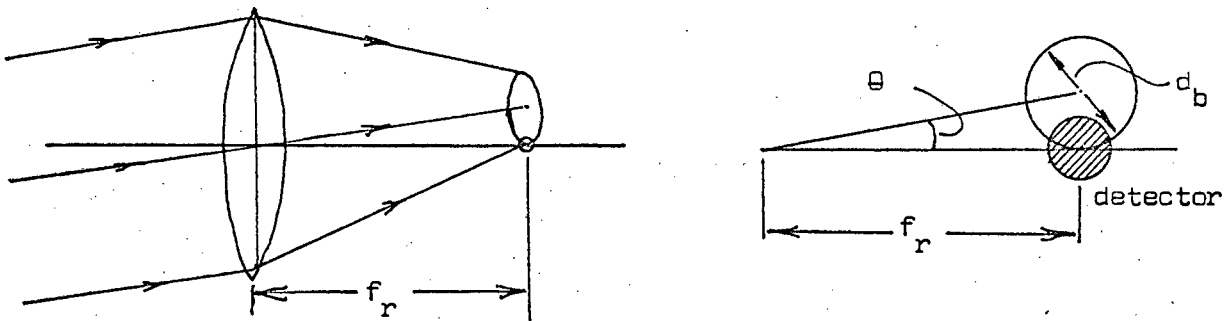
The term in parentheses corresponds to the power gain due to the use of the lens. For the critically-focused lens with minimum spherical aberration, this reduces to $A_c = 200 (F.no.)^4 . TA_d$. Typically, $T = 0,9$, giving $A_c = 180 (F.no.)^4 . A_d$.

Thus as a result of the blurring caused by spherical aberration, a large-aperture lens actually produces less received power than a smaller lens. However, an infinite reduction of lens aperture does not result in infinite P_r . The limit is reached when the areas of the blurred spot and of the detector diode are equal. This would be the optimal situation for a fixed communications link, but would create serious aiming problems for a portable system, as the acceptance

/ angle

angle would be extremely small. As with the transmitter optics, a certain amount of defocusing can be used to broaden the beam angle, at the expense of reduced H_i and hence P_r .

The beam angle for the detector at the focus can be estimated, provided that the image spot is several times larger than the detector. Consider a non-axial parallel beam of radiation producing an off-centre blurred image :



If the lateral displacement of the blurred image away from the axis is $\frac{1}{2} d_b$, then approximately half the detector area will be irradiated, and the received power will be halved. The corresponding angular displacement θ is half the total angle perceived by the detector through the lens.

$$\theta = \arctan \frac{d_b}{2f_r} \div \frac{d_b}{2f_r} \text{ (radians) for small } \theta.$$

This approximation introduces less than 1% error for θ up to 0,17 radian, or about 10° . Substituting for d_b in terms of F. no. and f_r gives $\theta = \frac{0,035}{(F.no.)^2}$ rad. or $\theta = \frac{2}{(F.no.)^2}$ degrees. Hence an F. number of 1,0 will give a 2° beam half-angle with the detector at the focus.

In applications where a large field of view is essential, two solutions are possible. The lens can be dispensed with altogether, resulting in a very wide acceptance angle, but very low sensitivity. To increase sensitivity, a detector with a large active area is required. Silicon photodiodes are available with areas of the order of 1 cm^2 , while maintaining low NEP values, circa $10^{-13} \text{ W/Hz}^{\frac{1}{2}}$. These large-area detectors are however expensive, costing around R50 to R100.

/ To obtain

To obtain an intermediate value of beam width, a combination of large-area detector and lens can be used. The circle of confusion of the lens must be small relative to the detector size, so that all the radiated power gathered by the lens falls on the detector. Thus the received power would be $P_r = HA_r T$, and the acceptance half-angle would be $\theta = \arctan \frac{d_d}{2f_r}$, where d_d = detector diameter. The acceptance angle is not increased by defocusing.

All of the above configurations are the results of a trade-off between acceptance angle, or ease of aiming, and receiver sensitivity in terms of minimum detectable power. But the use of a lens at the receiver has one decided advantage: it increases the signal photocurrent substantially, while leaving the background current almost unchanged. The background photocurrent does not increase because the average background radiance can be considered as constant, and although the effective detector area increases, the effective background radiation source area is reduced by the cut in acceptance angle.

A further matter to be considered when using a lens is its effect on optical interference filters. Since these devices are designed to work with parallel light normal to the filter surface, the optical bandwidth will be increased. Typically the increase is of the order of 50 to 100 %, and filters must be selected accordingly.

Summary:

The use of a converging lens at the receiver results in increased sensitivity and reduced beam angle, with background photocurrent essentially unchanged. A combination of high sensitivity and a large angle of acceptance requires the use of a large-area detector, with associated higher costs. Care must be taken in the choice of lens - the biggest is not necessarily the best, because of the spherical aberration limit.

Choice of the System Configuration.

After an investigation of the literature had been undertaken, it was apparent that research has been concentrated on high bit-rate systems for telephonic and data communications. Some work has also been done on portable and hand-portable systems for military voice communications.¹ These military systems have suffered from a philosophy that placed performance above all other criteria, with the result that their cost is unrealistically high for the non-military user. Thus there appeared to be a potential demand for a relatively low-cost optical voice communications link, with a range of the order of hundredaf metres.

Initial feasibility studies suggested that such a voice link could be made hand-portable, and at a cost which, while not competitive with the ubiquitous 27 MHz walkie-talkie, would still be low compared to available optical communicators. The basic instrument could be modified by use of different optics to give a narrow-beam portable link with ranges of over 1 km. Uses for this system could be found in land surveying (in conjunction with infra-red distance-measuring instruments, which lack a voice facility), on large civil engineering works such as dams and bridges, and in secure military and non-military communications. A further application would be as a cheap low bit-rate computer data link in situations where telephonic links are unavailable or inconvenient.

Once the decision had been taken to place emphasis on the cost aspects of the design, certain overall trends became clear. At the receiver, large-area photodetectors and avalanche photodetectors would not be justified, since their cost would be an order of magnitude higher than readily-available small-area p.i.n. photodiodes. To achieve a reasonable range with a small-area detector, it would be necessary to use a lens at the receiver to present a large capture area to the incoming radiation. Thus the field of view of the receiver would be limited to a few degrees, but tests carried out with a dummy transceiver indicated that hand-held aiming of the beam to within a degree should be quite practical, while fixed-base operation would present no problems.

1. Refs 22, 29, 30.

Choice of the infrared sources for the transmitter reduced to two types of commercially-available device, depending on the modulation scheme used. A pulsed drive might favour the use of the relatively cheap single-hetero junction laser diodes made by RCA. Peak output power of these devices is about one watt, and cost is in the region of R10. Simpler forms of modulation such as intensity modulation would require the use of an emitter capable of continuous room-temperature operation. For voice communications with low modulation rates, the preferred device would be a silicon-doped l.e.d. with integral lens, giving as narrow a beam as possible (to keep the cost and size of optics down). The l.e.d. device cost alone would be about 1/10 that of the laser diode, and the lens would be cheaper. Thus the l.e.d. would have a distinct advantage if there were no marked difference in system performance.

Within the cost-dominated limitations of the design, it would be desirable to include an audio squelch circuit, to mute the loudspeaker in the absence of a received signal, thus conserving battery power and reducing the operator's task. Since the transceiver would operate off dry batteries, both to minimise cost and for convenient replacement of the power source in the field, reduction in power consumption would be important. Savings on current drain could be used either to extend the battery life or to reduce the size of the batteries, and hence the size of the instrument as a whole. A system whereby the power supply to part of the receiver would be pulsed at a low duty cycle in the absence of incoming signal could give large savings in standby power drain. A further addition to the basic design in the interests of ease of use could be an automatic level control or compressor on the microphone circuit, to give a high average modulation level and to prevent overmodulation. All of these circuits should be capable of implementation with a relatively low parts count and total cost, yet should make the system considerably more usable and useful.

Modulation of the Optical Carrier.

As outlined previously, many of the possible means of modulating an optical carrier are not suitable for use in a portable, single-channel voice link, particularly when low cost is a major design criterion. The usable modulation schemes fall into three groups : analogue, pulse-analogue, and digital. The first group comprises direct intensity modulation and frequency modulation, the second, pulse width (PWM) and pulse position modulation (PPM), and the third has only one member, delta modulation. PAM (pulse amplitude modulation) is excluded since it confers no real advantage over intensity modulation, while being more complex. PCM likewise requires too many components to be viable.

Intensity modulation is easily the simplest of all the systems to implement, since it requires no modulator or demodulator other than the infrared source and detector. Unlike all the other modulation schemes, however, it makes no use of the large available bandwidth to improve the received signal-to noise ratio. In addition, the receiver frontend has to handle frequencies in the $1/f$ noise region. The design of a low-noise frontend is thus considerably complicated, since use must be made of devices (in this case, j.f.e.t.s) selected for low $1/f$ noise, at a cost which is typically an order of magnitude higher than standard devices. Another difficulty posed by intensity modulation is that of high current drain in the l.e.d. driver. This must operate in class A, and hence requires an average drive current equal to the peak modulation current. The sliding-bias class A amplifier ¹ is a possible means of combating the problem, but it does involve additional circuitry, thus somewhat nullifying the main advantage of intensity modulation, namely, its simplicity.

Frequency modulation is the oldest of the modulation schemes which trade bandwidth for S/N ratio enhancement. Wideband FM can offer significant improvements in S/N ratio, in the region of 20 dB in the case of broadcast FM ². The

1. Ref. 7

/ improvement ...

2. Ref. 12

improvement is proportional to $(B/f_m)^2$, where B is the demodulator input bandwidth and f_m is the modulating signal bandwidth³. Thus one should be able to increase the transmission bandwidth without limit to obtain greater output S/N. This however is only true if the input S/N is above a certain threshold value. Since for a given carrier power, input S/N is inversely proportional to bandwidth, an increase in transmission bandwidth results in an increase in demodulator input noise and a reduction in maximum range. Thus a compromise is necessary between maximisation of range and of S/N ratio. FM differs from the other wide-band systems examined here in that the carrier frequency is typically much higher than the transmission bandwidth, so that the receiver is a relatively narrow-band system. As will be seen later, this can confer certain practical advantages in an optical receiver design.

Pulse analogue systems have been used in optical communicators, particularly those using diode lasers, where PPM is the most practical modulation scheme⁴. Restrictions on duty cycle (to 0,1 % or less) and practical limits on minimum pulse duration limit the maximum repetition rate to tens of kilohertz. For a modulation bandwidth of 3 kHz, a sampling rate of 8 kHz is sufficient, allowing a 125 ns pulse time, corresponding to a transmission bandwidth of at least 8 MHz. These figures reveal the greatest drawback to the use of currently-available laser diodes, namely, the very high bandwidth which they require, even for transmission of a single voice channel. This great bandwidth results in a high value of minimum detectable signal, which effectively nullifies the gain due to the laser's high peak output power. In fact, Schwartz (Ref. 13) shows that for a given ratio of transmission bandwidth to signal bandwidth, the S/N improvement given by FM is nearly 14 dB greater than that for PPM.

3. Ref. 13

4. Ref. 22

$$\text{For FM : } \frac{S_o}{N_o} = \frac{3}{4} \left(\frac{B}{f_m} \right)^2 \frac{S_c}{N_c}$$

$$\text{For PPM : } \frac{S_o}{N_o} = \frac{1}{32} \left(\frac{B}{f_m} \right)^2 \frac{S_c}{N_c}$$

(formulae adapted from Schwartz, section 6-5.)

Since, in addition, the circuitry required to generate and detect PPM is more complex than that for FM, it can be seen that FM will be the preferred system, except where very narrow pulses are mandatory. Pulse width modulation suffers from most of the problems of PPM, without the ability to use short pulses, and is thus not to be considered seriously. A treatment of PWM can be found in Black (Ref. 11).

There remains only delta modulation⁵. This is a digital modulation scheme in which the signal is sampled, and a pulse is transmitted when the new sample is greater than the previous sample. The hardware required to implement the basic delta modulation system is simple, but the basic system has various drawbacks. The quantization S/N ratio is small, unless the sampling rate is very much more than twice the signal bandwidth. Attempts to reduce quantization noise have resulted in several types of companded delta modulation, but these schemes achieve their improved S/N ratios only at the expense of a greatly increased circuit complexity⁵. Delta modulation would seem to be more suited to multichannel telephonic links, where it occupies less bandwidth than the equivalent PCM system, than to the single-channel low-bandwidth link required here.

Thus of the wide bandwidth, S/N ratio enhancement modulation schemes, FM is the best choice for the low-cost optical voice link. A comparison between FM and intensity modulation shows that although IM is intrinsically simpler, all of the practical advantages lie with FM. FM makes better use of the available bandwidth to provide a low-noise channel, while avoiding the use of exotic devices and circuits, and allowing efficient modulation of the optical source. IM might be viable in a design where initial cost was the only criterion, and performance

5. Refs. 12,44

6. Ref. 44

would be a secondary consideration, but in this case, FM is the most desirable modulation scheme.

The FM System as Applied to the Optical Channel.

Frequency modulation in optical communications can mean two very different things. In high bandwidth optical links, the actual frequency of a gas laser can be varied, with detection at the receiver using an optical heterodyne technique. This, it has been claimed, results in better penetration of haze and atmospheric turbulence, since the other major modulation schemes used in high bandwidth systems, intensity modulation and phase modulation, are much more susceptible to fading or distortion in the atmospheric path.¹ In low bandwidth optical voice links, frequency modulation usually refers to the modulation of the frequency of a subcarrier, which in turn intensity modulates the optical source. The subcarrier frequency is of the order of tens or hundreds of kilohertz, possibly megahertz. Since the percentage deviation is small, the receiver is an optical intensity detector followed by a narrow-band amplifier and FM demodulator. It is this latter category of FM system which will be discussed here.

1. The FM Detector.

The FM detector requires a certain minimum S/N ratio at its input for proper operation. Above this first threshold value, the S/N ratio at the detector output is higher than at the input, and the difference increases with increasing input S/N ratio. Above a second threshold value, usually 2 to 3 dB higher than the first, the maximum S/N ratio improvement has been achieved, and the ratio of S/N(out) to S/N(in) is constant. An approximate analysis due to Betts (Reference 12) gives the S/N ratio for the case of a single-frequency modulation at 100 % deviation :

$$S/N(\text{out}) = 3(\Delta f/B_m)^2 \cdot B_{if}/2B_m \cdot S/N(\text{in})$$

where Δf = maximum frequency deviation
 B_m = modulation bandwidth
 B_{if} = noise bandwidth at demodulator input

The first input S/N threshold is taken as occurring when the peak signal amplitude is four times the r.m.s. noise amplitude, corresponding to a C/N power

ratio of 9 dB. In designing an FM system, account must be taken of the fact that the detector input noise power is proportional to B_{if} , which in turn depends on the modulation index $\Delta f/B_m$. Thus for a given B_m , an increase in B_{if} results in an increase in S/N ratio improvement at the demodulator, but the sensitivity decreases, because more signal is needed to reach the first threshold level. Optimisation of the modulation bandwidth, modulation index, demodulator input bandwidth relationships is required to produce the best system performance.

2. The FM Modulator.

To frequency modulate the subcarrier, it is convenient at low frequencies to use some kind of controllable oscillator, commonly a voltage-controlled oscillator or VCO, in which the frequency of oscillation is a function of an applied control voltage. It is important that the modulation voltage, corresponding to frequency deviation, and modulation frequency, affecting rate of change of subcarrier frequency, should not exceed the design values. Failure to ensure this will result in distortion of the demodulated signal at the receiver, as frequency deviations will occur which are outside the receiver bandpass characteristic, or which do not fall within the linear portion of the receiver detection curve.

The ideal VCO transfer function is a straight line. For FM purposes, it is required that there should be an offset, corresponding to the subcarrier frequency f_c . The slope of the transfer function is a measure of how much modulating voltage is needed for the desired maximum frequency deviation.

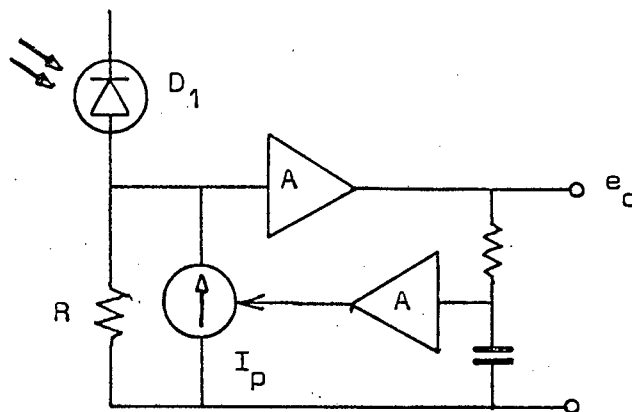
$$f_{vco} = kV_c + f_c$$

3. The Optical Detector.

The narrowband frequency characteristic of the FM system, when compared with the pulsed modulation systems, is a great advantage in overcoming one of the major potential problems in atmospheric optical link design. Background radiation results not only in shot noise, but also in a d.c. current

/ in the optical

in the optical detector. Even in a small-area detector, the magnitude of this current can be as much as 100 microamps. Since a requirement for low thermal noise is that the detector load resistor be as large as possible, the problem arises of saturation, not of the detector, but of the subsequent circuitry. For example, if $i_p = 0,1 \text{ mA}$ and $R = 10^6 \text{ Ohm}$, the offset voltage will be 100 volts, which is greater than the reverse voltage rating of most photodiodes, and far in excess of any likely power supply voltage. In wideband receivers, or those which must have a response down to very low frequencies, the only satisfactory cure is to provide a current sink which exactly balances the d.c. component of the photocurrent. This can be done by applying feedback with a low-pass characteristic, such that the feedback is inoperative at frequencies of interest.

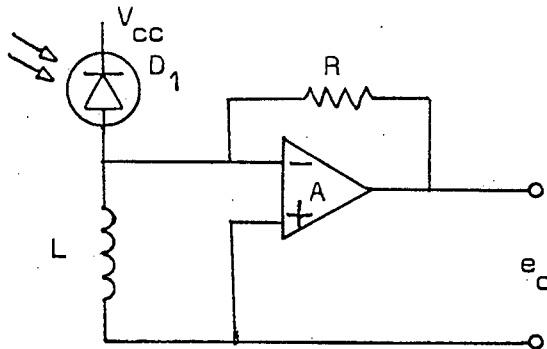


Feedback system

Such a system adds considerably to the cost and complexity of the receiver frontend, and in addition degrades performance, because of noise generated by the current sink. In an FM receiver, the background photocurrent problem can be neatly circumvented by the use of a low-loss inductance in parallel with the load resistor. The inductor provides a low-resistance d.c. path, much lower than R . At the subcarrier frequency, the inductor appears as a high inductive impedance in parallel with a large loss resistance $R_p = Q\omega L$. Provided that the Q value is high enough, R_p will be large compared to R , and will not contribute significantly to the total noise. The inductance is itself noiseless, and can be tuned out, if

/ necessary

necessary, by a parallel capacitor. This inductor d.c. bypass technique is more easily applied if the first stage of the receiver is a transresistance amplifier. The inductor can then be connected across the amplifier inputs, appearing as a very high impedance in parallel with the feedback load resistor.



Inductor bypass system

Since the photodetector is a square-law device, the 12 dB S/N ratio required at the FM demodulator will correspond to a 6 dB S/N ratio at the input to the photodetector (ignoring the effects of noise in the preamplifier). The minimum useful optical signal power will be four times the effective optical noise power in a noise bandwidth determined by the bandpass amplifier response. Since this last will probably be at least a four-pole bandpass network, the noise bandwidth will be approximately equal to the demodulator input bandwidth.

It has previously been shown that use of a junction f.e.t. in the frontend amplifier can result in a noise figure very close to unity. The detector shot noise and load resistor thermal noise thus determine the minimum useful signal power. In an atmospheric link, the dominant noise source will probably be the shot noise due to background radiation, since the background photocurrent is typically several orders of magnitude larger than the detector dark current, and use of a transresistance preamplifier reduces thermal noise. Any reduction in the relative levels of background and signal photocurrents will result in an improved S/N ratio. An obvious means of achieving this is to use an optical bandpass filter, centred

/ on the transmitter...

on the transmitter wavelength. Such a filter is the interference filter, consisting of a number of quarter-wave plates. Reflections at the interfaces result in destructive interference at wavelengths other than those desired. At near-infrared wavelengths (900nm) a bandwidth of 5 nm can be readily achieved. This corresponds approximately to the spectral characteristics of junction lasers, but is an order of magnitude smaller than that of typical non-coherent l.e.d.s.. Hence a system using a l.e.d. source is more susceptible to background radiation as a result of its broader spectral response, since clearly an interference filter for use with such a system must be 50 nm wide. Interference filters have two drawbacks, the major problem being one of cost, as a filter the size of a large coin may cost several tens of rands. If the filter is placed directly in front of the detector, a small section a few millimetres square will suffice, but if a lens is being used at the receiver, the converging beam falling on the filter will degrade its bandwidth substantially. To compensate for this effect, a narrower-bandwidth filter must be used, at higher cost.

A cheap and quite effective substitute for the interference filter in a l.e.d.-based system is the absorption filter. Suitable dyes in a gelatin or glass medium can absorb about 80 % of the total incident background radiation, while attenuating the wanted wavelengths by less than 20 % . The filter can for these purposes be regarded as an optical low-pass filter with a sharp cutoff, low pass-band attenuation, and high stopband attenuation. Since the silicon photodetector's response falls sharply at wavelengths above one micron, the filter and detector characteristics combine to give a bandpass response about 200 nm wide. More detail on the background radiation problem and the use of interference and absorption filters to combat it will be found in the appendices.

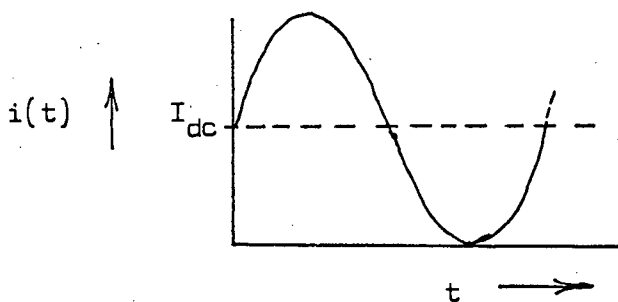
4. The Transmitter.

There are several ways in which the emitter in an FM system

/ can be

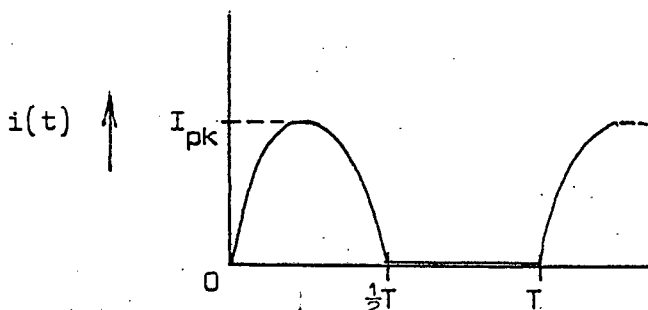
can be driven. Use of the most efficient drive waveform or circuit will minimise the largest consumer of power in the whole optical communicator, the light source. Since all junction emitters have a low forward resistance, at sub-ampere current levels most of the power dissipated in the emitter itself goes into the $V_f I_f$ term. The forward voltage V_f is invariably lower than the supply voltage, so that more power must be thrown away in the drive circuit. Hence the most useful measure of the driver efficiency is the ratio of subcarrier frequency current, producing useful optical power output, to d.c. supply current.

Because the receiver has a narrow electrical bandpass characteristic, it is unaffected by the harmonic content of the transmitted subcarrier waveform. Thus one is free to choose the most efficient or most convenient type of drive for the output stage. In the case of the commercially available laser diode, with 0,1 % maximum duty cycle, the only possible waveshape is a very narrow pulse. Continuous-wave emitters can be driven in a number of ways. In a class A output stage with sinewave drive, the peak value of the fundamental current is equal to the average d.c. current.



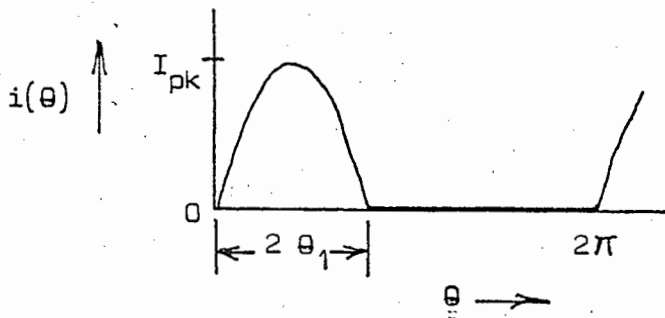
Class A
Peak $I_1 = I_{dc}$

Greater efficiency can be achieved with a class B drive, in which the current waveform is a half-wave rectified sinusoid. The peak fundamental component is half the total current's peak value, and so is $\pi/2$ times the d.c. current.



Class B
Peak $I_1 = 0,5 I_{pk}$
 $= 1,57 I_{dc}$

The class B drive is a special case of the class C drive, in which the conduction angle of the sinusoidal current waveform falls in the range 0° to 180° .

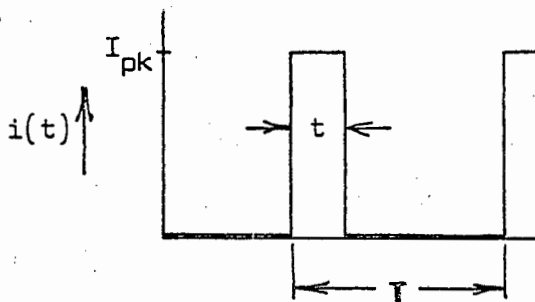


Class C

$$\frac{I_1}{I_{dc}} = \frac{\theta_1 - \sin\theta_1 \cos\theta_1}{\sin\theta_1 - \theta_1 \cos\theta_1}$$

I_1/I_{dc} varies between 1.57 for $\theta_1 = 90^\circ$ (class B) to 2.0 for $\theta_1 = 0$. The smaller the conduction angle, the greater the efficiency.

Use of a square-edged pulse drive waveform produces efficiencies very similar to those for class C, although the peak currents may be higher.



Pulsed Drive

Pulse period = T

Pulse duration = t

$$I_{dc} = \frac{t \cdot I_{pk}}{T}$$

By Fourier analysis, peak $I_1 = 2 / \sin(\frac{\pi t}{T}) I_{pk} = 2(\frac{T}{\pi t}) \sin(\frac{\pi t}{T}) I_{dc}$.

The expression for I_1 is of the form $\sin x / x$, and thus has a maximum at $x = 0$ (i.e. $t = 0$) of $2I_{dc}$. At a 50% duty cycle ($t = \frac{1}{2}T$), $I_1 = (4/\pi)I_{dc} = 1.27I_{dc}$. Pulsed drive is therefore inferior to class C drive for duty cycles much greater than zero. However, the relative ease with which constant duty cycle pulse waveforms can be generated and manipulated makes a low duty cycle pulsed drive scheme attractive for l.e.d.s as well as for laser diodes.

A problem that may occur with silicon-doped l.e.d.s is their considerable light output risetime. This can be treated as if it were the result of a lossless first-order lowpass filter interposed between the driver and the emitter. The corner frequency of this filter is given by :

$$f_c = \frac{0.35}{t_r} \text{ Hz, with } t_r = \text{risetime in seconds.}$$

If attenuation of the fundamental frequency occurs because of risetime effects, the loss will be the same for all of the drive schemes. Thus the low duty cycle pulsed drive is still the best overall choice for the silicon-doped GaAs source.

Development of a Practical System.

(i) Major System Parameters.

1. Choice of Emitter and Detector.

At the start of work on the design, no really suitable low-cost emitters were available, although reports of new designs had been received. This meant that initial work was concentrated on receiver development. An obsolete high-cost l.e.d. (TIXL 12) in a test-bed pulse generator was used for this phase. When the anticipated cheap sources arrived, an efficient though slow silicon-doped device was selected. This, the Monsanto ME7124, is rated at 3 mW output for only 50 mA of forward current, with a beam half-angle of 4° . Maximum continuous I_f is 100 mA, and pulses of up to 6 amps are allowed. Optical output risetime is 500 ns, resulting in a 3 dB reduction of modulation level at 700 kHz. The very narrow beam angle means that a small, cheap lens can be used to further collimate the beam. Use of a standard panel-mount style encapsulation allows a variety of mounting configurations. Cost of the ME7124 is less than R1,50 one-off, dropping to less than R1,00 in production quantities. Such a combination of desirable qualities seems too good to be true, and in fact it later emerged that the anticipated output power is not achieved in practice.

A small-area p-i-n photodiode detector represents a good compromise between cost and performance requirements. Development work was carried out using a Hewlett-Packard 5082-4207 diode with an $8 \times 10^{-3} \text{ cm}^2$ active area and 2,5 nA dark current, two of these devices having been salvaged from previous work on optical communications in the department. The high cost of the H-P diodes (about R45) made them impractical, and they were replaced by the Monsanto MD1, costing R6 one-off and R4 in lots of 100 or more. Its specifications are similar to those of the H-P device, with $5,8 \times 10^{-3} \text{ cm}^2$ area and 3,2 nA dark current. Full specifications of the MD1 and the ME7124 are included in the appendices.

2. The Power Supply.

One of the most fundamental decisions to be taken, once the actual modulation scheme and type of emitter and detector have been selected, is the nature of the power source. Batteries of some kind must be used, but should they be primary cells or secondary (rechargeable) cells? How many cells will be needed? Should one use a special high-capacity battery? A final choice of the exact specification required will need details of power consumption in transmit and receive modes, and desired operating battery life. But an initial decision on the general type of battery to be used must be made before the system can be designed. One major source of power drain will be the infrared emitter during transmission. An average current drain of up to several hundred milliamps might be expected, since the ME7124 can draw up to 100 mA continuously. However, only about 2 volts is required to bias the GaAs emitter, so there is an advantage in using as low a battery voltage as possible, so as to achieve a higher ampere-hour rating in the same volume. To simplify the design of any high-frequency circuitry in the system, a larger battery voltage is preferred. As a compromise between these conflicting requirements, an intermediate value of about 5 to 6 volts should suffice. An alternative would be to use a slightly higher value, say 9 volts, and to drop the supply to critical circuits with one or more voltage regulators. A switched-mode regulator would minimise power losses.

(a) Primary cells.

The most familiar primary battery is the zinc-carbon or Leclanche cell. Because of its universal availability and low cost, it is a very strong contender for use in any cheap portable equipment. It can deliver high currents, has a low output impedance and good shelf life. Its disadvantages are poor low-temperature performance, relatively low energy density, and a drooping discharge voltage curve. More recent additions to the range of primary cells are the zinc-chloride and alkaline types, both with the same 1.5 volt potential as the Leclanche,

/ but with

but with up to twice the energy density and better low temperature performance. They can therefore be used as a substitute for Leclanche cells in applications where the extra performance warrants the extra cost.

Three other types of primary cell are not really suitable for this task, despite some advantages. They are the mercury-oxide, silver-oxide, and divalent silver-oxide cells, all having a flat discharge characteristic, with essentially constant battery voltage over the useful life. All have high energy density, but poor low-temperature capacity, and their most serious problem is an inability to supply large (50 mA or more) currents.

