ERRORS - A POSITIVE APPROACH

by

Aubrey Max Sandman

A thesis submitted in partial fulfilment of the requirements for the Degree of Doctor of Philosophy.

THE CITY UNIVERSITY, LONDON

Centre for Information Engineering

November 1989

CONTENTS

Reducing Amplifier Distortion, Wireless World, Oct.1974, pp.367-371.

Class S, Wireless World, Sep.1982, pp.38-39.

ACKNOWLEDGEMENTS

I wish to express my thanks to Professor A.C.Davies for his encouragement and immense patience without which this thesis would not gave been finished, to B.Sapir for his co-operation and vital role in programming the compression system, and to D.Styles for his help around the laboratory as well as to Helen Shepherd for her help with Wordstar and MS-Dos. In addition S.Omarouayache suggested a more elegant method of solution to the basic equation for the second vehicle's velocity in the section on the Railway signalling scheme, for which he also has my thanks, as well as helping in the task of printing out this thesis.

ABSTRACT

The object of this thesis is to show that a deliberate consideration of the function of errors within a system can produce novel and useful results.

In the polynomial data compression system, cubic splines are applied, with the aid of an error controlled algorithm, in such a way as to produce not just compression, but much better fidelity of reproduction. This results in the reproduced waveform being free of visual discontinuities, which is not the case for the conventional zero and first order data compression systems.

The control of error is basic to this algorithm's operation.

The measurement of error is then applied to amplifiers, the measured error being applied in such a way as to combine the advantages of both negative feedback and feedforward.

Of particular interest is the filtering employed to ensure a better stability margin.

In a redesign of a known amplifier, due to Blomley, the deliberate generation of gross error is used to overcome certain problems in the original design with the aid of subsequent measurement and error manipulation.

Some experimental results are presented.

Then follows a discussion of the current 'Block' Railway Signalling System in which the problems of the present system are discussed and a much improved system is proposed, in outline.

This treats the problem of Railway Signalling as involving very reliable communications, as well as detecting errors as they arise.

The prompt detection of errors is central to the proposed scheme.

1 INTRODUCTION

'Things should be made as simple as possible -
but no simpler.'

-After Einstein

This thesis presents some of my contributions to the subject of using errors in a positive manner and includes several distinct topics investigated over a period of time.

The common theme linking these topics together is the positive contribution of error manipulation in achieving improved performance. The use of error in engineering (for example in steam governors, negative feedback amplifiers and servomechanisms) is well established and not in itself novel.

However although in such applications feedback is used to reduce error, this is not by an arbitrary predetermined amount.

In some of the examples chosen to illustrate a more positive approach to the use of error in engineering, errors are measured and used to reduce the overall resultant error arising which, as reiteration is possible, in principle can be by any arbitrary amount. In the case of the variable sampling rate data compression method the error controls the sampling rate.

I believe that the phrase which opens this chapter (which Professor Albert Einstein put in a much more complex way), is a profound truth and it has certainly tempted me to present the subject matter of my third-order polynomial method of compression (ref.1.2) as basis for the whole of this thesis since I am quite prepared to be judged on its novelty and significance.

However although the earlier paper was entirely my own work (ref.1), the second, more comprehensive, paper was a cooperative effort, in which the practical work of writing Pascal programs and organising the production of results was done by B.Sapir under my direction.

For this reason I have included two sections one, on 'Cross Coupled Error-Takeoff Amplifiers' and the other on 'Results from practical implementation of the proposed amplifiers' which deal with errors applied to amplifiers and which are entirely my own work.

Although they may appear unrelated, they share a common theme: they deal with error reduction either by the manipulation of the errors which arise naturally in a circuit using a novel form of feedforward, which I describe as 'Crosscoupled Error-takeoff', or alternatively they reduce the errors by deliberately generating gross distortion, measuring it continually and automatically and then in a very accurate way using the resultant error and the original deliberate distortion to together produce a very low distortion output waveform as in the Blomley circuit.

Thus the error signal is an essential aspect of achieving improvement in performance and so fits into the overall theme of the thesis.

These two sections on amplifiers together with the papers on data compression are what I am including as the 'heavyweight' part of the thesis.

There then follows a further section, on a proposed Railway Signalling scheme. This forms the 'lightweight' part of the thesis since the ideas it contains are not backed up by any detailed evaluation or practical experiments although I hope it will be of interest in showing a further example of the constructive benefits of errors when appropriately applied to engineering problems.

The proposed Railway Signalling system firstly tries to establish very reliable communications by applying feedback and labelling to each piece of information. If however an error still occurs it provides means for it to be instantly detected and action taken.

Having established such a reliable network, it then uses a much improved basis of computation to work out the maximum permitted speed of a train.

The method of approach used in analysing the circuits in this thesis leans heavily on one of the most significant advances in circuit design, that of Computer Aided Analysis, in this case in the shape of a commercial Analog analysis package called ECA2.

This enables accurate models to be employed and complex circuits to be accurately analysed with the minimum of approximations, in a swift and above all reliable way. That is to say the results can be relied on to be close to those obtained from practical implementation.

The circuits analysed during this thesis, including those which I claim to have originated, are far too complex to enable the frequency response at the band edge to be determined by hand calculation, without making simplifications which would render the results useless.

For this reason, a detailed mathematical analysis does not contribute to understanding the results, and is therefore not included in the thesis.

References

1.1. A.M.Sandman, "Bandwidth and signal-data compression using maximum specified error as control", Electronics Letters, Vol 21, No. 16, August 1985, pp. 673-674

1.2. A.Sandman and B.Sapir "Third order polynomial its use in data compression", Signal Processing 1988, Vol.15. pp 405-418.

Oscillators and negative feedback amplifiers not only share the famous formula G=A/1+A.B, with a sign difference, but also the physical method of operation, in which a proportion of the output is fed back to the input. In practice neither A or B are necessarily constant with frequency.

Small wonder that amplifiers can oscillate !

(G is the closed loop gain, A the Forward Gain and B the feedback factor).

This would not matter if the operation within the desired frequency band of interest were not gravely affected, at the design stage, by the necessity to avoid oscillation at any frequency , whether within or without the band of interest, since oscillation seriously increases distortion. The consequence of all this is to seriously reduce the obtainable feedback factor at the high frequency band edge. In other words because of stability considerations an arbitrary feedback factor is not obtainable at the high frequency band edge, the designer having to make do with some much lower figure than the one he might wish for on grounds of for example, distortion reduction.

The reason for this is twofold.

Firstly since $G=A/(1+A.B)\approx(1/B).(1-(1/A.B))$,(binomial approximation when $A.B>>1$) it follows that A should be as great as possible, since the gain becomes more independent of A as A increases, which since A is not well controlled ,is desirable.

Now, secondly, (fig. 2.11) the gain of an operational amplifier (op. amp.) falls off with frequency and so A.B decreases with frequency from which it follows that 1/A.B becomes more significant as the frequency increases, which as already pointed out is not a good thing.

However any attempt to solve this problem by increasing A within the passband runs into trouble because this usually means that outside the passband on the high frequency side, A also increases and at high frequencies phase shift effects start to become important (unlike at low frequencies when, for an op.amp. the phase shift is close to 90 degrees).From this it follows that the value of A.B can be such as to cause oscillation.

Hence the limitation of negative feedback in which an arbitrary reduction in distortion cannot be obtained unless conditionally stable amplifiers are used, with all their great practical disadvantages. Since for any given circuit there is a maximum value of A beyond which the circuit becomes unstable as well as developing a very peaked frequency response immediately before A becomes great enough to cause this instability.

When the transcontinental American telephone service was first introduced thick copper cable was employed to reduce losses since the numerous analogue repeaters which would have enabled lossier lighter weight cable to be used had much too high nonlinear distortion.

Let us take a hypothetical example and assume a 3000 mile link, with repeaters every 30 miles giving 100 repeaters.

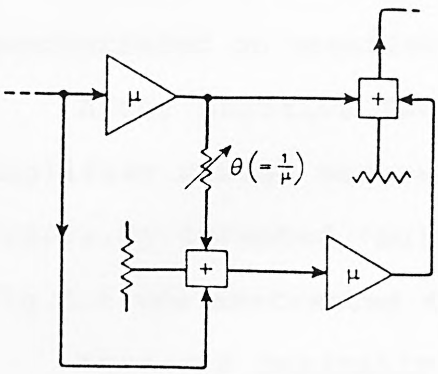
For an overall distortion of 10% this would mean that each amplifier would need to have a distortion of 0.1% or better. With valve amplifiers and no negative feedback, especially using valves of the nineteen twenties to thirties, this would be impossible to achieve with any reasonable output voltage such as 10 volts peak into 600 ohm lines.

The first attempt by Black (ref.2.1), to solve the problem was his invention, feedforward, in which he utilised the fact that distortion is, as will be explained later, one of the easiest of all measurements to make regarding an amplifiers performance. He simply subtracted the input from a suitable proportion of the output to be left with the distortion only.

It was at this point that difficulties occurred. What he tried to do next was to subtract this distortion signal from the principal or main amplifier output (Fig.2.1) with the aid of a subsidiary amplifier, so as to be left with an undistorted output voltage. The trouble was that both amplifiers varied in gain for all the reasons that amplifiers, especially valve amplifiers, vary in gain, such as ageing, power supply variations and so on.

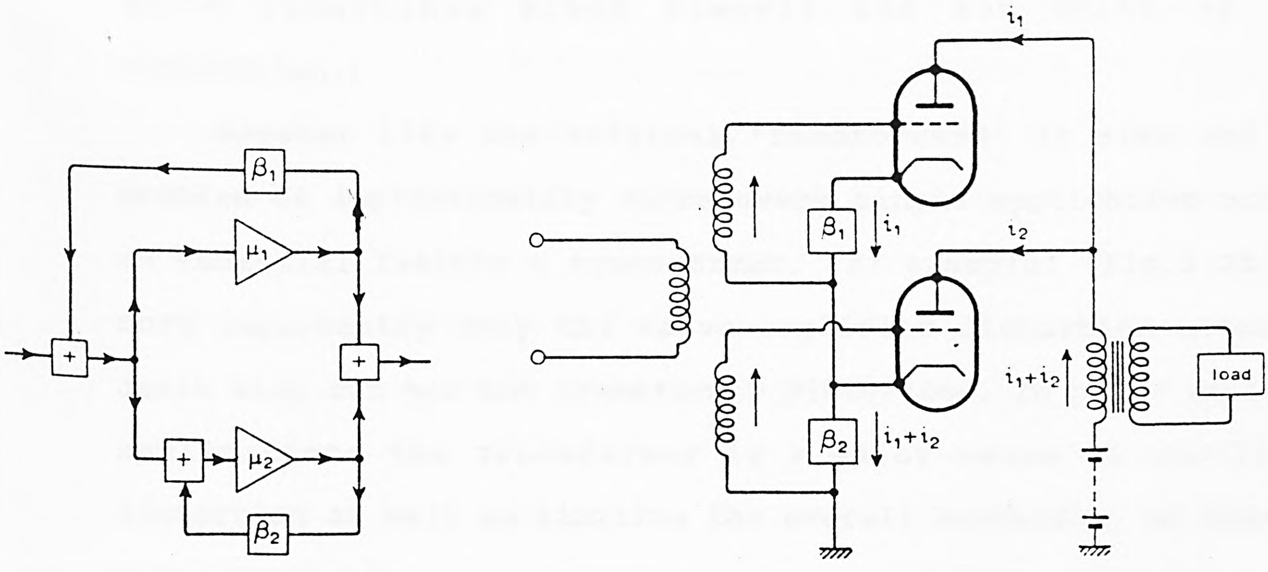
The result of all this was, quite simply, to make the system completely impractical, since although, in a manner similar to a bridge it is possible to 'balance out 'the distortion, after a very short time the gains of both the principal and subsidiary amplifiers would change by quite uncontrolled amounts and the balance would need readjusting. Of course in practice this could not be done, with the result that so far from reducing distortion such a system could easily increase it.

I imagine that it was with a sense of relief that Black invented negative feedback (whilst on a ferry boat in New York (ref.2.6)).



BLACKS FEEDFORWARD

FIG. 2.1



MCMILLAN'S MULTIPLE LOOPS

FIG 2.2

Certainly he abandoned feedforward from then onwards and concentrated on negative feedback.

After negative feedback the next significant advance in amplifier design occurred in the nineteen fifties when McMillan (ref.2.2) invented 'multiple loops' (fig.2.2).(Both fig.2.1 and fig.2.2 are abstracted from (ref.2.4)).

This was basically 'feedforward' but with gain stable, as distinct from gain unstable, sub-amplifiers. The gain stable sub-amplifiers used negative feedback, which had not been available to Black as he had yet to invent it! (Retrospectively it seems quite remarkable Black himself did not think of this combination.)

However like the original 'feedforward' it also had the problem of impracticality since every single application used as an essential feature a transformer, for example, (fig.2.2b) and more importantly only the valve amplifier distortion alone was dealt with but not the transformer distortion. In power amplifier applications the transformer is a major cause of non-linear distortion as well as limiting the overall bandwidth, so this was a serious weakness.

(Until the advent of the transistor all commercial audio amplifiers used valves and transformers in their output stages. With the arrival of the transistor the loudspeaker coil impedance made a good direct impedance match with the amplifier so eliminating the need for a matching transformer).

To the best of my knowledge the idea next surfaced in 1973 (ref.2.3) in a short contribution in Wireless World by myself, entitled 'Error addon', (later changed to 'Error takeoff'), in which an amplifier which did not use transformers was described.

This was followed by Walker's 'Current dumping' in 1975, (ref.2.5), which as well as having a tendency to instability also had a somewhat obscure method of operation which was the subject of much debate (ref.2.7). In addition an arbitrary degree of distortion reduction was not possible.

The subject of this section is an amplifying system which not only does not employ transformers but also overcomes the problem, in one manifestation, of the 'Cross coupled error takeoff amplifier', the bridge version , of not having a low output impedance. A low output impedance is of course a desirable feature if the load impedance varies widely with frequency as it does with a loudspeaker coil , for example ,since the output voltage is unaffected by this variation. (See the section on Practical Implementation fig.3.5).

It is a classified as a 'Modern amplifier' since it can defeat the limits of the 'Nyquist diagram', by which I mean the limitation that the necessity to never enclose -1 on the Nyquist diagram of a negative feedback circuit places on the amount of gain that can be employed before closing the loop.

This is, perhaps, a good point to emphasise that feedforward and its modern derivatives are free from this limitation since although their sub-amplifiers, in their modern guise, may employ negative feedback and are, of course subject to the Nyquist diagram, the overall final configuration of a 'feedforward' circuit, because there is no feedback, is not.

Error takeoff

Although having the disadvantages of a floating output and an error amplifier which had to carry the full load current (IL) (fig.2.3), whilst still contributing little output power itself, the original 'Error addon '(fig.2.3) amplifier had the substantial advantage of low output impedance as well as having an easy to understand method of operation. Due to the small voltage swing the error amplifier does not itself add significant distortion. Although the current swing is substantial this does not affect distortion due to the effect of local negative feedback around amplifier B in fig.2.3.

Mode of operation

The output with respect to earth of the top amplifier consists of the 'pure ' output voltage -Vin*R2/R1 plus Vd, the error voltage.

Now the input Vin balances exactly against the output Vin*R2/R1 at p leaving a measurement of Vd, value Vd*R1/R1+R2, at p. This measurement of error consists of a signal arising from finite gain, phase change, non-linearity, noise etc. This simple measurement is one of the easiest to make as Black realised so very clearly and is a true measure of the error. Needing only two resistors, it is immensely superior to, say, distortion measurement using a spectrum analyser which requires complex tuned filters.