There remain only two commercially available primary battery types : magnesium and lithium. The magnesium cell has a 2 volt rating, with energy density twice that of the Leclanche and high current ability. The discharge curve is better than that of the Leclanche, and low-temperature performance is fair. The lithium cell is the most exciting of the new primary cells. It has a 2,8 volt cell voltage, energy density up to seven times the Leclanche cell's, low output impedance, high current capability, excellent low-temperature capacity, down to -40°C , extended shelf life, and a flat discharge curve. Given greater availability and low cost, the lithium cell would sweep the field. However, with only one American company marketing the cell, it cannot be considered for an optical voice link.

(b) Secondary cells.

Four types of rechargeable batteries suitable for portable equipment are currently available. The most commonly used is the nickel-cadmium cell, characterised by a 1,2 V cell voltage, good low-temperature capacity, flat discharge curve, long life, and energy density slightly less than that of Leclanche cells. In larger-capacity batteries, the gelled-electrolyte lead-acid cell has a cost advantage. Its cell voltage is 2,1 V , its energy density half that of the Leclanche, its shelf life very good, and it has a sloping discharge curve. Silver-cadmium and silver-zinc cells offer higher energy densities, but short cycle lives

/ at higher

at higher unit cost, as much as twice that of NiCd cells.

Cost of all these secondary cells is much higher than zinc-carbon primary cells of corresponding capacity. This is unimportant in an application where recharging facilities are always close at hand, but in a portable communications system the cheap primary cell is the better choice. Therefore the design should be based around the use of conventional zinc-carbon Leclanche cells, with the size and number to be determined by the system power requirements. Design of circuitry must take into account the sloping discharge curve of the zinc-carbon cell, so that critical functions are not affected by a drop in supply voltage.

(c) Zinc-carbon battery data.

The following data relates in particular to batteries made by Union Carbide (Eveready). Other conventional cells of similar size should have like characteristics.

<u>Size</u>	<u>Mass (gm)</u>	<u>I_{sc} (A)</u>	<u>R_{int} (Ohm)</u>	<u>I_{max} (mA)</u>	<u>Life (hrs) at</u>	<u>I (mA)</u>
AA (penlight)	17	5,4	0,29	25	40	15
					15	30
C	40	3,3	0,47	80	100	25
					40	50
D (torch)	95	5,5	0,28	150	105	50
					45	100

I_{sc} = short-circuit current

R_{int} = internal resistance of fresh cell

I_{max} = recommended maximum continuous current drain

Life specified with constant resistive load, at indicated starting drain, to a final voltage of 1,0 volt.

A final voltage of one volt per cell can only be tolerated if 5 or more cells are used, so that if only four cells are used, the end voltage must be

/ higher

higher, say 1,1 to 1,2 volt, and the battery life will be correspondingly shorter. Since one can expect a transmit current drain of about 100 mA, a compromise choice would be the medium-sized C cell, as used in many transistor radios and tape recorders. While the recommended maximum continuous drain is 80 mA, one should be able to exceed this figure by a considerable amount in an intermittent-duty application such as the communicator. Since heavy-duty variants of the standard cells are available with similar capacities but lower internal impedances, it may well be possible to use AA-sized cells of this type, with reduced battery life but a considerable saving in bulk.

With low cost and simplicity as the major criteria, the decision was made to opt for a four-cell battery of six volts nominal. All circuits would have to be designed for minimum sensitivity to reductions in battery voltage, with the battery pack end-of-life voltage being determined by the point at which performance starts to degrade seriously. A final value of about 4,5 volts was expected, corresponding to a 25 % drop in voltage. It was anticipated that a review of battery requirements would be necessary once the circuit design had been finalised.

3. Modulation Scheme Values.

The aim being to provide a communications-quality voice link, the modulating bandwidth was effectively predetermined at the standard 3 kHz. It was felt that any gain in sensitivity by a reduction in bandwidth would be more than offset by the loss in intelligibility. A second constraint resulted from the 500 ns rise-time of the ME7124 emitter. To avoid waste of power in the l.e.d. driver, the subcarrier frequency must be limited to 500 kHz or less. At 500 kHz the loss is nearly 2 dB, and lower frequencies are preferable in this respect.

Since FM offers the prospect of considerable S/N improvement at the demodulator, one would wish to adjust the modulation index - demodulator input bandwidth relation to get the greatest advantage from this property. The ratio of output to input S/N ratio for single-tone modulation is ¹ :

$$\frac{S/N(\text{out})}{S/N(\text{in})} = 3(\Delta f/B_m)^2 \frac{B_{if}}{2B_m}$$

$B_{if}/2B_m$ corresponds to the number of significant sidebands, and is determined by the modulation index. According to Betts, the n th pair of sidebands are insignificant if their amplitudes are less than 1 % of the unmodulated carrier amplitude, when $J_n(\beta) < 0,01 J_0(0)$, where $J_n(x)$ is the Bessel function of order n , and $J_0(0) = 1$. Thus for sidebands to be significant, one must have : $J_n(\beta) \geq 0,01$. The values of $J_n(\beta)$ are best evaluated by use of a graph or table, such as can be found in the references. Betts provides a useful table of number of sideband pairs against modulation index, which will serve as a rough guide. Although the relationship is not linear, the number of sidebands increases monotonically with β , so that the expression for S/N ratio improvement has no maximum value, in theory. But increasing modulation index to gain more S/N improvement requires a larger B_{if} , and the demodulator input noise power is proportional to B_{if} . The minimum detectable signal needed to reach the second S/N threshold value will also increase, and as a result the maximum range of the system will be reduced. To approach the question of how to choose the value of modulation index, it is necessary to examine what S/N is required at the demodulator output. For a good quality speech link, a S/N ratio of at least 20 dB, and preferably more is desirable. Now speech has an approximate negative exponential amplitude distribution, and the probability of peak amplitudes exceeding four times the r.m.s. value is very small. Thus if the design peak frequency deviation Δf is produced by a speech amplitude four times the r.m.s. speech level one can say that :

$$\frac{S/N(\text{out})}{S/N(\text{in})} = 3\left(\frac{\Delta f}{4B_m}\right)^2 \frac{B_{if}}{B_m} \quad \text{for speech.}^2$$

To allow a slight margin for error, assume that maximum range is reached when the input S/N falls to the second threshold value. Now if S/N(in) is 12 dB at the second threshold, and S/N(out) must be 20 dB or more, then the required noise

improvement is at least 8 dB.

$$\text{Hence } \frac{3}{16} \beta^2 \frac{B_{if}}{B_m} = 6,3 \quad (8\text{dB})$$

$$\beta^2 \frac{B_{if}}{B_m} = 33,7$$

Now if $\beta = 2$, B_{if}/B_m must be at least 8, corresponding to 4 pairs of significant sidebands, so that

$$\beta^2 \frac{B_{if}}{B_m} = 32$$

This represents a good approximation to the optimal system, based on the assumptions set out above. The required peak deviation is 6 kHz, and the demodulator input bandwidth is about 25 kHz. At this stage, the subcarrier frequency f_c was still to be decided. In theory, any frequency in the range 50 to 500 kHz would suffice, but real modulators and demodulators imposed restrictions on choice of f_c . In general, the larger the percentage frequency deviation, the greater the distortion generated in modulation and demodulation. To reduce this distortion requires either more complex and costly design or the use of the highest available subcarrier frequency.

Operation at an f_c near 500 kHz offered another practical advantage. One might be able to use as tuned elements in the transmitter and receiver, ordinary 450 kHz i.f. transformers. The types used in smaller AM transistor radios are extremely compact, have integral tank circuit capacitors, and can readily be loaded resistively to achieve the desired bandwidth. It was therefore decided to set the subcarrier frequency at 450 kHz. No local oscillator or mixer would be used at the receiver, since tuning of the subcarrier should be unnecessary.

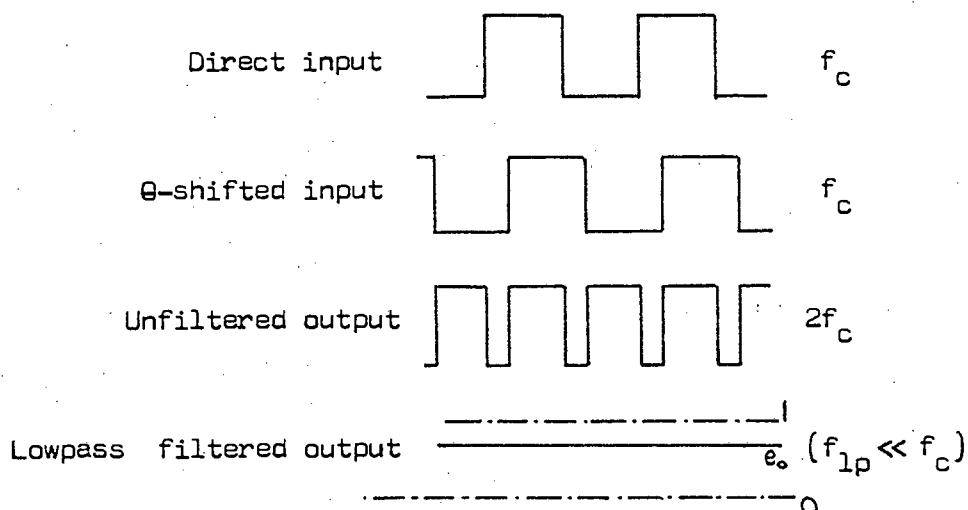
(ii) The Receiver.4. The Receiver Demodulator.

The traditional FM demodulators are the Foster-Seeley discriminator and the ratio detector. Both of these are means of realising a type of demodulator in which FM is converted to AM and then envelope detected. The conversion to AM relies on the combination of two signals at f_c , one of which has been given a frequency-dependent phase shift of about 90° . These demodulators perform their phase shifting by means of a multiple-wound transformer. Such transformers are manufactured for use at the common 4,5 and 5,5 MHz TV and 10,7 MHz sound FM i.f. frequencies, but would have to be specially designed and produced for 450 kHz. Since special parts mean high cost, these conventional circuits must be discarded. The pulse counting FM detector must likewise be dismissed as being too complex and in consequence drawing too much power, although possibly the use of CMOS logic would enable the development of a low-power demodulator of this type. Another possibility which has only become commercially feasible in recent years is the phase-locked loop demodulator. Integrated circuit p.l.l.s are available, but (with one exception) they suffer from high power consumption (100 mW or more) and require at least 10 V d.c. supply. This last characteristic is incompatible with the design supply voltage, and rules out the use of an i.c. phase-locked loop. Various attempts were made to produce a low-power, low-cost p.l.l., but it was concluded that a satisfactory design would require too long a development time.

Thus it was decided to make use of one of the multiple-function i.c.s produced for use in broadcast FM and TV receivers. In general, these circuits consist of a number of i.f. gain stages, with typically 55 to 65 dB of voltage gain, which provide a limited i.f. signal to the demodulator, which in turn is followed by a buffer or an audio preamplifier. The demodulator on these i.c.s is a quadrature detector, broadly similar in principle to the transformer-type demodulators. But the complex transformer is replaced by a balanced product detector, and only one single-tuned circuit is required. The balanced product

/ detector.....

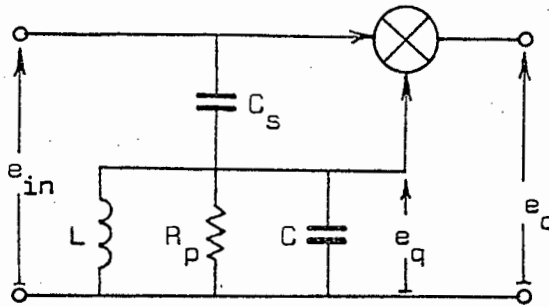
detector is easily realised in i.c. form, although it would be expensive to build using discrete components. Since one of the product detector's inputs is the direct, hard-limited output of the gain stages, the relation between phase shift and average output voltage is linear. The basic output waveform from the product detector is a squarewave at twice the input frequency, with a duty cycle dependent on the phase difference between the inputs. When the inputs are in phase, duty cycle is 100 % , when they are 90° out of phase, duty cycle is 50 % , and when they are 180° out of phase, duty cycle is zero. In practice, the phase shift is restricted to a region around 90° to preserve linearity. In order to demodulate FM signals, one input to the product detector must have a phase shift that is 90° at the centre frequency, and that is linearly related to frequency over a range either side of centre frequency. Then a given Δf will result in a proportional phase shift θ , and the duty cycle of the product detector output will deviate from 50 % . By using a low-pass filter at the output, one can remove the $2f_c$ component and leave only the demodulated voltage due to the frequency shift Δf .



The corner frequency of the lowpass filter will determine the bandwidth of the demodulated signal, and of any unwanted noise accompanying it. It should therefore be equal to the highest transmitted modulation frequency, in this case 3 kHz. Clearly, linear demodulation depends on a phase shift that is linear with frequency over

/uency over

uency over the range used. The phase shift produced by a single L-C tuned circuit is only approximately linear in the region around 90°.



$$e_q = e_{in} \frac{s^2 C_s / (C + C_s)}{s^2 + s(R(C + C_s))^{-1} + 1/L(C + C_s)}$$

From the expression for e_q one can calculate the phase shift :

$$\begin{aligned} \theta &= \arctan \frac{f_r f / Q}{(f_r^2 - f^2)} \\ &= \pi/2 - \frac{Q(f_r^2 - f^2)}{ff_r} + \frac{1}{3} \left\{ \frac{Q(f_r^2 - f^2)}{ff_r} \right\}^3 + \dots \end{aligned}$$

valid for θ between 45° and 135°, with $f_r^2 = 1/2\pi L(C + C_s)$

$$\Delta f = |f_r - f|$$

$$Q = R_p / 2\pi f_r (C + C_s)$$

For a typical FM system, the percentage deviation is small (in this case $f_c = 450 \text{ kHz} = f_r$ and $\Delta f = 6 \text{ kHz max.}$). Thus $(f_r^2 - f^2)/ff_r$ is much less than unity, and one can regard all distortion as being due to the cubic term in the arctan expansion¹. On this basis one can derive the percentage harmonic distortion:

$$\begin{aligned} \text{THD}(\%) &= 33,3 \left\{ \frac{Q \Delta f}{f_r} \left(2 + \frac{\Delta f}{f} \right) \right\}^3 \\ &= 33,3 (2Q \Delta f / f_r)^3 \quad \text{since } \Delta f / f \ll 2 \end{aligned}$$

While output is proportional to phase shift, and hence to Q , distortion is proportional to Q^3 . Thus for all likely values of Q , distortion increases faster than output voltage. For a maximum acceptable value of harmonic distortion, one can calculate a maximum value of Q , given Δf and f_r .

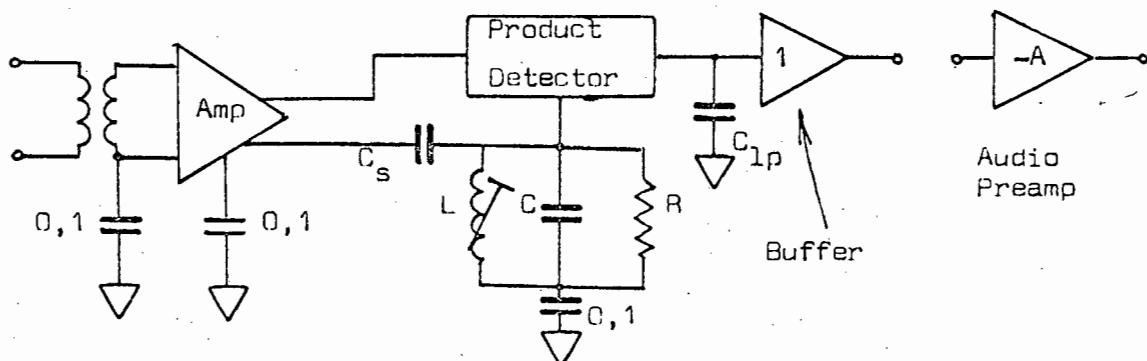
1. Reference is National Semiconductor Linear Applications, AN-54, on the LM373.

$$Q_{\max} \approx \frac{f_r}{2\Delta f} (\text{THD}(\%)/33,3)^{1/3}$$

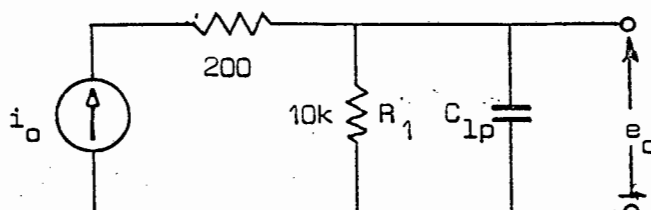
For a voice communications link, a maximum distortion of 1% should be adequate.

With $\text{THD} = 1\%$, $Q_{\max} = f_r/6,4\Delta f$. In the system being developed, $f_r = f_c = 450 \text{ kHz}$, $\Delta f = 6 \text{ kHz}$, and so $Q_{\max} = 12$.

The actual device used in the design of the receiver was the National Semiconductor LM1351N, a second-sourced version of the Motorola MC1351, intended for use in the sound section of a TV set. The 14-pin package contains three limiting differential amplifier stages with a total gain of 65 dB, a balanced product detector with buffered input and output, an audio preamplifier, and a d.c. bias network. Current drain of the circuit, excluding the audio stages, is between 12 and 16 mA over the 4,5 to 6 volt supply range. To provide the 90° phase shift, a subminiature 455 kHz i.f. transformer with integral tuning capacitor was used. With slug tuning, and a high unloaded Q, and good screening properties, this represents a very convenient solution.



Since the inputs to the product detector are buffered, loading on the tuned circuit should be entirely due to R. The output of the product detector is at a high impedance, and a small capacitor C_{lp} is all that is required to implement the lowpass filter function.



The lowpass corner frequency $f_{lp} = 1/2\pi R_1 C_{lp}$. Putting $C_{lp} = 4700$ pF gives $f_{lp} = 3,4$ kHz, and the 900 kHz $2f_c$ component is suppressed by about 48 dB. Further filtering will be provided by the audio sections that follow. The recovered audio voltage can be calculated if the peak current output of the product detector is known. With a six volt supply, the peak-to-peak output current I_{p-p} is about 0,55 mA. A phase shift of 180° corresponds to a current swing of 0,55 mA, about a mean value of 0,275 mA at $\theta = 90^\circ$. If the peak frequency deviation is Δf , resulting in a phase shift θ_p , where

$$\theta_p \approx \frac{\pi}{2} - \frac{Q(f_r^2 - f^2)}{ff_r} \quad \text{radian}$$

then the resulting change in current is :

$$\begin{aligned} i_o &= (0,5 - \theta_p/\pi) I_{p-p} \quad \text{peak} \\ &= (0,5 - \theta_p/\pi) I_{p-p}/\sqrt{2} \quad \text{r.m.s.} \end{aligned}$$

Since the load resistance is 10 kilohm, the recovered audio voltage is :

$$e_o = 10^4 I_{p-p} (0,5 - \theta_p/\pi) / \sqrt{2} \quad \text{V r.m.s.}$$

In this case, $f_r = 450$ kHz, $\Delta f = 6$ kHz, $Q = 12$, and $I_{p-p} = 0,55$ mA, giving $e_o = 0,4$ volts r.m.s. This is more than sufficient to drive a small audio amplifier without further amplification.

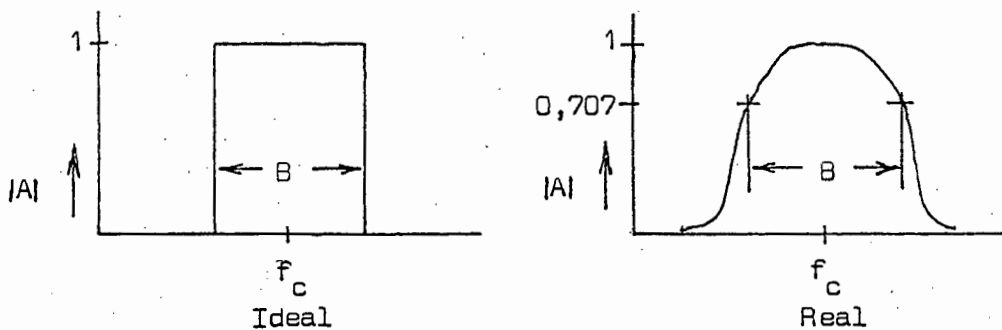
When overdriven, the limiting gain stages preceding the demodulator provide an approximately square waveform of about 170 mV peak-to-peak. To drive the output to within 3 dB of the 170 mV level, an input of 80 microvolts r.m.s. is required (manufacturer's specification). This high sensitivity simplifies the design of the early stages in the receiver by reducing the extra gain needed. Only one real disadvantage results from the use of this integrated circuit. Its power consumption is greater than is necessary for operation at 450 kHz, because it is intended for use at 4,5 and 10,7 MHz. However, this problem is not peculiar

/ to this i.c. ...

to this i.c., as all of the similar limiter-detector packages run at power levels of the same order. Since the LM1351N is very cheap for the functions which it performs (R1,50 one-off), it is a good compromise solution to the demodulator problem.

5. The Receiver Bandpass Filter.

The ideal bandpass response of a top-hat shape in the frequency domain cannot be realised in practice. One must therefore determine what deviations from the ideal can be accepted, and design the filter accordingly.



The real filter does not have a flat amplitude response in the passband. In an FM system, small deviations in amplitude can be tolerated, an upper limit of 2 dB peak-to-peak ripple being typical. Phase response should be as linear as possible. Nonlinear phase will result in distortion due to relative phase shift of the modulation sidebands, but this is not easy to analyse. A further defect of the real filter is its non-infinite attenuation outside the passband. This increases the noise power at the filter output, thus reducing the receiver sensitivity. It is therefore important to eliminate the excess noise power.

To quantify the effect of the filter response on noise, one can calculate the equivalent noise bandwidth B_n . This corresponds to the bandwidth of an ideal top-hat bandpass filter which has the same output noise power as the real filter.

$$B_n = \frac{\int_0^\infty A^2 df}{A_0^2} \quad \text{where } A_0 = \text{maximum passband gain}$$

For a single L-C bandpass filter, B_n is $\pi/2$ times the bandwidth between -3 dB points, corresponding to 2 dB more noise power than the ideal, even if a 3 dB

/ drop in

drop in amplitude response were acceptable. Thus the single-tuned filter is not close enough to the ideal for this application.

Another point of difference between the real and perfect filters is the steepness of the attenuation at the edges of the passband. In a conventional radio i.f. bandpass filter, rapid stopband attenuation is essential as a means of discriminating against unwanted signals at adjacent frequencies. In the optical receiver, however, such interference is very unlikely, so that the filter form factor (ratio of -6 dB to -60 dB bandwidths) is in itself unimportant. Naturally, the faster the cutoff, the lower the value of B_n/B , but as will be shown, a four-pole (two-stage) bandpass characteristic is very close to the ideal in this latter respect. The use of a higher-order filter is therefore not justified, especially in view of the cost factor.

An overcoupled double-tuned bandpass filter, with coupling coefficient $k = 1,65 k_c$ has an amplitude ripple in the passband of about 1 dB, and a bandwidth between -1 dB points of 1,86 times the -3 dB bandwidth of a single stage ($=f_r/Q$). The noise bandwidth of this filter is 1,23 times the -1 dB bandwidth, resulting in an increase of noise over the ideal square bandpass response of just less than 1 dB. If a 2 dB passband ripple can be tolerated, corresponding to $k = 2 k_c$, the bandwidth between -2 dB points is 2,45 times the single-stage bandwidth. The noise bandwidth is 2,50 times the single-stage bandwidth, or 1,02 times the -2 dB bandwidth. This latter configuration combines all of the desired bandpass characteristics in an easily realisable form. The tuned circuits can be implemented using AM 455 kHz i.f. transformers with integral tuning capacitors. A trimmer capacitor between the live ends of the tuned circuits will enable the setting-up in production of the coupling coefficient k .

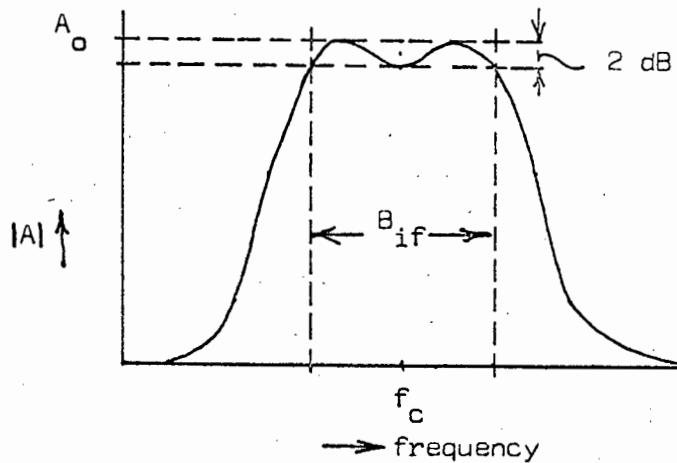
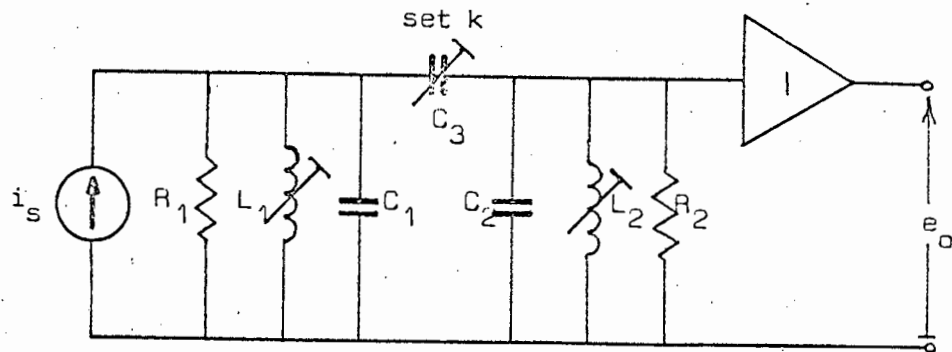
With a 2 dB passband ripple ($k/k_c = 2$), $f_c = 450$ kHz, and $B_{if} = 25$ kHz, the tuned circuit loaded Q is given by :

$$2,45 f_c / Q = B_{if}$$

$$Q = 2,45 f_c / B_{if} = 44$$

Hence $k_c = 0,023$ ($=1/Q$)

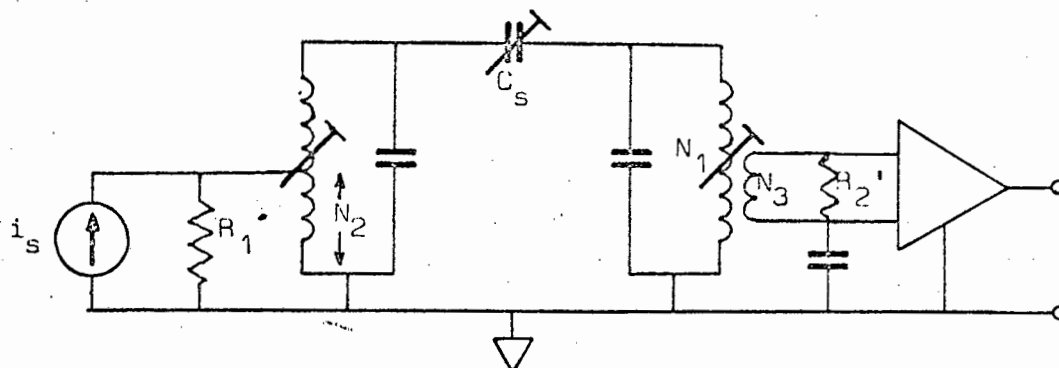
and $k = 0,045$



The filter must be driven by a current source, which can conveniently be a common-emitter stage, and the load presented by the following gain stage must be included in the calculation of R_2 , so that the actual value of the external resistor R_2 will be larger than R_1 . In fact the input impedance of the limiting amplifiers of the LM1351 is low, of the order of 5 kilohm, and if the LM1351 is connected directly to the tuned circuit, it will not be possible to achieve the desired high Q . But this problem has already been overcome by the designers of the AM i.f. transformers, and there is a low-impedance tap on the coil which can be used to interface to external circuits. In addition, there is a second winding on the transformer at a still lower impedance, which can be used to provide d.c.

/ isolation

isolation. This second winding is a convenient means of driving the LM1351, the inputs of which must be connected by a d.c. path, and must also be allowed to float above d.c. ground.



The commercial AM i.f. transformers chosen for the filter are a subminiature type made by Hitachi. The turns ratio $N_1:N_2:N_3$ is 31,7:8,3:1, or, referred to the input, 3,8:1:0,12. Thus the input impedance of the amplifier on the output winding can be much lower than the reflected tank circuit impedance, although the insertion loss of the filter will be correspondingly greater. To produce the desired Q factor, R_1' must be 3300 Ohms and R_2' , 82 Ohms. Since the required value of C_s is about 9 pF, one can use a variable capacitor of approximately 5 to 15 pF to set up the coupling coefficient k . The transfer impedance of the filter can now be evaluated. For the simple overcoupled double-tuned case, the maximum value of transfer impedance in the passband is one half of the parallel load resistance R_1 . Then the output voltage e_o is $R_1/2$ at the response peaks. In the more complex case used here however, the maximum transfer impedance is $(N_3/N_2)R_1'/2$, due to the voltage transformations at the input and output. The assumption has been made that the filter is symmetrical, i.e. that both tuned circuits have equal Q's.

Now the current source drive to the filter will be provided by a transistor, with output resistance much greater than R_1' . Thus for an input voltage to the transistor e_{in} , an a.c. component of collector current $g_m e_{in}$ will flow. One can compute a passband voltage gain for the filter-driver combination :

$$\text{Max. } \frac{e_o}{e_{in}} = \frac{g_m R_1' N_3}{2N_2}$$

With 2 dB passband ripple, the mean passband gain is 1 dB down on this value :

$$\text{Mean } \frac{e_o}{e_{in}} = \frac{g_m R_1' N_3}{2,24N_2}$$

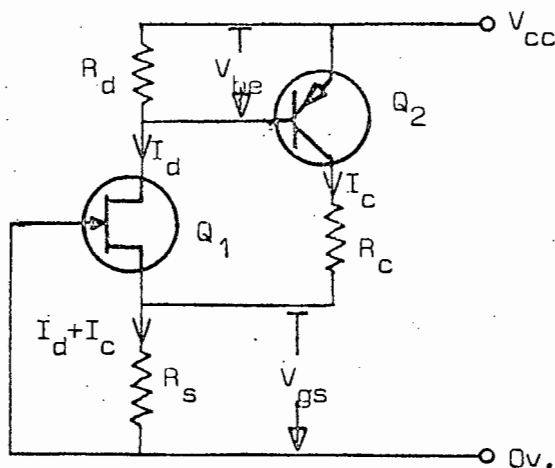
In the final design, with $R_1' = 3300$ ohms, the gain of the bandpass filter and its driver transistor is approximately $0,18 g_m$, with g_m expressed in mS (mA/V).

6. The Frontend Circuit.

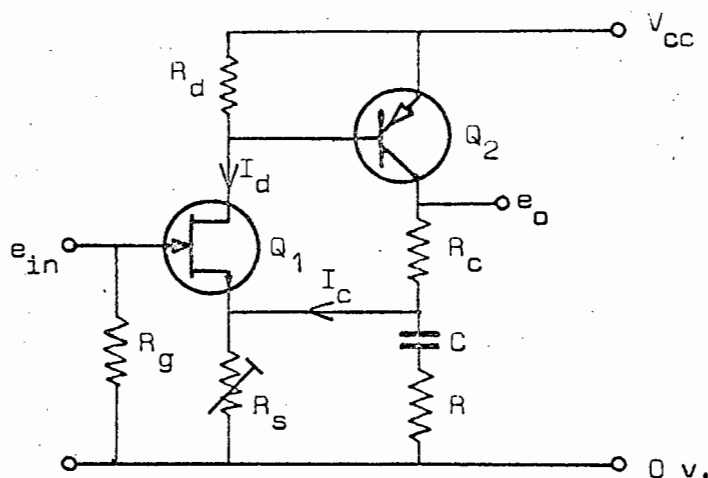
The earlier general treatment of frontend amplifier requirements has defined quite clearly the desired characteristics of the design. It remains only to produce a circuit having these characteristics, while capable of running off a battery supply that will drop from 6 volts to 4,5 volts, and yet drawing the minimum current from the battery. If possible, the design should be simple, cheap, and easy to set up in production. It should be noted that a considerable problem results from the two requirements of having to use a f.e.t. as the input device and of keeping the battery voltage low. The gate-source characteristics of j.f.e.t.s of a given type may vary over a range of 3:1 or more. Thus the operating gate-source bias voltage may be anywhere between 1 and 3 volts, and the drain current will change proportionally. Two approaches to the solution of this problem exist : to design a circuit which can tolerate even these gross performance deviations, or to select j.f.e.t.s for a limited spread of characteristics, and thus produce a simpler or more efficient design. The latter approach was chosen, as it seemed that production-line testing of devices could be done at low cost with simple instrumentation, on a go-no go basis.

The d.c. operating point for the first device is stabilised by use of a f.e.t.-bipolar feedback pair configuration. The bipolar transistor's base-emitter voltage and the f.e.t.'s drain load resistor define the f.e.t. drain
/ current

current, provided that the f.e.t.'s I_{dss} is greater than the desired I_d value. The second-stage current is dependent on the f.e.t.'s source resistor and gate-source voltage corresponding to the desired I_d .



If R_S is a variable resistor, then I_C can be set up independently to any desired value. As the circuit is shown, however, the maximum gain realisable is small, because the gain is approximately $(1 + R_C/R_S)$, and R_C cannot be made much larger than R_S , except by severely reducing I_C , which in turn reduces the open-loop gain and the high-frequency response. But R_S can be bypassed at signal frequencies, either completely with a suitable capacitor, or partly, with a series R-C combination. The closed-loop a.c. gain is then approximately $(1 + R_C/R')$, where $R' = R_S \parallel R$.



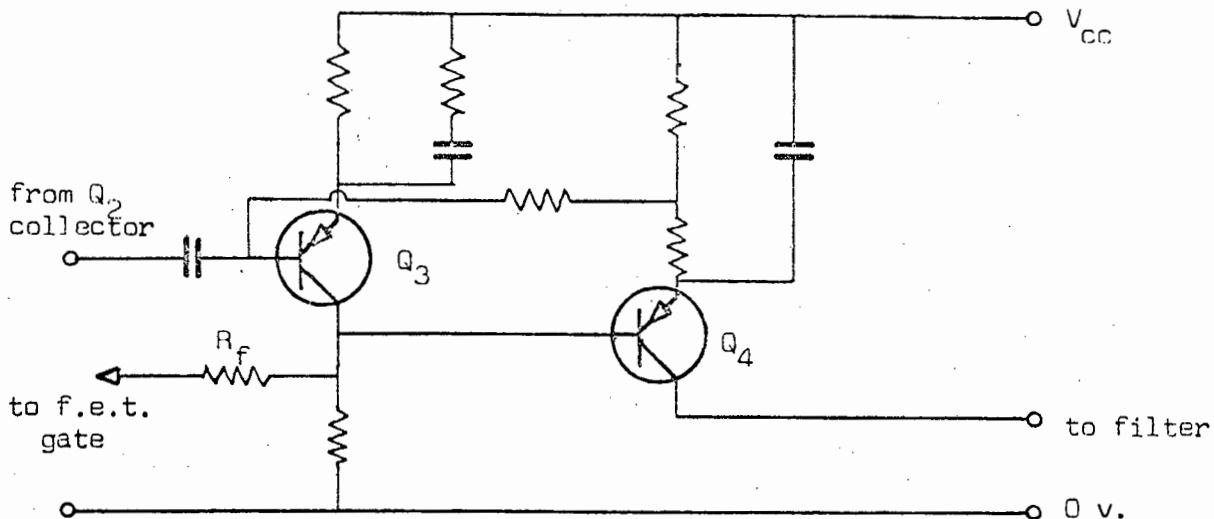
Open-loop a.c. gain of the first stage is $g_{m1}(R_d \parallel h_{ie2})$, and of the second stage, $g_{m2}(R_c + R' \parallel R_{in3})$. Two high-frequency breakpoints are likely to dominate the frequency response : one at the input of Q_2 caused by $(R_d \parallel h_{ie2})$ and C_{m2} , where $C_{m2} = C_{be2} + (1 + A_{v2})C_{bc2}$, the other at the collector of Q_2 caused by the collector load and C_{bc2} plus the input capacitance of the third stage. The Q_2 input breakpoint can be moved high in frequency by reducing R_d . This has the added benefit of increasing I_d , and hence g_{m1} . Although the gain of Q_1 will not increase, the input noise voltage will be reduced, because R_{nv} is inversely proportional to g_m . However, one does not wish to make I_d very large, because power consumption should be minimised. A compromise value in the vicinity of 1 mA is achieved by making $R_d = 680$ ohms. The collector current of Q_2 could be chosen purely on noise considerations, since it is possible to select an optimum I_c for a known value of source resistance¹, which in this case is the f.e.t. drain resistor R_d .

$$I_c(\text{opt.}) = \frac{(\beta_o)^{\frac{1}{2}}}{40 R_d} \quad \text{where } \beta_o = \text{d.c. current gain}$$

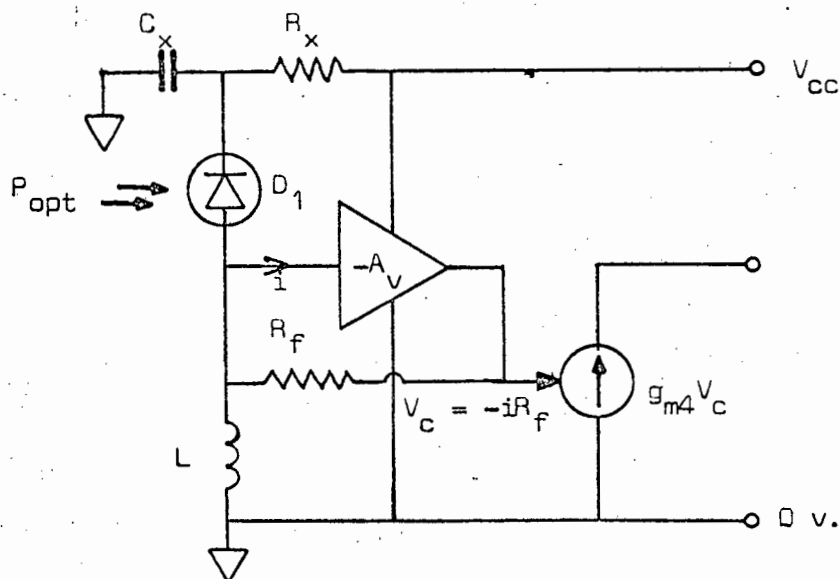
For a typical low-noise device, $\beta_o \approx 300$, resulting in $I_c(\text{opt.}) \approx 0,65$ mA. However, little degradation of noise performance results if the current is slightly reduced, to say 0,25 mA, in the interests of lower power consumption.

A moderate value of a.c. closed-loop gain for this feedback pair, of the order of 20 to 30 dB, is sufficient to render the noise contributions of subsequent stages negligible, while retaining enough loop gain to provide a low output impedance and broad frequency response. Another gain stage is now necessary to raise the open-loop gain of the transresistance amplifier as a whole to a level of the order of 50 dB, and to provide the desired phase-inverting characteristic. Thereafter, a further transistor is needed to drive the bandpass filter. It is convenient to make these two transistors into another feedback pair, with d.c. feedback only, so as to reduce the effect of changes in supply voltage. The use

of p-n-p devices has the advantage that the outputs are ground-referenced.



To reduce power drain, Q_3 and Q_4 are run at collector currents of about 0,3 mA. The gain of Q_3 is set to just over 30 dB at midband, and Q_4 's g_m of 12 mS gives a bandpass filter gain, to the LM1351 input, of about 6 dB. The frontend can now be represented as an inverting gain block with feedback resistor, driving a voltage-controlled current source (Q_4).



D_1 : MD1 photodiode, $V_r \neq V_{cc}$

L : d.c. input bypass choke

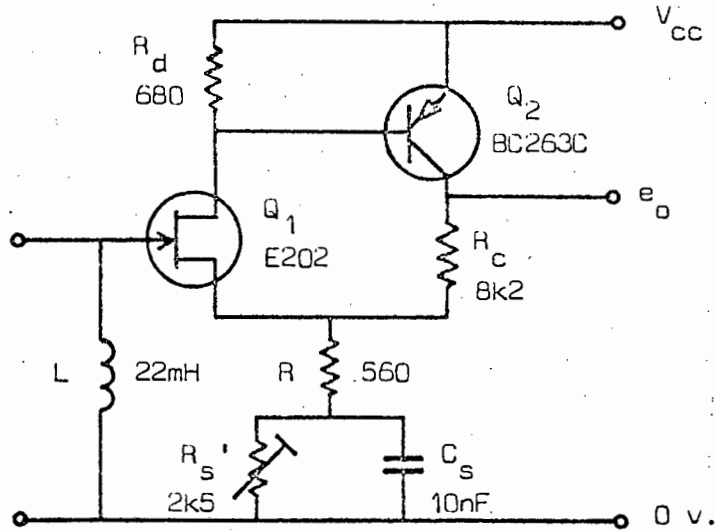
R_f : load resistor for D_1

C_x, R_x : D_1 's bias supply filter

$g_{m4} V_c$: a.c. component of Q_4 collector current

7. Component Values and Detailed Analysis.

(a) Q_1, Q_2 feedback pair.



To minimise frontend noise figure, Q_1 must have a very low gate leakage current and high g_m . In addition, it must have a low pinch-off voltage, so that the operating V_{gs} will be low, and the voltage across R_c can be made large, to increase second stage gain. The Siliconix E202 is a device which meets these requirements. It has a maximum gate leakage current of 100 pA at 25°C, and typical values are probably at least an order of magnitude lower. Transconductance is not notably high, at 1 mS minimum, but devices with low V_p and high g_m do not exist. Gate-source cutoff voltage varies between -0,8 and -4,0 volts, though typical values lie in the range -1 to -2,5 V. The E202, despite its outstanding potential as a low-noise amplifier at mid-frequencies, is very cheap, at 36 cents in single quantities, or 20 cents in lots of a thousand.

Since the voltage gain of the first stage is slightly less than unity, it is important that Q_2 should have a low input noise voltage. The device selected is the BC263C, a p-n-p version of the familiar BC109C. It is cheap, readily available from several manufacturers, and has good current gain at low I_c . Its high frequency characteristics are less than ideal, as a result of a rather high /collector-base.....

collector-base capacitance, but the gain-bandwidth product holds up well at low currents. A transistor with better h.f. performance would be expected to have a higher input noise, as a result of its lower current gain, so that these imperfections must be accepted.

The d.c. bypass inductor L , is a Philips microchoke of 22 mH, with a Q of 40 at 100 kHz, series resistance of 130 ohms, and stray capacitance of 3 pF. The coil is wound on a Ferroxcube tube, and is magnetically shielded by a second Ferroxcube tube, the whole encapsulated in an axial-lead package 10 mm by 6,5 mm diameter. Its cost is relatively high, at about R1,40 each, but the savings when compared to an active current source with feedback are considerable. The variable resistor R_s' provides for a range of V_{gs} values in Q_1 . Setting-up procedure involves measuring the voltage drop across R_c and adjusting R_s' until V_{Rc} is $0,25 R_c$ ($I_c = 0,25$ mA). The minimum value of V_{gs} allowed for is $(I_d + I_c)R$, where $I_d = V_{be}/680 \div 0,9$ mA. The values of R and R_c derived later give a range of V_{gs} from 0,65 V to about 2 volts. This should permit the use of a large percentage of E202s. R and R_s' are placed in series so that the closed-loop gain is substantially independent of the setting of R_s' , and no signal voltage appears across a potentially noisy variable resistor. Because of the restriction on V_{cc} and the range of V_{gs} that must be accommodated, the d.c. voltage across R_c must be limited to about 2 V, so that Q_2 will not saturate as V_{cc} falls. If $I_c = 0,25$ mA and the voltage drop is 2 V, then R_c must be 8200 ohms. Designing for a feedback factor of $1/15$ (-24 dB), the required value of R is $8200/(15-1) \div 560$ ohms. The capacitor C_s must have a reactance that is low compared to R . A value of 10 nF was chosen, giving $X_c = 35$ ohms at 450 kHz, and resulting in an error of less than 1 % in the feedback factor.

The open-loop gain of Q_1 is given by $g_{m1}R_d$, since h_{ie2} is much larger than R_d , and will not load it. $A_{v1} = -g_{m1}R_d = -0,68$, assuming $g_{m1} = 1$ mS. The midband collector load of Q_2 is the parallel combination of Q_3 's input impedance and R_c and R in series.

Hence $A_{v2} = -g_{m2}(R_c + R) \parallel R_{in3}$

where $R_{in3} = \beta_3(1/g_{m3} + R_{e3}) = 55 \text{ kilohm}$

and $g_{m2} = 39 \text{ mS per mA of } I_{c2}$
 $\approx 10 \text{ mS}$

Thus $A_{v2} = -75$ at midband, and the total open-loop gain for this feedback pair is :

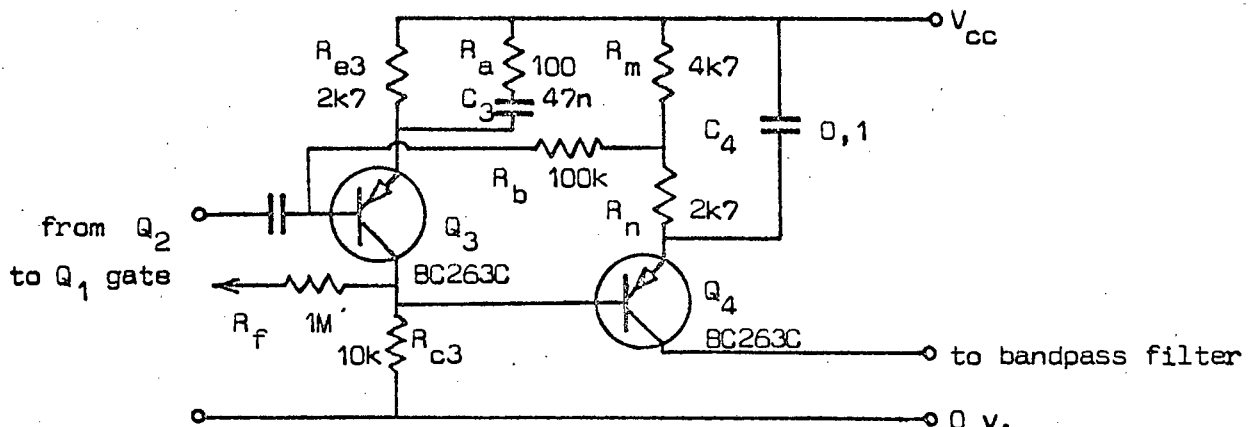
$$A_{vol} = 51 \text{ (34 dB)}$$

The closed-loop gain is thus :

$$A_{vc1} = \frac{A}{1-A\beta} = 12 \text{ (21,6 dB)}$$

The closed-loop output impedance is approximately 2 kilohms. The upper frequency response limits are determined by two R-C breakpoints, at the input and output of Q_2 . The input breakpoint is given by $f_1 = 1/2\pi R_d C_{m2}$ where C_{m2} is given by $C_{be} + C_{bc}(1-A_{v2})$, with $C_{be} \approx 8 \text{ pF}$ and $C_{bc} \approx 4 \text{ pF}$. The output breakpoint is at $f_2 = 1/2\pi R_{L2} C_{m3}$, where $R_{L2} = 7500 \text{ ohms}$. Analysis of the high-frequency behaviour is complicated by the fact that C_{m3} falls (because A_{v3} falls) at large f . Thus it is necessary to know the response of the third and fourth stages before the first and second stage frequency response can be calculated. Since in fact the only area of interest is that in the vicinity of 450 kHz, one need not evaluate gain and phase shift at other frequencies.

(b) Second feedback pair.



Choice of transistors for these stages is not very critical, since their noise contribution is small. However, Q_3 should be a high-gain device, so as to reduce loading on the second stage due to h_{ie3} . The BC263C used for Q_2 has the necessary gain, and can be used here as well. The filter driver has no special requirements, and any BC263 will suffice. Since the gain required of Q_3 is quite high, one needs a considerable fraction of V_{cc} across R_{c3} (the higher the voltage across the load resistor, irrespective of its value, the higher the midband gain). Taking a nominal 3 volts across R_{c3} and choosing $I_{c3} = I_{c4} = 0,3 \text{ mA}$, gives $R_{c3} = 10 \text{ kilohms}$.

$$A_{v3} = -(R_{c3} \parallel h_{ie4}) / (R_a + 1/g_{m3})$$

Now h_{ie4} will be 25 kilohm, and $g_{m3} = 12 \text{ mS}$. Hence $A_{v3} = -7140 / (R_a + 83,3)$. Making $R_a = 100 \text{ ohms}$ gives $A_{v3} = -39$ (32 dB) at midband.

The feedback bias network R_{e3} , R_b , R_m , R_n , sets up the currents in Q_3 and Q_4 in conjunction with the voltage drop across R_{c3} . Capacitors C_2 , C_3 , and C_4 are convenient standard values giving adequately low impedances for decoupling.

The gain at high frequency of Q_3 is affected by loading due to the Miller capacitance of Q_4 . The maximum collector load seen by Q_4 will be 3300 ohms near 450 kHz, resulting in a voltage gain of 40 and a Miller capacitance of 170 pF. The upper -3 dB frequency at Q_3 's collector will then be $1/2\pi R_{c3} C_{m4} \approx 100 \text{ kHz}$. At 450 kHz, the collector load of Q_3 will be $425 - j2000 \text{ ohms} \approx 2000 \angle -78^\circ \text{ ohms}$, and the gain will be $-11 \angle -78^\circ$ (21 dB).

Returning to the first two stages, one can evaluate A_{v2} at 450 kHz :

$$\begin{aligned} Q_2 \text{ collector load impedance} &= 7500 \text{ ohms} \parallel C_{m3} \\ &= 7500 \text{ ohms} \parallel 52 \text{ pF} \\ &= 3400 - j3700 \text{ ohms} \\ &= 5000 \angle -48^\circ \text{ ohms} \end{aligned}$$

$$\text{Thus } A_{v2} = -50 \angle -48^\circ \text{ at 450 kHz}$$

Hence Q_1 's drain load is $Z_d = R_d \parallel C_{m2} = 680 \text{ ohms} \parallel 228 \text{ pF} = 620 \angle -24^\circ \text{ ohms}$, with the result that $A_{v1} = -0,62 \angle -24^\circ$ and $A_{v01} = 31 \angle -72^\circ$.

$$\text{Hence } A_{vcl} = \frac{A}{1-A\beta} = \frac{31/-72^\circ}{1+2/-72^\circ} = 12,4/-22^\circ$$

Thus the total gain of the transresistance amplifier (Q_1 to Q_3) is -135 (43dB), with a phase lag of 100° . The virtual earth impedance presented by the feedback resistor R_f at the input is:

$$\begin{aligned} Z_e &= R_f / (1-A_v/\theta) \\ &= R_f (0,0074/-80^\circ) \end{aligned}$$

Alternatively, one may consider the virtual earth admittance:

$$\begin{aligned} Y_e &= 1/Z_e \\ &= 24,4/R_f + j133/R_f \end{aligned}$$

The parallel combination of Y_e and the admittances of the d.c. bypass inductor and the various capacitances will determine the input frequency response.

(c) The complete frontend.

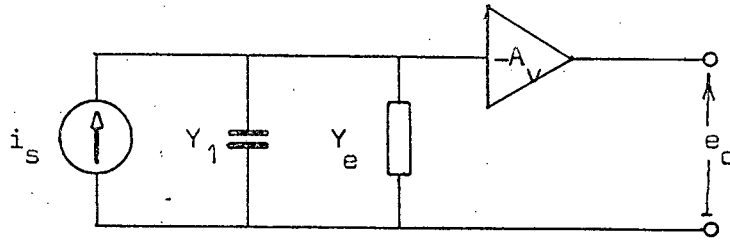
Components not yet specified are the photodiode bias decoupling network C_x and R_x , and the feedback resistor R_f . Since the anticipated maximum value of photocurrent is of the order of 10 microamps, R_x can be 10 kilohms and still only reduce V_{bias} by 0,1 volt. C_x should be an effective short at r.f., and also if possible at audio frequencies. Using a tantalum bead electrolytic of 10 microfarads satisfies both criteria, giving a time constant of 100 ms, while its low inductance gives a self-resonant frequency in the megahertz range.

The feedback resistor should in theory be as large as possible, to minimise thermal noise current. The upper limit for R_f is determined by the bandwidth requirement, the gain of the transresistance amplifier at the desired upper -3dB frequency, and the input capacitance.

This last is the sum of several components: the diode capacitance, the choke stray capacitance, the j.f.e.t. input capacitance, and mounting capacitances. These amount in total to some 20 to 25 pF, and resonate with the bypass choke at

/about....

about 200 kHz. At 450 kHz, the input susceptance is capacitive, and the input circuit can be represented for bandwidth calculations as:



$$Y_1 = 25 \text{ pF} \parallel 22\text{mH} \parallel Q\omega L = 5,36 \times 10^{-7} + j5,46 \times 10^{-5} \text{ (taking } Q = 30 \text{ at } 450 \text{ kHz)}$$

$$Y_e = 24,4/R_f + j133/R_f$$

Total input admittance is:

$$\begin{aligned} Y_t &= Y_1 + Y_e \\ &= (24,4/R_f + 5,36 \times 10^{-7}) + j(133/R_f + 5,46 \times 10^{-5}) \text{ S.} \end{aligned}$$

$$\text{Now } e_o/i_s = A_v/Q/Y_t$$

One expects that as R_f increases, the Y capacitive component will start to dominate Y_t , and the transfer impedance e_o/i_s will no longer be equal to R_f , as it is for small values of R_f . The -3 dB value of e_o/i_s occurs when R_f is 10^6 ohm, giving:

$$e_o = 7,1 \times 10^5 i_s$$

Increasing R_f to 10^7 ohm results in a transfer impedance of magnitude 2×10^6 ohm, 14 dB less than R_f . The sensitivity of such a large-valued resistor to inductively-coupled interference from other parts of the circuit is large, and since for R_f greater than the bypass choke parallel loss resistance $Q\omega L$, the thermal noise power will be increasingly affected by the $Q\omega L$ component, very large values of feedback resistor are not desirable. Hence a choice of 10^6 ohm for R_f represents a good compromise value.

/The.....

The overall gain of the frontend plus bandpass filter combination can now be found.

$$\text{Frontend transfer impedance } Z_{ft} = nZ_{f1}g_{m4}R_1'/2,24$$

$$\text{where } Z_{f1} = \text{transresistance transfer impedance} = 7,1 \times 10^5 \text{ ohm}$$

$$g_{m4} = Q_4 \text{ transconductance} = 12 \text{ mS}$$

$$R_1' = \text{tank circuit load resistance} = 3300 \text{ ohms}$$

$$n = \text{stepdown ratio of output coupling winding on second tank circuit} = 0,12$$

$$\text{Hence } Z_{ft} = 1,5 \times 10^6 \text{ ohms}$$

The signal photocurrent required to drive the LM1351 gain stages to limiting can be evaluated. Quoted sensitivity for the LM 1351 (output 3 dB below maximum) is 80 microvolts r.m.s., and the input voltage is given by:

$$e_{in} = i_s Z_{ft} \quad \text{where } e_{in} = \text{LM1351 input voltage}$$

$$i_s = \text{signal photocurrent}$$

Hence the limiting value of signal photocurrent is:

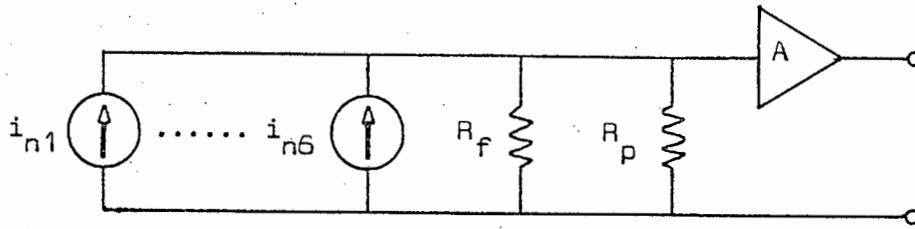
$$\begin{aligned} i_s &= e_{in}/Z_{ft} \\ &= 5,3 \times 10^{-11} \text{ A.r.m.s.} \end{aligned}$$

$$\approx 50 \text{ pA r.m.s.}$$

Bearing in mind the number of approximations and assumptions used, single-decimal accuracy is really all that one can expect.

(d) Frontend noise performance.

As described in the section on the bandpass filter, the effective noise bandwidth of the receiver is equal to the filter bandwidth, namely, 25 kHz. With all component values finalised, the noise model for the frontend developed earlier can be applied.



Device data: Photodiode MD1 $i_d = 3,2 \text{ nA max at } V_r = 6\text{V}.$

J.f.e.t. E202 $i_g = 100 \text{ pA max at } 25^\circ\text{C}.$

$g_m = 1\text{mS min at } V_{gs} = 0$

Q2 BC263C input noise $e_{n2} = 1,6 \times 10^{-9} \text{ V/Hz}^{\frac{1}{2}}$ for $I_c = 0,25 \text{ mA}, R_s = 680 \text{ ohm}.$
 $= 2,5 \times 10^{-7} \text{ v}.$

Assume that $i_s = i_b = 0$ (minimum shot noise). Then the photodiode shot noise is:

$$i_{n1}^2 = 2qBi_d = 2,6 \times 10^{-23} \text{ A}^2$$

Thermal noise is caused by the feedback resistor R_f and the choke parallel loss resistance $R_p = Q\omega L \approx 1,9 \times 10^6 \text{ ohms}.$

$$i_{n2}^2 = 4kTB/R' \text{ where } R' = R_f \parallel R_p$$

$$= 6,1 \times 10^{-22} \text{ A}^2$$

The j.f.e.t. noise sources are the gate leakage noise i_{n3} and the channel noise $i_{n4}.$

$$i_{n3}^2 = 2qBi_g$$

$$= 8 \times 10^{-25} \text{ A}^2$$

$$i_{n4}^2 = 2,67kTB/(g_m R'^2)$$

$$= 6,3 \times 10^{-25} \text{ A}^2$$

Referred noise due to the second stage is:

$$i_{n5}^2 = (e_{n2}/A_{v1} R')^2$$

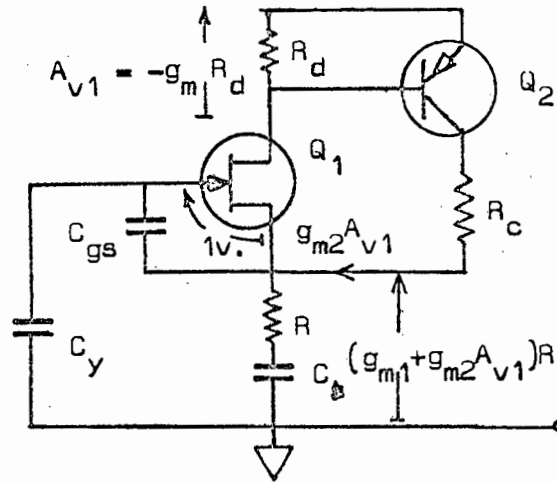
$$= 3,2 \times 10^{-25} \text{ A}^2$$

Finally, one must evaluate the high frequency noise component. To do this

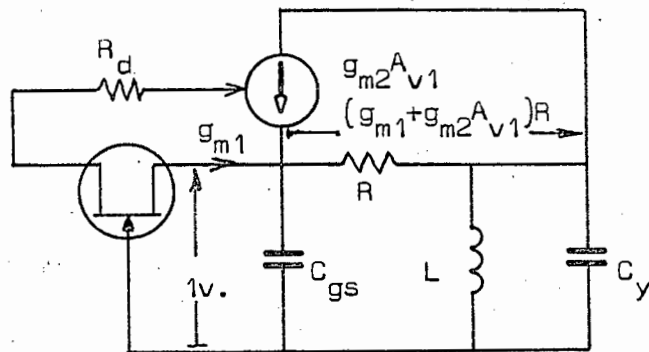
/it.....

it is necessary to calculate the total effective gate-source capacitance of Q_1 . Because of the feedback at Q_1 's source, the photodiode capacitance and other stray input capacitances do not appear directly across the gate-source terminals. In fact, as the following analysis shows, they are reduced in value.

Consider the circuit of the feedback pair:



This can be redrawn, with the gate as the reference node:



The effective value of C_x transferred to gate-source is $C_x = C_y(g_{m1} + g_{m2}A_{v1})R$. But C_x can be considered as two capacitors in parallel, one, C_1 , carrying the g_{m1} current of the f.e.t., and the other, C_2 , carrying the $g_{m2}A_{v1}$ current of Q_2 .

$$C_x = C_1 + C_2$$

$$\frac{C_1}{C_2} = \frac{g_{m1}}{g_{m2} A_{v1}}$$

Hence

$$C_x = \left(\frac{g_{m2} A_{v1} + g_{m1}}{g_{m1}} \right) C_1$$

One can express C_1 in terms of C_y : $C_1 = g_{m1} R C_y = 0,56 C_y$

Now C_y is made up of the parallel combination of the photodiode capacitance C_d , the stray input capacitance C_s and the d.c. bypass inductance L (this is only valid for a narrow band of frequencies), and can be found by summing susceptances :

$$\omega C_y = \omega C_d + \omega C_s - 1/\omega L$$

If $C_d = 14$ pF, $C_s = 6$ pF, $L = 22$ mH and $\omega = 9\pi \times 10^5$ ($f = 450$ kHz), then

$$C_y = C_d + C_s - 1/\omega^2 L = 14 \text{ pF}$$

Hence $C_1 = 8$ pF

Add to C_1 the f.e.t. gate-source capacitance of 5 pF, and $C_{in} = 13$ pF.

$$\begin{aligned} \text{Then } i_{n6}^2 &= 1,2 \text{ kTB} (\omega C_{in})^2 / g_m \\ &= 1,6 \times 10^{-22} \text{ A}^2 \end{aligned}$$

The total mean square input noise current is $i_{nt}^2 = 8,0 \times 10^{-22} \text{ A}^2$, and the resulting noise figure for the frontend is :

$$F = \frac{i_{nt}^2}{i_{n1}^2 + i_{n2}^2} = 1,26 \text{ (1,0 dB)}$$

Nearly all the excess noise is due to the i_{n6}^2 term, leading to the conclusion that an increase in g_m or a reduction in C_{in} would be desirable in improving the system performance. However, these noise calculations were based on the premise that there was no background photocurrent. If the noise figure is as low as 1 dB under such conditions, one can expect frontend noise to be negligible in normal daylight usage, with i_b of the order of microamps.