Having measured this error voltage it will now be applied to reduce distortion. In the circuit of fig.2.3 the attenuated error voltage Vd*R1/R1+R2 is restored to very close to Vd at the lower subsidiary amplifiers output (this restoration is not exact due to tolerances on R1 and R2 as well as finite gain in B) and is

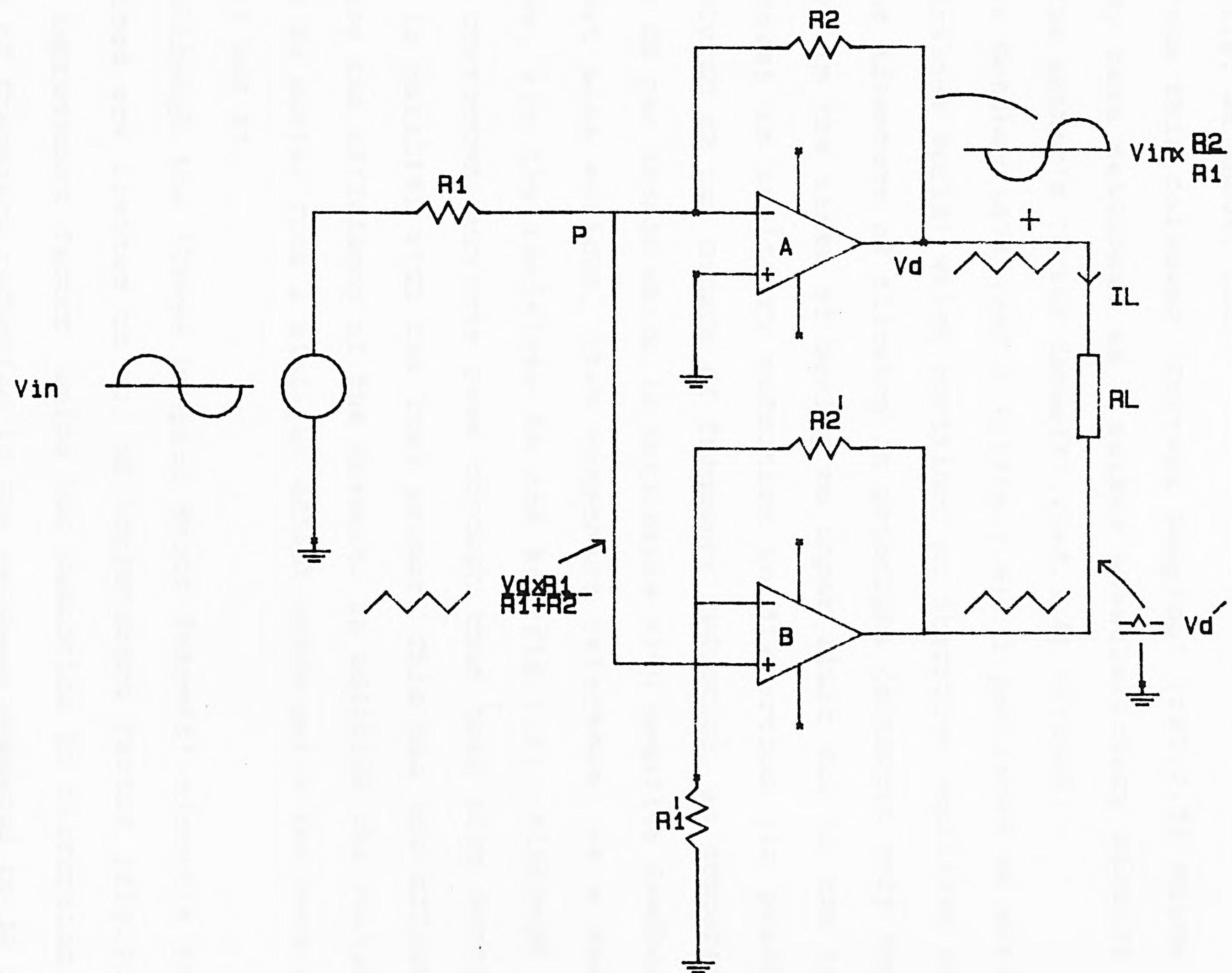


FIG. 2.3 ERROR ADDON

applied in such a way as to subtract or 'takeoff' the error which appears across RL, the load resistor.

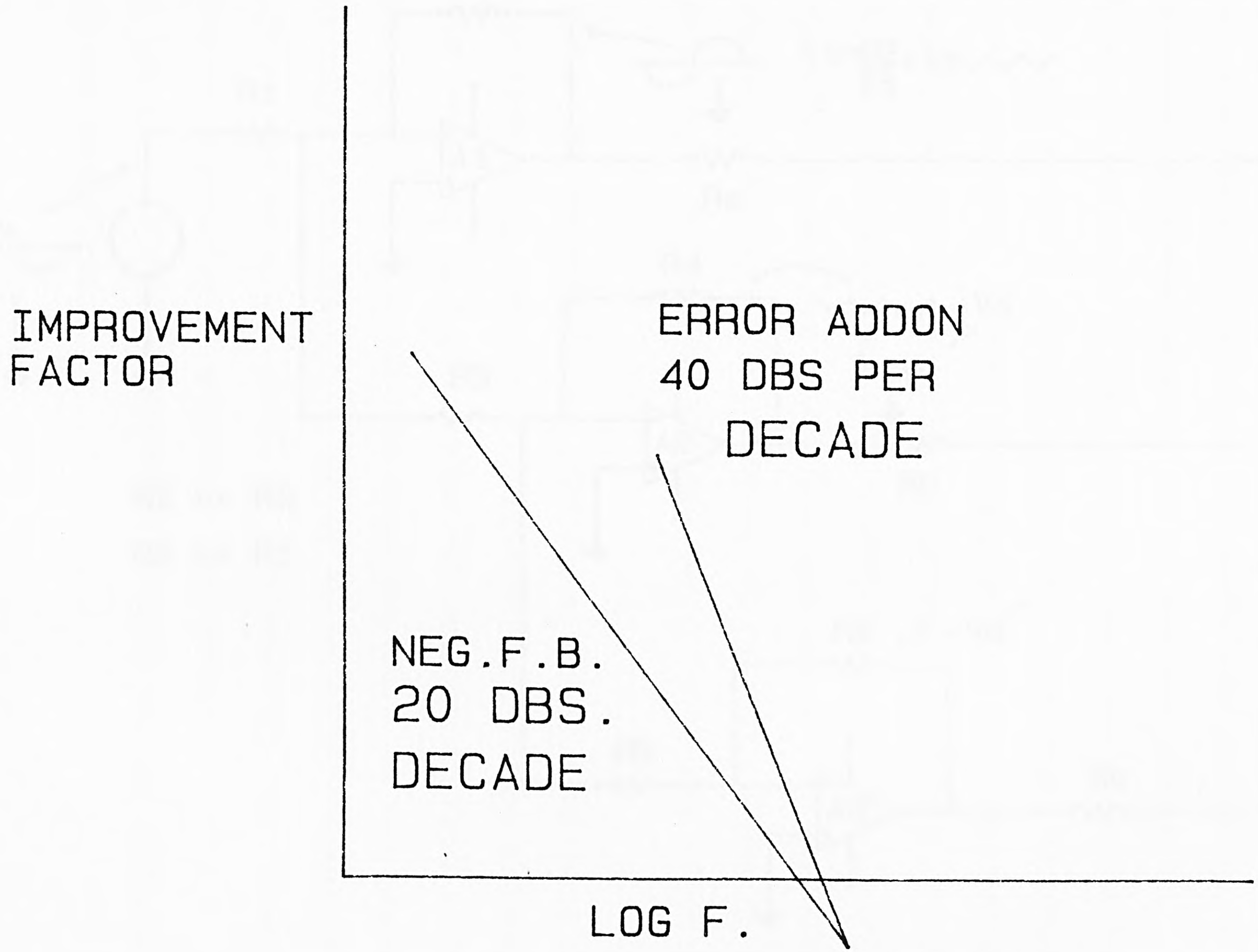
In the pass band as the frequency drops the performance (including distortion reduction) improves at 40 dB per decade instead of the 20 dB per decade of a negative feedback circuit (Fig.2.4.) until the point at which the gain becomes constant typically, at about, 10Hz.

From this followed 'Current Dumping' (ref.2.5) which has already been mentioned as a rather unsatisfactory circuit and then the author's 'Error takeoff',(ref.2.4) circuit.

In October 1974 (ref.2.4)(fig.2.5). I published an article in 'Wireless World' which contained an iterative amplifier which had the advantage of allowing in principle (although only imperfectly, in the sense of having an upper limit due to non ideal resistors) an arbitrary reduction in distortion (in practice probably 60 dB per decade of frequency reduction, as compared to the 20 dB per decade which is obtainable with negative feedback).

But more serious, than component tolerance, as a design problem, are the resistors Rb and Rc.(fig.2.5). Although the error correcton currents pass through them they also serve as loads in parallel with the load proper. This has the effect of reducing the efficiency of the circuit. In addition the resistors Rb and Rc suffer from a similiar effect which calls for more gain from A2 and A3.

Although the 'Cross Coupled Error Takeoff' circuits to be described are limited to 40 dB improvement factor (fig.2.4), (the improvement factor being the reduction in distortion per decade of frequency reduction in the passband compared to 20 dB



IMPROVEMENT FACTOR
ERROR ADDON
40 DBS PER
DECADE
NEG.F.B.
20 DBS.
DECADE
LOG F.

FIG. 2.4

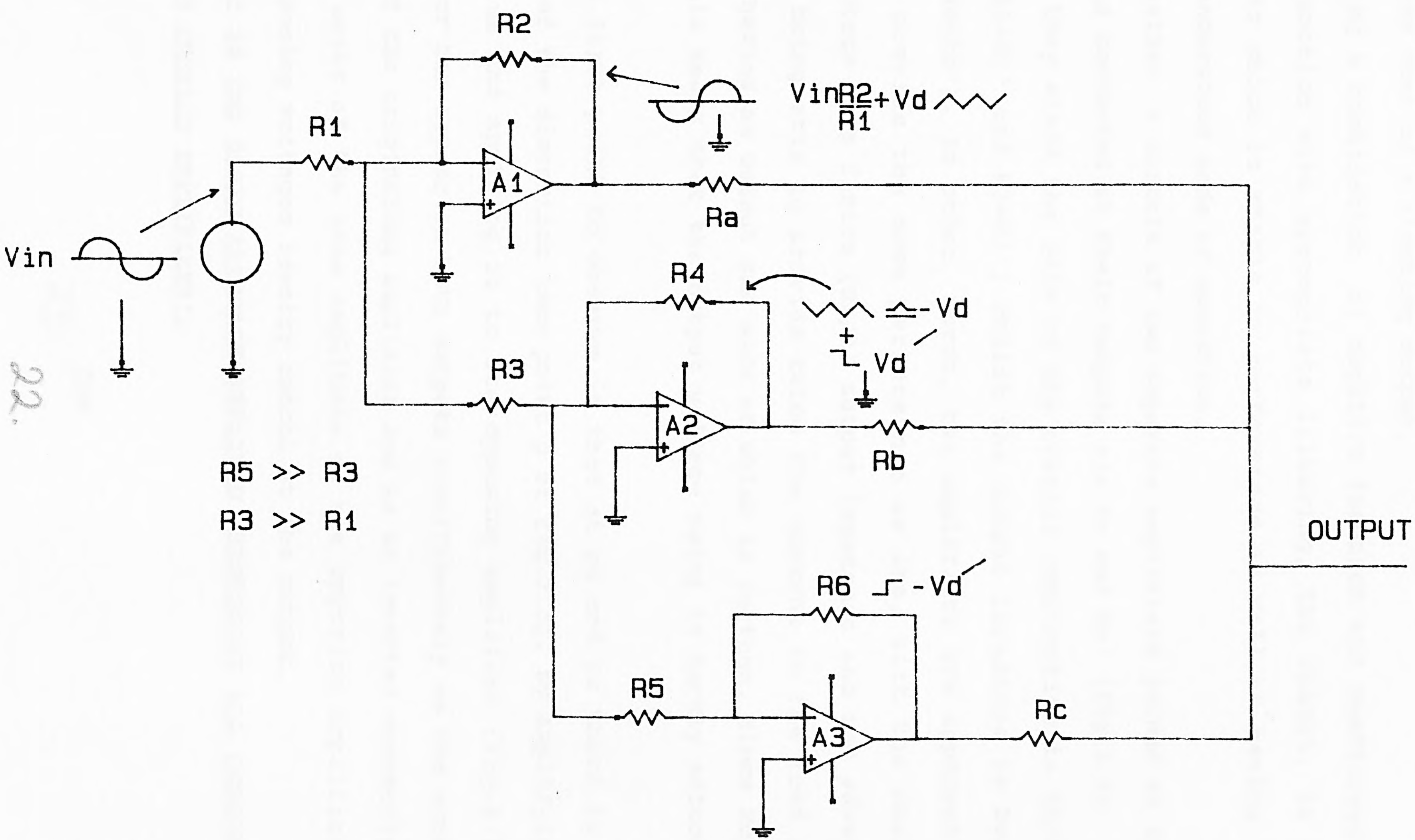

Vin
R1
R2
Vin R2/R1 + Vd
A1
Ra
R4
-Vd
+
Vd
R3
A2
Rb
R5 >> R3
R3 >> R1
OUTPUT
R6
-Vd
R5
A3
Rc

FIG. 2.5

for negative feedback), either are partially free of loading effects due to a finite output impedance or are completely free but at the cost of a floating output.

Using a combination of negative feedback and feedforward, in conjunction with appropriate filtering, the result, is an amplifier which is stable in performance as well as having an easily understood mode of operation.

Consider a circuit of two separate amplifiers joined at the input and connected at their outputs via Ro and Ro' (fig.2.6).

As they stand the gain of the overall combination is R2/R1 (since R1=R1' and R2=R2') whilst the output impedance is Ro/2 (since Ro=Ro'). In other words, two amplifiers are apparently used to provide the same performance as one, with the small disadvantage of a finite (Ro/2) output impedance and the advantage of being able to provide twice the current in the load as well as having an output one side of which is earthed. Since Ro/2 << RL this means that the output voltage swing is hardly affected.

The first point to observe is that at p1 and p2 there is a measure of the distortion (see point p of fig.2.3). By amplifying each value and applying it to its opposing amplifier (fig.2.7) this error can appear at both outputs simultaneously as the error proper of the originating amplifier and as an inverted correction of the error of the same amplitude at the opposing amplifier. These opposing voltages ideally cancel at the output.

THIS IS THE BASIC APPROACH EMPLOYED THROUGHOUT ALL DESIGNS OF 'CROSS COUPLED AMPLIFIERS'.

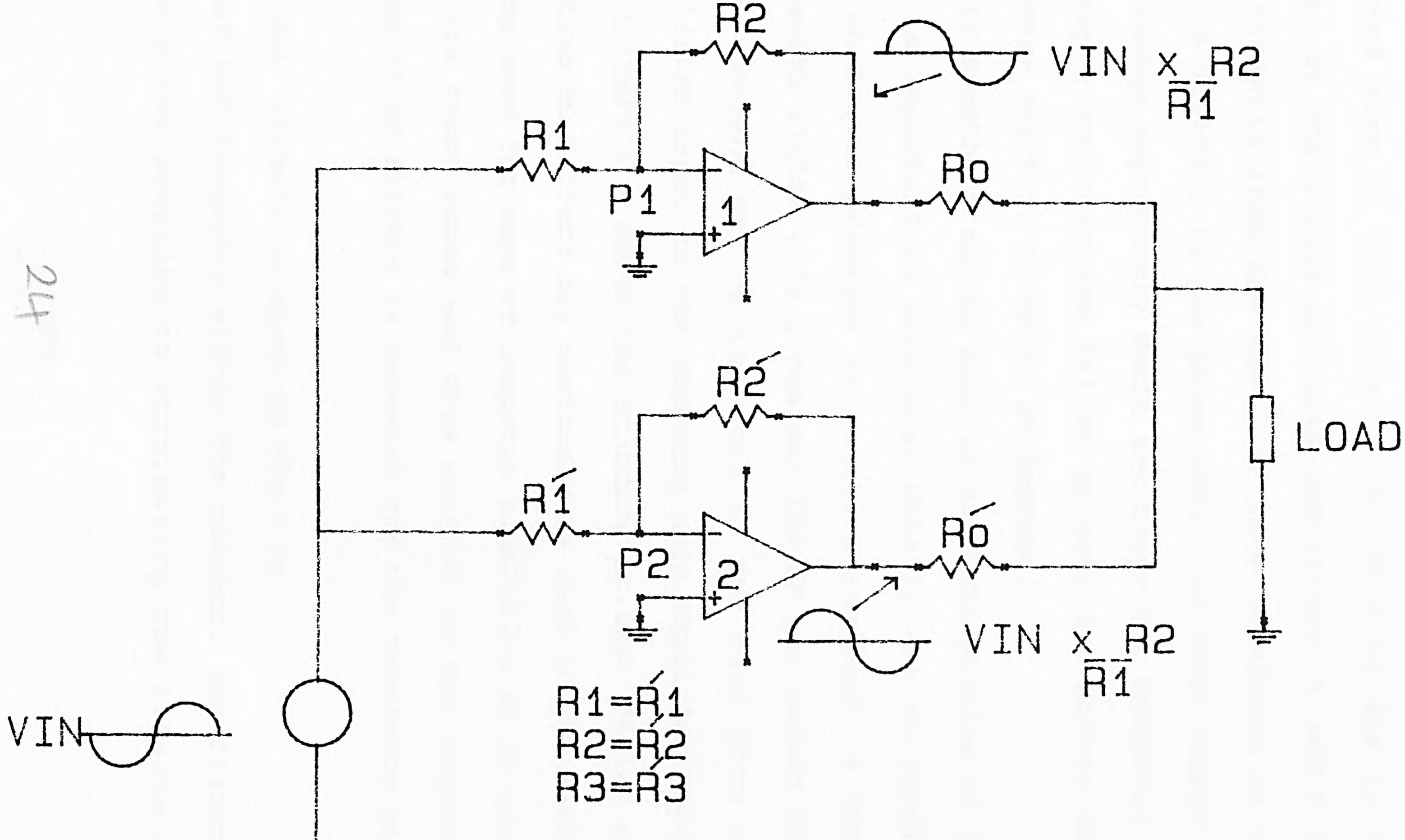

R1
R2
P1
1
Ro
VIN x R2/R1
R1'
R2'
P2
2
Ro'
VIN x R2/R1
LOAD
VIN
R1=R1'
R2=R2'
R3=R3'

FIG. 2.6

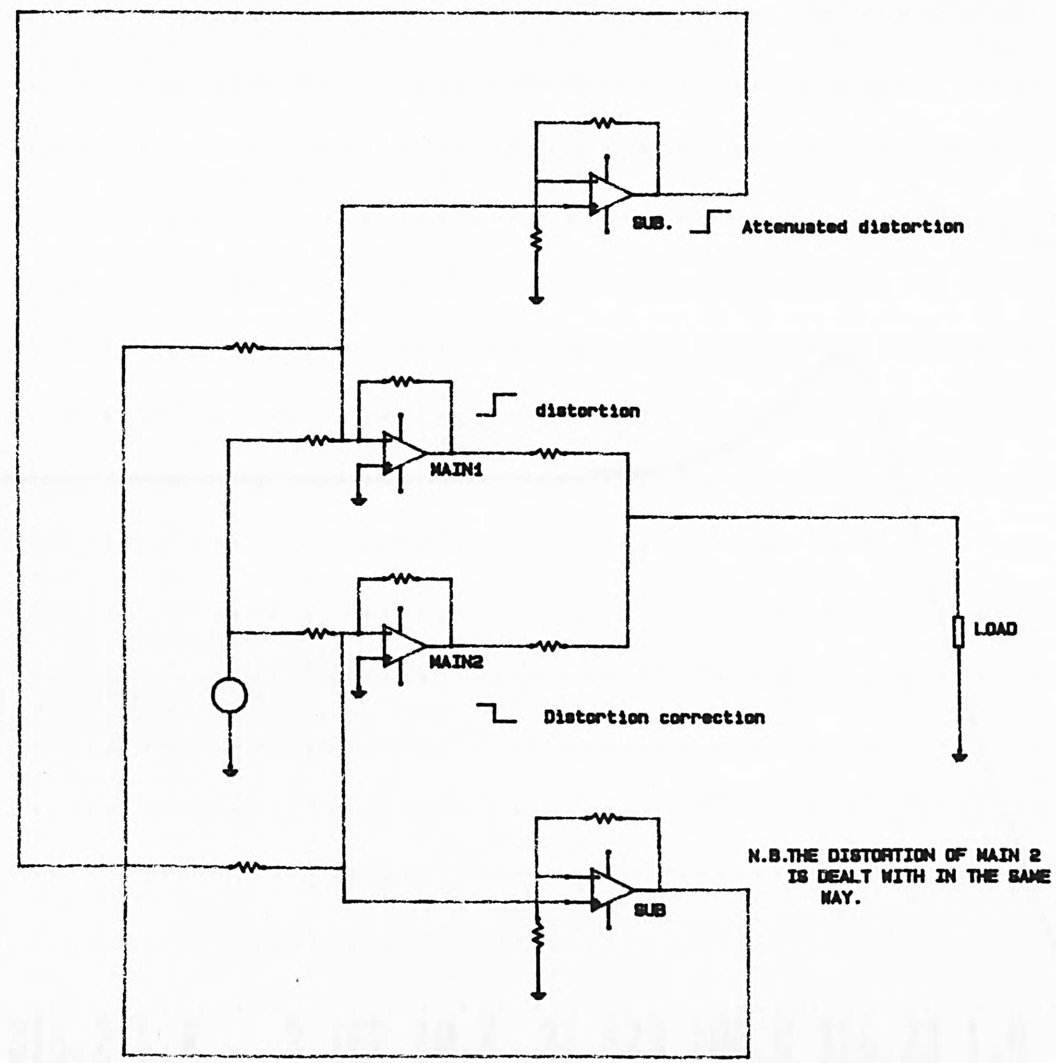
What follows are the results a computer simulation study using a 'Spice' (Simulation Program with Integrated Circuit Emphasis) like Analogue Circuit design package called ECA2.

As the circuit stands in fig 2.7a and using realistic models for the operational amplifiers the response has a pronounced peak at the pass-band edge of 4 dB (fig. 2.7b) This is due to the 'virtual earth' at the inputs of 'main amplifiers' 1 and 2 departing significantly from the ideal of zero impedance at the band edge. It is doubtful if the phase shift has much effect as this does not change appreciably until ten times the frequency at which the op.amp's own gain has fallen to zero dB. Before this happens, the phase shift is close to 90 degrees.

Fortunately something can be done to cure the problem of the gain peak. It is important to note that this is not as readily possible with analogous problems in negative feedback. A three section Butterworth filter with a nominal 120 dB per decade cut-off above its passband may be inserted in the feed from one subsidiary amplifier input to the opposing main amplifier input. (Although LC filters are shown the standard op.amp. Sallen and Key-type of active RC filters may obviously be used in practice.) (LC Filters were used for ease of computer simulation as in practice they require fewer nodes and this speeded up the computations.) This use of LC filters is repeated for the remaining pair of amplifiers.

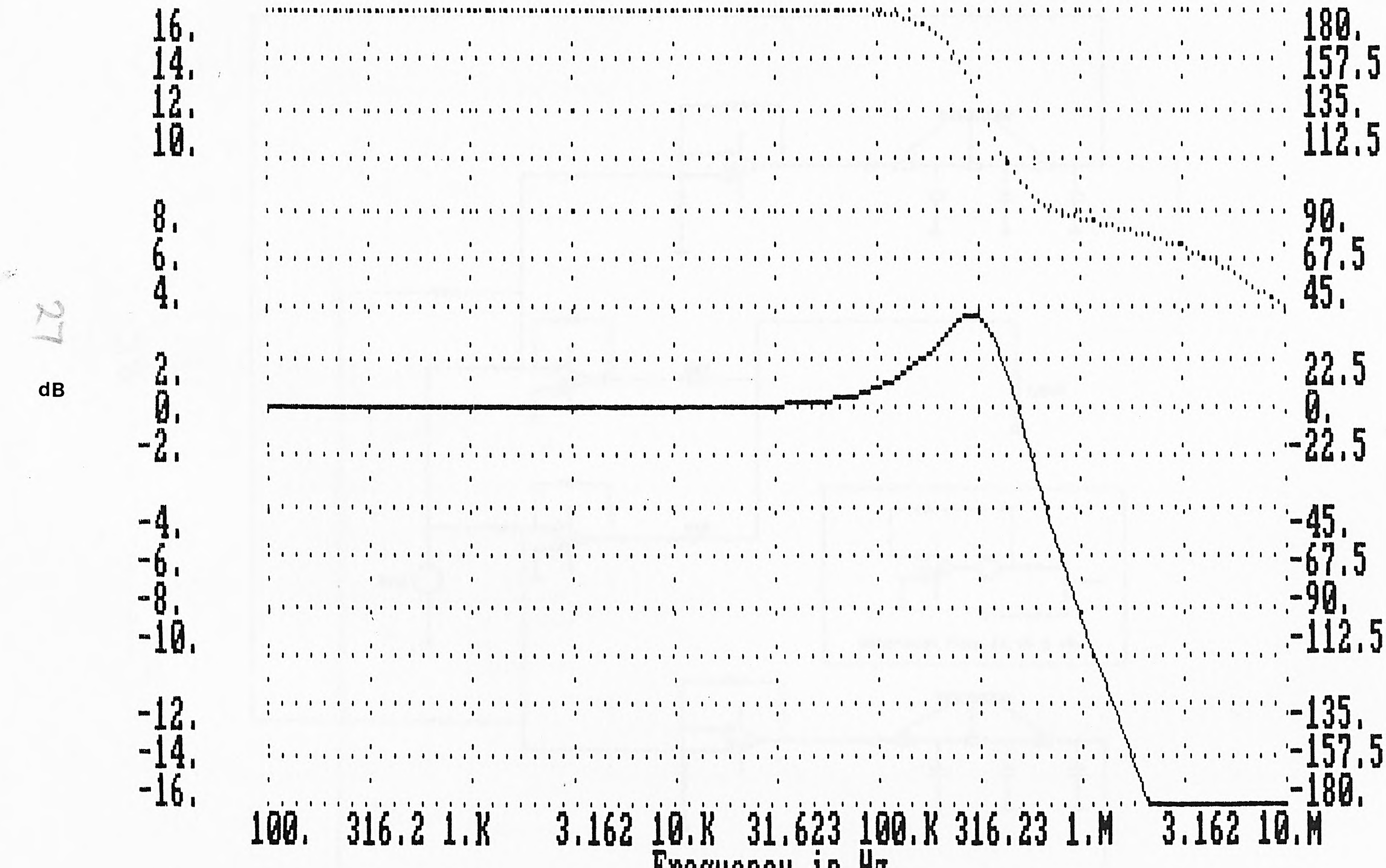
The resultant circuit is shown in Fig.2.9a.

Above a defined frequency within the passband the filters disconnect the cross coupling so eliminating the effects of



ERROR TAKEOFF - SINGLE ENDED VERSION.

FIG. 2.7A



dB
16.
14.
12.
10.
8.
6.
4.
2.
0.
-2.
-4.
-6.
-8.
-10.
-12.
-14.
-16.
180.
157.5
135.
112.5
90.
67.5
45.
22.5
0.
-22.5
-45.
-67.5
-90.
-112.5
-135.
-157.5
-180.
100. 316.2 1.K 3.162 10.K 31.623 100.K 316.23 1.M 3.162 10.M
Frequency in Hz
-->

FIG. 2.7B

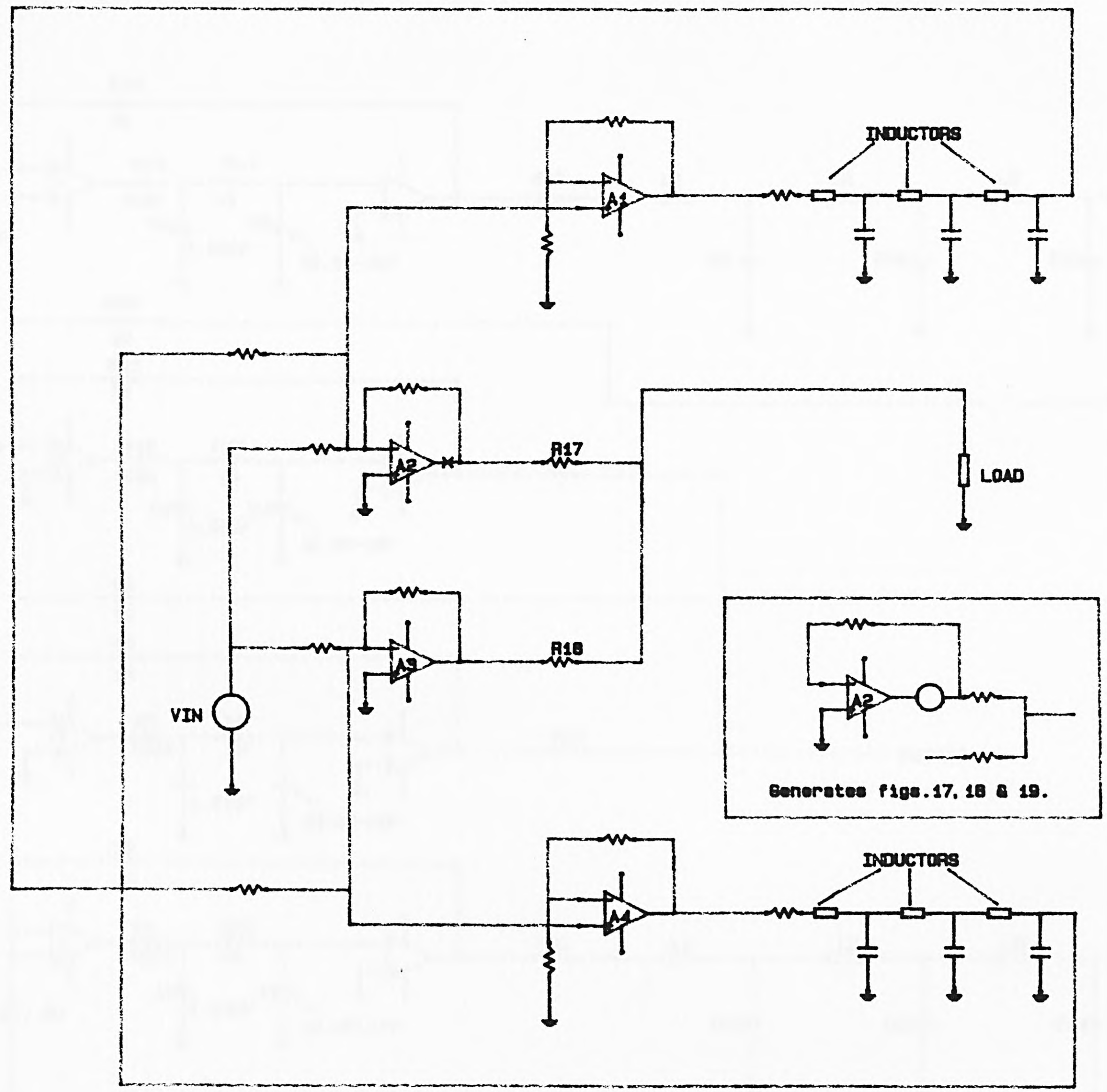

INDUCTORS
A1
A2
R17
LOAD
A3
R18
VIN
A2
Generates figs.17, 18 & 19.
INDUCTORS
A4

FIG. 2.9A

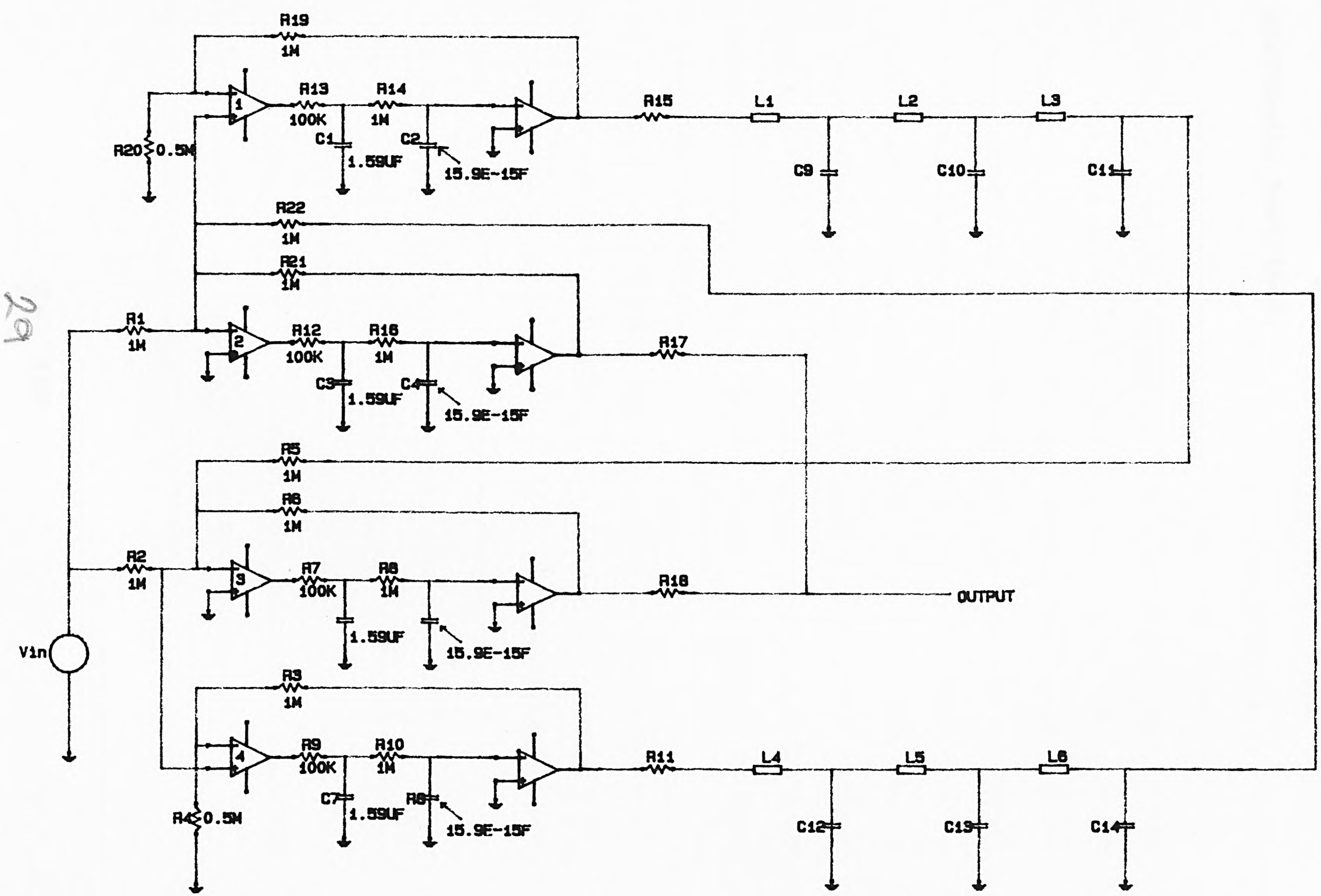

R19
1M
R13
100K
R14
1M
R15
L1
L2
L3
R20
0.5M
C1
1.59UF
C2
15.9E-15F
C9
C10
C11
R22
1M
R21
1M
R1
1M
R12
100K
R16
1M
R17
C3
1.59UF
C4
15.9E-15F
R5
1M
R6
1M
R2
1M
R7
100K
R8
1M
R18
OUTPUT
1.59UF
15.9E-15F
Vin
R3
1M
R9
100K
R10
1M
R11
L4
L5
L6
R4
0.5M
C7
1.59UF
R6
15.9E-15F
C12
C13
C14

FIG. 2.9B

finite gain, which are the cause of the deviation from virtual earth conditions at the amplifiers band edge which cause the 4 dB peak (fig.2.7a the circuit without filters). Within the passband of the filters the circuit performance is close to the idealised operation previously described.