Note that the minimum value of i_{nt} is 28 pA r.m.s., compared with a -3 dB limiting sensitivity for the frontend of 50 pA r.m.s. Since a 12 dB S/N ratio is

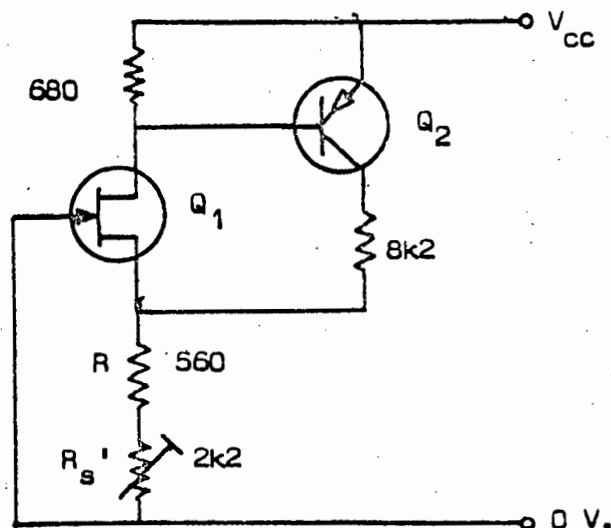
/ required

required for good FM reception, the sensitivity is sufficient for all possible conditions of use. In fact, with significant amounts of background radiation, the noise alone will cause the LM1351's amplifiers to go into limiting. The background photocurrent required to reach the -3 dB condition is 0,2 microamp.

(e) Performance at reduced supply voltage.

The frontend must operate satisfactorily on a 4,5 volt supply. Some degradation of performance is probably inevitable, unless one were prepared to make sacrifices in other areas, such as power consumption. However, the use of d.c. feedback should reduce the sensitivity of performance to supply voltage variations. The circuit is analysed at $V_{CC} = 4,5 \text{ V}$, and the results compared with the values obtained for $V_{CC} = 6 \text{ V}$.

(i) Q_1 and Q_2 .

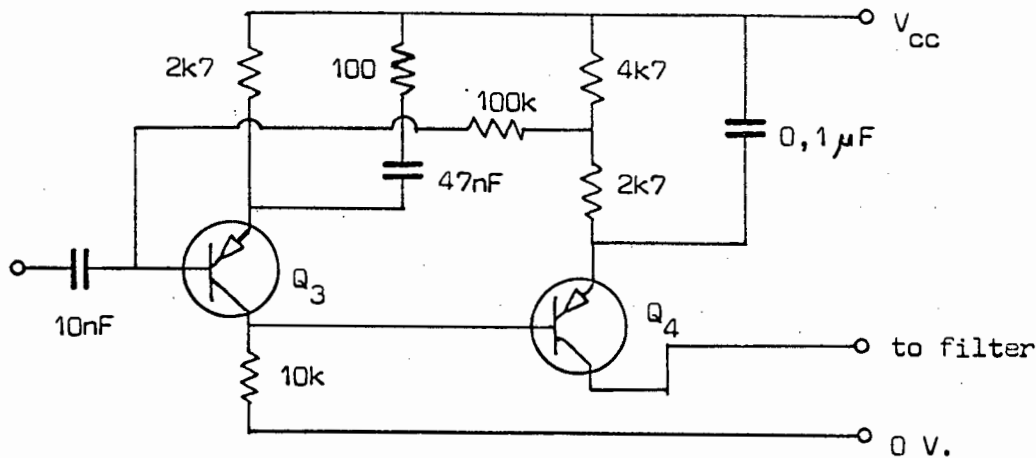


For $V_{CC} = 6 \text{ V}$, $I_d = 0,9 \text{ mA}$, and R_s' is adjusted to give $I_{C2} = 0,25 \text{ mA}$. The d.c. voltage at the collector of Q_2 is 2 volts above that at the source of Q_1 . Thus for V_{gs} of Q_1 up to 2 volts, Q_2 will not go into saturation at $V_{CC} = 4,5 \text{ V}$, and the only changes in circuit behaviour will be the result of the device capacitances varying with junction voltage. The most severely affected will be the collector-base capacitance of Q_2 , since the junction voltage can be zero under low battery conditions. The zero-bias value of C_{bc} is about 12 pF, so that the Q_2 input breakpoint would move down in frequency by a factor three. However,

/ a reverse

a reverse bias of only 0,5 V is sufficient to reduce C_{bc} to 6 pF, so that under most conditions, with $V_{cc} + V_{gs}$ less than 6,5 volts, the high-frequency performance will not be seriously affected, Any reduction in bandwidth will be in part offset by the applied feedback.

(ii) Q_3 and Q_4 .



For $V_{cc} = 6,0 \text{ V}$, $I_{c3} = 0,29 \text{ mA}$ and $I_{c4} = 0,33 \text{ mA}$.

For $V_{cc} = 4,5 \text{ V}$, $I_{c3} = 0,22 \text{ mA}$ and $I_{c4} = 0,27 \text{ mA}$.

V_{cc} (V)	I_{c3} (mA)	I_{c4} (mA)	Midband A_{v3} (dB)	g_{m4} (mS)	V_{cb3} (V)	V_{cb4} (V)
6,0	0,29	0,33	32	13	1,7	3,0
4,5	0,22	0,27	30	11	1,1	1,9

The collector-base voltages of Q_3 and Q_4 do not fall below 1,1 volt for V_{cc} greater than 4,5 V. Thus the collector capacitances of these transistors will not be greatly affected by falling V_{cc} , and the gain of Q_3 at 450 kHz will fall only slightly. In addition, Q_3 's substantially constant C_{bc} means that the h.f. gain of Q_2 , which is largely determined by the Miller capacitance of Q_3 , will not suffer at low V_{cc} .

The overall gain and noise characteristics of the frontend will be degraded by the drooping supply voltage. However, the design used has reduced the loss in performance to tolerable levels. The 4,5 volt minimum supply voltage places

/ restrictions

restrictions on the selection of devices for Q_1 . At a drain current of 0,9 mA, V_{gs} of Q_1 must lie between 0,7 V and 2,0 V. A simple test jig for the production selection of f.e.t.s is described in the appendices.

8. The Receiver Lens.

As mentioned in the section on choice of the system configuration, the low-cost criterion demanded the use of a small-area photodiode with an external lens to present a larger capture area. An acceptance angle of the order of one to two degrees either side of the optical axis can be expected. This narrow field of view is compatible with a hand-held communicator, providing operation from a moving platform is not required.¹ Cost and weight considerations dictate the use of a plastic, rather than glass, lens. Absorption of near infrared radiation by commonly-used plastics is about 10 %, which is comparable with the loss in glass. The advantage of a plastic lens is that it can be moulded directly, and does not need to be ground. Aspheric lens surfaces can thus be generated if required, and biconvex lenses with different radii of curvature for each face can be produced very cheaply with quite large F. numbers. A locally-made lens with radii of curvature chosen to minimise spherical aberration was selected as fulfilling all the major requirements. Such a lens must be used with the more convex side directed towards the transmitter for minimum aberration, so that the angles of refraction at the two surfaces are approximately equal.²

The chosen lens has a diameter of 58 mm and a focal length of 82,5 mm, the corresponding F. number being 1,42. For the case of minimum spherical aberration of a single lens, the diameter of the blurred spot is given by :

$$d_b = 0,067 \frac{f_r}{(F.no.)^2} = 1,9 \text{ mm.}$$

1. Refs. 22,30

2. Ref. 8, p.258.

The beam half-angle $\theta = c_b/2f_r$ radian

= 0,67 degree for the critically-focused case. This is rather smaller than is desirable for ease of aiming, so some defocusing is needed to increase the beam half-angle to 1° . The diameter of the defocused spot is then:

$$\begin{aligned} d_b' &= 2 f_r / 180 \text{ mm} \\ &= 2,9 \text{ mm} \end{aligned}$$

The photodiode active area is $0,58\text{mm}^2$, and the lens capture area is given by:

$$\begin{aligned} A_c &= \frac{d_r^2 A_d T}{(d_b')^2} & \text{where } A_d &= \text{diode active area} \\ &= 360 A_d & T &= \text{lens transmission coefficient} \\ &= 2,1\text{cm}^2 & &\neq 0,9 \\ & & d_r &= \text{receiver lens diameter} \end{aligned}$$

The lens provides a power gain of 360 times, or about 26 dB. If the receiver were to be used in a semi-fixed application, the lens could be critically focused for maximum range, and that would give:

$$\begin{aligned} A_c &= 840 A_d \\ &= 4,9 \text{ cm}^2 \end{aligned}$$

This corresponds to a lens power gain of 29 dB. Alternatively, for greater ease in aiming, the acceptance angle could be increased to $\pm 2^\circ$, with a corresponding drop in capture area, by further defocusing.

$$\begin{aligned} \theta &= 2^\circ \text{ gives } d_b'' = 5,8 \text{ mm} \\ \text{and } A_c &= 91 A_d \\ &= 0,53 \text{ cm}^2 \end{aligned}$$

Thus a small mechanical adjustment can vary the field of view over a useful range, so that the tradeoff between acceptance angle and sensitivity can be optimised for a number of potential applications.

9. The Receiver Audio Amplifier.

To complete the basic optical receiver, all that is now required is a means of making the demodulated signal audible. The choice of headphones or a loudspeaker is somewhat dependent on the configuration of the system. If a single unit similar to a stick-type walkie-talkie were used, a combination speaker/microphone would be ideal. A head-mounted set would be an obvious application for the use of a headset, probably with boom microphone. Another possibility is a cine-camera shape, in which case wither a loudspeaker or headphones could be used.

Because the coupling to the ear is much closer, a headset requires less driving power for the same level of intelligibility than a loudspeaker. It might be possible to drive a headset from a single-ended class A amplifier (in the interests of simplicity), without incurring excessive current drain. The loudspeaker on the other hand needs a class B or just conceivably a class D amplifier, to reduce the standing current requirement and to obtain reasonable efficiency. Since class B audio amplifiers are available in monolithic i.c. form, they are the most cost-effective choice for both types of transducer. A problem with many of the integrated amplifiers is a limited capability for operation at low supply voltage. One of the most widely-used small amplifiers, the LM380 by National Semiconductor, will only operate down to 8 volts, and others are specified down to 6 volts, but do not work well due to severe crossover effects. The silicon transistor base-emitter voltage of about 0,65 volt creates difficulties in the design of low-voltage power amplifiers. If a complementary Darlington output stage is used, the voltage drop across four base-emitter junctions (2,6V) must be subtracted from the supply voltage to give the approximate peak-to-peak output voltage. Such voltage losses severely restrict the power output at V_{cc} of the order of 6V.

/An.....

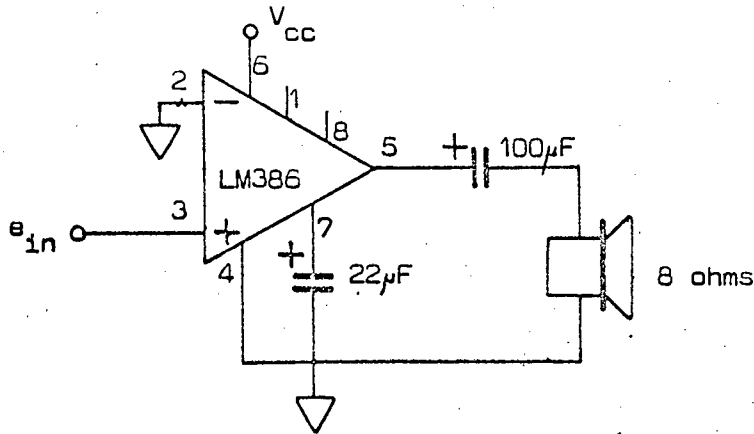
An amplifier with this defect is the Motorola MFC4000B. Its quasicomplementary output stage throws away 2,2 volts of potential signal swing, so that for $V_{CC} = 6V$, the maximum output power into 8 ohms is 50 mW. While 50 mW can be louder than one would expect, the overall performance of this amplifier was inadequate and initial tests with it were discontinued.

The amplifier used in the first prototype was a physically larger, European-sourced device, the TBA820. Housed in a 14-pin dual in-line package, it operates off 3 V to 16 V supplies, and delivers up to 2 watts into 8 ohms with a 12 volt supply. With $V_{CC} = 6 V$, 500 mW into 8 ohms is possible. Quiescent current is about 3,5mA at 6 volts, input resistance is very high (5 Megohm), and distortion at powers below the clipping level is less than 1%. But the number and size of external components required is excessive. The typical application circuit uses two resistors and seven capacitors, three of the latter being relatively large value electrolytics. For communications-bandwidth use these can be reduced in size, but a simpler circuit would be preferred.

Towards the end of the design process, a number of new audio amplifier circuits appeared. Amongst these was a device which was very nearly ideal. In an 8-pin mini-dip package, the National LM386 is specifically designed for low-voltage, low-power consumer applications. It requires in its simplest form only one external part, namely the output coupling capacitor. Voltage gain is set internally to 20, quiescent current is 3 mA, and a minimum of 250 mW can be delivered into 4 or 8 ohms at $V_{CC} = 6V$. The 8 ohm load is preferred, since it keeps dissipation low, thanks to an efficiency at full output of about 50%. Distortion is typically less than 0,5% at 100 mW output, which is quite adequate for a communications-quality channel. Clever design has removed the need for an external bootstrapping capacitor. Apart from the output capacitor, which for a 300 Hz minimum modulation frequency can be 100/microfarad, the only optional external component that one would

/like.....

like to add is a power supply bypass capacitor, for which a tantalum bead would be ideal. The bypass capacitor ensures a 50 dB attenuation of supply ripple at the output.



A potential design problem results from the direct-coupled input. The input bias current is 250 nA, and if the source resistance is high, the input offset voltage can be as much as 12,5 mV. Since the gain is 20 even at d.c., this will result in an offset voltage of 0,25 V at the output. The available output voltage swing will be reduced because the quiescent output voltage will no longer be $\frac{1}{2} V_{CC}$. For the output offset to be negligible, the d.c. source resistance must be 10 kilohms or less.

Tests of this power amplifier showed typical efficiency at maximum output of 48 %, and output power of 100 mW at $V_{CC} = 4,5$ V, up to 220 mW at $V_{CC} = 6$ V, with 8 ohm load, and maximum input voltage of 70 mV r.m.s. The frequency response is very broad, with the upper -3 dB point at about 400 kHz. It will be important to prevent carrier feedthrough to the power amplifier if one is to avoid unnecessary dissipation in the output stage. Further power dissipation reduction will be effected by the squelch and power-saving circuitry.

10. Receiver Power Saving Circuitry.

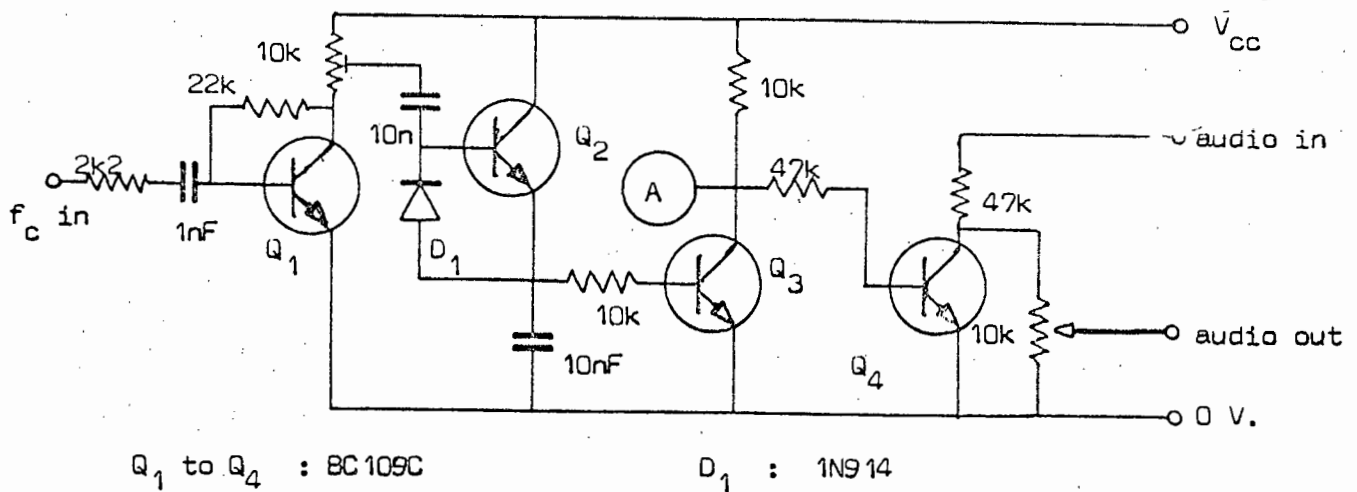
In many applications requiring a single-channel voice link, the receiver must be active for long periods during which no signals arrive. It cannot be turned off altogether, or new messages will not be detected, but some means of

/reducing

reducing the battery drain during these quiescent periods could extend the battery life substantially. One can be certain that there will be no output from the transmitting end except when a message is being sent, since the power drain of the transmitter is high. Thus a test for whether an incoming signal is present is the level of the subcarrier frequency at the output of the LM1351's limiting amplifier. In the absence of transmitted subcarrier, the amplifier output will be band-limited random noise, while with signal the output is a 50 % duty cycle squarewave. Since in the absence of carrier an FM detector's output is noisy, it is desirable to suppress the audio amplifier input under these conditions, both to conserve power and to remove an annoyance to the user. A peak detector can distinguish between the two cases, provided that the noise is not causing the amplifier to limit continuously. The d.c. output of the peak detector can then be used to drive circuits which control the more power-hungry parts of the receiver. The simplest of these circuits would be an analogue switch at the demodulator output, to prevent the noisy output caused by absence of subcarrier from reaching the audio amplifier. Since the output level of the LM1351 limiter is low (170 mV p-p), a gain stage is needed to drive the peak detector. If the gain is made variable, the squelch (audio suppression) threshold level can be controlled.

(i) The squelch circuit.

The gain stage is as simple as possible, yet is stable against supply voltage changes, and has low current drain. A transistor-aided diode peak detector ^{attack} combines a fast/and slow decay with negligible power consumption. A third transistor provides d.c. gain and a phase inversion so that it can drive the analogue switch. This last is not required to have any special characteristics, since it is only controlling a small a.c. voltage. A single transistor shunt switch is sufficient. The demodulator is expected to produce a maximum output of 400 mV r.m.s., and only 70 mV is required at the power amplifier, so some attenuation of the signal in the 'on' state is desirable.



In the quiescent state, current drain is about 0,6 mA, mostly due to Q₁. In the active state, Q₃ adds a further 0,6 mA to this. The voltage at A is high (near V_{cc}) with no subcarrier input. When the a.c. voltage at the base of Q₂ reaches about 0,7 V p-p, Q₂ turns on, and charges up the capacitor in its emitter circuit, allowing Q₃ to turn on and pulling down the voltage at A. Q₄ then turns off, and the audio line is enabled.

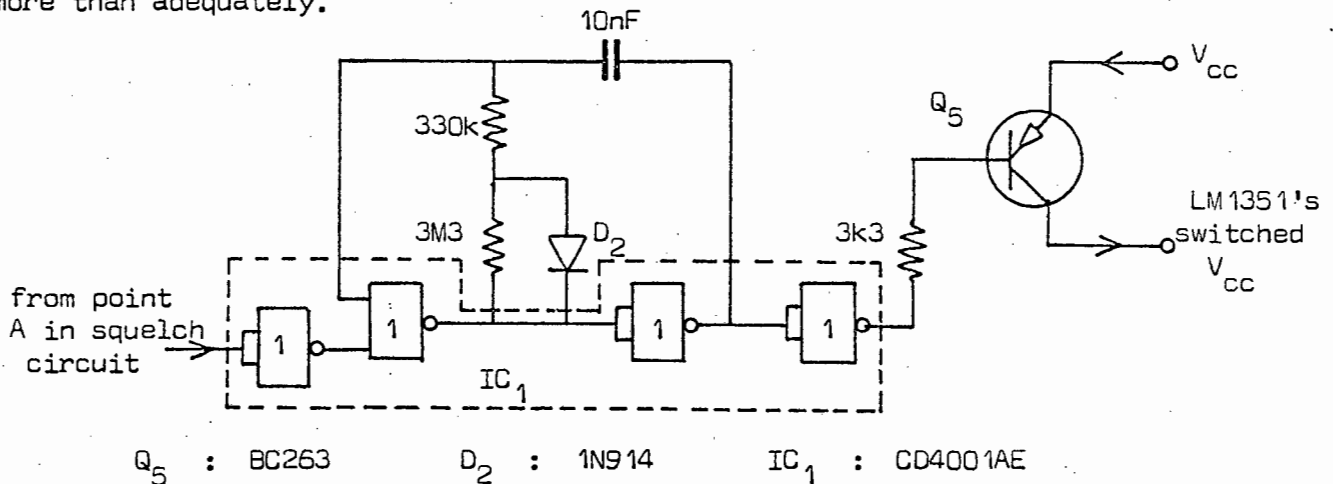
(ii) More power saving.

The receiver frontend draws about 1,8 mA, the audio amplifier 3 mA, and the squelch circuit 0,6 mA under no-signal conditions, making a total of about 5,5 mA. But the LM1351 alone draws in the region of 15 mA. If one could reduce this current drain, the total power consumption might be reduced by 50 % or more. The simplest approach to reducing the LM1351's current drain is to switch it off most of the time, and turn it on occasionally to see whether any signals are being received. When a signal is received, the squelch circuit can be used to hold the power to the LM1351 on until transmission ceases. To ensure that part of a message is not lost, the receiver must activate itself several times a second. But since the squelch circuit requires only a few cycles of subcarrier to change state, the on-time can be small, of the order of milliseconds.

What is then required is an astable multivibrator with a low duty cycle

/ which

which can drive a power switch and which can be gated off by a d.c. voltage. Low power consumption and simplicity are prerequisites. A discrete transistor version was built, but it was far from ideal, and needed large capacitors for timing. The obvious candidate for any low-powered logic circuit is CMOS, and a circuit was evolved using one gate package and a bipolar transistor which met the requirements more than adequately.



The astable oscillates at just over 40 Hz with a 9 % output duty cycle (on time of 2,3 ms). Gating is derived from the squelch circuit, so that when the point A goes low, the oscillator stops with Q₅ turned on. Oscillator power drain is effectively zero, and the only additional current is drawn by Q₅'s base current when it is on, this current being 1,5 mA or less (sufficient to saturate Q₅ for I_C = 15 mA). Thus when there is no incoming subcarrier, the astable runs, and it and Q₅ and the LM1351 draw a combined total of 1,6 mA average current. Under these conditions, the total receiver current drain is 7 mA, compared to 20 mA without the power saving circuit. Considering the small number and cost of the components involved, the gain in battery life should be ample justification for their inclusion in the design.

11. Receiver Circuitry - General.

The receiver section and its anticipated performance has been described. Certain additional components such as supply decoupling elements have not been

/ shown

shown. These will be found on the circuit diagram for the complete optical communicator, and have to a large extent been determined empirically. Less time is consumed in measurement of ripple levels on the actual circuit than in the calculation required, and the results achieved have greater reliability. The effects of lead inductance and capacitive coupling are much easier to cure than to predict.

Because of the very small signal levels present in the receiver frontend, and the large voltages in the squelch circuits, extreme care in layout is necessary. Construction on double-sided printed circuit board with a ground plane will minimise stray capacitance, and arrangement of circuits in a logical sequence should reduce the length of p.c. tracks. Some simple, easy to produce mounting for the photodiode will be necessary, so that its position relative to the lens can be adjusted. Once again, it will be essential to keep lead lengths on the photodiode short.

In the design of the receiver, every effort has been made to produce a cost-effective design, but one which does not sacrifice performance for trivial cost reductions. As many common parts as possible are used, and the only devices which may be hard to procure are the audio amplifier (because it is a new product) and the photodiode. Power consumption has been kept low, so that the receiver can remain on standby in semi-fixed applications for long periods. Use of a lens confers high sensitivity, while avoiding the cost of a large-area detector, so that the total cost of all receiver parts in sample quantities is about R20, falling to about R12 for production quantities. The design requirements have been met.

(iii) The Transmitter.12. The Transmitter Modulator.

The general specification of the modulator is well defined : a voltage-controlled oscillator with offset, centre frequency of 450 kHz, maximum deviation ± 6 kHz, control voltage frequency response from 300 to 3000 Hz. All the design constraints relating to supply voltage, power consumption, and complexity must be taken into account, and in addition it is required that the centre frequency shall not drift from its nominal value by more than $\pm \frac{1}{2} \%$. Frequency stability is necessary because the receiver bandpass is fixed, and cannot be tuned to the transmit frequency. The main cause of frequency drift will be variations in temperature, over a probable range of 0° to 50° C. (a typical commercial equipment temperature specification).

The voltage-controlled oscillator can be implemented in a number of ways : one of the various controllable function generator i.c.s would meet most of the performance requirements, but its power supply demands would be excessive, both in voltage and in current (the CMOS CD4046 was not yet available). An astable multivibrator with variable period could be constructed with discrete devices or with logic gates, or an L-C oscillator with varicap tuning could be used. The astable multivibrator idea is attractive, but difficult to effect within the power supply limits. It also has an inherent tendency for frequency to drift with temperature, as a result of the transistor base-emitter voltage temperature dependence, and it is extremely susceptible to power supply voltage variations. An all-discrete astable multivibrator VCO was constructed and temperature-tested, and performance was outside the allowable limits. In addition, elaborate circuitry was necessary to overcome the supply voltage dependence problem.

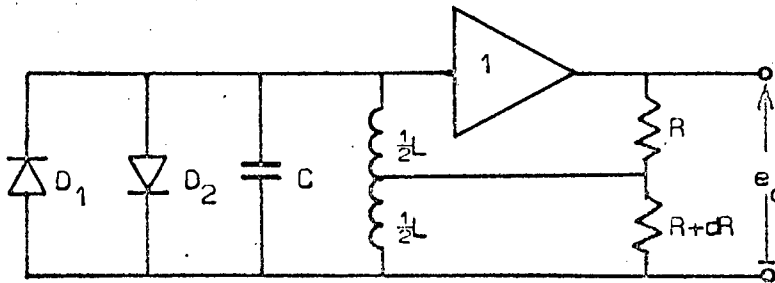
The final design resorted to the use of a varicap-tuned L-C oscillator. Choice of a suitable temperature coefficient for the fixed tank circuit capacitor can balance the positive temperature coefficient of the varicap diode, but other

/ problems ...

problems have to be solved. The diode capacitance law is a nonlinear one, as is the frequency-capacitance relation of the oscillator. The result is only approximate linearity, and even for the 6 kHz deviations required, the harmonic distortion is somewhat higher than is desirable. However, since the distortion is mainly low-order harmonics, it can more easily be tolerated than is the case for higher-order odd harmonics such as are produced by amplifier crossover distortion.

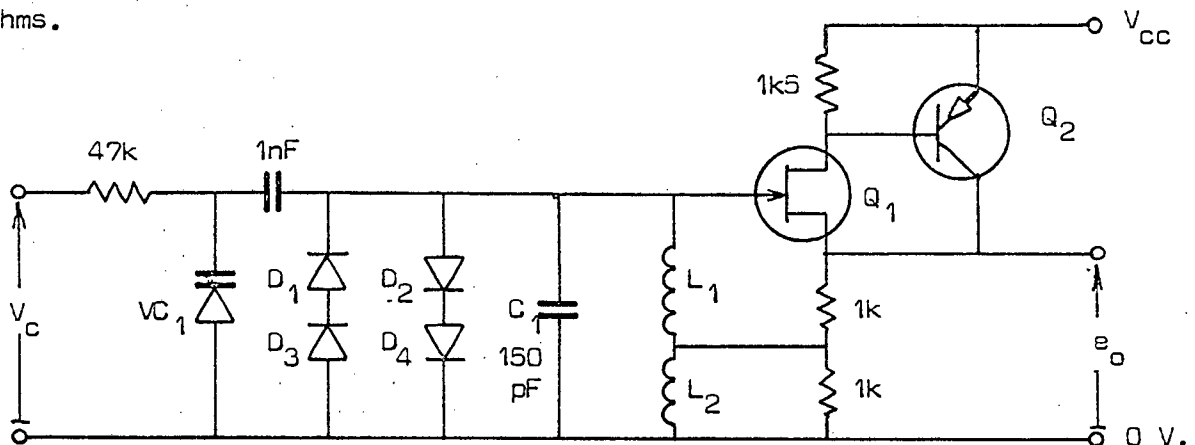
The waveform generated by the L-C oscillator is sinusoidal, whereas the most convenient type of drive for the transmitting l.e.d. is a low-duty-cycle pulse waveform. Some circuitry is therefore required to transform the sinewave from the oscillator into a pulse waveform. Such circuitry should produce a constant duty-cycle regardless of input amplitude, and should be able to fail safe, so that if the VCO should cease oscillation, the l.e.d. will turn off, rather than on. This is necessary not only to conserve battery power, but more importantly to prevent destruction of the l.e.d. itself, since the peak pulse current will be greater than its maximum continuous current rating.

Some stabilisation of the amplitude of oscillation may be necessary, depending on the type of circuit used to perform the pulse shaping. For the sake of linearity, it is important to avoid any clipping of the sinusoidal waveform due to excessive amplitude or low power supply voltage. A simple method of amplitude stabilisation involves the use of diodes connected in parallel with the tuned circuit. At voltages below the diode conduction voltage, they present a high impedance to the tank circuit, and its Q is high, but when the diodes conduct, the tank circuit Q drops sharply because of the low on-resistance. Some nonlinearity of the waveshape results, the peaks being somewhat rounded, but the effect is small, and is not serious in this application. The amplitude resulting from the use of clamping diodes will be of the order of 1 V p-p or more, depending on the diodes used. Since the amplitude of oscillation at the tuned circuit will be large, no gain is required of the driving amplifier. A tapped-coil drive will provide the necessary loop gain.



The Basic Oscillator.

To maintain a high loaded Q on the tuned circuit, the input impedance of the amplifier must be large. One could use a f.e.t. source follower alone as the buffer, but the gain would be very dependent on individual f.e.t. characteristics. In addition, the output impedance of the source follower is relatively high. The use of a f.e.t.-bipolar feedback pair in a unity gain configuration overcomes both of these disadvantages, with only two extra components. A loop gain of typically 30 dB results in a gain error of less than 5 % and an output impedance of under 100 ohms.



The Complete Modulator (VCO).

Q_1 : E202 Q_2 : BC263 D_{1-4} : 1N914 VC_1 : Motorola MV2115 varicap
 L_1 and L_2 are contained in a Hitachi type 20/3 oscillator coil intended for use in an AM radio. L_1 is greater than L_2 , so that the feedback factor from the output to the top of the tank circuit is greater than unity. The coil is slug-tuned, so that f_c can be adjusted. Diodes D_1 to D_4 have a 0,6 volt forward voltage, resulting in an output voltage at 450 kHz of 2,4 V p-p.