A simple theoretical approach.

The performance at the edge of the passband is not readily calculated using a manual computational approach since, because of the low amplifier gains at the band edge, the impedances of the 'virtual earths ' at the inputs to the appropriate amplifiers have departed from the ideal of zero to some finite value which produces interaction between the two main amplifiers. Although an iterative, algebraic solution could be derived, it would require a numerical solution to tell us much of value. For this reason, as already mentioned, a commercial electronic circuit simulation package was used. This is a modern approach which lifts a heavy algebraic and numerical computation burden from the designers shoulders, much increasing the efficiency of the design process compared to pre-computer times.

However, it is instructive to investigate the operation using conventional circuit analysis of the system for frequencies well within the passband where ideal virtual earths are a reasonable assumption, equivalent to high amplifier gain.

This has been done by myself for relatively simple circuits (ref.2.4) and the result for the simple Error Addon circuit of fig.2.3 is that the distortion is reduced k.A2 (where k is a constant) times compared to a conventional negative feedback circuit consisting of R1, R2 and amplifier one alone.

A very significant improvement over negative feedback.

SOME COMPUTER RESULTS.

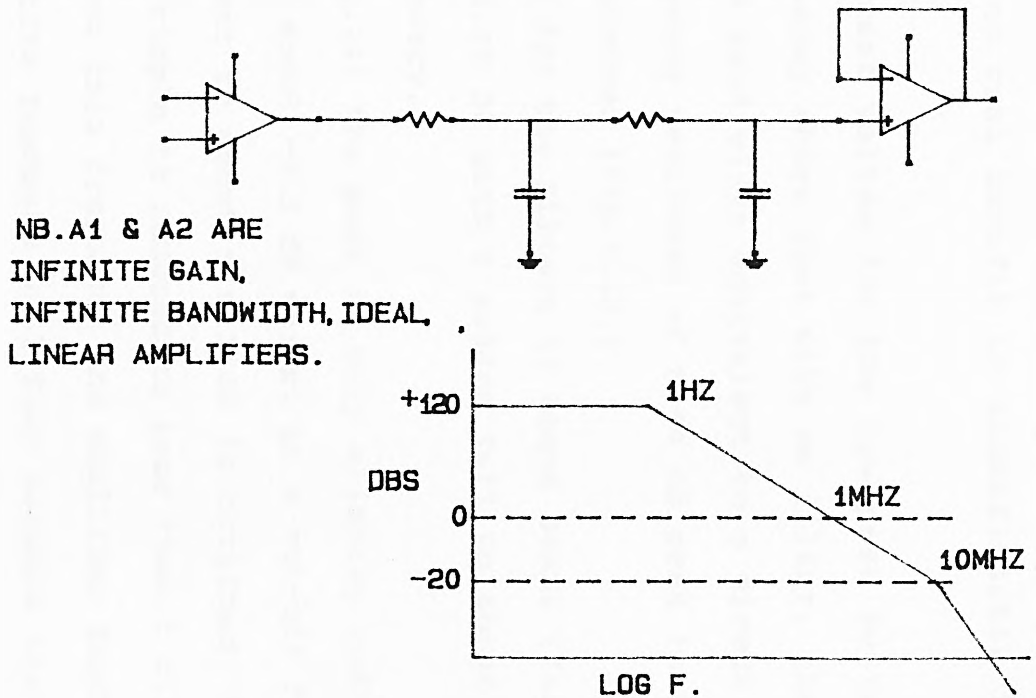
The system under study calls for two lines of investigation, the first is the performance from input to output, and the second is the reduction of error with reducing frequency.

A computer aided design study is ideal for both these purposes since ideas can be tried out reasonably quickly using realistic models, whereas due to the slowness and difficulty in setting up and solving the equations manually only simplified models are feasible. The results as the collective experience of industry has shown can be relied on. Fig.2.11, shows a realistic model for the op.amps. having two break points in the frequency vs. gain plot.

In addition, a well tested computer package such as ECA2 has the important practical advantage that provided the list of components, their associated nodes and their parameter values are entered correctly (which is in my experience is quite feasible for circuits of these types) the results can be relied on. In addition they can be presented graphically in a good quality format and easily printed out.

Of the greatest importance is the fact that vastly more calculations can be done within a reasonably short time, indeed the development of the 'cross coupled ' circuit, would, as far as getting reliable graphical results is concerned have been impossible without ECA2.

The results of using ECA2 to analyse the circuit of fig.2.9a (shown in more detail in fig.2.9b.) give the overall frequency



OP. AMP. COMPUTER MODEL.

FIG. 2.11

response , assume an infinite load, and are shown in figs.2-12,13,14 & 15. They will be discussed shortly.

As for the sake of simplicity the operational amplifiers (fig.2.9b) are assumed to have zero output impedance and the load to be infinite the values of R17 and R18 may have any value.

They are taken entirely arbitrarily as 1 M ohm but could as easily be 1 ohm.

A slightly more realistic model with a realistic load impedance would unnecessarily add to the difficulty of understanding the behaviour with no real benefit in significantly improved accuracy.

Choosing different values for the low pass Butterworth filter cut-off frequency shows that with no filter, that is to say an infinite pass band width equivalent to a direct connection, the result already mentioned of a +4 dB peak in the frequency response is obtained (fig.2.12.).

With a passband for the filters of about 30kHz (fig.2.13.) the peak is about +0.25 dB with a sudden fall to about -0.6 dB before a partial recovery.

At 10kHz (fig.2.14) the peak is only slightly reduced but the fall is now only about -0.3 dB,whilst at a cut-off frequency of 3kHz an improvement to about ± 0.1 dB is obtained. Also, at 3kHz (fig.2.15) the ripple is reduced to less than ± 0.1 dB. It can be seen that above this frequency the amplifier functions as a conventional negative feedback amplifier because the filters eliminate the cross-coupling

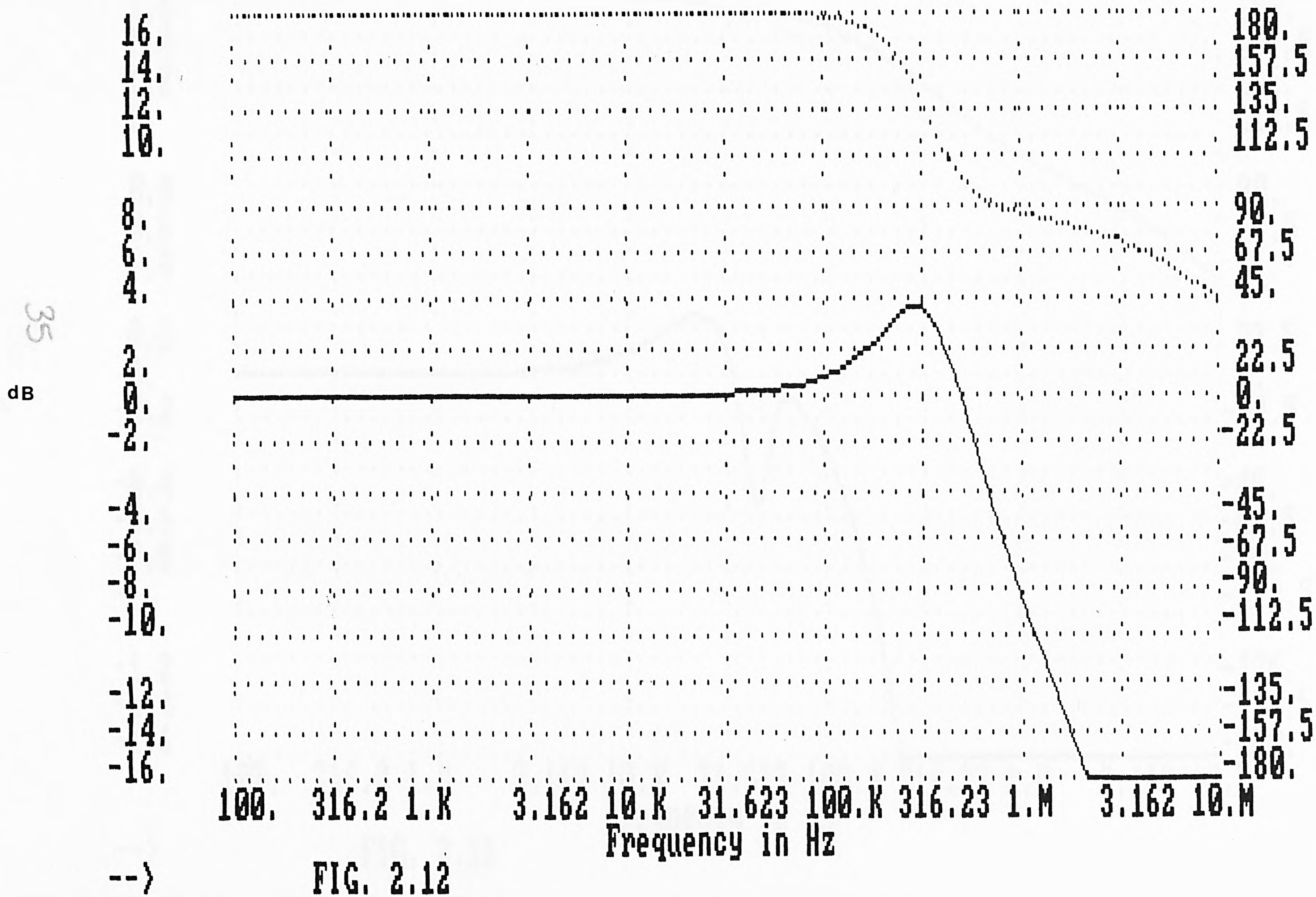

dB
16.
14.
12.
10.
8.
6.
4.
2.
0.
-2.
-4.
-6.
-8.
-10.
-12.
-14.
-16.
180.
157.5
135.
112.5
90.
67.5
45.
22.5
0.
-22.5
-45.
-67.5
-90.
-112.5
-135.
-157.5
-180.
100. 316.2 1.K 3.162 10.K 31.623 100.K 316.23 1.M 3.162 10.M
Frequency in Hz
-->

FIG. 2.12

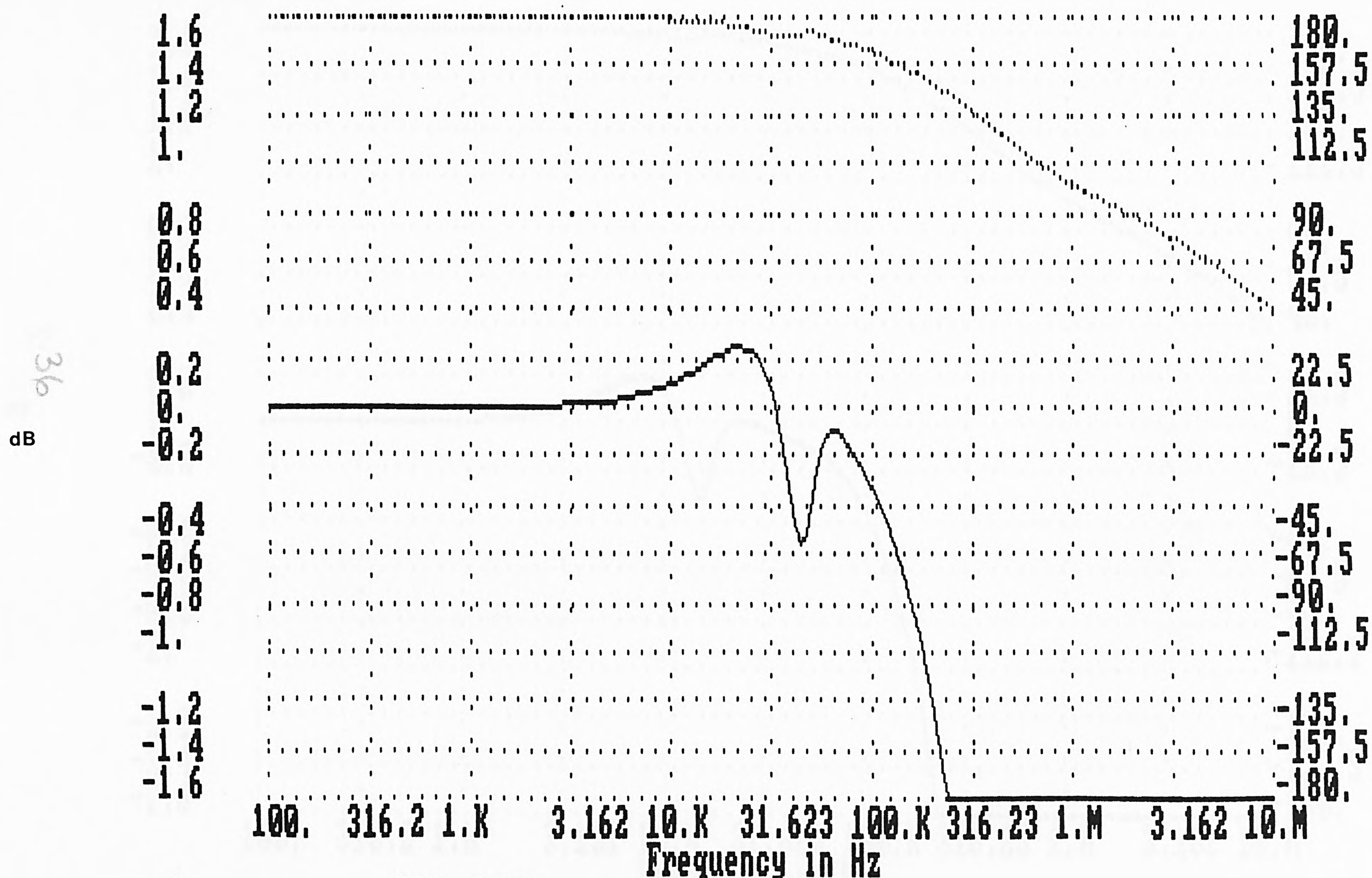

dB
1.6
1.4
1.2
1.
0.8
0.6
0.4
0.2
0.
-0.2
-0.4
-0.6
-0.8
-1.
-1.2
-1.4
-1.6
180.
157.5
135.
112.5
90.
67.5
45.
22.5
0.
-22.5
-45.
-67.5
-90.
-112.5
-135.
-157.5
-180.
100. 316.2 1.K 3.162 10.K 31.623 100.K 316.23 1.M 3.162 10.M
Frequency in Hz
-->

FIG. 2.13

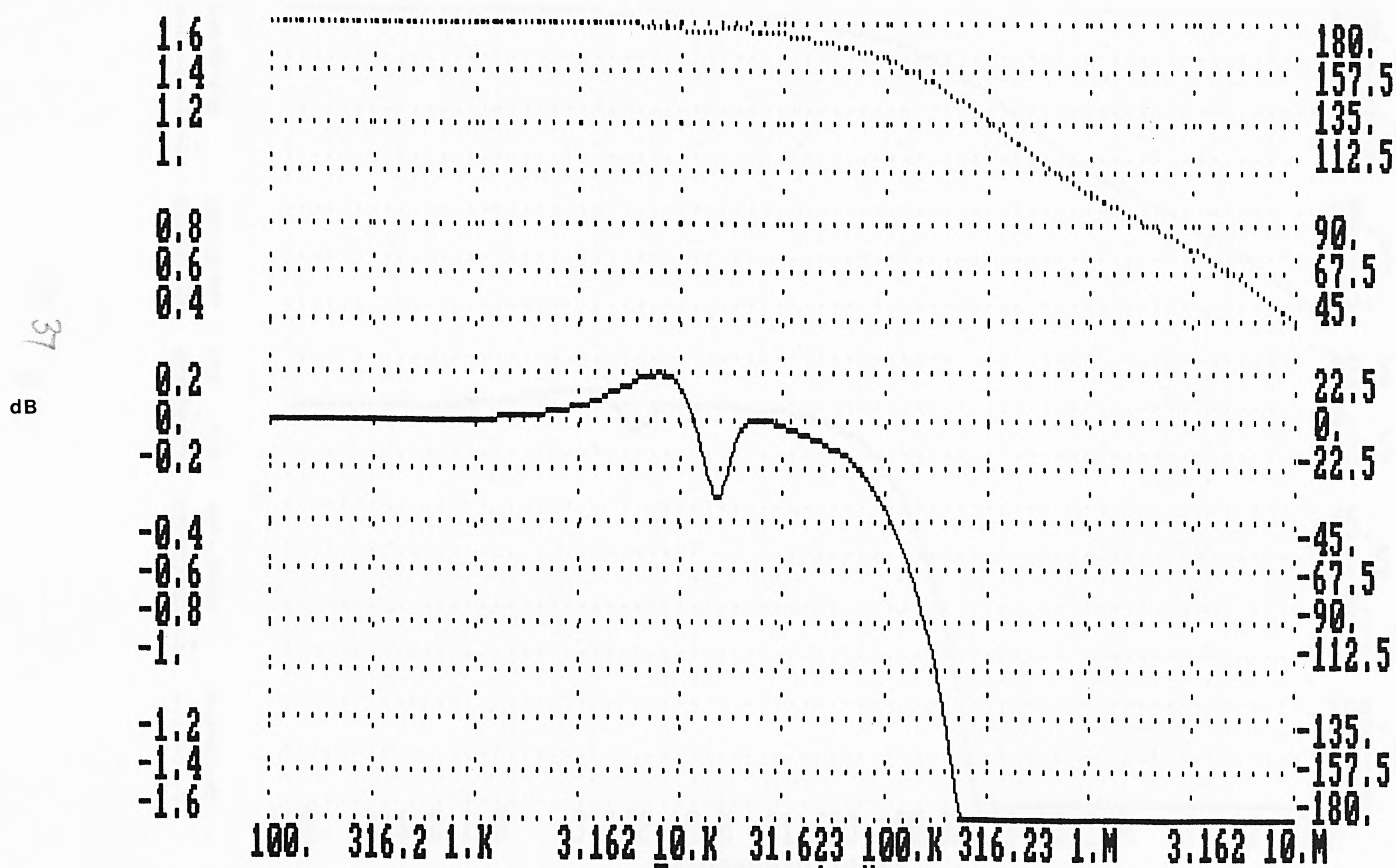

dB
1.6
1.4
1.2
1.
0.8
0.6
0.4
0.2
0.
-0.2
-0.4
-0.6
-0.8
-1.
-1.2
-1.4
-1.6
180.
157.5
135.
112.5
90.
67.5
45.
22.5
0.
-22.5
-45.
-67.5
-90.
-112.5
-135.
-157.5
-180.
100. 316.2 1.K 3.162 10.K 31.623 100.K 316.23 1.M 3.162 10.M
Frequency in Hz
-->