Selection of devices for Q_1 is necessary, so that at 0,4 mA I_d , V_{gs} lies

/ between.....

between 1,2 V and $V_{cc} - V_{be} - e_o(pk)$. This latter voltage has a minimum value for $V_{cc} = 4,5$ V of about 2,5 volt.

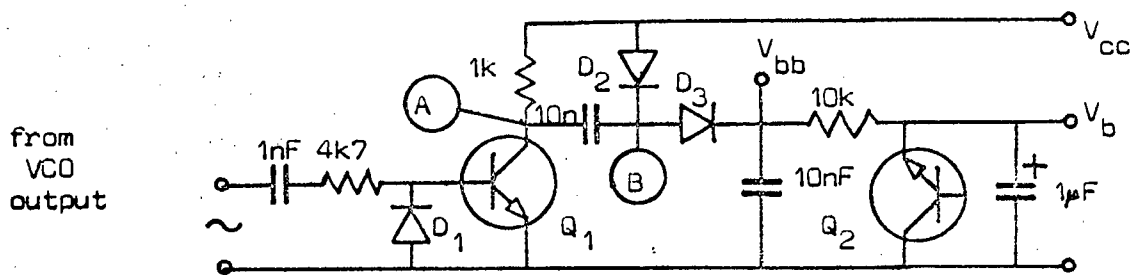
To justify the use of an L-C oscillator as a limited-range VCO, one must consider the tuned-circuit capacitance-frequency and the varicap capacitance-voltage relations. Tuned circuit centre frequency is $f = (2\pi LC)^{-\frac{1}{2}}$. For the varicap, junction capacitance $C_j = C_o (V_r + \phi)^{-\frac{1}{44}}$, where C_o = zero bias capacitance, V_r is the reverse voltage, and ϕ is the contact potential, equal to 0,6 V. The resulting expression for f in terms of V_r is complex, and to evaluate its linearity analytically is not easy. A more practical technique is to evaluate the expression for f with given constants at several values of V_r , and to fit a polynomial curve to the data obtained. V_r can be replaced in the polynomial by the sum of a d.c. bias voltage V_b and a sinusoidal voltage $V_1 \sin \omega_m t$. It is then possible to evaluate the distortion at varying modulation levels and bias voltages. However, there are practical difficulties involved in obtaining a good polynomial fit to a set of data points which are so close together, since one is limited by the accuracy to which one can calculate. At peak frequency deviations of ± 6 kHz (around 450 kHz), one can expect a well-configured L-C type VCO to produce distortion of the order of several percent. In fact, measured distortion at the receiver demodulator output for peak frequency deviations of about 6,5 kHz was just over 4 %, mostly 2nd and 3rd harmonic. At typical modulation levels, one would expect distortion to be 2 % or less, almost entirely 2nd harmonic. Thus while this type of VCO is not linear, its nonlinearities can be tolerated in this application, where its potential centre-frequency stability is advantageous.

In the interests of linearity, it is desirable to use a high-capacitance varicap diode, so as to reduce the voltage swing required for the desired modulation depth. However, it is also necessary to add a fixed capacitor to cancel the varicap diode's temperature coefficient, which is approximately +280 ppm per °C. The largest cheap varicap available was the Motorola MV2115, nominally 100 pF at a reverse voltage of 4 volts. To balance the diode temperature coefficient, the fixed capacitor, which had to be 150 pF to give $f_c = 450$ kHz, must have an N150 (-150 ppm / per °C)

per °C) characteristic.

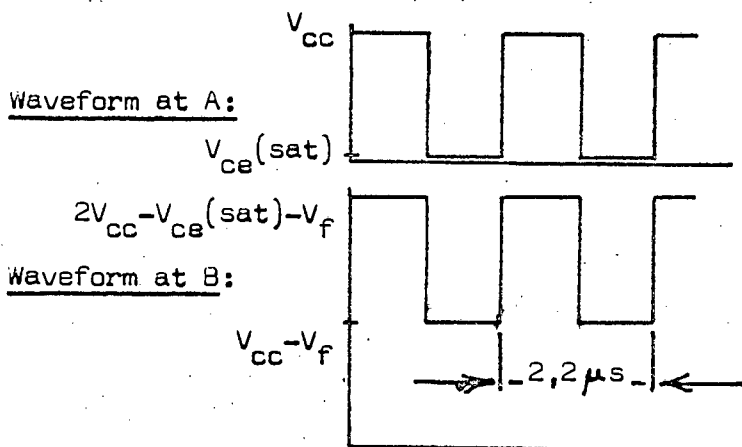
Without temperature coefficient compensation, the frequency drift with temperature due to capacitance variations alone would be nearly 1 % over a 50° C range whereas the compensated circuit will be almost independent of temperature.

The preceding treatment of the varicap-controlled VCO has assumed that a fixed bias voltage is available to set up the centre frequency. The bias voltage must be high enough to avoid the most nonlinear portion of the voltage-capacitance characteristic, near zero, yet it must be stable against changes in supply voltage. Since the minimum value of V_{cc} is only 4,5 volts, one cannot use a zener-stabilised fraction of V_{cc} as V_b . A d.c.-to-d.c. converter is required to produce a voltage greater than V_{cc} , which can in turn be used to provide the varicap bias voltage. Since the varicap draws only leakage current, of the order of nanoamps, the converter can have very limited current capability, just sufficient to drive the voltage regulator. This makes it possible to use a diode-capacitor voltage multiplier, and avoid the need for a special transformer. A number of circuits, self-oscillating and otherwise, were considered, and a very simple design resulted.



Q_1, Q_2 : BC109C (Q_2 selected for BV_{eb} between 5,5 and 7,0 V).

D_1 : 1N914 D_2, D_3 : AAZ15 (or any diode with V_f less than 0,5 V).



$V_{ce}(sat) = Q_1 \text{ saturation voltage}$

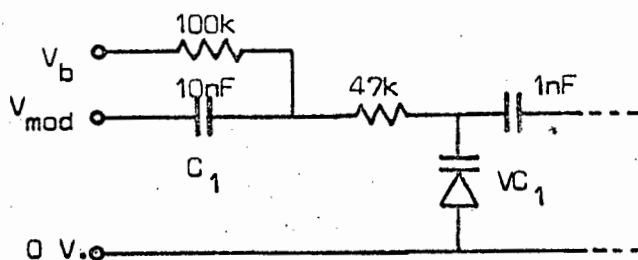
$V_f = \text{forward voltage (of } D_2)$

C_1 and D_1 bias Q_1 so that input voltages of 1,5 V p-p or more will cause it to turn hard on and off, with the collector voltage a squarewave. R_1 limits the base current, to prevent excessive loading of the driver, which is in fact the VCO itself. Diodes D_2 and D_3 and their associated 10 nF capacitors form a voltage doubler rectifier, with the output voltage V_{bb} approximately twice V_{cc} . Running at 450 kHz makes the use of capacitors as small as 10 nF in the doubler possible. Some voltage is lost in Q_1 , D_2 and D_3 , so that :

$$V_{bb} = 2V_{cc} - V_{ce}(\text{sat}) - 2V_f = 2V_{cc} - 1 \text{ volts, under}$$

no-load conditions. The load presented by Q_2 , which is used as a voltage reference, increases rapidly as V_{bb} goes above its breakdown voltage, so that for large V_{cc} , V_{bb} is considerably less than the value given by the simple formula above. The breakdown characteristic of Q_2 's base-emitter junction is sharp, much sharper than that of a conventional avalanche breakdown diode, so that very small reverse currents are sufficient to establish the breakdown voltage. Connecting the collector-base diode in series, forward-biased, reduces the temperature coefficient of V_b . Measurements showed that V_b varied by about 0,15 % over the 4,5 to 6 V range of V_{cc} , so that VCO drift due to power supply voltage variation is negligible. Temperature drift of frequency due to V_b drift can be compensated for by selection of a suitable temperature coefficient for the fixed capacitor in the VCO, should this prove necessary.

One must isolate the bias voltage source from the compressor output, to prevent loading of the latter. Since the varicap draws no d.c. current, a large-value resistor can be used for this purpose.



The coupling capacitor C_1 is chosen to give a highpass filter with

/ corner

corner frequency of 160 Hz. This prevents excessive low-frequency signals from reaching the modulator.

Current drain of the VCO itself is typically 1,3 mA, while the bias generator draws a maximum of just over 3 mA, for a total of about 4,5 mA at $V_{cc} = 6V$, falling to about 3,5 mA at $V_{cc} = 4,5V$. The overall complexity of the circuit, although high compared with integrated circuit VCOs, is low by comparison with astable multivibrator VCOs and their associated power supply regulation.

13. Microphone Preamplifier and Level Compressor.

The r.m.s. value of speech waveforms is one quarter of the peak value.¹ This results in a low average modulation index and consequently a smaller amount of S/N ratio improvement at the demodulator than if the modulation index were maximum. The intelligibility of the communications channel can be raised by increasing the r.m.s.-to-peak ratio of the modulating waveform. This can be done by amplitude compression of the voice signals. Such compression constitutes a form of distortion, and would not be accepted on a broadcast system. When used with voice signals only, the slightly unnatural sound is an acceptable consequence of an increased S/N ratio. Choice of a suitable time constants in the compressor feedback loop can to a large extent restore a natural sound to the compressed speech, at the expense of some loss in modulation level. The use of a compressor has a second advantage, in that it renders less critical the actual speech level and microphone position, giving a constant output amplitude for a wide range of input levels.

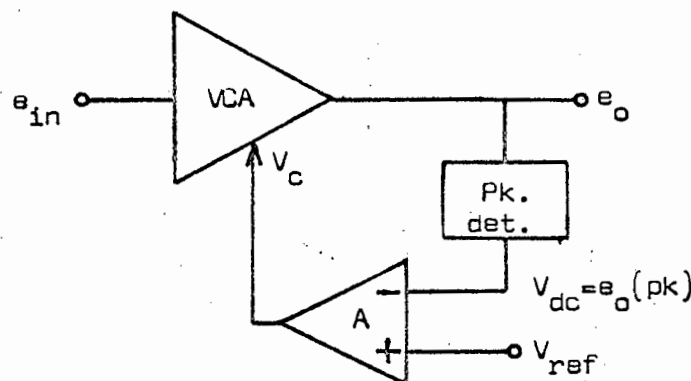
Several types of microphone could be used in the communicator. Of these, the crystal microphone has a capacitive source characteristic, and must be used with a high load resistance. Physical size is another problem, ranging up to 30 mm.

/diameter.....

diameter and more, for the active element alone. A possible advantage of the crystal is its high output voltage, typically tens or hundreds of millivolts. The capacitor microphone, thanks to the development of the electret capacitor, has ceased to be solely a high-cost device used by sound recordists. Availability is still a problem, however, and the cost advantage lies with the common dynamic or moving-coil microphone. This last has a low source impedance, low output voltage (of the order of 1 mV), and can be obtained in small packages suitable for mounting in the case of the communicator.

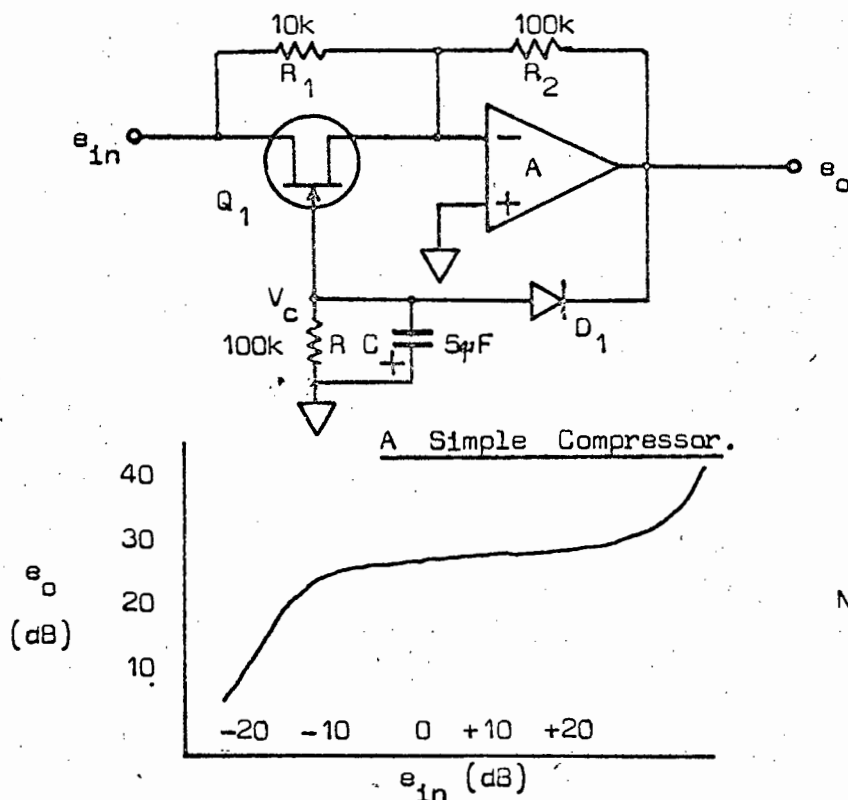
The microphone chosen is a Japanese dynamic, type SM-1901B, with 2 kilohm nominal source impedance, and sensitivity quoted as -70 dB (referred to 1 V/microbar). At a typical operator-to-microphone separation of 0,1 metre, male speech sound pressure levels lie in the range of 50 to 90 dB. Since 0 dB SPL corresponds to 2×10^{-4} microbar, one can expect the output voltage from the microphone to be in the range 0,02 to 2 millivolts r.m.s., or more if closer miking is used. Frequency response of this microphone is restricted to communications bandwidth, 300 Hz to 4 kHz, thus simplifying the task of filtering out unwanted signals at the modulator input.

The compressor consists of a gain-controlled amplifier, peak detector, and d.c. amplifier in a feedback loop, such that an increase in output results in a change in the d.c. gain voltage tending to reduce the output.



As the gain control element, a junction f.e.t. used as a voltage-controlled resistor provides a wide control range. The minimum useful j.f.e.t. channel resistance occurs for $V_{gs}=0$, and is usually referred to as $r_{ds(on)}$. The value of $r_{ds(on)}$ is typically not more than a few hundred ohms, and when the gate-source junction is sufficiently reverse-biased, the channel is pinched off, and has a very high resistance, of megohms or above. Thus a control range of 60 dB should be feasible, provided that the amplifier's maximum gain is high enough.

The d.c. amplifier in the feedback loop compares the peak detector's output with a d.c. reference voltage, V_{ref} . This reference voltage will thus define the peak output voltage from the compressor, since if $e_o(pk)$ exceeds V_{ref} , the gain of the controlled amplifier will drop. Simpler compressor designs omit the d.c. amplifier, and feed the peak detector's output straight back to the VCA. The result is that the output level increases appreciably with increasing input, but at a much reduced rate. The increase in output is necessary to provide a larger control voltage for the VCA. If a d.c. amplifier is included in the feedback loop, the increase in output for a given change in e_{in} is reduced by the d.c. gain, and thus is usually negligible.



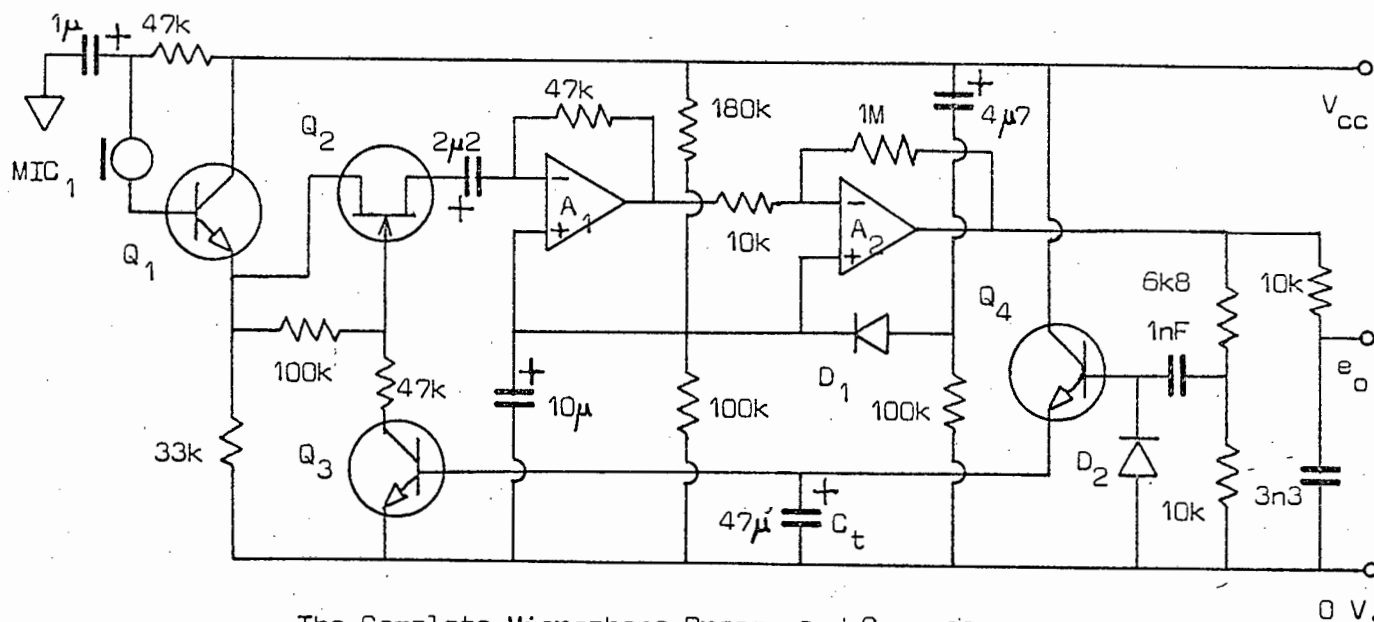
This simple circuit shows some of the characteristics and problems of a typical f.e.t.-type compressor. The peak detector has a fast attack time (limited only by the amplifier output current) and a slow decay, with time constant $RC = 0,5$ second. Thus large-amplitude transients are clamped by the compressor after one or two cycles, while the long recovery time ensures that the gain does not rise to its maximum value during short breaks in speech. Short recovery times give rise to an unpleasant "breathing" effect, as the background noise level increases rapidly during breaks in speech, and is then suppressed by the start of the next word. The time constant required depends on the amount of compression available, and for high-gain systems may be several seconds.

Two problems result from the use of a f.e.t. as the control element. The input resistance of the amplifier is R_1 in parallel with the f.e.t. channel resistance and this last varies over a wide range, so that under some conditions the input resistance will be unacceptably low. Also, for the f.e.t. to act as an ohmic resistor, the voltage across the channel must be small, typically ± 100 mV or less. Larger values of V_{ds} lead to severe distortion, and cannot be tolerated. This effect is particularly noticeable for r_{ds} large, when e_{in} is large and the gain of the VCA is low. The cure used in the simple compressor was to place R_1 in parallel with Q_1 , so that at high values of e_{in} , the peak detector would lose control, and e_o would rise with e_{in} , but would be undistorted. The range of V_{ds} for which r_{ds} is linear can be increased by allowing a fraction of the a.c. component of V_{ds} to appear across the gate-source junction. For $v_{gs} = \frac{1}{2}v_{ds}$, the useful range of v_{ds} is approximately doubled.

The signal amplifier in the compressor determines the compression range available. If the gain required for the maximum allowable value of input voltage to drive the output to limiting A_x , and the amplifier's maximum gain at the operating frequency is A_y , then the range of input signals for which e_o is constant is A_y/A_x . Thus for a large compression range, A_y must be large relative to A_x . It is

/convenient.....

convenient to implement the signal amplifier with an operational amplifier, but the resulting restricted gain-bandwidth product may make two stages necessary.



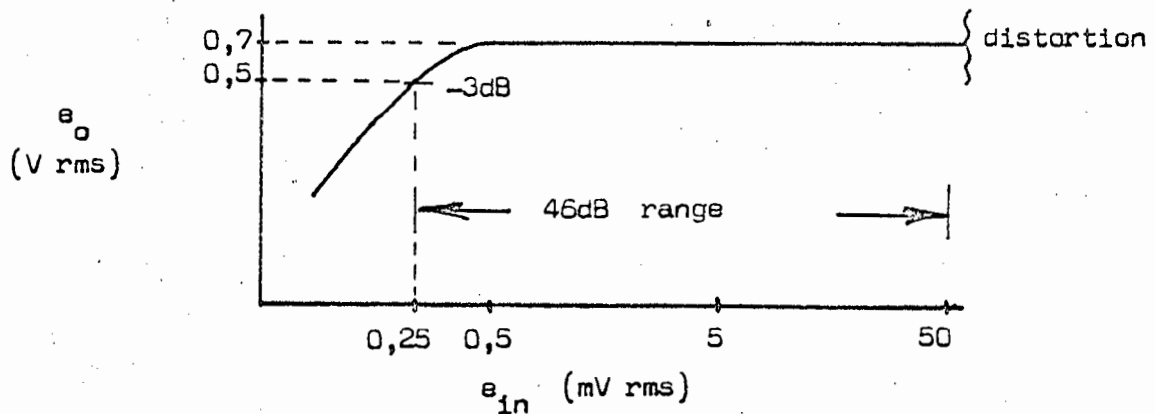
The Complete Microphone Preamp and Compressor.

A_1, A_2 : each $\frac{1}{2}$ of LM358N dual op amp. Q_1, Q_3, Q_4 : BC109C Q_2 : E202
 D_1, D_2 : 1N914 MIC₁ : SM1901B

The microphone is buffered by Q_1 , which draws its base current through the microphone, to reduce component count. An R-C network removes supply ripple at Q_1 's base, while Q_1 's emitter is close to V_{CC} to allow a large range of V_{gs} for Q_2 . One third of the input voltage appears across the gate-source junction of Q_2 , to increase the maximum usable value of e_{in} . The signal gain is provided by a dual low-power operational amplifier, and for $r_{ds(on)} = 500$ ohms, the combined maximum gain is 80 dB. Since the output voltage range of the LM358 is from 0 V to $(V_{CC} - 1.5V)$, the inputs of A_1 and A_2 are biased at $0.36V_{CC}$ to ensure sufficient output voltage swing. C_2 the op amp bias supply, while C_1 and D_1 provide a fast turn-on for the bias voltage, by allowing the voltage on C_2 to rise to $(V_{CC} - 0.6)/3$ at the moment of switching on. The time taken for C_2 to charge up through the bias resistor network would be of the order of seconds, while this modification puts the op amps in the

/active.....

active region after less than a millisecond. The peak detector comprising D_2 , Q_4 and C_t sets the output voltage to 2 volts peak-to-peak, and the d.c. gain is provided by Q_3 , the base-emitter voltage of which constitutes the reference voltage, V_{ref} . The compressor attack time is not more than one cycle of the input voltage waveform, and the decay time is variable, depending on the preceding input level, but sufficiently slow as to be unnoticeable on normal speech.



Compressor Gain Characteristic.

The output of the compressor goes via a 5 kHz lowpass filter (10k and 3.3 nF) to the modulator, which itself has a lowpass characteristic with f_{-3dB} at 3.4 kHz. High frequencies which might result in sidebands outside the receiver bandpass characteristic are thus heavily attenuated.

As usual, the variability of f.e.t. V_{gs} creates design problems. To cater for the expected range of pinch-off voltages, the input buffer was biased at as high a voltage as possible. This results in a minimum voltage at Q_1 's emitter, for $V_{cc} = 4.5V$, of about 3.8 volts. The minimum output voltage of the d.c. amplifier Q_3 is very nearly zero, with the result that the maximum usable value of $V_{gs}(\text{off})$ is $3.8 \times 100/47$, or approximately 2.6V. A wider range of $V_{gs}(\text{off})$ can be accommodated at the expense of large-signal handling capacity by shorting out the 47 kilohm /resistor.....

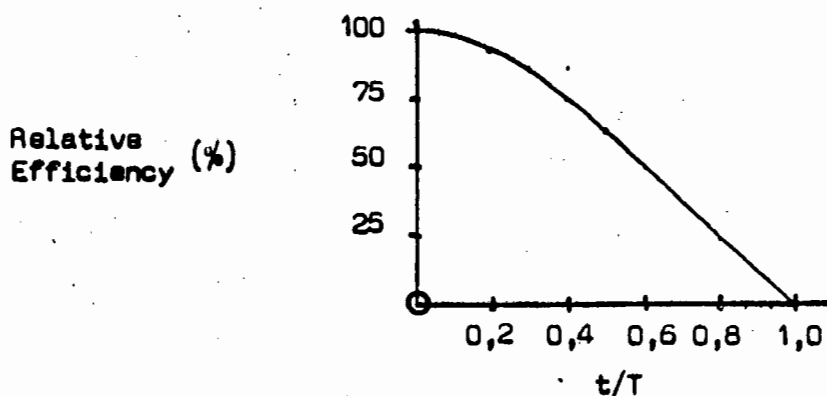
resistor at the collector of Q_3 . The maximum value of $V_{gs}(\text{off})$ would then be 3,8V, but the maximum input voltage would drop to about 25 mV.

Power consumption in the compressor was minimised by choice of components. An earlier version of the design used the LM3900 quad Norton op amp to provide the gain stages and the d.c. amplifier, but when the LM358 became available, it enabled a considerable reduction in the compressor current drain, to a total of 1 mA, compared with the previous figure of 7,5 mA.

14. Transmitter Pulse Shaper.

The output of the VCO is an approximate sinewave of 2,7V p-p. It is desired to produce a large optical output at 450 kHz with the minimum power consumption and quantity of electronic components. As has been shown in an earlier section, the efficiency of the output stage is maximum for low conduction angle class C or low duty cycle pulsed operation. Since class C has several disadvantages, such as higher power dissipation in the driving transistor, when used to drive a diode load, the pulsed drive technique was preferred. Some means of deriving a fixed, low duty cycle pulse train from a sinusoidal input is required. It is also necessary to determine what duty cycle will give a good compromise between transistor switching speed and current rating. For optimal efficiency, the duty cycle must be very low, but to achieve a useful output at the fundamental frequency with such a waveform, the peak current must be very high. Thus a very fast-switching transistor is needed to produce the low duty cycle, while the same device is expected to carry a current of the order of amps. Such performance is possible only with expensive, exotic transistors, or by avalanche mode operation, which requires the generation of a substantial high-voltage supply, with its attendant loss of power. If one accepts a higher duty cycle of the pulsed waveform, then the peak current drops, and the switching speed requirement is less severe, because the pulse duration is greater. Since the ratio of peak fundamental
/frequency....

frequency current to average d.c. current is a $\sin(x)/x$ function of the duty cycle (x = half of the conduction angle), it can be shown that at a duty cycle of 20%, the efficiency is 93,5% of that at a zero duty cycle.



Because of the $\sin(x)/x$ behaviour, the efficiency falls slowly as the duty cycle increases from zero, but at an increasing rate. Thus a 30% duty cycle results in 86% efficiency, and a 50% duty cycle gives 63,5% efficiency. There is little advantage to be gained from the use of a duty cycle much smaller than 20%, and since this corresponds to a manageable pulse duration of 440 ns, it was chosen as the design value.

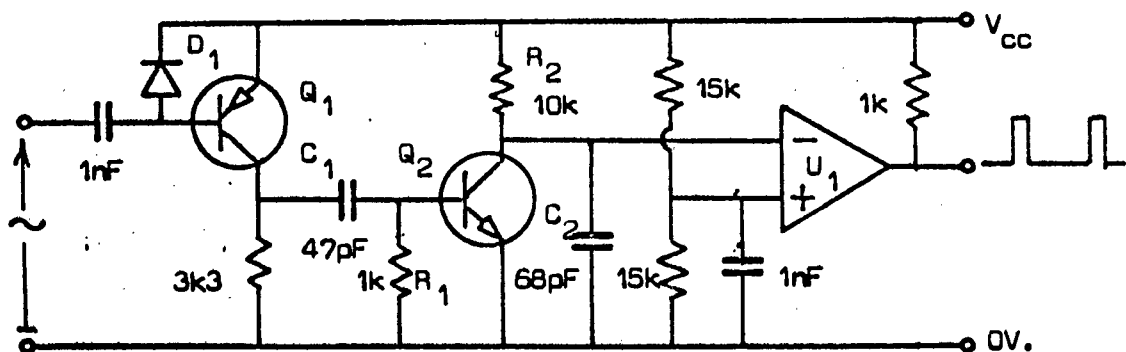
The peak current in the l.e.d. will be five times the average value in this case, so that it is necessary to avoid the possibility that the l.e.d. driver should remain permanently on, as destruction of the l.e.d. is then likely. The maximum continuous rating of the ME7124 is 100mA, so that a safe value for the average l.e.d. current should be in the region of 60 to 70 mA. The devices in the early breadboarded versions of the transmitter have run for considerable periods at these current levels, although long tests at higher currents were not attempted. It was felt that a relatively small gain in range would not be justified by a probable decrease in reliability. With 70 mA of average current, the peak current will be 350 mA, and the r.m.s. fundamental frequency current will be 93mA. Since the peak current is less than half an amp, one can use a relatively small-geometry, and hence high-speed, transistor. Devices with gain-bandwidth products of hundreds

/of.....

of megahertz, and saturated switching times of tens of nanoseconds, can be considered, and are available at low cost. While the problem of switching 350 mA, 440 ns pulses is not an easy one, it is much easier than that posed by marginally more efficient 3,5 A, 44 ns pulses.

The task of the pulse shaper is now well defined. It must take 2,7 volt peak-to-peak sinewaves at 450 kHz and give 20% duty cycle pulses (440 ns long) which can drive an output stage that handles 350 milliamp peak currents. The duty cycle must be independent of supply voltage, and, if possible, of input voltage. It need not be independent of frequency, since the changes in frequency are small. The fact that a 20% duty cycle implies a 440 ns pulse duration suggests the possibility of a monostable as the pulse shaper. The only objections to this are practical ones. It was not possible to find or to create a monostable, discrete or i.c., which would produce a sub-microsecond pulse with supply-independent duration, and yet draw only a few milliamps from a six volt supply.

A rather unusual circuit was designed which would produce the desired pulse duration, based on the fact that the capacitor voltage in an R-C network, connected across the supply, is a fixed fraction of the supply voltage at a given time after power is applied, irrespective of the magnitude of the supply voltage. A comparator was used to compare a fraction of V_{CC} with the voltage on a capacitor which was completely discharged at the start of each cycle. The result is a pulse train with fixed pulse duration at the comparator output.



A Workable Pulse Shaper Design.

Q_1 : BC263 Q_2 : BC109 D_1 : 1N914 U_1 : LM311N comparator

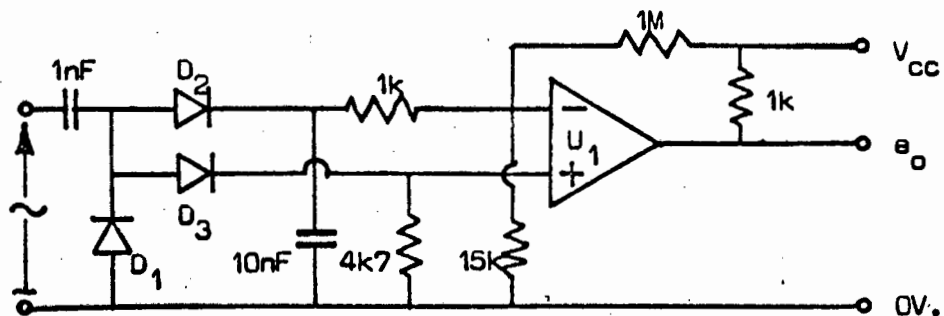
Q_1 squares the VCO output, and drives a differentiator, R_1, C_1 , which turns on Q_2 very briefly on the upstroke of each pulse. The use of a PNP device for Q_1 ensures a fast rising edge, since it can source large currents without drawing a large standing current. If the voltage on the timing capacitor C_2 is V_2 , then:

$$V_2 = V_{CC}(1 - e^{-t/R_2 C_2})$$

At the end of the pulse $V_2 = \frac{1}{2}V_{CC}$, so that $R_2 C_2 = \frac{-440 \times 10^{-9}}{\ln 0,5} = 640 \text{ ns.}$

Hence $R_2 = 10\text{k}$ and $C_2 = 68 \text{ pF.}$

This circuit works well, with pulse duration changing by only 2% over the supply voltage range. However, another odd circuit was designed in which the parts count and the power consumption were reduced, and since its only flaw is not serious in this application it is this latter design which is preferred.



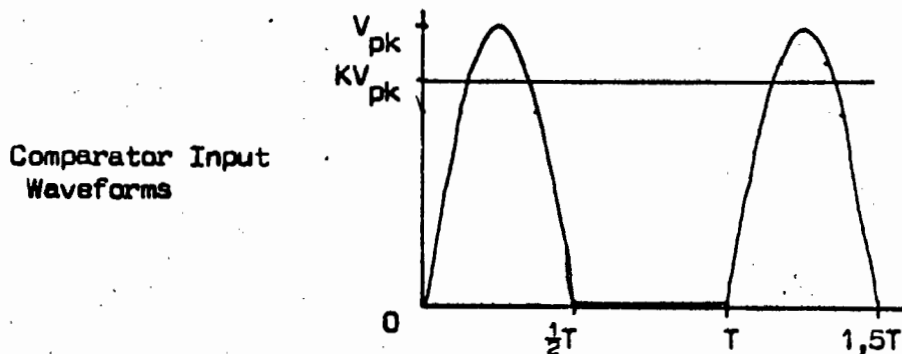
Final Pulse Shaper Design.

$D_{1-3} : 1N914 \quad U_1 : \text{LM311N comparator.}$

The incoming signal, which must be sinusoidal, can be of any amplitude greater than 1,8V p-p, and any frequency above 200 kHz. It is clamped to ground by D_1 , and rectified by D_2 and D_3 . The comparator compares the unsmoothed half-wave rectified sinusoid with a fraction of its peak voltage. When the unsmoothed waveform exceeds the smoothed voltage, a fixed duty-cycle pulse results. If the diodes were ideal (i.e. $V_f=0$), this would be exactly true for all input voltages, but even with real diodes, at the duty cycle required, the errors due to amplitude changes are of

/the.....

the order of 10% per volt, and can thus be ignored.



In the ideal diode case, for a 20% duty cycle, $K = 0,81$. With real diodes, and impure sinewaves, such as those produced by the VCO, the required ratio is best found empirically. The 1 megohm resistor shown connected between V_{cc} and the inverting input of the comparator is a fail-safe device. It ensures that when no input signal is present, the comparator's output stays low, thus protecting the output stage.

This circuit gives a duty cycle that is entirely independent of supply voltage, and which has very small sensitivity to input voltage fluctuations in the vicinity of the VCO output voltage (2,4V p-p). The only significant power consumed by the circuit is that due to the comparator, which must drive the output stage.

15. The Output Stage.

The performance of the transmitter output stage is one of the critical factors in the system design. An efficient output stage increases the effective optical output power and minimises battery drain, and hence the size of the battery pack. The l.e.d. itself has a considerable (500 ns) light risetime, with the result that the fundamental frequency optical output at 450 kHz is 1,5 dB below the low-frequency value. The l.e.d. driver should have rise- and fall-times appreciably shorter than 500 ns. to prevent further reduction of the output. A figure of 100 ns for the driver t_r and t_f should be sufficient, while lower values are not undesirable

/The....

The optical output should be independent of supply voltage, or the range of the system may be appreciably reduced at low battery voltages. This can be ensured by the use of a switched constant-current l.e.d. driver. However, the high source impedance of such a driver results in another difficulty. The l.e.d. junction capacitance, C_d , which is large (hundreds of pF), is charged to the diode forward voltage V_f during the current pulse. At the end of the pulse, the l.e.d. is isolated as the current source turns off, and the light output decays slowly due to the stored charge on C_d . If the light output is not zero at the start of the next pulse, the fundamental frequency output will be reduced. There are two solutions to the problem: the use of an active device to short out the diode during the time between pulses, or provision of a resistive path for the junction charge. The latter method is by far the simpler, and is adequate for the speeds required in this case. At the end of the current pulse, $V_f = 1,8V$ (at $I_f = 350 \text{ mA}$), and the voltage then decays exponentially. At $V_f = 1,1V$, the l.e.d. current has dropped to about 1% of the peak value, and the time taken to reach this level is given by:

$$\begin{aligned} t_{\text{off}} &= -RC_d \ln(1,1/1,8) \\ &= 0,49 RC_d \end{aligned}$$

Choosing a value for the parallel resistor R of 470 ohms results in about 1% of the pulse current flowing in R ($I_R = V_f/R = 3,8\text{mA}$). The current decay time t_{off} can be as high as 150 ns and still degrade the light faltime by less than 5%. Thus C_d can be as large as 600 pF (it is not specified by the manufacturer), before loss of output power becomes significant.

The large peak current and high di/dt values in the l.e.d. driver circuit will tend to produce considerable ripple on the power supply. A fresh four-cell battery pack will have a series resistance of between 1,2 and 2 ohms, resulting in ripple of up to 700 mV peak-to-peak, without decoupling. With partly discharged batteries, the situation is worse. A current of 350 mA for 440 ns corresponds to a

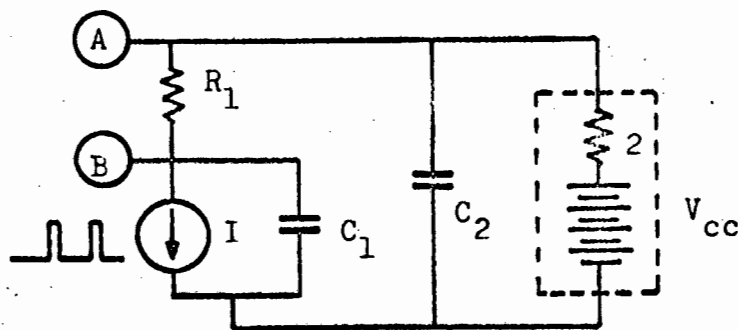
/charge.....

charge of $1,5 \times 10^{-7}$ Coulomb. If this charge is supplied by a capacitor connected in parallel with the battery pack, and the maximum acceptable ripple is 50 mV p-p, then the capacitor value is given by:

$$C = Q/V_{\text{ripple}}$$

$$= 3 \text{ microfarad}$$

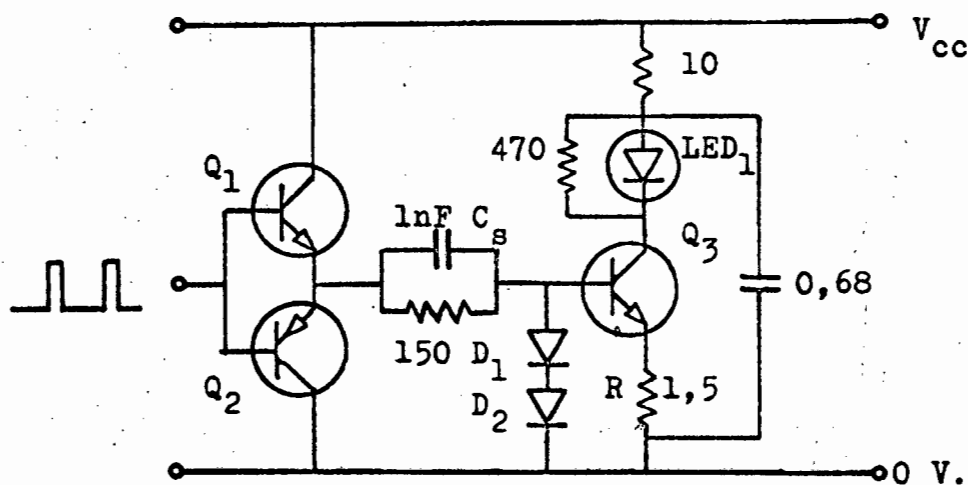
Use of a 33 microfarad tantalum bead electrolytic capacitor would then reduce the ripple to a negligible 5 mV. But this is in fact not so, because of appreciable series resistance and inductance in tantalum capacitors. A non-electrolytic capacitor of 3 microfarads is unacceptably large and expensive, so an alternative solution is required. An R-C network can be used to decouple the l.e.d. circuit from the supply. Most of the charge required during each cycle is supplied by the decoupling capacitor and this capacitor will be recharged between pulses through the resistor, so that the current drain from the battery is nearly constant.



With minimum V_{CC} of 4,5V and l.e.d. V_f of 1,8 V, and an allowance of 2 V for the l.e.d. driver, one can accept a loss of 0,7 V across R_1 , which thus becomes 10 ohms. Since the l.e.d. is current-driven, the ripple level at B can be fairly large. Some smoothing is required to remove inductive di/dt spikes, and that is provided by C_1 , which must be a suitable low-inductance r.f. capacitor. If the current risetime is 50 ns, di/dt is 7 amps per microsecond, and a 10 nH inductance will generate a 70 mV spike. Thus layout is an important factor in /determining...

determining the ripple produced, and care must be taken to ensure that the pulse current does not flow in the ground return line of the rest of the system. To fulfil the contradictory requirements for C_1 of large capacitance, small size and cost, and low inductance, a stacked-foil construction polycarbonate capacitor of 0,68 microfarad was used. This results in a ripple voltage at B of about 180 mV p-p, which is quite acceptable. The ripple waveform will be essentially triangular, with inductive spiking at the peaks. R_1 and C_2 form a lowpass filter for this ripple, and if C_2 is a 33 microfarad tantalum, then the ripple at A will be of the order of millivolts or less.

One can now consider the problem of creating a constant current source or sink which can be turned on and off very rapidly, which does not require a large controlling current, and which will operate with 2 V across it. It is convenient to use a current sink, rather than a source, as the drive for it can then be ground-referenced. While non-saturated switching logic in the control stage offers the highest speeds, the lower power consumption of a saturated switch in a low duty cycle application is more important. A simple switched current sink can be constructed with one NPN transistor as the control element, two diodes to provide a reference voltage, and a resistor to set up the current. To reduce the drive current requirement, a complementary emitter follower buffer can be used.



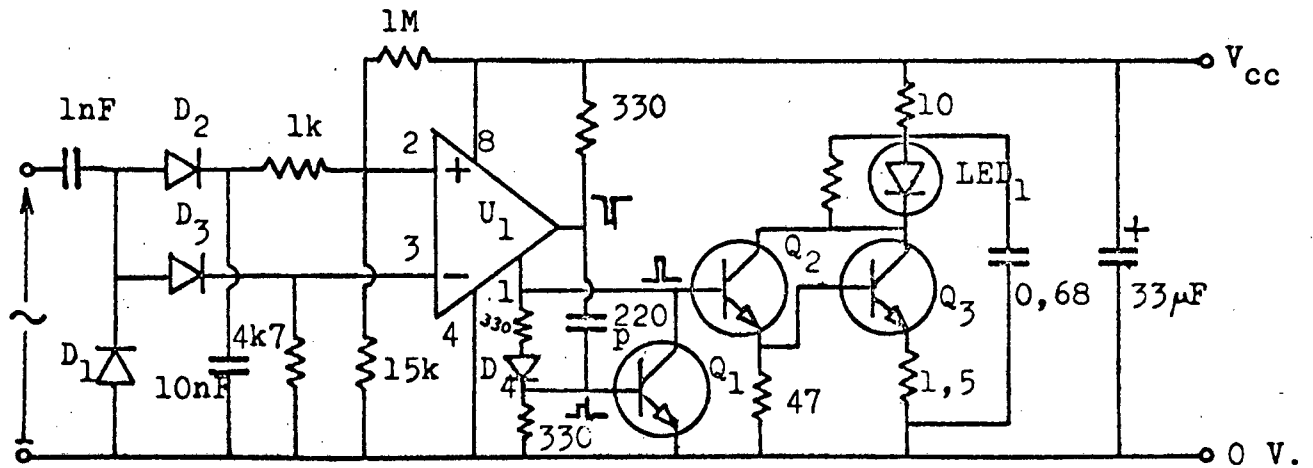
Simple L.e.d. Driver.

Q_1 : BC109 Q_2 : BC263 Q_3 : 2N2219A D_1, D_2 : 1N914 LED $_1$: ME7124

The diodes clamp Q_3 's base at about 1,3 V, and at $I_C = 350\text{mA}$, Q_3 's base-emitter voltage is 0,8 V, so that 0,5 V appears across R, to set the current level. The current will be dependent on temperature, as a result of the -2mV/deg. temperature coefficient of the silicon junction voltage. Thus a ± 25 degree temperature range will produce a $\pm 10\%$ current change, a spread which is acceptable. The capacitor C_s speeds up the turn-on and turn-off times of Q_3 , to between 50 and 70 ns. The drive current drain is 3,4 mA, and the output current changes by about 5% over the supply voltage range. When combined with the duty cycle shaping circuit, the total "wasted" current drain is 11,5 mA, and the performance is satisfactory. However, means were sought to reduce the power consumption, and possibly the number of components as well.

A considerable waste of current occurs at the comparator output, where an average current of about 4,8 mA flows in the 1 kilohm pull-down resistor, because of the low duty cycle. The LM311 can be used to drive a ground-referred load, although the positive voltage swing is then restricted, and at $V_{CC} = 4,5\text{ V}$, the high state output is only 2,5 V. This results in a better than 50% reduction in the current drawn by the LM311. But the limited voltage swing prevents the use of the emitter follower drivers, and then the lack of an active turn-off for Q_3 results in a long storage time and a slow falling edge for the current waveform. An idea was borrowed from an output stage design that had been rejected as too complex, and the two diodes were replaced by a diode and a transistor. By making available both outputs of the LM311, an inverted-logic pulse is produced which can be capacitor coupled to the base of this transistor, to turn it on at the end of the current pulse and so extract Q_3 's base charge, ensuring a fast turn-off.

Power was still being wasted in the base current required by Q_3 . Replacing Q_3 by a Darlington pair reduced the drive current by almost two orders of magnitude, without appreciable performance degradation.



The Pulse Shaper and L.e.d. Driver.

D_{1-4} : N914 Q_1, Q_2 : BC109 Q_3 : 2N2219A LED₁: ME7124 U_1 : LM311N

The use of U_1 to drive a ground-referred load has necessitated a polarity reversal at its inputs. Q_1 performs the functions of voltage reference and turn-off current sink. D_4 prevents the positive-going pulse occurring at Q_1 's base at the end of the l.e.d. current pulse from being coupled through to Q_2 's base. The collector current of Q_2 flows in the l.e.d., so that the wastage of drive current is minimised. The current waveform, monitored at the emitter of Q_3 , shows rise- and falltimes of less than 50 ns, even at $V_{cc} = 4,5V$. Total supply current drain of the above circuit, excluding the l.e.d. current is about 4,5 mA at $V_{cc} = 6V$, when average l.e.d. current is 74 mA, and drops to 3,4 mA at $V_{cc} = 4,5 V$, when $I_{led} = 68mA$. Thus the optical output will drop by 8 % when the power supply drops 25%, which is no better than acceptable. The efficiency on the other hand is good. At $V_{cc} = 6V$, with the average l.e.d. current 74 mA, the rest of the transmitter will draw less than 11 mA. This compares very favourably with the previous breadboarded version, with similar performance, where the two current drains were 76 mA for the l.e.d. and 32 mA for the rest of the circuit. The new design reduces power consumption by about 20%, without sacrificing output power.

16. The Transmitter Lens.

Although the Monsanto ME7124 l.e.d. has a remarkably narrow emitting angle ($\theta_1 = 6^\circ$), it is not sufficiently directional for use in a moderate-range

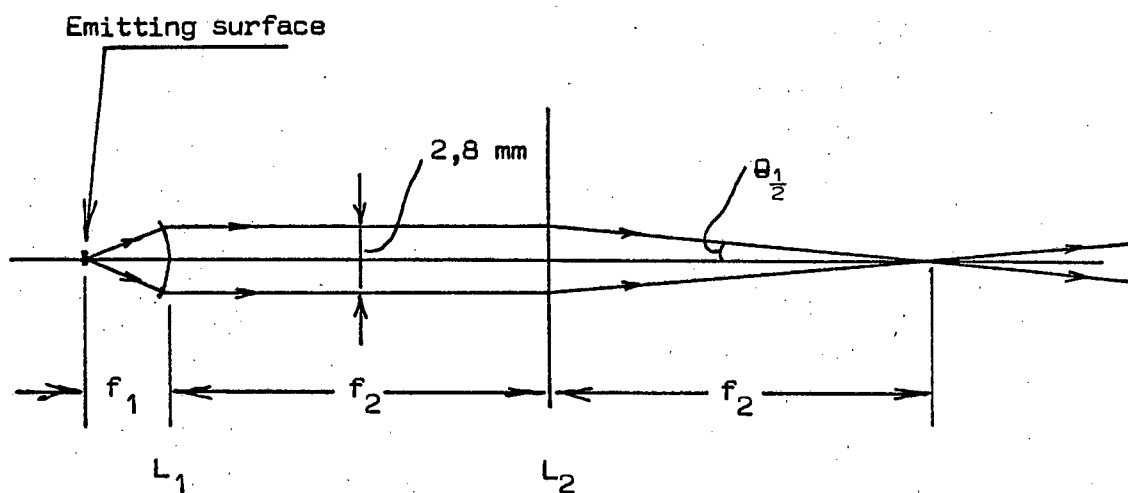
/voice...

voice link. However, it does enable the use of a small-aperture lens as a collimator, without severe losses of optical power. A value for the transmitted beam half-angle of between 1 and 2 degrees should provide a good compromise between the conflicting requirements of high radiant intensity ($\theta_{\frac{1}{2}}$ small) and ease of aiming ($\theta_{\frac{1}{2}}$ large). To select the lens which will produce such a beam angle, one must know more about the emitter and its epoxy lens. The GaAs emitting area appears to be a square flat surface, about 0,6 mm on a side, and it is positioned at the focus of a lens formed by the epoxy encapsulation, of focal length 4 mm and diameter 2,8mm. The beam width is given by:

$$\theta = \arctan \frac{\frac{1}{2}(\text{length of side})}{\text{focal length}}$$

$$= 4,3^\circ$$

Since the emitting surface is in the focal plane, rays from any point on the surface emerge parallel from the epoxy lens. Thus radiation from the centre of the emitter emerges as a "bundle", 2,8 mm in diameter, parallel to the lens axis. If the focus of the collimating lens coincides with the front face of the epoxy lens, this paraxial radiation will be brought to the far focus of the external lens, from which point it will diverge once more. The angle of divergence is the required beam angle, $\theta_{\frac{1}{2}}$.



The radiation from the epoxy lens L_1 , which is not parallel to the axis is collimated by the external lens L_2 , and emerges from L_2 off-axis, but parallel to the axial radiation, so that the total beam angle is unchanged. If the l.e.d.

/is....

is moved inside the focus of L_2 , then the beam angle will increase, since the off-axis radiation will no longer be completely collimated by L_2 . Thus if a range of beam angles is required, it is necessary to design for the smallest value of θ_1 , so that defocusing can provide larger values. In this case, the minimum value of θ_1 will be 1° .

$$\theta_1 = \arctan \left(\frac{L_1 \text{ radius}}{f_2} \right)$$

Hence
$$f_2 = \frac{L_1 \text{ radius}}{\tan \theta_1}$$

$$= 80 \text{ mm.}$$

The diameter and F number of the external lens are determined by the polar characteristics of the emitter. In this case, the half-power points are at $\pm 4,3^\circ$, but a considerable fraction of the total radiated power is emitted outside this cone. Thus a compromise is necessary between the need to capture as much of the radiated power as possible and the desire to reduce the cost and size of the external lens. If the angle subtended at the emitter by L_2 is made twice the half-power angle, most of the emitted power should be collected, while the diameter of L_2 is acceptably small.

$$\begin{aligned} \text{Diameter of } L_2 &= 4f_2 \tan 4,3^\circ \\ &= 25 \text{ mm.} \end{aligned}$$

Since a plastic lens from a local manufacturer's standard product line had characteristics very close to the design values, it was the obvious choice, since neither focal length nor diameter are critical in this application. The actual values of f_2 and diameter are 82 mm and 30 mm respectively. Thus the external lens intercepts a cone of half-angle = $(15/82) = 10,4^\circ$. When one attempts to calculate how much radiated power is captured by the lens, one again faces the problem of insufficient and confusing data. Monsanto quote a "Total External Output Power" for the ME7124 of 3mW at $I_f = 50 \text{ mA}$. But they also state that on-axis intensity is 243,6 mW/sr, and that "into cone @ $\frac{1}{2}$ power points @ $I_f = 50 \text{ mA}$ ROP = 3 mW." This last statement can be taken as meaning that the emitted power into

/the...

the cone of half-angle 4° is 3 mW. Since there is considerable emission outside this cone, the total output power would then be expected to be greater than 3 mW. This assumption is reinforced by calculations based on the polar characteristics provided. Although as a result of the narrow beam angle the graph is somewhat approximate, it can be used to evaluate the expression for emitted power into a given cone for known axial radiant intensity:

$$P(\theta) = 2\pi \int_0^\theta J(\theta) \sin\theta \, d\theta$$

For $\theta = 4^\circ$, and $J(\theta) = 243,6 \text{ mW/sr}$, the calculated emitted power is 2,5 mW, which is a fair approximation to the manufacturer's value. Thus one can evaluate the power collected by the external lens using this formula, for $\theta = 10^\circ$. The result is a figure of 6,2 mW at $I_f = 50 \text{ mA}$. Given the beam angle of the radiation from the external lens, being approximately 1 degree, one can calculate the new transmitted radiant intensity in mW/sr. The solid angle of emission is $\Omega_e = 2\pi(1 - \cos\theta) = 10^{-3} \text{ steradian}$, so that the radiant intensity is:

$$\begin{aligned} J_e &= P/\Omega_e \\ &= 6,5 \text{ W/sr at } I_f = 50 \text{ mA} \end{aligned}$$

Now the transmitter l.e.d. driver produces a current of 93 mA r.m.s. at 450 kHz. The l.e.d.'s light output is down by 1,5 dB at this frequency, and the lens transmission loss is about 10%, so that the actual transmitted radiant intensity at 450 kHz is:

$$J_e = 9,2 \text{ W/sr (at 20% duty cycle, } I_{dc} = 70 \text{ mA)}$$

Predicted Performance of the FM Communicator Design.

The calculated values of transmitted radiant intensity and of minimum receivable signal power can be used to predict the maximum range of the system under varying background radiation conditions. The ultimate maximum range occurs when the background radiation is zero. Operation at night in the absence of artificial lighting is a good approximation to this condition. The zero-background input noise current at the receiver frontend is calculated as 28 pA r.m.s. For FM reception, a minimum detector input S/N ratio of 12 dB is required, so that the minimum useable received signal current is 112 pA. The optical signal power needed to produce this current is $112/0,35 = 320$ pW. For an acceptance half-angle of 1° , the effective capture area of the receiving lens is $2,1 \text{ cm}^2$, so that the irradiance at the receiver must be $H_r(\text{min}) = 320/2,1 = 150 \text{ pW/cm}^2$. At a distance r from the transmitter, the signal irradiance is $H_t = J_e/r^2$. Hence the maximum range is:

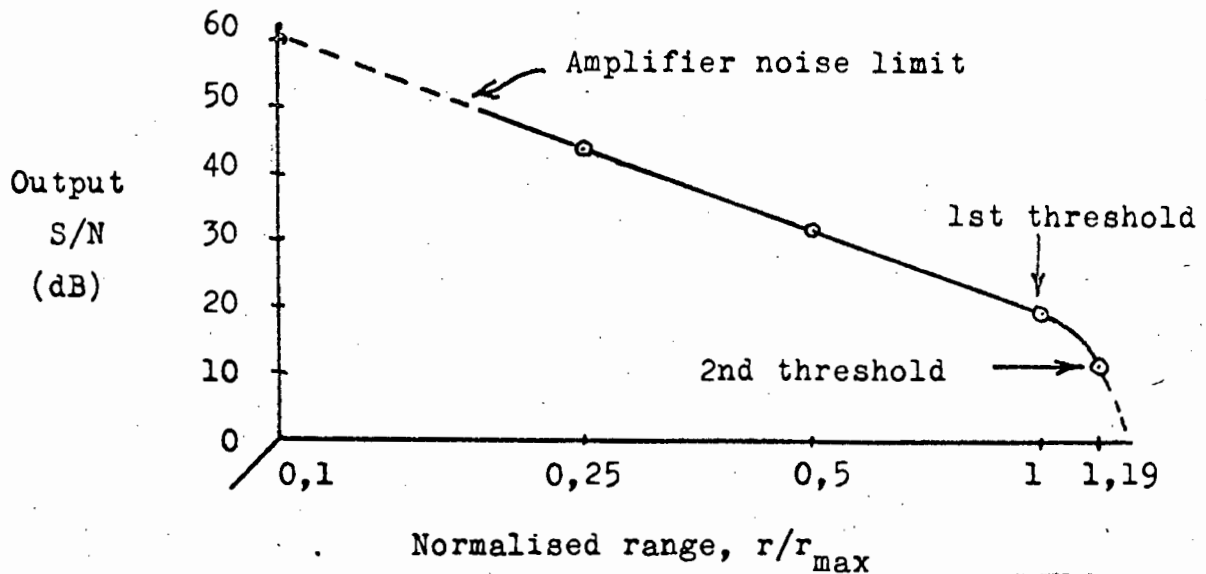
$$r_{\text{max}} = (J_e/H_r(\text{min}))^{1/2}$$

The value of J_e derived from the manufacturer's data of $9,2 \text{ W/sr}$ results in a maximum range in darkness of $2,5 \text{ km}$. In normal daylight, the sensitivity of the receiver is determined by the background radiation level. Using an infrared absorption-type filter, one can expect a typical background photocurrent of $0,5 \text{ mic-roamp}$. This results in an input noise current of 63 pA , corresponding to a value for $H_r(\text{min})$ of 340 pW/cm^2 . The maximum range is then $1,6 \text{ km}$. Without the filter, background photocurrent will be about five times greater. Thus $H_r(\text{min})$ will increase to 760 pW/cm^2 , and the range will be just over 1 km . It is worthy of note that in background radiation limited systems, the range is inversely proportional to the fourth root of the background radiation power.

As a measure of the quality of the voice link, one can plot the demodulator output S/N ratio against normalised range. A maximum S/N ratio improvement at the demodulator of 8 dB is used assuming normal voice modulation, producing an r.m.s.

/frequency.....

frequency deviation 25% of maximum. In practice, the compressor would be expected to improve on this figure.



If the modulation waveform is a single tone at 100% deviation the S/N ratio is improved by a further 9 dB at ranges less than $r_{\max}$. Range is very sensitive to beam angle. As an illustration of this, if the beam angle of both the transmitter and the receiver were doubled, to 2° , the value of $H_r(\min)$ would quadruple, and J_θ would fall to one quarter of its previous level. The range would thus also be reduced to 25% of its value at 1° beam angle.

Breadboard Testing.

The performance of the breadboarded version of the final design was tested and the results compared with the design values. Numerous discrepancies were found, some being small, and not worthy of mention, but others resulting in varying amounts of loss in performance. As has been shown, the two parameters which ultimately determine the range of the system are the minimum acceptable signal irradiance at the receiver and the emitted radiant intensity at the transmitter. Since the maximum usable range more than any other characteristic determines the suitability of the optical communications link for any given application, a check of these parameters is essential.

The value of $H_r(\min)$ is affected by several variables. Of these, the lens capture area is dependent on the detector size, which is well controlled, the lens focal length and aperture, also well defined, and the actual setting-up of the detector at the focus. This last is the only area of the lens performance which is likely to cause problems, and since it is merely a matter of providing an adjustment for the desired beam angle, it can be disregarded. Once the capture area is fixed, $H_r(\min)$ is determined by the photodiode sensitivity S , which is tolerably well controlled, and by the minimum input current needed by the frontend amplifier. Since the techniques used in the frontend design involved a number of approximations, and since the circuitry is quite complex, the frontend sensitivity may differ considerably from the design value.

To test the frontend, a current source was simulated by driving a calibrated sinewave voltage source (a Hewlett-Packard model 651B test oscillator) through a 1 megohm resistor. Thus 1 V r.m.s. from the oscillator gave 1 microamp of signal current at the frontend. In fact, these measurements revealed a value for the limiting sensitivity (demodulator input 3 dB below maximum) of 200 pA or more, instead of the 50 pA predicted. Since the voltage required at the detector input had been well defined in earlier tests, this meant that the overall transfer

/ resistance

resistance of the receiver was less than expected. To locate the source of the error, the a.c. voltage at the input to the LM1351's limiting amplifier was checked, and this revealed that the gain of the i.c. limiting amplifier was much less than expected from the National Semiconductor data. Instead of 65 dB, a gain of some 35 dB was measured. An explanation for this poor performance was sought, and one cause was found to be insufficient decoupling at pin 6 of the LM1351, used as a ground reference for the input. A larger decoupling capacitor cannot be used, since the pulsed supply operation requires a rapid turn-on of the gain stages, and the capacitor must be charged through a resistor of several kilohms. It was however found to be possible to use pin 5 of the i.c. as the reference point, with better results. A 6 dB improvement in gain due to this modification gave a limiting sensitivity of 100 pA at the input. Since the gain of the LM1351 is so much lower than the design value, one would have expected a greater difference between the predicted and measured values of sensitivity. However, the frontend transfer impedance is actually higher (by 8 dB at $V_{CC} = 6$ V) than the design $1,5 \times 10^6$ ohms, thus offsetting some of the loss in gain. In a search for the cause of the reduced gain of the LM1351, use was made of the original Motorola MC1351 data. This had clearly been the source of National's data for their second-sourced version. The Motorola data was considerably more detailed, and revealed what National had not, namely that the gain of the amplifiers drops from a typical 65 dB at 4,5 MHz to about 55 dB at 1 MHz. Gain at frequencies below 1 MHz is not specified, but extrapolation of the curve suggests a value of about 50 dB at 450 kHz. Since, in addition, these figures refer to operation at $V_{CC} = 12$ V, the measured value, though disappointing, is credible. The limiting value of demodulator input voltage at $V_{CC} = 12$ V is higher than the 150 mV p-p at 6 V, with the result that the specified limiting sensitivity at the i.c. input of 0,08 mV r.m.s. rises to only 0,4 mV r.m.s. at 450 kHz and 6 V.

The value for limiting sensitivity at the frontend of 100 pA r.m.s. does not in fact appreciably degrade the performance of the system, since the predicted

/ minimum

minimum input noise current of 28 pA r.m.s. is less than 12 dB below it. Thus there is sufficient sensitivity for operation under all background radiation conditions, although in darkness ($i_b = 0$) there is very little gain margin.