FIG. 2.14

37


db
1.6
1.4
1.2
1.
0.8
0.6
0.4
0.2
0.
-0.2
-0.4
-0.6
-0.8
-1.
-1.2
-1.4
-1.6
180.
157.5
135.
112.5
90.
67.5
45.
22.5
0.
-22.5
-45.
-67.5
-90.
-112.5
-135.
-157.5
-180.
100. 316.2 1.K 3.162 10.K 31.623 100.K 316.23 1.M 3.162 10.M
Frequency in Hz
--)

FIG. 2.15

38

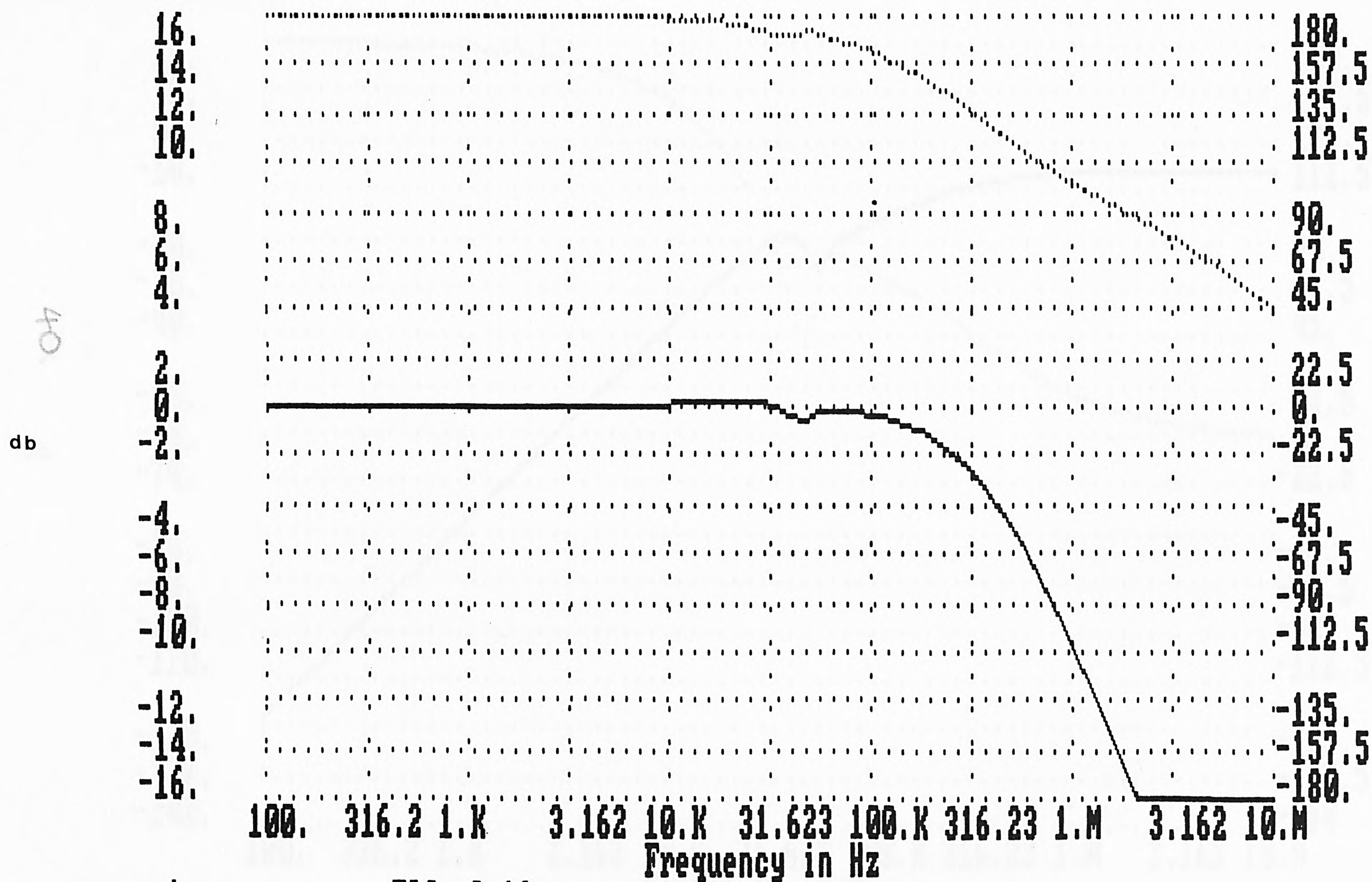
As fig 2.16 (Fig 2.13 replotted with extended Y ordinate) clearly shows, outside the immediate vicinity of the band edge the performance is unaffected.

By inserting an oscillator directly between one amplifier output and its feedback resistors and keeping its amplitude constant while varying the frequency a plot may be made of the effect of error against frequency (figs.2.17,2.18 & 2.19).

This measurement enables an exact forecast to be made of the effect on harmonic distortion which the circuit has since, by Fourier analysis, a waveform, may be analysed as consisting of a collection of sine waves each of which is subject to distortion, each sine wave may be analysed separately in the sense that because of the amplifier distortion it will have added to it a series of harmonics. Each harmonic member of the series will have its value reduced by the amount given by the value of the improvement factor, for example in fig.2.17 at 1 kHz. the second harmonic of 500 Hz. wave generated by amplifier distortion will have its value reduced by 76 dB.

Examining these figures which are associated with figs.2.13, 2.14, and 2.15, (fig.2.13 corresponds to fig 2.17, fig 2.14 to 2.18 and fig.2.15 to 2.19) the fundamental improvement due to the 'Cross Coupling is obvious.

Below about 30kHz. for fig.2.17, 10kHz for fig.2.18 and 3kHz.for fig.2.19 the improvement is 40 dB per decade and above it is 20 dB per decade.


16.
14.
12.
10.
8.
6.
4.
2.
0.
-2.
-4.
-6.
-8.
-10.
-12.
-14.
-16.
db
180.
157.5
135.
112.5
90.
67.5
45.
22.5
0.
-22.5
-45.
-67.5
-90.
-112.5
-135.
-157.5
-180.
100. 316.2 1.K 3.162 10.K 31.623 100.K 316.23 1.M 3.162 10.M
Frequency in Hz

--> FIG. 2.16

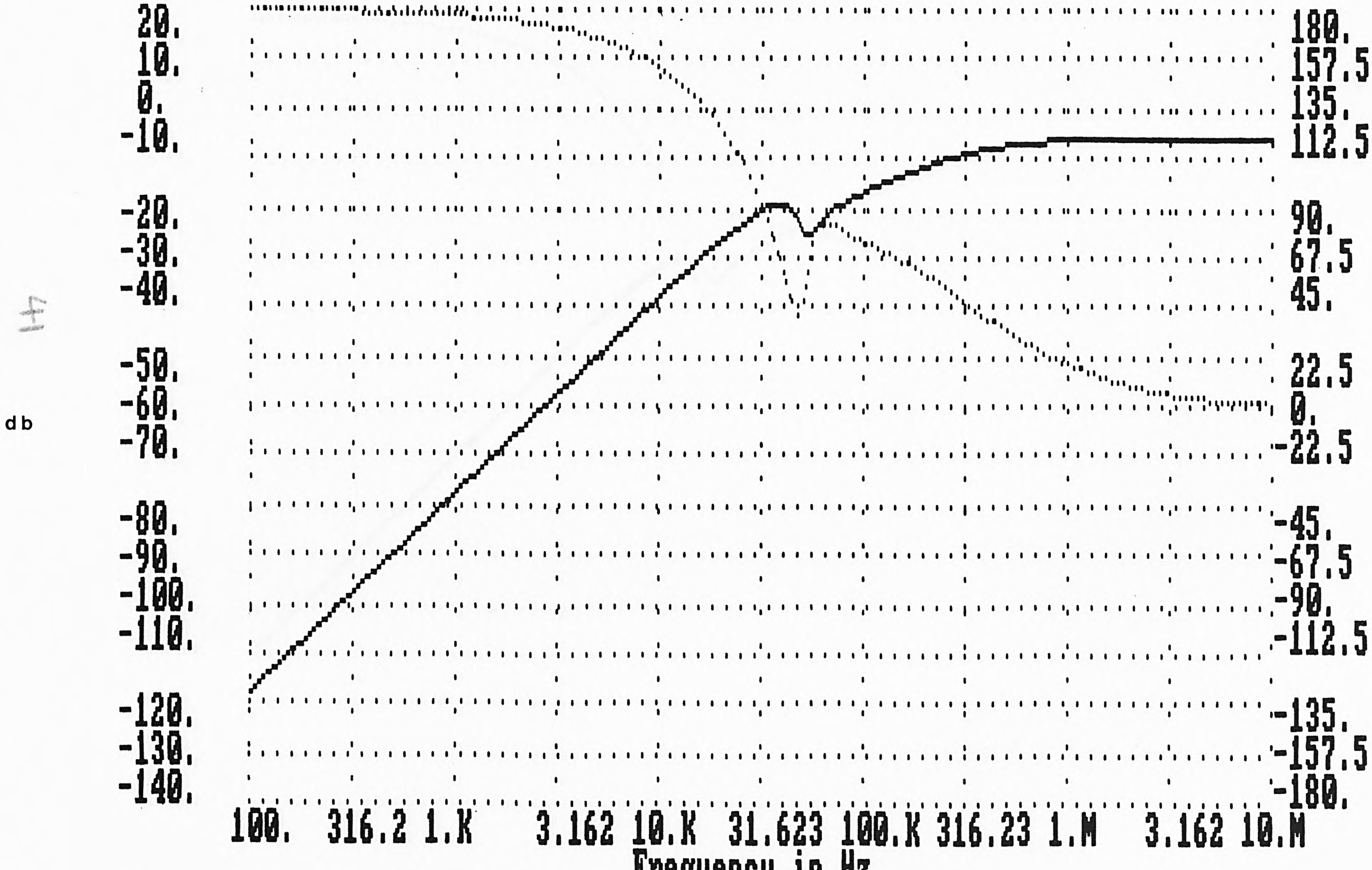

db
20.
10.
0.
-10.
-20.
-30.
-40.
-50.
-60.
-70.
-80.
-90.
-100.
-110.
-120.
-130.
-140.
180.
157.5
135.
112.5
90.
67.5
45.
22.5
0.
-22.5
-45.
-67.5
-90.
-112.5
-135.
-157.5
-180.
100. 316.2 1.K 3.162 10.K 31.623 100.K 316.23 1.M 3.162 10.M
Frequency in Hz
-->

FIG. 2.17

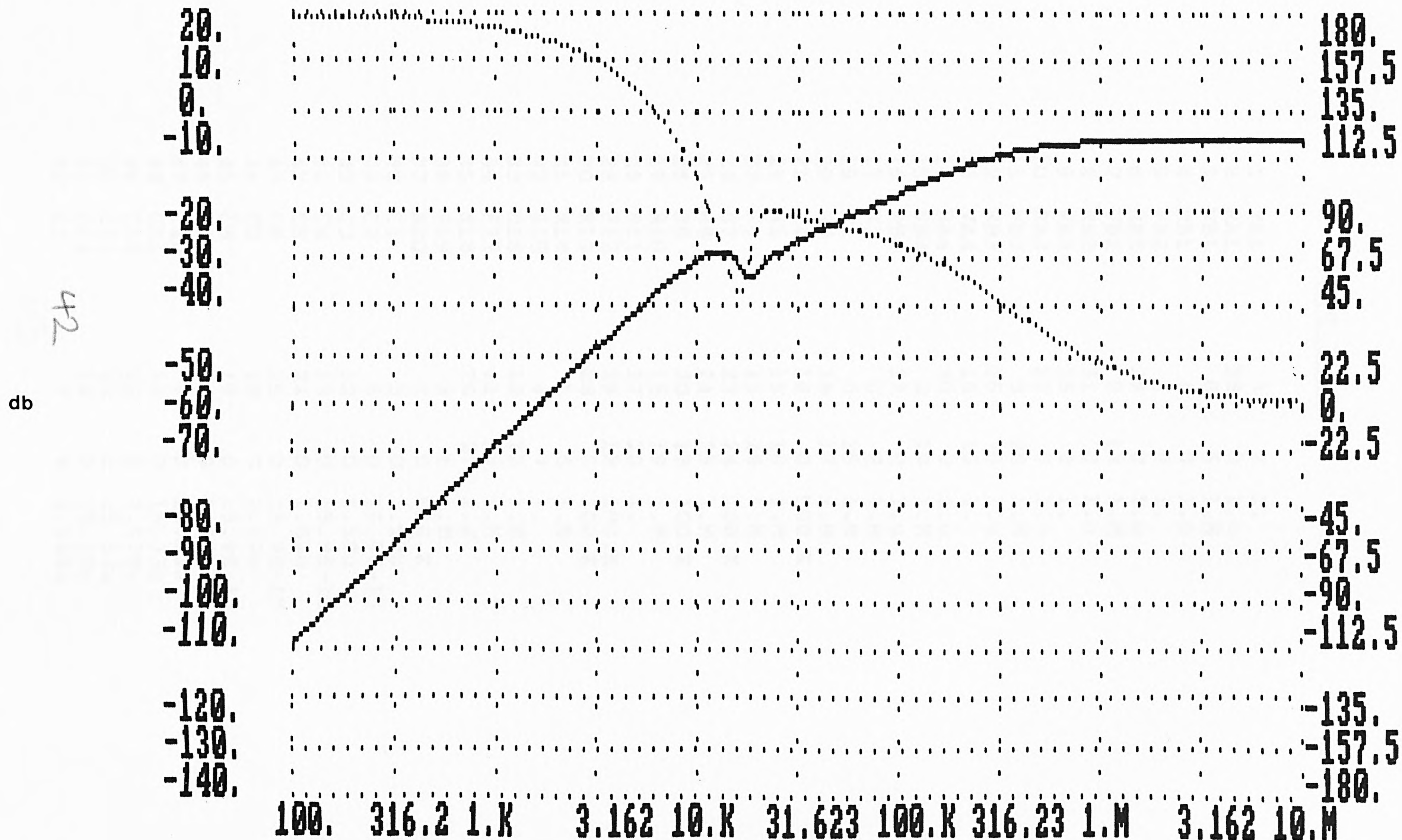

db
20.
10.
0.
-10.
-20.
-30.
-40.
-50.
-60.
-70.
-80.
-90.
-100.
-110.
-120.
-130.
-140.
180.
157.5
135.
112.5
90.
67.5
45.
22.5
0.
-22.5
-45.
-67.5
-90.
-112.5
-135.
-157.5
-180.
100. 316.2 1.K 3.162 10.K 31.623 100.K 316.23 1.M 3.162 10.M
Frequency in Hz
-->

FIG. 2.18

FIG.2.18

1	VA1	3	2	1.
2	RA1	28	0	1.u
3	VB1	0	6	1.M
4	RB1	4	0	1.u
5	VA2	2	0	1.
6	RA2	29	0	1.u
7	VB2	0	13	1.M
8	RB2	31	0	1.u
9	VA3	15	0	1.
10	RA3	23	0	1.u
11	VB3	0	25	1.M
12	RB3	26	0	1.u
13	VA4	16	15	1.
14	RA4	30	0	1.u
15	VB4	0	19	1.M
16	RB4	17	0	1.u
17	R1	0	2	1.M
19	R2	0	15	1.M
20	R3	16	17	1.M
21	R4	16	0	500.K
22	R5	15	10	1.M
23	R6	15	26	1.M
24	R7	23	24	100.K
25	R8	24	25	1.M
26	R9	30	18	100.K
27	R10	18	19	1.M
28	R11	17	20	1.
29	R12	29	12	100.K
30	R13	28	5	100.K
31	R14	5	6	1.M
32	R15	4	7	1.
33	R16	12	13	1.M
34	R17	11	27	1.M
35	R18	26	27	1.M
36	R19	3	4	1.M
37	R20	3	0	500.K
38	C1	5	0	1.59u
39	C2	6	0	15.9E-15
40	C3	12	0	1.59u
41	C4	13	0	15.9E-15
42	C5	24	0	1.59u
43	C6	25	0	15.9E-15
44	C7	18	0	1.59u
45	C8	19	0	15.9E-15
46	C9	8	0	7.579u
47	C10	9	0	15.53u
48	C11	10	0	15.33u
49	C12	21	0	7.579u
50	C13	22	0	15.53u
51	C14	14	0	15.33u
52	L1	7	8	2.588u