Once the sensitivity of the receiver had been established, it could be used to test the transmitter output power. In a simple bench-top test, both transmitter and receiver were used without lenses, with a known small separation (order of hundreds of millimetres) between and detector. As a first approximation, the photodiode (which at this stage was a Hewlett-Packard 5082-4207) was assumed to have a sensitivity of 0,3 A/W. Since the detector area was also known, one could calculate the irradiance required to produce a certain voltage, at, for example, the input to the LM1351. (It was found that when a probe was connected to the output of the LM1351 gain stages, the radiated power produced was often sufficient to send the receiver into oscillation. It was thus safer to use the input as a reference point in these tests.) Since the separation is known, one can find the transmitted radiant intensity:

$$J = \frac{i_s r^2}{A_d S} \text{ W/sr, where } \begin{aligned} i_s &= \text{photodiode signal current} \\ r &= \text{separation in mm} \\ A_d &= \text{photodiode area in mm}^2 \\ S &= \text{photodiode sensitivity in A/W} \end{aligned}$$

The tests were conducted using the design output stage, driven by a General Radio Model 1340 pulse generator, and this was adjusted to give the desired 20% duty cycle, and $I_{dc} = 70 \text{ mA}$. Based on the Monsanto data for the ME7124, one would have expected an axial radiant intensity at 450 kHz of about 380 mW/sr. In fact, the measured value was an alarming 20 mW/sr. The discrepancy between data and measurement was so large that one suspected a flaw in the measurement technique.

To remove any doubt on this issue, further tests were carried out using a calibrated radiation detector. This took the form of a Hewlett-Packard Model 8330A radiant flux meter and its associated Model 8334A radiant flux detector. The

detector is a thermopile, with an active area of 10mm^2 , covering wavelengths from 300 to 3000 nanometers. The meter indicates irradiance in W/cm^2 , down to 3 microwatts/ cm^2 full scale. The meter frequency response is very limited, so that one can only measure average values of irradiance. Thus the ME7124 was removed from circuit and driven with direct current. With a known emitter-detector separation, sufficiently large to ensure uniform irradiance of the detector (a matrix 4,3 mm square), one can measure the radiant intensity of the emitter. Use of a turntable enables off-axis measurements, so that the polar characteristics can be checked. Readings were taken on several devices, at different currents and emitter-detector separations. The results confirmed the measurements made earlier, in that at $I_f = 50\text{ mA}$, the axial radiant intensity, quoted by Monsanto as 243,5 mW/sr (four figure accuracy?), is in fact between 10 and 13 mW/sr .

The normalised polar characteristics of the device are substantially similar to those claimed by the manufacturer in the revised data sheet, with $\theta_1 = 6^\circ$. When the actual radiant intensity values are used to calculate the power output into a cone of given semiangle, however, the figures obtained are once again much lower than expected.

$$P(\theta) = 2\pi \int_0^\theta J(\theta) \sin\theta \, d\theta$$

$$P(4^\circ) = 0,1\text{ mW}$$

$$P(10^\circ) = 0,19\text{ mW}$$

$$P(20^\circ) = 0,29\text{ mW}$$

Since the radiant intensity at 20° off axis is only 0,5 mW/sr (about 5% of the axial value), and that at 45° is unmeasurable with the test gear used, one can assume that the total output power will not be much greater than that enclosed by the cone of 20° semiangle. Thus the total power is also well below minimum specified value, which is 1 mW at $I_f = 50\text{ mA}$. The power emitted into a cone of 10° half-angle is that which will be captured by the external lens of the transmitter. This power is focused into a beam of 1° semiangle, so that the radiant intensity in such a beam at $I_f = 50\text{ mA}$ is 195 mW/sr . In the actual transmitter, allowing for a higher average current and for a loss of about 10% in the lens, the radiant

intensity component at 450 kHz is 275 mW/sr. The value predicted from the manufacturer's data was 9,2 W/sr, some 33 times higher. The predicted range in daylight, without the use of an absorption filter, was 1 kilometre, so that the new value of J reduces this to about 170 metres, if no other problems are encountered.

The difficulty posed by these out-of-specification devices is considerable. It appeared that the emitters tested had come from two separate batches, imported some months apart, so that the performance would seem to be typical, rather than an isolated case, or a batch that had escaped quality control. Some doubt is thus cast on Monsanto's measuring techniques or equipment. (The equipment used here is unlikely to be at fault, since two different sets of apparatus were used, and gave very similar results.) One suspects, from the fact that Monsanto's figures for axial radiant intensity for the ME7121 to 7124 family are all quoted to the nearest tenth of a milliwatt per steradian, with either three or four significant figures, that they were calculated rather than measured. Texas Instruments make a silicon-doped l.e.d., the TIL 31, with a specification that is broadly similar to that of the ME7124. It claims a typical 6mW total output at $I_f = 100$ mA, a beam semi-angle of 5° , and axial radiant intensity of 0,25W/sr. Texas do however quote the formula on which their radiant intensity figure is based. They have used:

$$\begin{aligned} J(0) &= P_{\text{total}} / \Omega_{\frac{1}{2}} \\ &= P_{\text{total}} / 2\pi(1 - \cos \theta_{\frac{1}{2}}) \end{aligned}$$

Use of the same formula with the Monsanto emitters does not produce consistent results, however, implying that Monsanto did not use the Texas formula. The question is whether this formula actually produces accurate results. One would not expect a high degree of accuracy, since it is based on the assumption that the errors due to two separate approximations cancel each other. The value of J is given as total emitted power divided by solid angle between half-intensity points. In fact, the power emitted into the solid angle $\Omega_{\frac{1}{2}}$ is appreciably less than P_{total} ,

and the intensity inside that solid angle is not uniform. The validity of the formula will be entirely dependent on the shape of the polar pattern of the emitter. A lack of time unfortunately prevented a comparison between the ME7124 and the TIL 31 being carried out. If the Texas Instruments device meets its specification, however, it will make an almost ideal replacement for the ME7124. It is likely to be more expensive, as a result of its glass and gold-plated metal encapsulation, and tighter general specifications, but should not be prohibitively so.

Other aspects of the transceiver performance were generally as expected. The microphone preamp and compressor, on which a considerable development effort had been expended, proved its efficacy by producing a continuous acoustic feedback when the link was being tested at short range (up to 15 metres) in the laboratory. At greater ranges, it became apparent that the quality of the received signal was quite adequate for a communications-bandwidth link. Aiming was found to be rather critical, but this was only to be expected. The squelch and power-saving circuits worked as expected, and the former provided a useful indication of aiming inaccuracy during reception of a signal, by cutting off the audio when the transmitter drifted out of the field of view. The audio power amplifier produced more than sufficient sound level prior to overload, and the current drains in the various operating modes were acceptably close to the expected values. It was felt that one would be justified in proceeding with the design and construction of a pair of prototypes, so that the communicator would exist not merely as an electronic design, coupled with expected performance figures, but as a manufacturable product.

Prototype design.

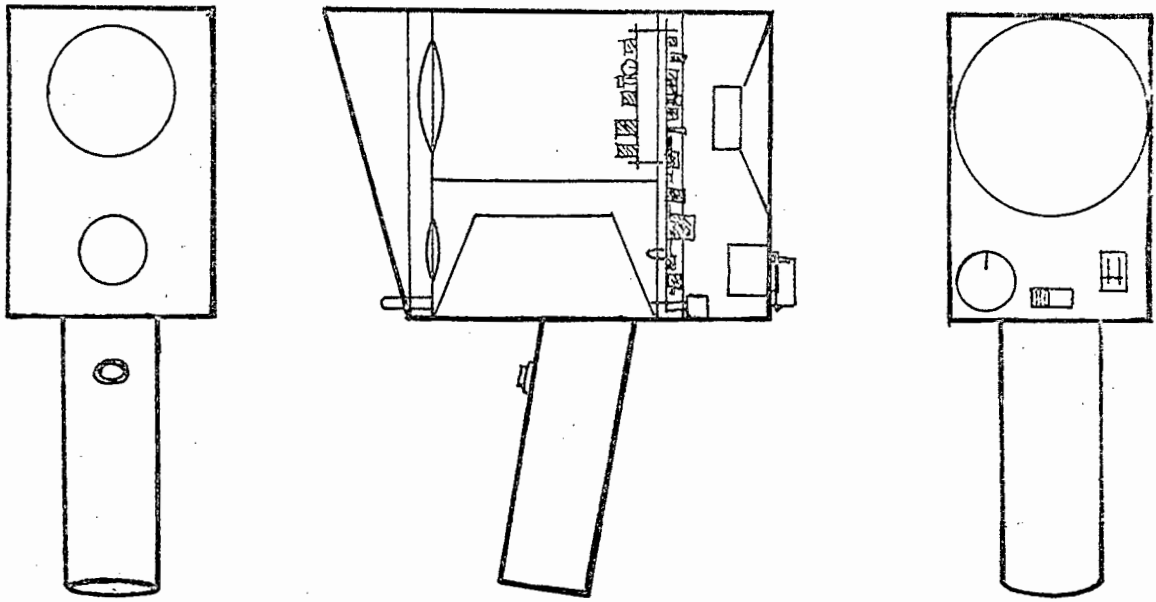
Experience at the breadboarding stage had shown that a considerable problem existed with layout-induced feedback in the high-frequency sections of the receiver. To avoid a recurrence of this in the final design, two steps were taken. Firstly, it was decided to mount all components on double-sided printed circuit boards, using a ground plane technique to minimise stray coupling effects. Secondly, the receiver frontend was to be physically separated from all other circuitry, in a screened box, on its own double-sided board. Its power supply would be derived from the common receiver supply, but would be decoupled with an R-C network (220 ohms and ten microfarads proved adequate), to prevent feedback through the power line.

The choice of a suitable mechanical configuration for the communicator was the first step in the design. Various layouts had been suggested: a helmet-mounted arrangement, with possibly a separate battery pack; a tube-shaped package to be held next to the head in the same way as similarly-shaped 27 MHz walkie-talkies; a gunstock mounted shape, and a cine-camera layout with pistol grip. The helmet-mounting idea is attractive, in that combined with VOX transmitter switching, it offers hands-off operation. However, it requires a broad transmit and receive beam angle, of the order of ten degrees, and that means both a higher order of transmitter power and the use of a large-area photodetector. These would raise the cost to unacceptable levels, so the head-mount is excluded. The tube-shaped package suffers from the same basic problem, namely, how to aim accurately when one cannot sight along the optical axis. One is left with two contenders, both capable of being aimed accurately. The gunstock-shaped design can probably be pointed with a higher degree of accuracy than the cine-camera, since a similar arrangement allows the taking of unblurred photographs with camera-lens combinations having a total field of view of only 6°. Thus aiming to within less than half a degree should be feasible. But the gunstock requires a fully-committed operator, since two hands must be used at all times. This and its considerable bulk combine to make it an unattractive proposition.

/The.....

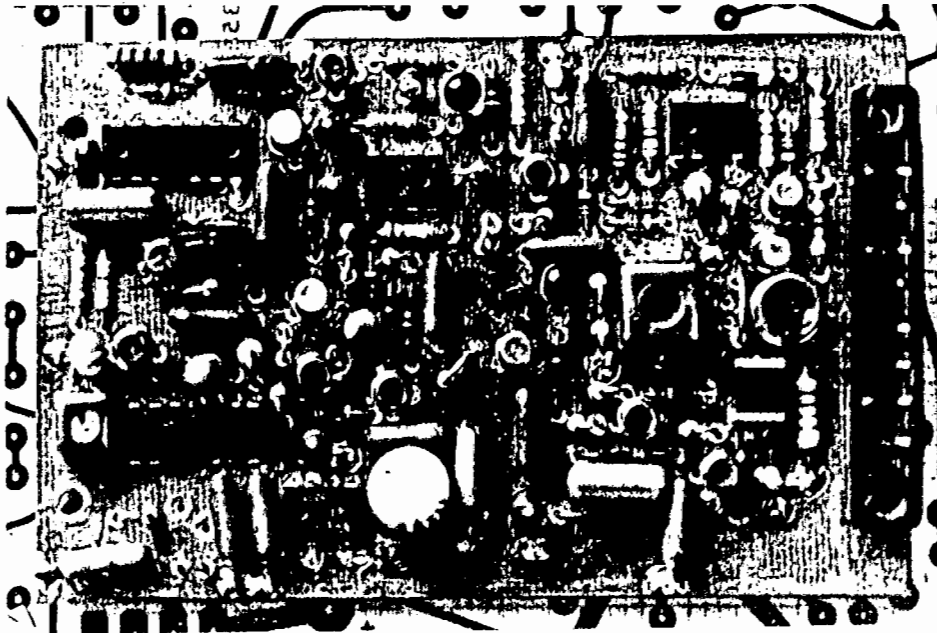
The remaining design, the 'cine-camera', is quite promising. It can be made fairly compact, can be aimed easily with a notch-and-post peepsight, and can be held steady and operated with one hand under most conditions. Two-handed operation can provide added stability when needed, for example, in windy conditions. Since two lenses are needed, they can be mounted one above the other to make a relatively narrow package. The front-to-back depth will be at least the focal length of the transmitter lens, plus sufficient room for the mounting of components. If AA (penlite) cells are used, the battery pack can be housed in the pistol grip, saving space and lowering the centre of gravity for better balance.

This defines the shape of the prototype fairly well. One still has to locate the various controls and the microphone and loudspeaker. Most of these can be positioned on the back panel, with the exception of a press-to-talk (PTT) switch on the pistol grip, for one-handed operation. The top and sides of the case can be extended forward of the lens panel to act as a lens hood, to exclude unwanted light from the receiving lens. The material to be used for these prototypes should be aluminium, being light and easy to work. One would however anticipate the probable use of plastic mouldings in a production version, as they would considerably reduce the unit cost, and further reduce the weight. Plastic moulding techniques could also be used to provide the means of mounting and of adjusting the focus of the lenses. These could be moulded integrally with a threaded lens mount, so that a corresponding screw thread on the body of the communicator would allow setting-up of the beam angles. On the prototype, this can be done by moving the circuit boards in relation to the body. For ease of servicing, the lenses and circuit boards would be assembled together on a common chassis, which would plug into the case of the communicator.



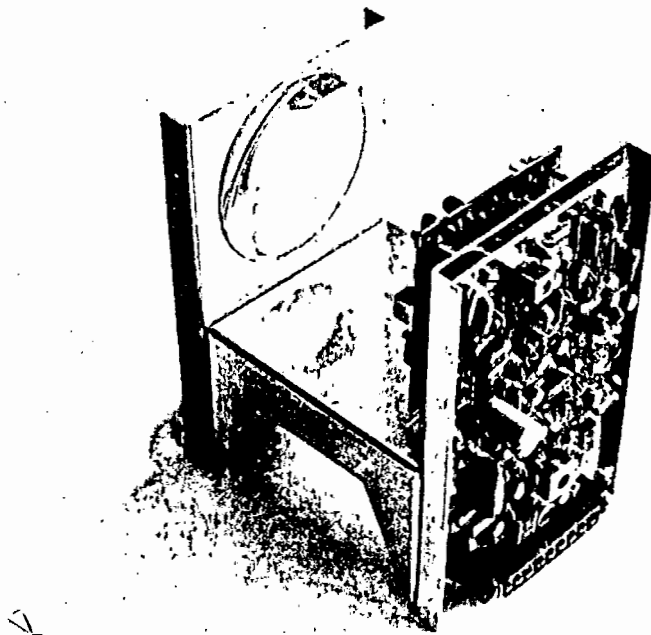
Prototype General Layout.

To keep costs down, one should keep the number of mechanical parts as small as possible. This is true in particular of circuit boards, where the greater the number of separate boards, the greater the number of interconnections. Thus it was decided to put all the transmitter components and all of the receiver components, except those in the frontend, on a single circuit board. The size of this board would determine the width and height of the communicator. Considerable care was taken with the layout of the main p.c.b., a satisfactory design being achieved at the third attempt. The receiver and transmitter sections are physically separated, the transmitting l.e.d. mounting on the reverse side of the board, facing the lenses.

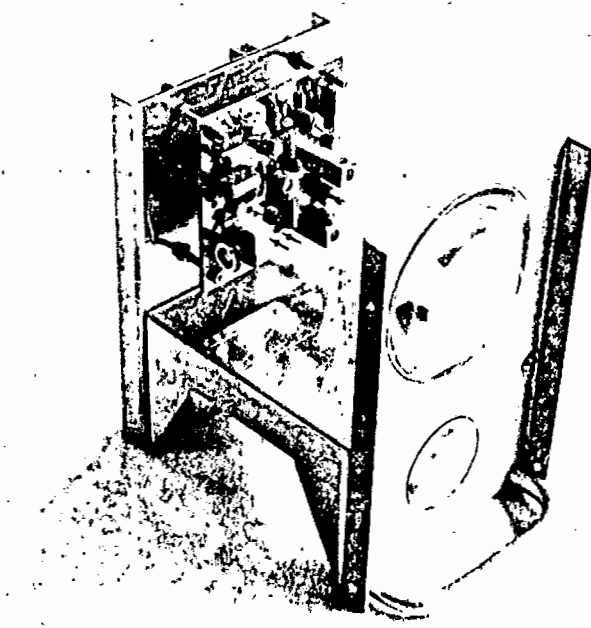


The Main Circuit Board.

The circuit board for the receiver frontend is considerably smaller and much simpler, with a lower component packing density. The main worry here was the possibility of feedback, so that the layout was kept relatively open. The photodiode is mounted onto the board with an insulating spacer, with the result that lead lengths in the critical input stage are short. All adjustments on both boards must be made with the plug-in chassis removed from the case, but care has been taken to make the preset Rs, Ls and Cs as accessible as possible.

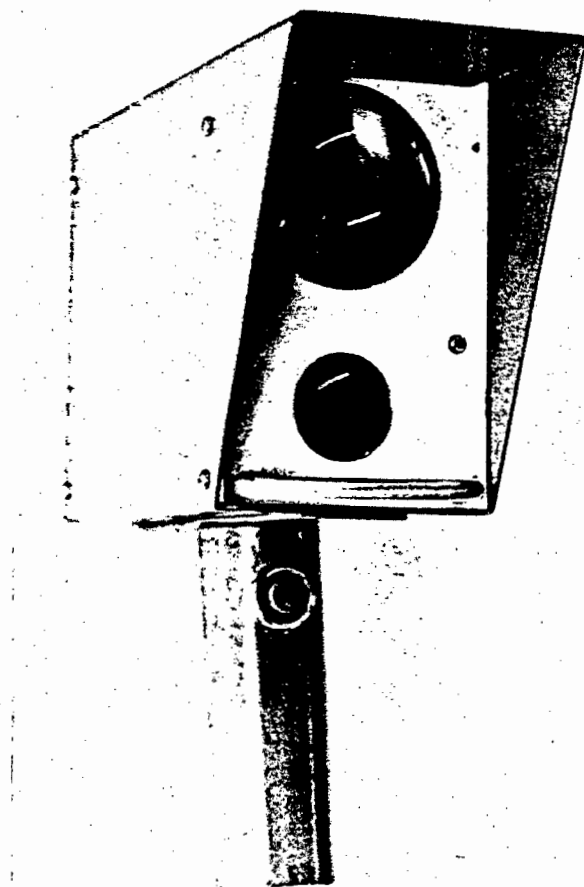


The Plug-in Chassis (Rear View).



The Plug-in Chassis (Front View)

The chassis components were designed around the lenses and the circuit boards, the result being an object slightly larger than an 8mm. cine-camera, but lighter, thanks to the plastic and aluminium constructions. The prototypes were completed without a sighting device, since plans for this had not been finalised.



A Completed Prototype.

The sight, of the notch- and-post type, would be mounted on the top surface of the communicator, and would be used like a revolver sight. If blurring of the notched backsight should prove a problem, a ring backsight could be substituted without difficulty.

Unfortunately, at this stage development was curtailed owing to lack of time. It was however established that the newly-designed prototypes functioned as expected, with performance comparable with that of the breadboarded circuitry used earlier. Further work on this design could be concentrated on three areas:

/testing.....

testing of different types of emitters, with the aim of improving the range; refining the mechanical design, reducing the bulk of the package and improving its ease of use; and finally, actual circuit redesign. Costs might be reduced by using integrated circuits to replace transistors. Any reduction in component count would be an advantage, and it might be possible to develop a simple transmitter modulator and l.e.d. driver using a combination of CMOS and discrete transistors. The possibility of having a custom-made potted bandpass filter made by a specialist manufacturer should be investigated. A factor which would radically affect the whole design would be the availability at moderate cost of a suitable large area detector. Such a device would provide a broad received beam pattern, and greatly enhance the ease of use of the system.

Future Developments.

As is apparent from the photographs, the use of lenses results in the generation of considerable amount of wasted space. The receiver lens could be made redundant by the advent of cheap, large-area photodiodes, the technology for which already exists. Circuit design would have to take into account their high capacitance and large d.c. photocurrents, but the advantages would be considerable. The transmitter has been waiting for a breakthrough in technology, since what is required is an emitter with a really narrow beam angle without the use of external lenses. This could in fact be achieved by using one of the small-area, high-radiance continuous emitters that are becoming available from firms such as Plessey encapsulated with a relatively long focal length epoxy lens. An exciting new alternative has recently been announced by Xerox, in the form of a diode laser with a diffraction grating engraved in the plane of the junction. The result is the concentration of emission at only one wavelength, and the production of a very narrow beam, about $0,35^\circ$ wide.

Either of these types of emitter, used with large-area photodetectors, should make possible a very compact, simple communicator. The cost would probably be higher than that of the design presented here, but the improved performance would more than justify it, in terms of ease of use and versatility. Such a system might well have applications in areas of high electrical noise, where conventional radio would suffer severe interference, for example in steel mills.

The future for point-to-point optical communications is a limited one. Its main advantage lies in its non-exclusive bandwidth occupancy. With the density of communications increasing continuously, and limited r.f. bandwidth available, there will undoubtedly be a market for communicators which do not use the existing wavelengths. Optical communications as a whole is a field which must expand rapidly, in order to provide for the new generation of data communications that is almost upon us.

BIBLIOGRAPHY.(i) Electro-optics.

1. "RCA Electro-Optics Handbook" (RCA 1968)
2. "Injection Electroluminescent Devices" - C.H. Gooch (Wiley 1973).
3. "Infrared System Engineering" - R.D. Hudson, Sr., (Wiley 1969).
4. Hewlett-Packard Optoelectronics Designer's Catalog (1975).
5. Hewlett-Packard AN-915: "Threshold Detection of Visible and Infrared Radiation With PIN Photodiodes." (1967).
6. Hewlett-Packard AN-909: "Electrical Isolation Using the HP 5082-4310".
7. B.Sc. Thesis at U.C.T. "Light Beam Walkie Talkie System Using A GaAs Diode and A Photodiode" - I.H. Telfer 1970 (unpublished).

(ii) Optics.

8. "Optics" - C.J. Smith (Arnold 1960).
9. "Fundamentals of Optics" - F.A. Jenkins and H.E. White (McGraw-Hill-Kogakusha 1957, 3rd Edition).
10. "A Treatise On Geometrical Optics" - R.S. Heath. (Cambridge University Press 1895, 2nd Edition).

(iii) Modulation and Noise.

11. "Modulation Theory" - H.S. Black (Van Nostrand 1953).
12. "Signal Processing, Modulation, and Noise" - J.A. Betts (English Universities Press 1970).
13. "Information Transmission, Modulation, and Noise" - Mischa Schwartz (McGraw-Hill 1959).
14. "Noise" - F.R. Connor (Arnold 1973).

(iv) General.

15. "Electronic Fundamentals and Applications" - Ryder (Prentice-Hall 1965, 3rd Edition).
16. "Network Analysis and Synthesis" - F.F. Kuo (Wiley 1966, 2nd Edition).
17. "Operational Amplifiers: Design and Applications" - Edited by G.E. Tobey, J.G. Graeme, and L.P. Huelsman (McGraw-Hill-Kogakusha 1971).
18. "Motorola High-Speed Switching Transistor Handbook" (Motorola 1963).

Also used were numerous data books and applications notes by National Semiconductor, Motorola, RCA, Siliconix, Philips, Fairchild, Texas Instruments, Signetics, Monsanto and others.

(v) Articles.

19. J.R. Kerr et al., "Atmospheric optical communications systems," Proc. IEEE, Vol. 58, pp.1691 - 1709, October, 1970.
20. N.S. Kopeika and J. Bordogna, "Background noise in optical communications systems," Proc. IEEE, vol. 58, pp 1571 - 1577, October, 1970.
21. F.E. Goodwin, "Review of operational laser communication systems," Proc. IEEE, vol. 58, pp. 1746 - 1752, October, 1970.
22. G. Schlisser and J. Insler, "Portable optical communicator rides laser for secure voice transmissions, "Electronics, vol 43, pp. 92-96, March 16, 1970.
23. R.J. Keyes et al., "Modulated infrared diode spans 30 miles,"Electronics, vol. 36, April 5, 1963.
24. L.M. Magill, "Optical link firms see wide horizons," Electronics, vol. 43, pp. 91-92, December 21, 1970.
25. T. Motoki et al, "Automatic acquisition and tracking system for laser communications," IEEE Trans. Com. vol. 20, pp. 847 - 851, October, 1972.

26. J.N. Kessler, "Fiber optics sharpens focus on laser communications," Electronics vol. 44 pp 46 - 52, July 5, 1971.
27. "Laser link hops 3,4 Miles, even in poor weather," Electronics vol. 45 p. 7E Mar. 13, 1972.
28. H. Melchior et al, "Photodetectors for optical communication systems," Proc. IEEE, vol. 58 pp 1466 - 1486, October, 1970.
29. "Laser link uses stabilised binoculars, " Aviation Week, vol. 92, pp. 42-43, June 29, 1970.
30. "Stabilised laser communicator operates from moving vehicle," electronics vol. 46, p. 34, May 24, 1973.
31. "C-W diode lasers at room temperature," Electronics, vol. 43, p 37, August 31, 1970.
32. L.A. Trinogga and D.F. Oxford, "J.f.e.t. noise figure measurement," Electronic Engineering, vol. 46, pp. 69 - 74, April, 1974.
33. A. van der Ziel, "Thermal noise in field-effect transistors," Proc IRE, vol. 50 pp 1808 - 1812, August, 1962.
34. A. van der Ziel, "Gate noise in field effect transistors at moderately high frequencies," Proc. IEEE vol 51, pp. 461 - 467, March 1963.
35. C.T. Sah, "Theory of low-frequency generation noise in junction-gate field-effect transistors," Proc IEEE, vol. 52, pp 795 - 814, July 1964.
36. S.E. Miller, "Optical communications research progress," Science, vol. 170 pp. 685 O 695, November 13, 1970. Discussion: vol. 174, p. 1045, December, 3, 1971.
37. S. Saito, "Optical communications in Japan," IEEE Trans. Com. vol. 20, pp. 725 - 730, August, 1972.
38. "Glass guide loss is cut to 4dB/km," Electronics, vol. 45, p. 30 September 11, 1972.

39. "FM video system shines through rain: Light line system," Electronics, vol 45 pp 42 - 44, November 20, 1972.
40. T.E. Price, "An atmospheric infra-red communication link," Electronic Engineering vol 46, pp 57 - 59, October, 1974.
41. J.A. Roberts and N.A. Jolly, "Audio noise: Why settle for more?" Electronics, vol. 43 pp 82 - 83, September 28, 1970.
42. "Practical laser data link, " Electronic Engineering, vol. 44 p 62, May 1972.
43. F.M. Klaasen, "High-frequency noise of the junction field-effect transistor" IEEE Trans. Electron Devices, vol. ED-14, pp 368 - 373, July 1967.
44. H.R. Schindler, "Delta Modulation," IEEE Spectrum, vol. 7 pp 69 - 78 October 1970.
45. "Semiconductor lasers expand optical communications capacity," Computer Design, vol. pp 46 - 47, May 1975.
46. "Light-beam gun and targets top Yule list," Electronics vol. 44 p. 13E, December 20, 1971.
47. P.H. Wendland, "Solid state combo senses light well enough to vie with tubes," Electronics, vol. 44 pp 50 - 54, May 24, 1971.
48. "Laser system runs at 1,5 Mb/s, " Electronics, vol. 48, p 169, March 20, 1975
49. C.A. Burrus, " Radiance of small-area high-current-density electroluminescent diodes, " Proc. IEEE, vol. 60, pp 231 - 232, February, 1972.

Appendix 1: Device Data.

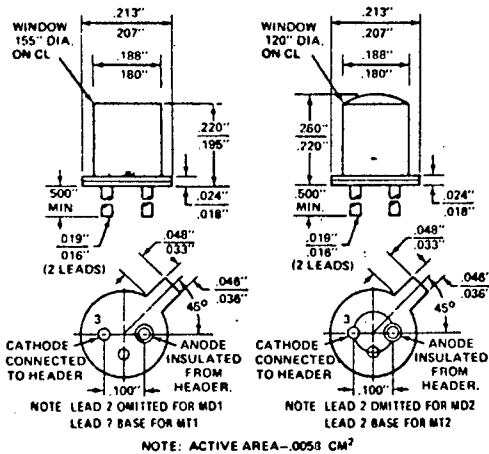
SILICON PIN PHOTODIODE

MD1 MD2

PRODUCT DESCRIPTION

The MD1 and MD2 are diffused planar silicon PIN photodiodes. Both are mounted on a standard TO46 header. The MD1 has a flat window at the top of a metal shielding can. The MD2 has a domed lens in the window position for optical gain.

PACKAGE DIMENSIONS



FEATURES

- Fast Response 0.5 nsec
- Responsive To GaAs Sources 4.0 $\mu\text{A/mW/cm}^2$
- Responsive To Tungsten Sources 1.6 $\mu\text{A/mW/cm}^2$
- Optional Flat Lens or Built-in Optics
- Standard Transistor Package For Easy Handling and Mounting.