53	L2	8	9	12.02u
54	L3	9	10	17.59u
55	L4	20	21	2.588u
56	L5	21	22	12.02u
57	L6	22	14	17.59u
58	VI	0	0	1.
59	RI	11	31	1.u
60	R21	2	11	1.M
61	R22	2	14	1.M

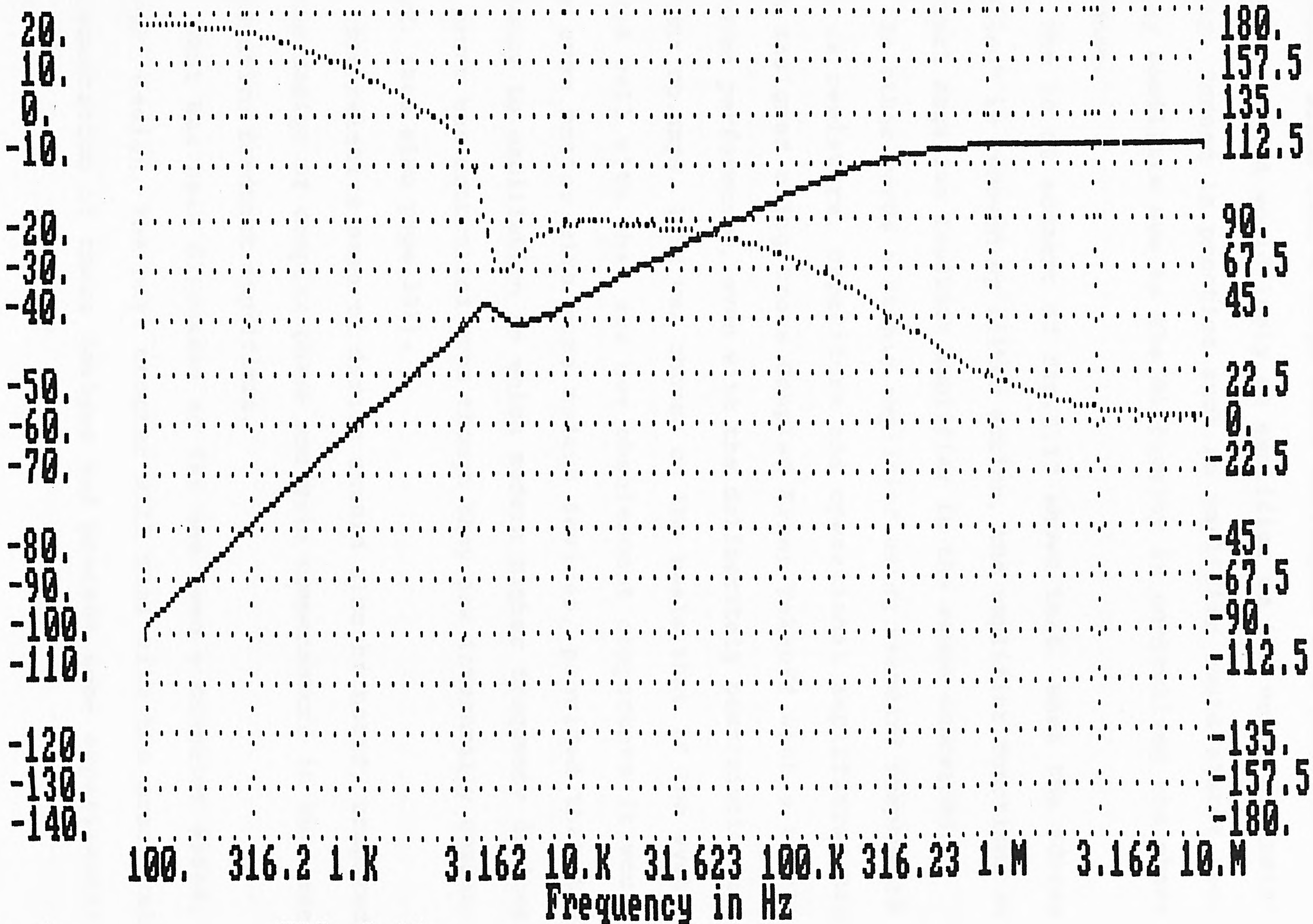

20.
10.
0.
-10.
-20.
-30.
-40.
-50.
-60.
-70.
-80.
-90.
-100.
-110.
-120.
-130.
-140.
db
180.
157.5
135.
112.5
90.
67.5
45.
22.5
0.
-22.5
-45.
-67.5
-90.
-112.5
-135.
-157.5
-180.
100. 316.2 1.K 3.162 10.K 31.623 100.K 316.23 1.M 3.162 10.M
Frequency in Hz
-->

FIG. 2.19

40 dB per decade is in practice unobtainable with Negative feedback, since such a slope would correspond to a phase shift of 180 degrees and would imply an amplifier on the verge of instability. Indeed in practice such an amplifier would almost certainly oscillate due to the difficulty in controlling the phase response.

The 20 dB segment of fig.2.17 shows that when the "Cross Coupling" is removed by filter action, the amplifier functions as a normal Negative feedback amplifier in the range concerned.

In other words a stable amplifier using standard components such as resistors, capacitors and operational amplifiers, has been designed using Cross Coupled Error Takeoff with a vastly improved performance, even with the deliberately pessimistic use of 1MHz op.amps. 1MHz was chosen on the basis that if the system worked well with what are now obsolescent components it would work even better with more modern devices, provided that the tendency to oscillation to which modern higher frequency devices are prone was controlled even though they are internmally compensated. (See also page 301).

The relative ease of design should also be noted, compared to the design of complex phase and gain compensators in the case of negative feedback amplifiers.

What has been discussed so far has been a computer aided, 'paper' design. The next chapter will deal with the practical implementation of these designs and present some experimental results.

CURRENT DUMPING

As the task of writing up was proceeding it occurred to me that it was quite possible to use ECA2 on Current Dumping (ref.2.5) and so present performance details in a comparable form while avoiding discussion as to how it did, or did not work.

The justification for a new scheme for Class B amplifiers may be questioned on the grounds that an existing scheme, Current Dumping, works perfectly well. However my investigations have shown that Current Dumping has defects which the new scheme overcomes.

It occurred to me (fig.2.20) that there is a distinct range in voltage between the values at which T1 ceases to conduct and T2 starts conducting and that within this range the circuit changes from the fig.2.20 structure to the fig.2.21 structure. I.E. the two transistors are effectively out of the circuit. (In my computer simulation I have substituted a voltage follower for the transistors during the region of conduction in the circuit of fig.2.21, for both convenience and also in fairness to Current Dumping, to demonstrate that any anomalous results cannot be attributed to a low gain transistor.

If the circuit did not suffer from cross over distortion the gain frequency response would be the same for both circuits, since cross over distortion is caused by a change in amplifier characteristics as the amplifier output changes range.

In a nutshell it is not!

ECA2 was used to plot the results shown in the following

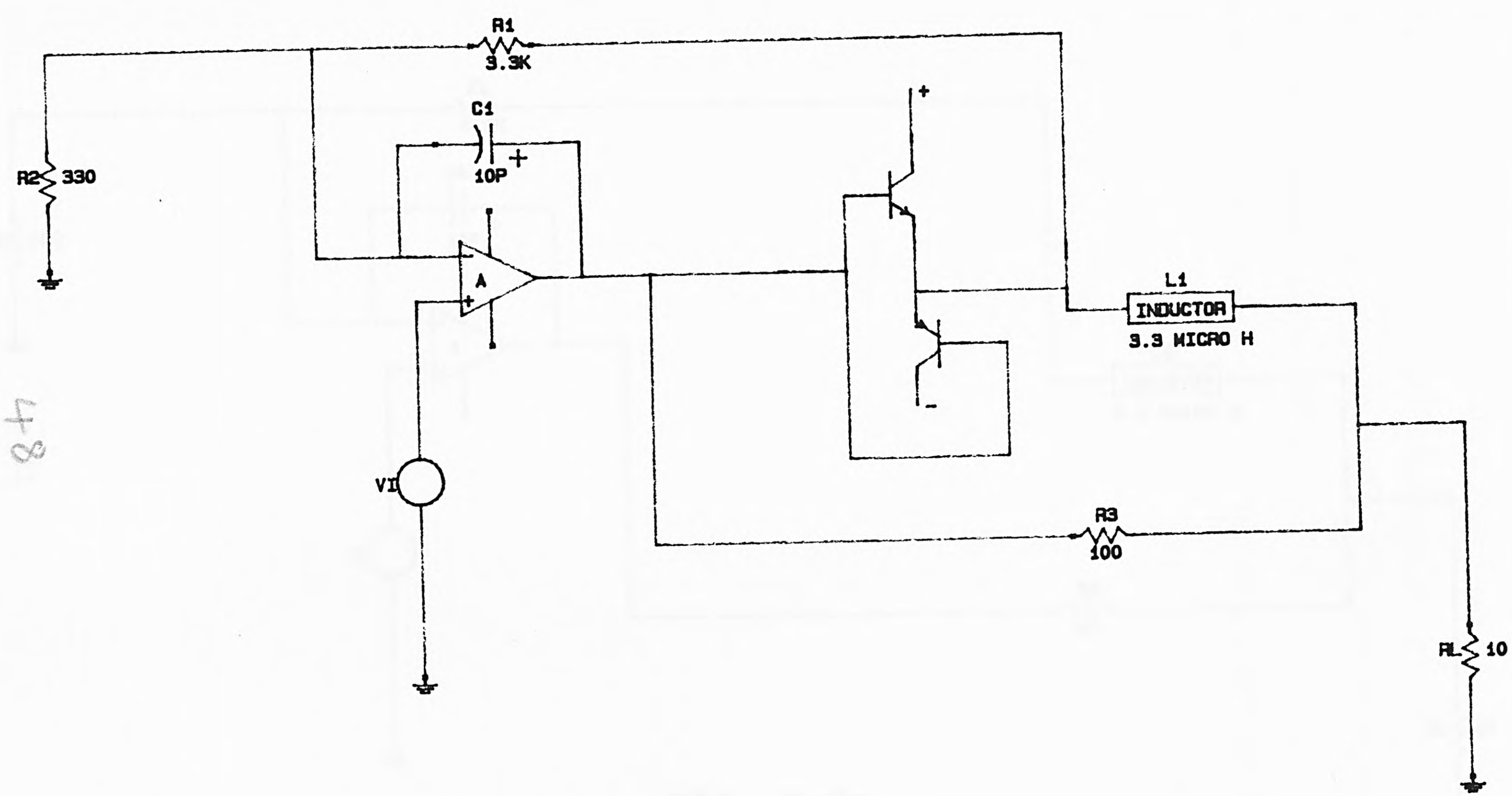


FIG. 2.20

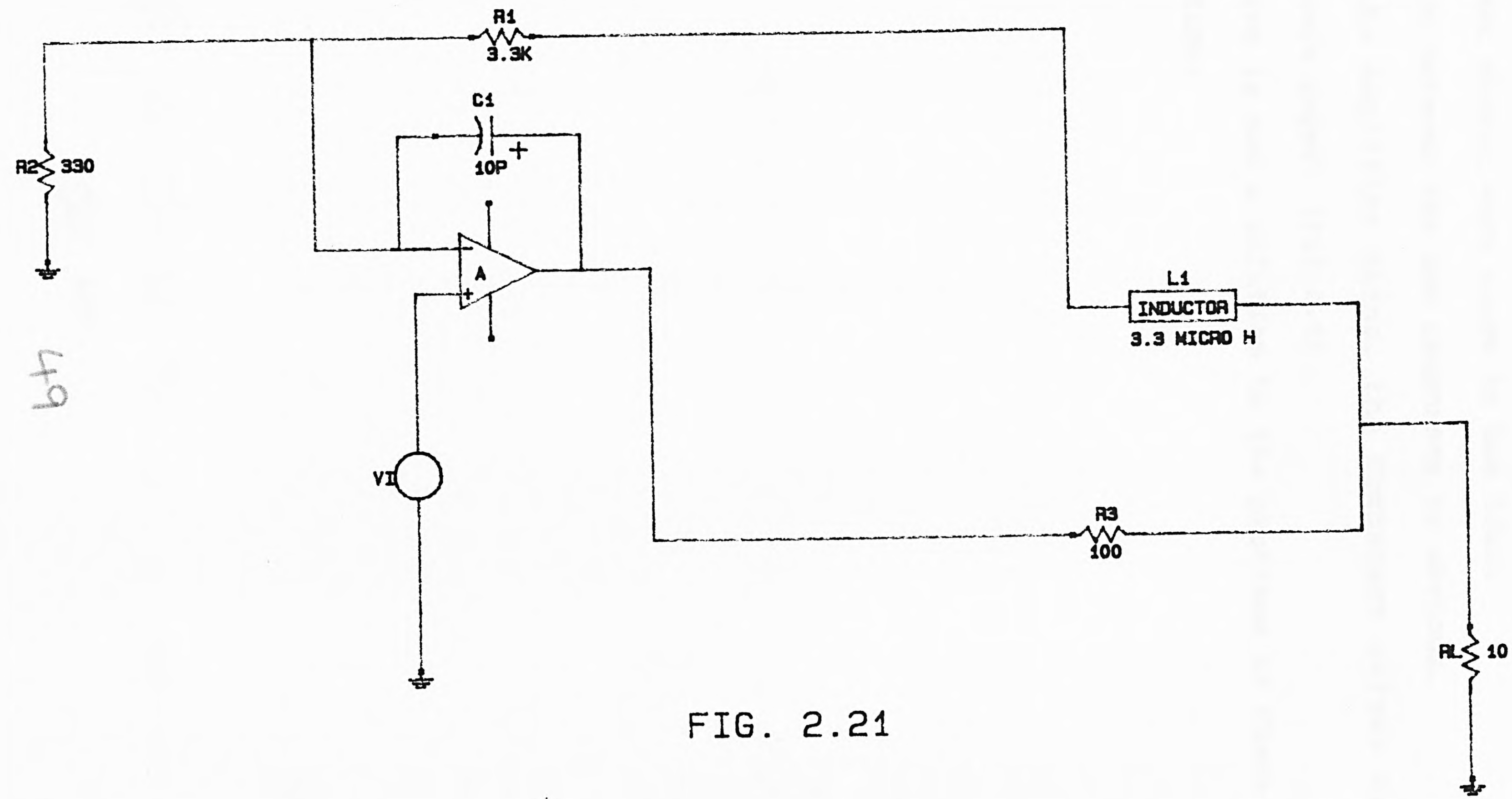


FIG. 2.21

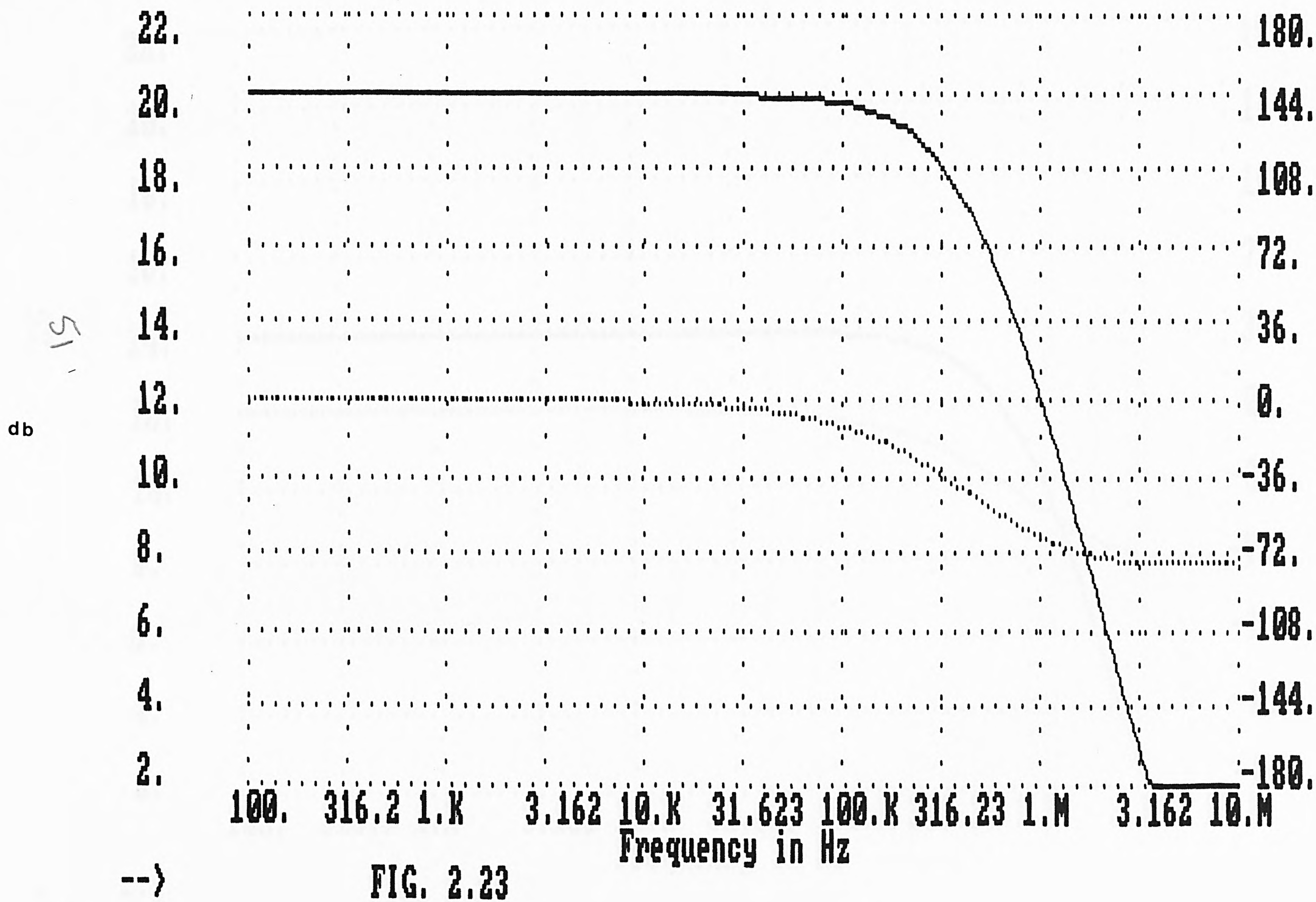
diagrams.

Fig.2.23 gives the results obtained from fig.2.20 and fig.2.24 gives those from fig.2.21. The gain ascribed to amplifier A is 100 and that to amplifier B is 1,000,000 in order that the voltage follower should work close to the ideal.

The difference between the two responses is obvious.

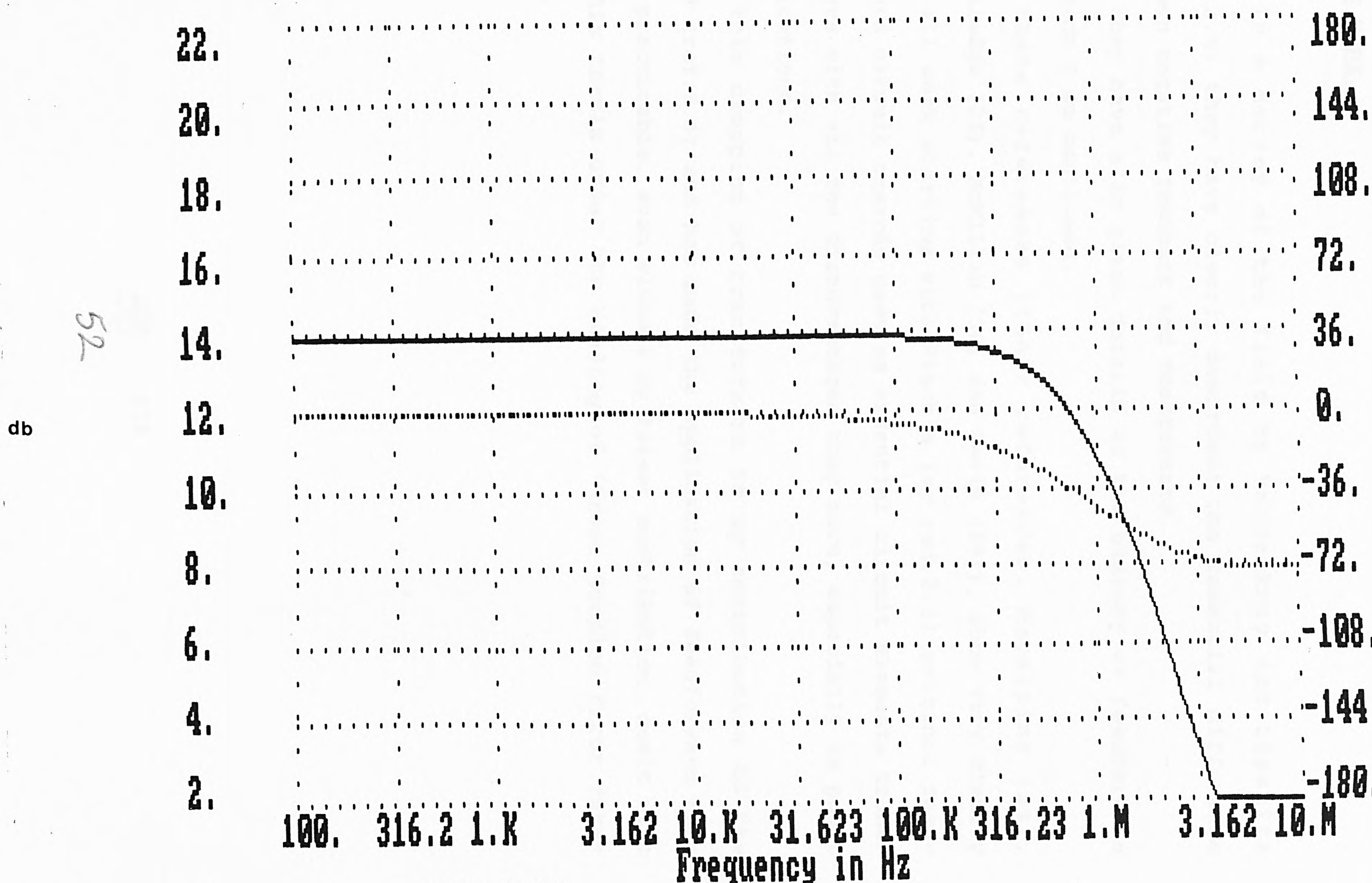
Apart from the amplifier gains, the component values are selected from Walkers paper. (ref.2.5).

Current dumping is not a solution to the problems of Class B cross over distortion.



db
22.
20.
18.
16.
14.
12.
10.
8.
6.
4.
2.
180.
144.
108.
72.
36.
0.
-36.
-72.
-108.
-144.
-180.
100.
316.2
1.K
3.162
10.K
31.623
100.K
316.23
1.M
3.162
10.M
Frequency in Hz
-->

FIG. 2.23



db
22.
20.
18.
16.
14.
12.
10.
8.
6.
4.
2.
180.
144.
108.
72.
36.
0.
-36.
-72.
-108.
-144.
-180.
100. 316.2 1.K 3.162 10.K 31.623 100.K 316.23 1.M 3.162 10.M
Frequency in Hz

--〉

FIG. 2.24

PRIOR WORK.

In a survey of the field by Vanderkooy and Lipshitz (ref.2.8) they have clearly described the essential difference between negative feedback and feedforward.

They have also given details of the history of feedforward in which I am mentioned.

These references (their references. Macalpine (24), Ketchledge (25), McMillan (20) and Davis (29)), show very clearly that all work starting with Black's (my ref.2.1) original feedforward circuit onwards uses as essential circuit elements transformers with all the disadvantages they have, especially in power applications.

This dropping of transformers is my contribution to the field (ref.2.4) and has made the application of feedforward much more practicable, even without my latest contribution, dealt with in this thesis under the heading of 'Cross Coupled Error Take-off'.

References.

2.1. Translating System, H.S.Black, U.S.Patent 2102671.

2.2. Multiple feedback Systems,McMillan,U.S.Patent 2748201,May 1956.

2.3. Reducing Amplifier Distortion by Error Addon, A.M.Sandman, Wireless World, Jan. 1973, p.32.

2.4. Reducing Amplifier Distortion, A.M.Sandman, Wireless World, Oct.1974, pp.367-371.

2.5. Current Dumping Audio Amplifier, P.J.Walker. Wireless World, Dec.1975. pp 560-562.

2.6. Pioneers, Atherton, Electronics & Wireless World, Feb. 1989, pp. 194-195.

2.7. Letters. Wireless World, April.1976, pp.54-55. July 1976, pp.60-62. Nov. 1976, pp.52-54

2.8. Feedforward Error Correction in Power Amplifiers, Vanderkooy & Lipshitz, J. Audio Eng. Soc., Jan.-Feb.1980. pp.2-16

3 RESULTS FROM THE PRACTICAL IMPLEMENTATION OF A CROSS-COUPLED AMPLIFIER

My initial plans did not include any experimental models on the grounds that modern circuit analysis packages could demonstrate all that was needed to know about a circuit and its operation. In addition the practical problems of actually getting a circuit to work would be avoided.

However in the course of assessing the simulation results the temptation to see 'Cross Coupled' Error Takeoff working proved too much for me.

It seemed to me that what was required was 'Cross Coupling ' combined with my 'Class S' amplifier scheme (ref.3. 1) which in turn was to be combined with an improved 'Class B' amplifier.

The attempt to implement this caused me considerably more practical problems than I had expected. This was because I had mistakenly assumed that the TL071 and TL074 op.amps. which I chose would be as well behaved in terms of stability as the now somewhat old fashioned 741 op. amps. It was only when I recognised that this was not so and replaced the TL071 and TL074's with 741's that a working circuit with 'Cross Coupling' was achieved.

There is no reason to believe that modern op. amps would not work even better. The problem is that modern components have a much higher frequency response and so need much better physical layout if oscillation is to be avoided as well as more care in the design of the output stages in the present case. As the object was to demonstrate that the principle worked, not to have the best possible results, 741 op.amps. were used.

However the extra effort was not wasted as one outcome was a circuit operating on a very good principle first described by Blomley (ref.3.2). Unlike the original circuit of Blomley, my development works without needing any setting up, a very worthwhile improvement.

In addition this effort produced an improved version of a much earlier attempt of mine to overcome 'Class B' bias problems by the use of automatic control of current.

These are all methods of reducing the value of errors in the outputs of amplifiers.

Class S

Although described at length in Wireless World (ref.3.1), (included in appendix), a brief description of the Class S amplifier is given below.

This amplifier is based upon two ideas.

1) That high impedance loads, for a given voltage swing, are easier to drive with low distortion than low impedance loads (fig.3.1, reproduced from ref.3.1) and

2) That with the aid of a subsidiary amplifier a high impedance load is provided to the basic amplifier which provides the voltage swing. In other words the low impedance load appears to be a high impedance load thanks to the action of the subsidiary amplifier. (fig. 3.2 & 3.3, ref.3.1).

Assuming that A2 in fig.3.3 of ref.3.1 has a very high gain (the standard assumption about op. amps.) then the voltage at junction of Rm and Rm' will be the same as at the junction of Rg and Rg'.Since Rm=Rm' and Rg=Rg'it follows that IL-IL' (the current drawn from A1 and its feedback resistors) is close to zero, ideally zero, and hence the impedance seen by A1 will be high. With idealised components this current would be zero and so the impedance would be infinite.

An important point which was overlooked in the original article was that the output, Vo, will appear across RL attenuated by Rm whether A2 has gain or not and whether its output impedance is zero or infinite (ignoring the very small effects of Rg and Rg') which is justified due to their high value compared to Rm and Rm'.

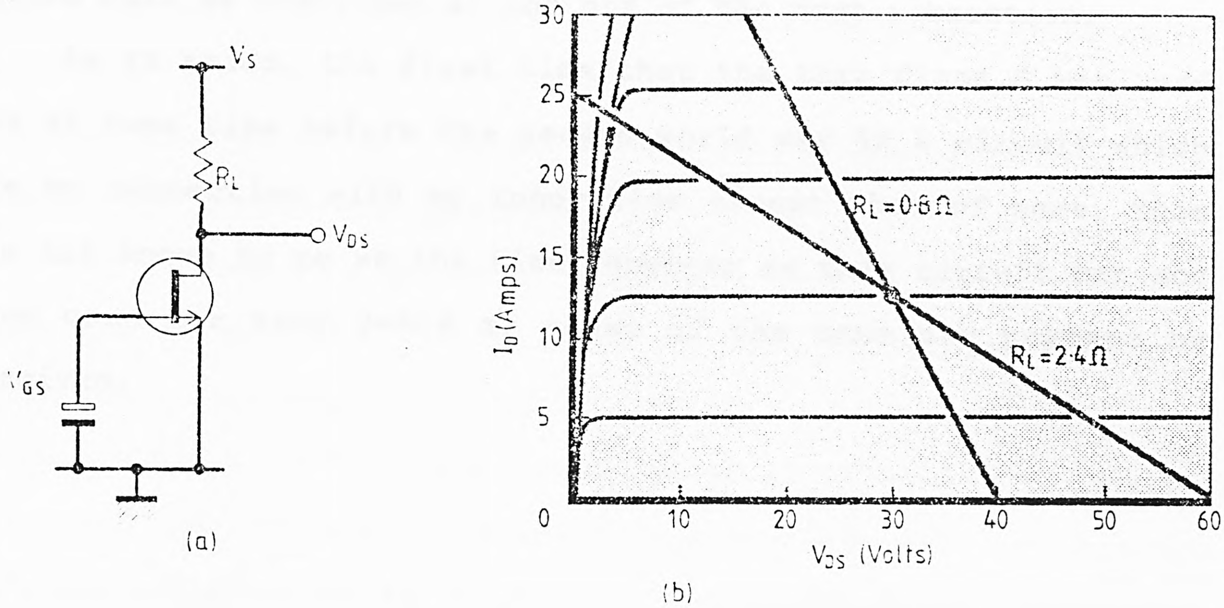


Fig. 1. *In the basic voltage-amplifier stage, a high-resistance load provides a larger swing with less distortion.*

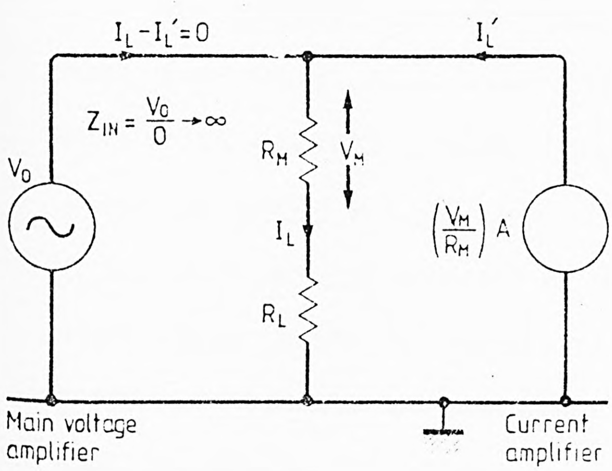


Fig. 2. *If the current from the current amplifier equals that into the load R_1, the voltage amplifier sees an infinite impedance. Voltage V_M controls the current amplifier.*

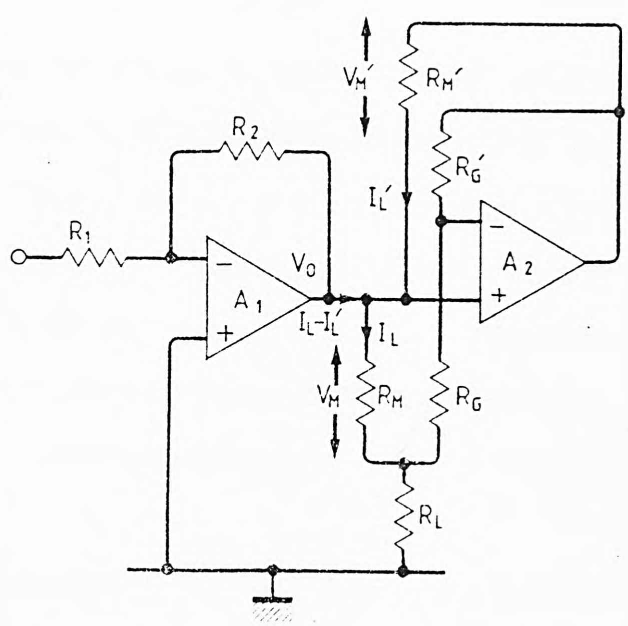


Fig. 3. *Basic circuit of Class A amplifier. A_2 is current amplifier.*

FIG. 3.1, 2, 3

From this it follows that if A1 has a sufficiently low output impedance the effects of A2 and its distortion may be ignored.

The results of using this circuit in a 'Cross Coupled' system will be described at the end of the next subsection.

As an aside, the first time that the term Class S was used was at some time before the second world war in a circuit which has no connection with my innovation except that of name. This was not known to me at the time. However as this circuit has not been used for many years my reuse of the name may perhaps be forgiven.

'Cross Coupled' balanced output circuit.

The overall scheme uses two 'Class S' amplifiers and four subsidiary amplifiers. (Fig.3.6).

At first glance the scheme seems too complex to work, what ever computer aided studies may say, since the complex transfer functions of real operational amplifiers and the effects of circuit strays often results in configurations with multiple loops being unstable in practice.

The answer to this is to constantly bear in mind the fact that the virtual earth point on an op. amp. is called this because it is already at virtually zero volts i. e. virtually at earth. Because of this the virtual earth point acts to isolate one part of the overall circuit from another and so prevent unwanted signal coupling which could otherwise cause instability.

However let us first briefly survey the results of an eca2 computer aided study.