These devices are recommended for applications in:

- high speed optical switching
- laser detecting
- optical encoding
- intrusion alarm or warning
- process control
- industrial control

ABSOLUTE MAXIMUM RATINGS

Maximum Storage and Operating Temperature	 -55°C to 150°C
Maximum Lead Solder Time @ 260°C (See Note 1)	 7.0 sec
Power Dissipation @ 25°C Ambient Temperature	 300 mW
Derate Linearly From 25°C	 2.4 mW/°C
Reverse Voltage	 50 volts

ELECTRO-OPTICAL CHARACTERISTICS (25°C Free Air Temperature Unless Otherwise Specified)

CHARACTERISTICS	MIN.	TYP.	MAX.	UNITS	TEST CONDITIONS
Breakdown voltage	50			V	$I_R=10 \mu\text{A}$
Dark current (see note 2)		2.5	10	nA	$V_R=20 \text{ V}$
Dark current at 100°C		200		nA	$V_R=20 \text{ V}$
Capacitance		8		pF	$V_R=20 \text{ V}, f=1.0 \text{ MHz}$
Sensitivity					
MD1 (see note 3)	.6	.8		$\mu\text{A/mW/cm}^2$	2875°K, $V_R=20 \text{ V}$
MD2 (see note 3)	1.2	1.6		$\mu\text{A/mW/cm}^2$	2875°K, $V_R=20 \text{ V}$
MD1 (see note 4)	1.5	2.0		$\mu\text{A/mW/cm}^2$	$\lambda=.9 \text{ microns}, V_R=20 \text{ V}$
MD2 (see note 4)	3.0	4.0		$\mu\text{A/mW/cm}^2$	$\lambda=.9 \text{ microns}, V_R=20 \text{ V}$
Rise time		0.5		ns	$V_R=20 \text{ V}, R_L=50 \Omega$

MDI MD2

TYPICAL ELECTRO-OPTICAL CHARACTERISTIC CURVES

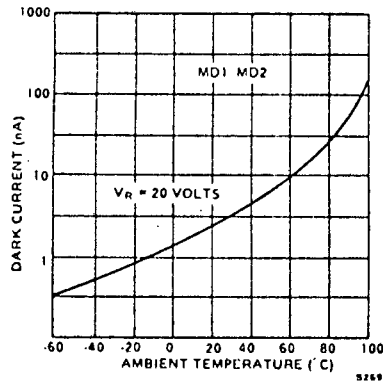


Figure 1 Dark Current vs. Temperature

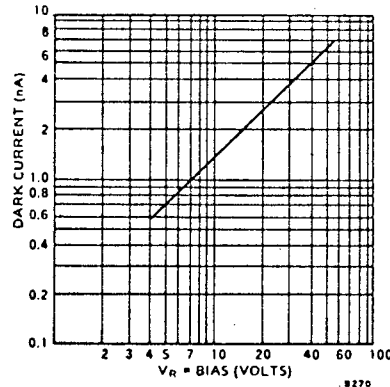


Figure 2 Dark Current vs. Bias Voltage

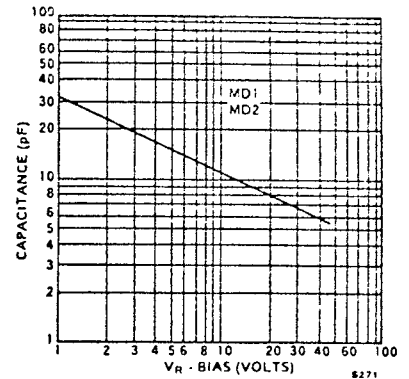


Figure 3 Capacitance vs. Bias Voltage

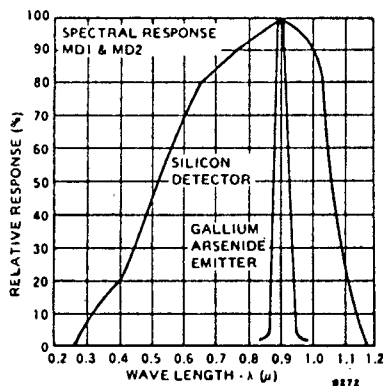


Figure 4 Spectral Response MD1 and MD2

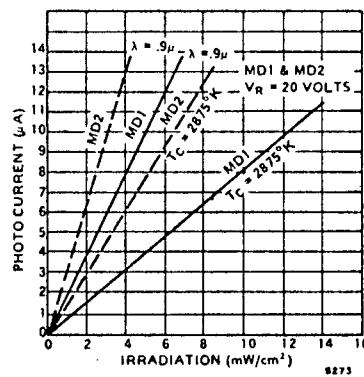


Figure 5 Photo Current vs. Irradiation MD1 and MD2

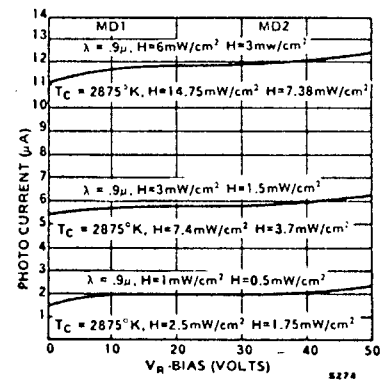


Figure 6 Photo Current vs. Bias Voltage

NOTE: The axis of maximum sensitivity shall be within a 10° cone with reference to the central axis of the device.

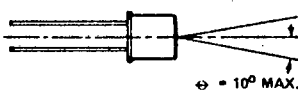


Figure 7 Angular Alignment

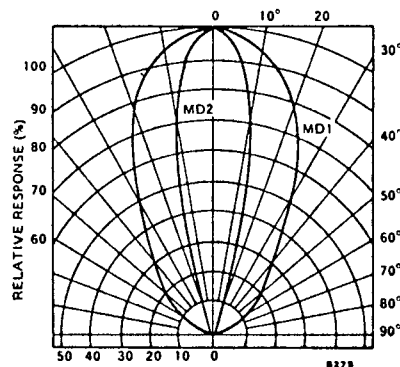


Figure 8 Angular Response

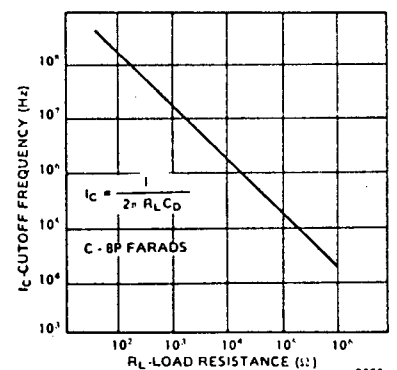


Figure 9 Cut-off Frequency vs. Load Resistance

(25°C Free Air Temperature Unless Otherwise Specified)

NOTES

1. The leads of the device were immersed in molten solder, heated to a temperature of 260°C, to a point 1/16 inch from the body of the device per MIL-S-750.
2. Measured under dark conditions $H \leq 1.0 \mu\text{W}/\text{cm}^2$.
3. Radiation flux intensity (H) of 5.0 mWatt/cm² as emitted from tungsten filament lamp operated at a color temperature of 2875°K.
4. Measured with a GaAs light source at .9 microns with a radiation flux density of 3 mWatt/cm².
5. Rise time - the time required for the current pulse to rise from 10% to 90% of its maximum amplitude.

Monsanto

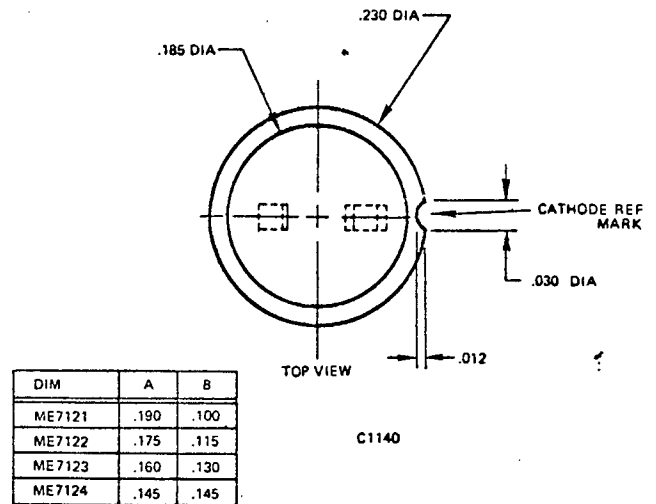
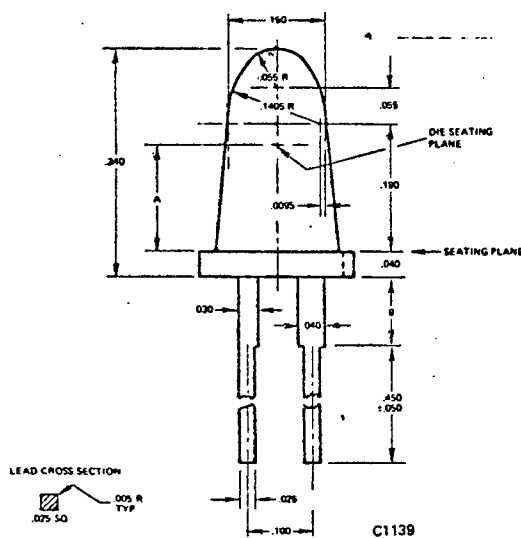
HIGH POWER INFRARED EMITTERS

ME7121, ME7122, ME7123, ME7124

PRODUCT DESCRIPTION

This family of high power liquid phase epitaxial IR Emitters is designed to accommodate all needs of the emitter detector relationship. Products range from a wide angle power spread for non-critical detector location to sharp-angle concentration of power for detectors located a significant distance from the emitter. The devices can be mounted with a plastic pop-in, furnished upon request.

PACKAGE DIMENSIONS



DIM	A	B
ME7121	.190	.100
ME7122	.175	.115
ME7123	.160	.130
ME7124	.145	.145

ALL DIMENSIONS IN INCHES
TOLERANCES = $\pm .010$ UNLESS SPECIFIED

ABSOLUTE MAXIMUM RATINGS

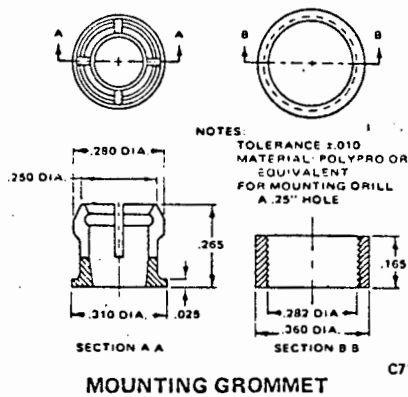
Maximum power dissipation @ 25°C ambient	150 mW
Derate linearly from 50°C	2.8 mW/°C
Maximum storage & operating temperature	-55° to 100°C
Maximum lead solder time @ 260°C (Note 3)	5 sec
Maximum continuous forward current	100 mA
Maximum reverse voltage	3.0 V
Peak forward current (PW - 1.0 μ sec, Duty Cycle = 0.1%)	6.0 A

ELECTRO-OPTICAL CHARACTERISTICS

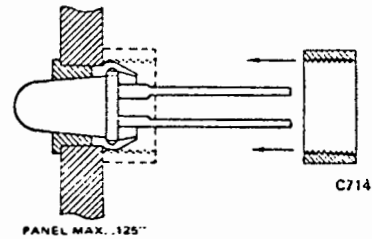
	TYPICAL HALF ANGLE (DEGREES)	TYPICAL ON AXIS INTENSITY (MW/STR.) @ 50 mA	
ME7121	17°	10.8	$\left\{ \begin{array}{l} \text{into cone @ 1/2 power points} \\ \text{@ } I_F = 50 \text{ mA} \\ \text{ROP} = 3 \text{ mW} \end{array} \right.$
ME7122	12°	26.4	
ME7123	8°	105.6	
ME7124	6°	243.6	

	MIN.	TYP.	MAX.	UNITS	TEST CONDITIONS
Total External Output Power (Note 2)	1.0	3.0		mW	$I_F = 50 \text{ mA}$
Peak Emission Wavelength		940		nm	$I_F = 50 \text{ mA}$
Spectral Line Half Width		50		nm	$I_F = 50 \text{ mA}$
Forward Voltage		1.4	1.8	V	$I_F = 50 \text{ mA}$
Light Turn On & Turn Off Time	-	500		nsec	50 Ω Load

ME7121 ME7122 ME7123 ME7124

PANEL
MOUNTING
TECHNIQUES

MOUNTING GROMMET



PANEL MOUNTING

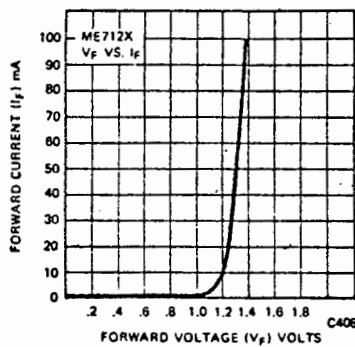
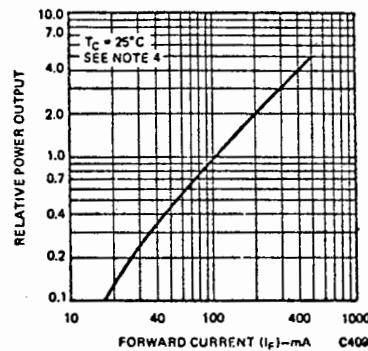
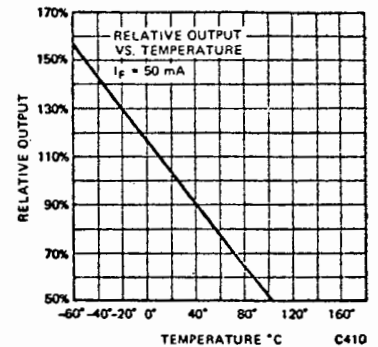
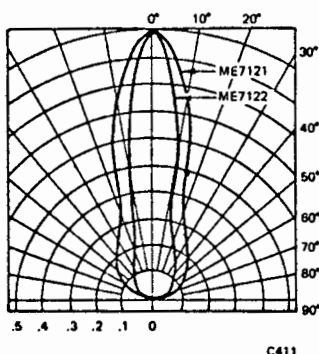
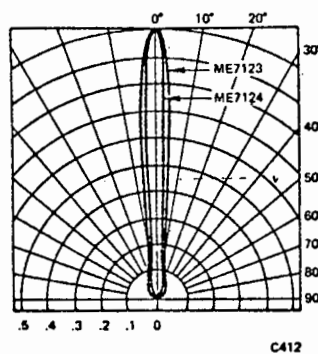
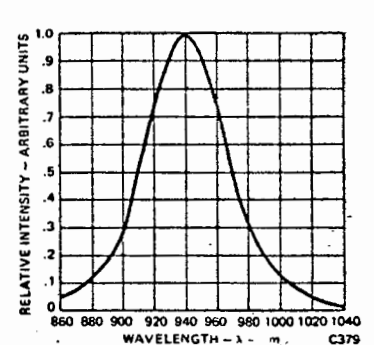
TYPICAL ELECTRO-OPTICAL CHARACTERISTIC CURVES
(25°C Free air temperature unless otherwise specified.)Fig. 1. I_f vs. V_f Fig. 2. ROP vs. I_f PeakFig. 3. ROP vs. Temperature
(Note 1)Fig. 4. Spatial Distribution
(ME7121 and ME7122)Fig. 5. Spatial Distribution
(ME7123 and ME7124)

Fig. 6. Spectral Distribution

NOTES

1. The curves in figure 3 are normalized to the power output at 25°C to indicate the relative efficiency over the operating temperature range.
2. The total external radiated power output measurements are made with a Centralab 110C solar cell terminated into a 100Ω impedance.
3. The leads of the ME7121, ME7122, ME7123, and ME7124 were immersed in molten solder, heated to 260°C, to a point 1/16 inch from the body of the device, per MIL-S-750.
4. This parameter is measured using pulse techniques $tw = 40\text{ }\mu\text{sec}$ duty cycle $\leq 10\%$.

Appendix 2: Background Radiation and Optical Filtering.

The sensitivity of the receiver in an atmospheric point-to-point optical link is usually limited by shot noise due to background radiation. In order to calculate the minimum receivable signal power, it is necessary to know the noise power due to background radiation. Since the use of filters will reduce background noise, data on their efficiency is also required. Data on typical reflectances at 930 nm is not available, so the most satisfactory means of evaluating the typical background radiation shot noise is simply to build a receiver lens plus detector mock-up, and use it to measure the d.c. background photocurrent I_b .

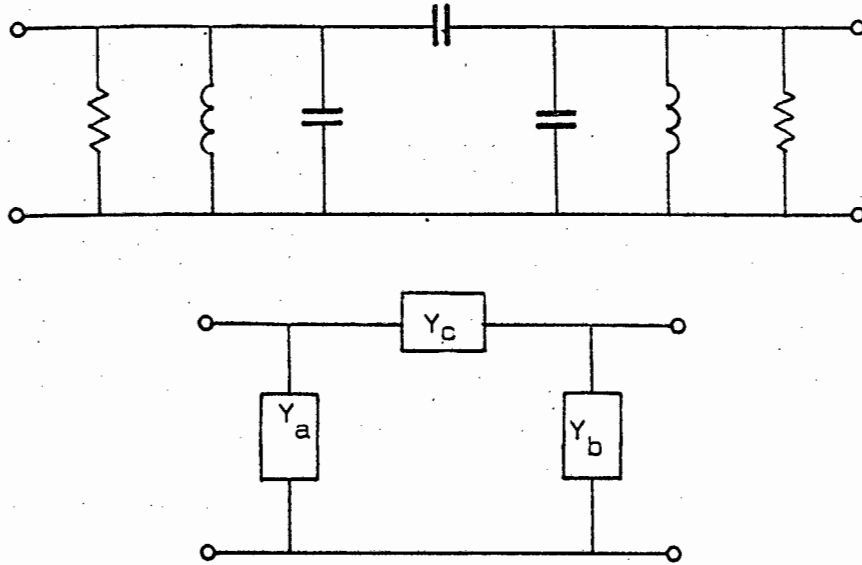
This was done, with the lens actually used in the prototypes, and the H-P 5082-4207 photodiode. A reverse bias of 6V was applied and the current measured with a d.c. microammeter. When the detector was aimed at the sun, an I_b value of about 130 microamps resulted. Use of a Kodak type 87C absorption filter reduced I_b to about 25 microamps. These high values are not typical of the conditions under which the optical link will operate. Readings were taken of various sunlit surfaces, and of the sky, with and without clouds. The range of I_b measured (without filter) was surprisingly small, and encouragingly low in value, from about 0.7 to 3 microamps. When the Kodak filter was used, the readings were too low to be reliably read on the equipment available, but seemed to be of the order of 20% of the unfiltered values.

An alternative theoretical method of evaluating the reduction in I_b due to the filter involves the use of the solar spectral irradiance curve (measured at the earth's surface), the filter response curve, and the photodiode sensitivity curve. The filter excludes wavelengths shorter than 800 nm, while the photodiode's sensitivity falls rapidly at wavelengths greater than 1 micron. By plotting the product of the three curves and evaluating the area under the resulting graph, one can establish the proportion of solar radiation passed by the filter which results in a background photocurrent in the detector. This was done, using solar spectral irradiance data from the RCA Electro-optics Handbook (Ref.1), and the results indicated a reduction of I_b to 22% of the unfiltered value. This compares well with the measured figure.

The use of an interference filter would further reduce the I_b contribution. The transmitted optical bandwidth is about 50 nm, and this is the effective filter bandwidth required. Since the filter would be passing non-parallel light, its bandwidth will be broadened, so that a nominal bandwidth of about 30 to 35 nm is necessary. Such a filter can be expected to reduce I_b to about 8% of its unfiltered value. The cost will, however, be considerably higher than that of the absorption filter, and since the system range is inversely proportional to the fourth root of the background power, the improvement in range would only be about 25%, compared with the 50% improvement obtained by going from no filter to an absorption filter. Thus the cheaper absorption filter should be adequate for the low-cost link application. If a narrow-band emitter had been used, the advantage to be gained by use of an interference filter would have been much greater, since I_b could be reduced to as little as 1 or 2% of its former value.

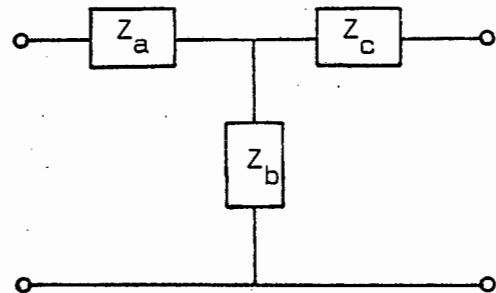
Appendix 3 : Double-tuned Bandpass Filter Circuit Analysis.

The chosen configuration can be represented as a pi-network :

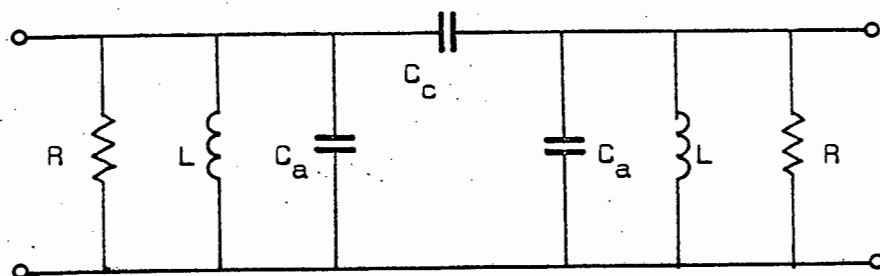


It is convenient to analyse the circuit by means of a delta-star transform, giving :

$$\begin{aligned} Z_a &= Y_b / \det Y \\ Z_b &= Y_c / \det Y \\ Z_c &= Y_a / \det Y \\ \det Y &= Y_a Y_b + Y_b Y_c + Y_c Y_a \end{aligned}$$



Since this filter circuit is symmetrical, we have $Y_a = Y_b$, so $\det Y = Y_a(Y_a + 2Y_c)$. Because the parallel-tuned circuit Y_a is a low admittance at resonance, a current drive is required to produce the desired bandpass response. Hence the gain is expressed as a transfer impedance, $e_o/i_s = z_{21} = Z_b = Y_c / (Y_a(Y_a + 2Y_c))$.



$$\begin{aligned}
 z_{21} &= \frac{sC_c}{(sC_a + 1/R + 1/sL)(s(C_a + 2C_c) + 1/R + 1/sL)} \\
 &= \frac{s^3 C_c}{(s^2 C_a + s/R + 1/L)(s^2 (C_a + 2C_c) + s/R + 1/L)}
 \end{aligned}$$

Now substitute $C = C_a + C_c$ and $C_c = kC$, so that $C_a = C(1-k)$.

$$\begin{aligned}
 \text{Rewrite } z_{21} &= \frac{s^3 kC}{(s^2 C(1-k) + s/R + 1/L)(s^2 C(1+k) + s/R + 1/L)} \\
 &= \frac{s^3 k/C}{(1-k^2)(s^2 + s/((1-k)RC) + \omega_o^2/(1-k))(s^2 + s/((1+k)RC) + \omega_o^2/(1+k))}
 \end{aligned}$$

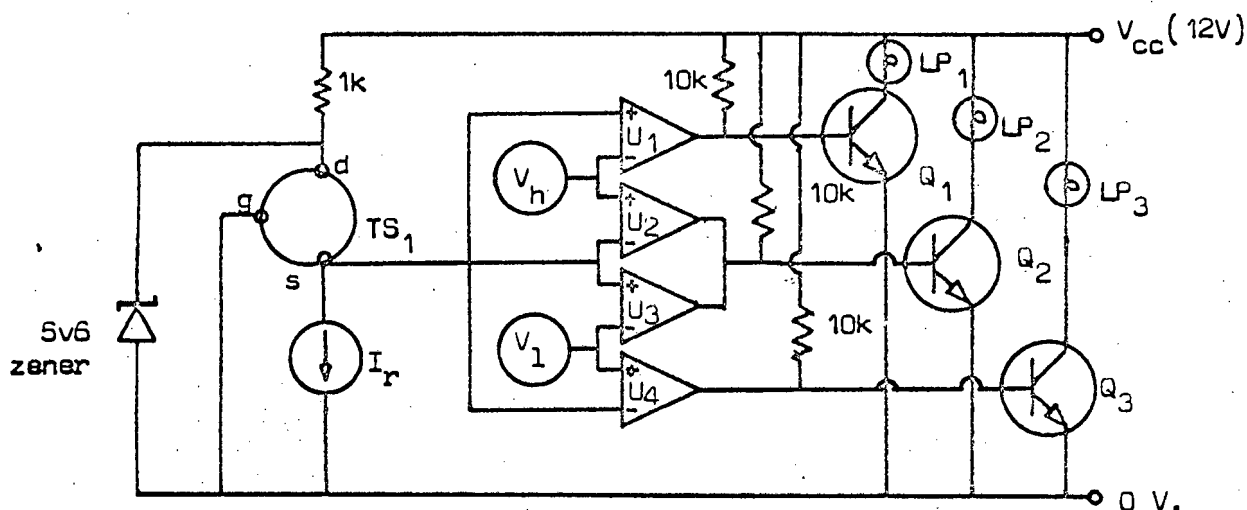
where $\omega_o^2 = 1/LC$. Now $1/RC = \omega_o/Q$, so :

$$z_{21} = \frac{s^3 k/C}{(1-k^2)(s^2 + s\omega_o/Q(1-k) + \omega_o^2/(1-k))(s^2 + s\omega_o/Q(1+k) + \omega_o^2/(1+k))}$$

where k is in fact the coefficient of coupling. For the case where $k \geq k_c$, z_{21} has a maximum value of $z_{\max} = R/(2(1-k^2)) \approx R/2$ for $k \ll 1$.¹

Appendix 4 : J.f.e.t. Selection Test Jig.

It is necessary to make a selection from the E202 f.e.t.s used in the communicator so that in specific circuits, the gate-source voltage lies within certain limits. The receiver frontend requires a f.e.t. that has V_{gs} between 0,7V and 2,0V for $I_d = 0,9$ mA. The test jig should be usable by an unskilled operator with minimal instruction, and thus must provide a simple go-no go indication. Devices can be divided into three categories: V_{gs} too low, V_{gs} within range, and V_{gs} too high. Use of coloured indicator lights, respectively red, green, and yellow, for example, should provide the desired simplicity of operation. A further requirement is that no light should be on when there is no f.e.t. in the test jig or when the device on test is open-circuit ($I_d=0$), or when it has a gate-source short ($V_{gs}=0$), since it would be convenient to use the V_{gs} test as a screening process for defective devices.



The Basic Tester

U_{1-4} : each 1/4 of a LM339N quad comparator

Q_{1-3} : any 200 mA NPN transistor, e.g. 2N2222, 2N3704

LP_{1-3} : yellow, green, and red (in that order) 12V 100mA panel lights.

TS_1 : test socket

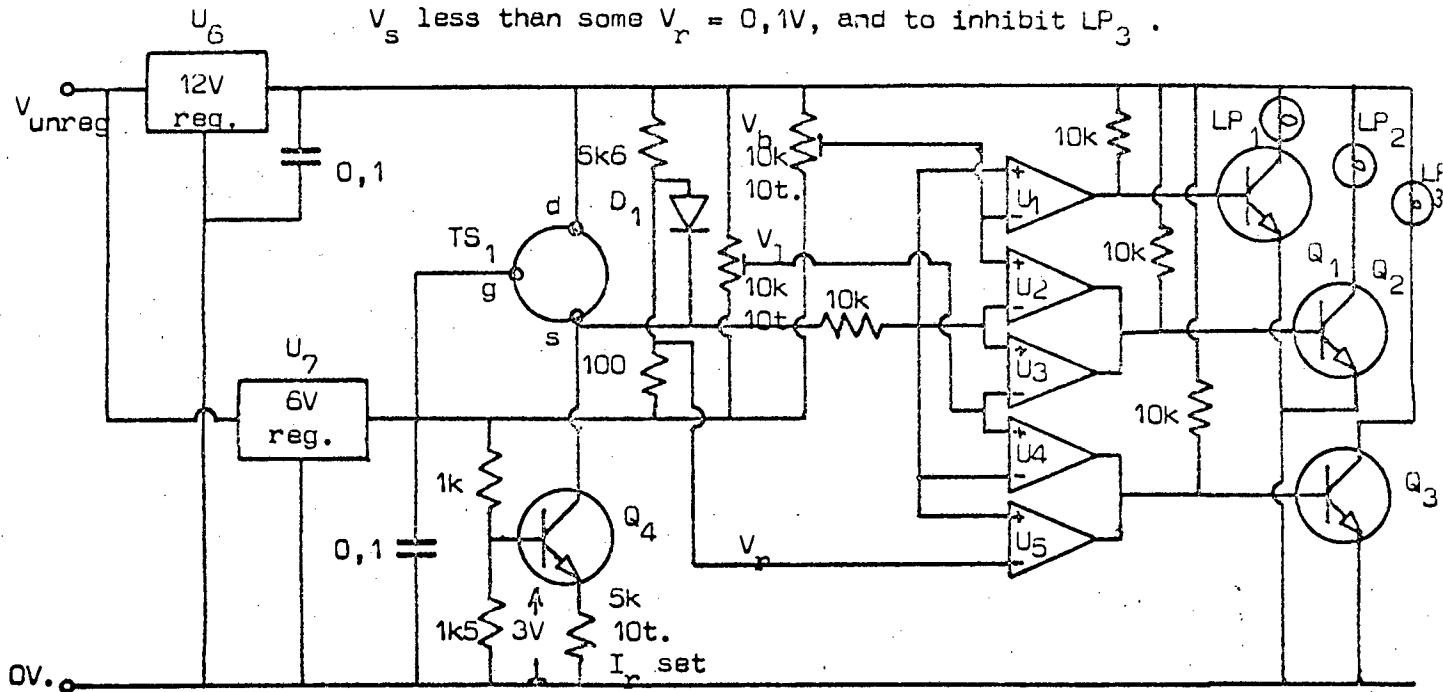
V_h : upper V_{gs} threshold

V_l : lower V_{gs} threshold

I_r : reference drain current

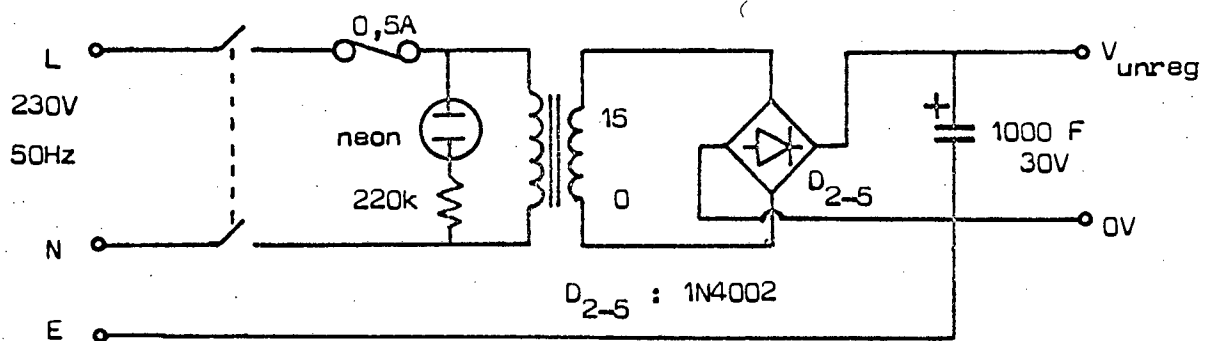
As shown, the circuit cannot detect $V_{gs} = 0$ or $I_d = 0$, and when $V_{gs} = 0$, or no device is inserted, LP_3 is on. V_h and V_l can be derived from the supply voltage, using multi-turn potentiometers, since the input current of the comparators is very small. The supply must be regulated, using a 3-pin monolithic regulator for convenience. V_{dg} is fixed at about 6V, to approximate the actual in-circuit value.

In the following circuit, I_r is derived from a negative supply rail, so that when $I_d = 0$, V_s will be pulled below 0V. Thus a fifth comparator can be used to test for V_s less than some $V_r = 0,1V$, and to inhibit LP_3 .



The Complete Tester.

- U_{1-4} : each 1/4 of an LM339N quad comparator } all V_+ at +12V.
 U_5 : LM311N comparator } all V_- and
 } gnd at 0V.
 Q_{1-3} : 2N2222 or 2N3704
 Q_4 : BC109C
 D_1 : any diode with low V_f and low leakage, e.g. OA47, OA5.
 LP_{1-3} : yellow, green, and red 12V 100mA panel lights.
 U_6 : 12V monolithic regulator, e.g. MC7812, LM340T-12, LM341P-12
 U_7 : 5V monolithic regulator, e.g. MC7806, LM340T-6, LM341P-6, LM342P-6



To ensure a high testing rate and corresponding low testing cost, the design of the test socket must be such that insertion requires no special lead-forming procedures. This can be achieved by providing a conic "funnel" above each socket contact to guide the device lead into the socket. This jig should enable testing rates of several hundred per hour, reducing the cost of testing to a few cents per device.