The circuit can be considered as being a conventional bridge amplifier (fig. 3.4) with a floating load but with cross coupling added to greatly reduce errors (fig.3.5) in a simple way. Instead of using a full equivalent for the 'Class S' output stages (fig. 3.6) an assumption is made that op. amps. 5 and 6 together with their pairs of output transistors and associated resistors, function ideally and increase the load impedance to such a degree that it can be ignored.

Under these conditions an analysis using the equivalent circuit of fig. 3.7 is justified.

Now from the work of the previous chapter it must be obvi-

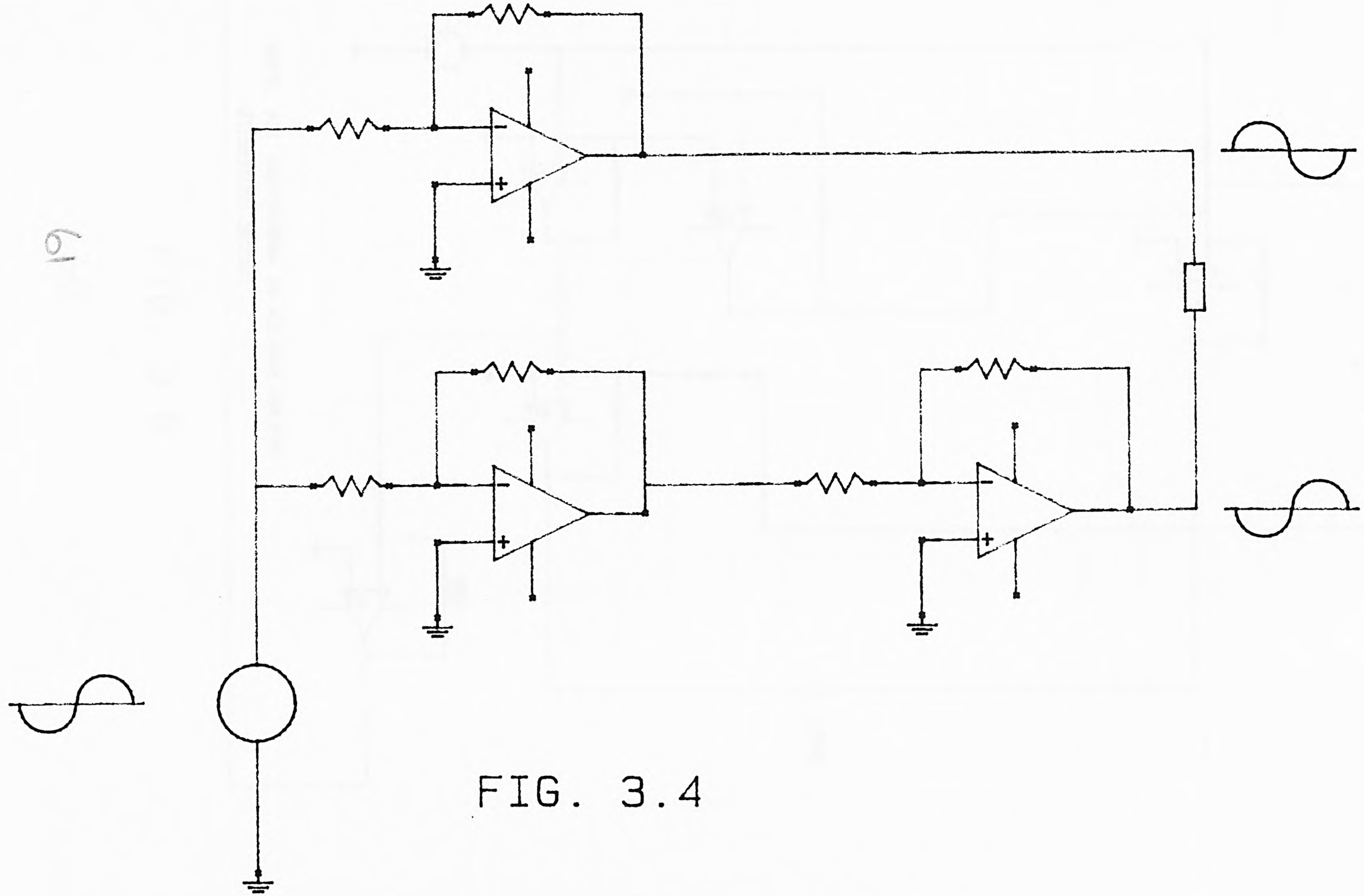


FIG. 3.4

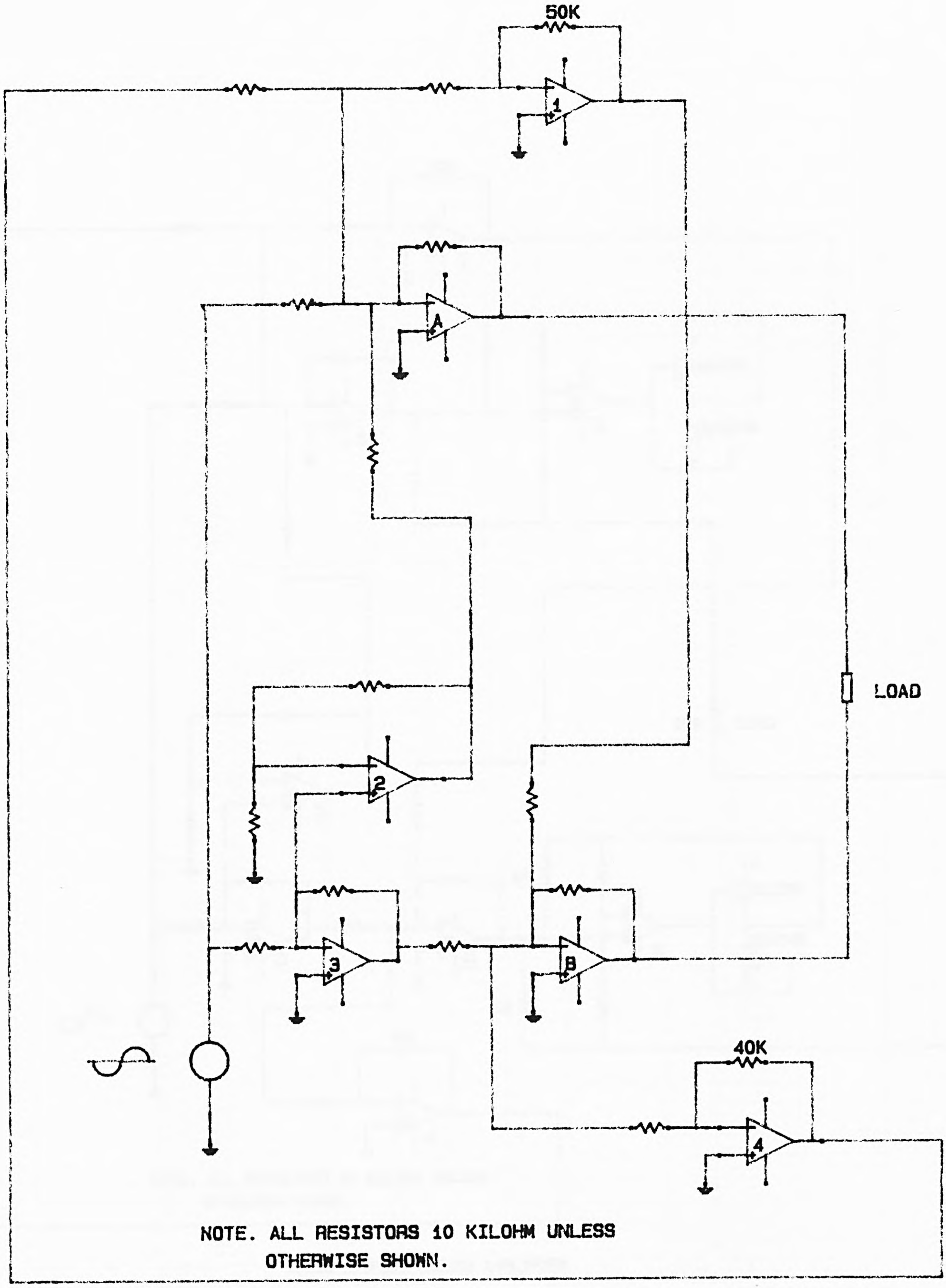
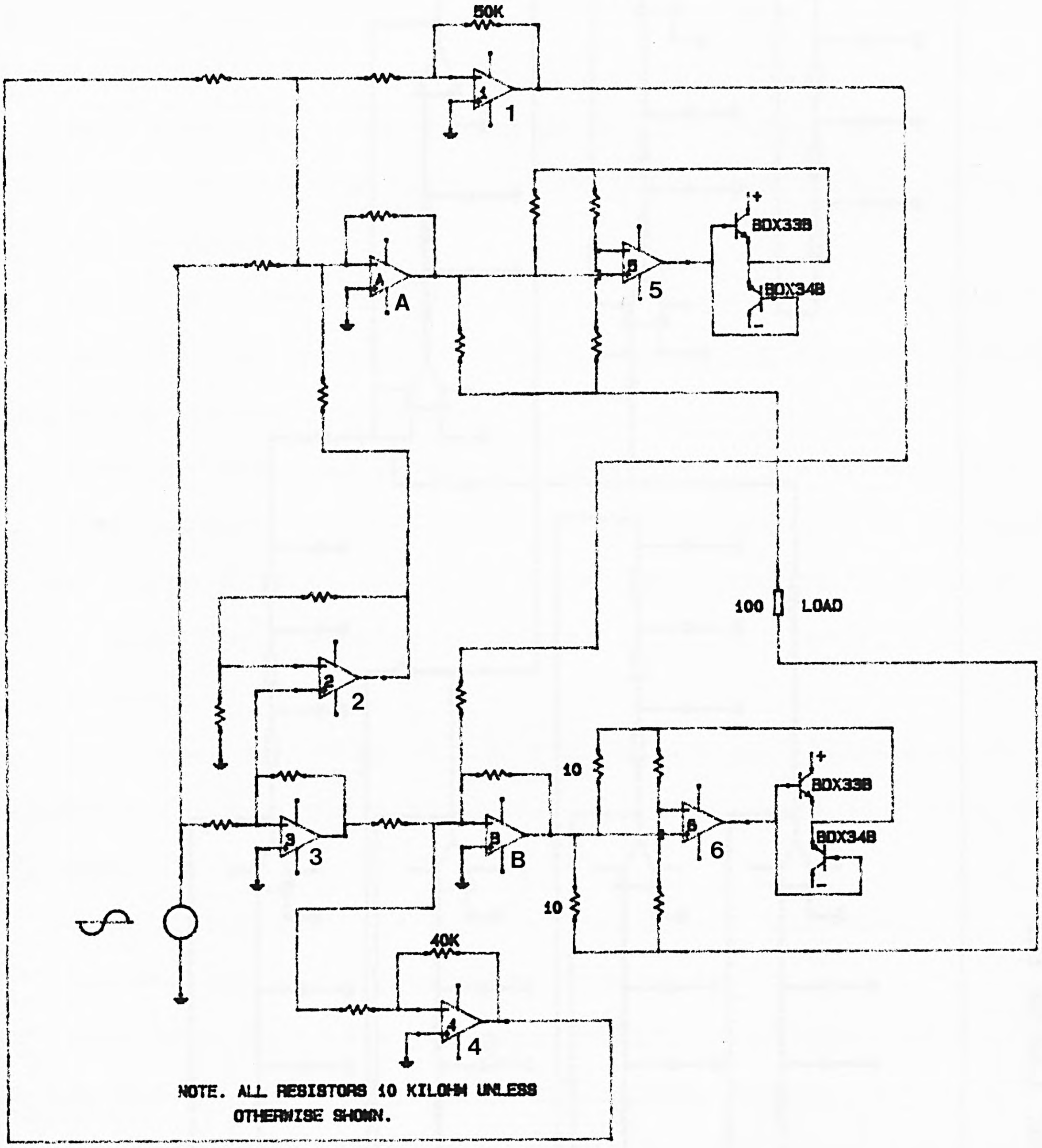

50K
1
A
LOAD
2
3
B
40K
4
NOTE. ALL RESISTORS 10 KILOHM UNLESS
OTHERWISE SHOWN.

FIG. 3.5



CROSS COUPLED BRIDGE AMPLIFIER

FIG. 3.6

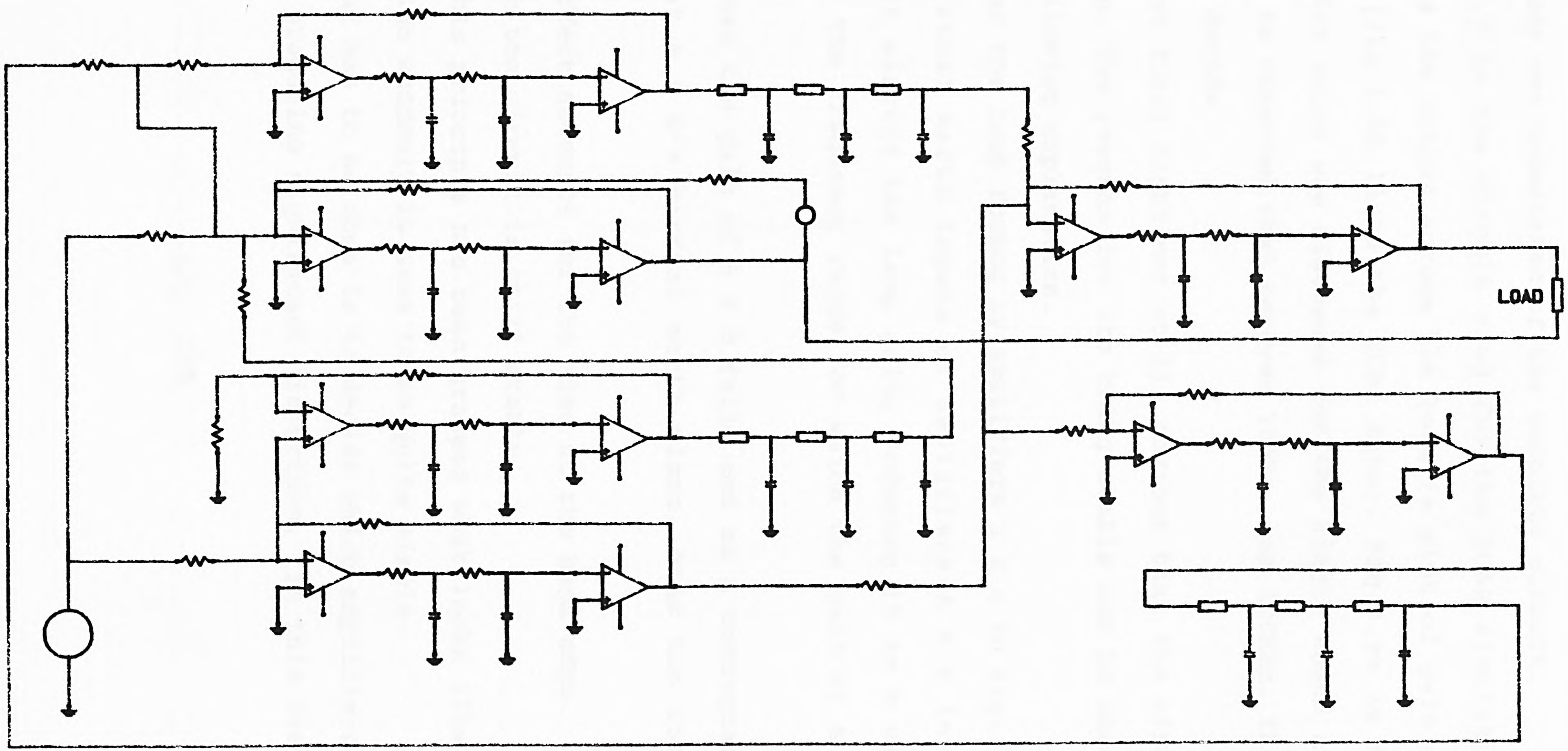


FIG. 3.7 CROSS COUPLED BRIDGE CIRCUIT USED FOR ANALYSIS

ous that, unlike a conventional negative feedback amplifier, the improvement factor is not 20 dB per decade but 40 dB and that the circuit is stable. However in order to demonstrate this a computer aided study was undertaken of the complete circuit.

Fig. 3.7 is the circuit used for the ECA2 simulation and Fig. 3.8a is the output across the load (a plot of gain against frequency) (fig.3.8b lists the ECA2 data). Fig.3.9a is the improvement plot which was discussed for the single ended amplifier. It will be observed that between 100Hz. and 100kHz. the slope is 40dB per decade.

It is at first sight not at all obvious that the circuit is not unstable. The reasons for its being stable can be understood from the following explanation.

Consider the loop formed by amplifiers 1 & 4 in Fig. 3.6.

The virtual earth inputs of amplifiers A & B in effect almost short circuit the loop gain, reducing it to a very low value, over the frequency range for which the gain of A & B is significant.

Only when the gain of A & B falls and as a consequence the impedance at A & B's virtual earth rises, does the loop gain increase.

This effect accounts for the rise at the band edge.

However the circuit is still stable.

Once this principle has been grasped what looks like a very complex design suddenly is seen to be quite simple.

All that has to be done is to decide which amplifiers in the circuit are producing significant distortion, (in this case it is

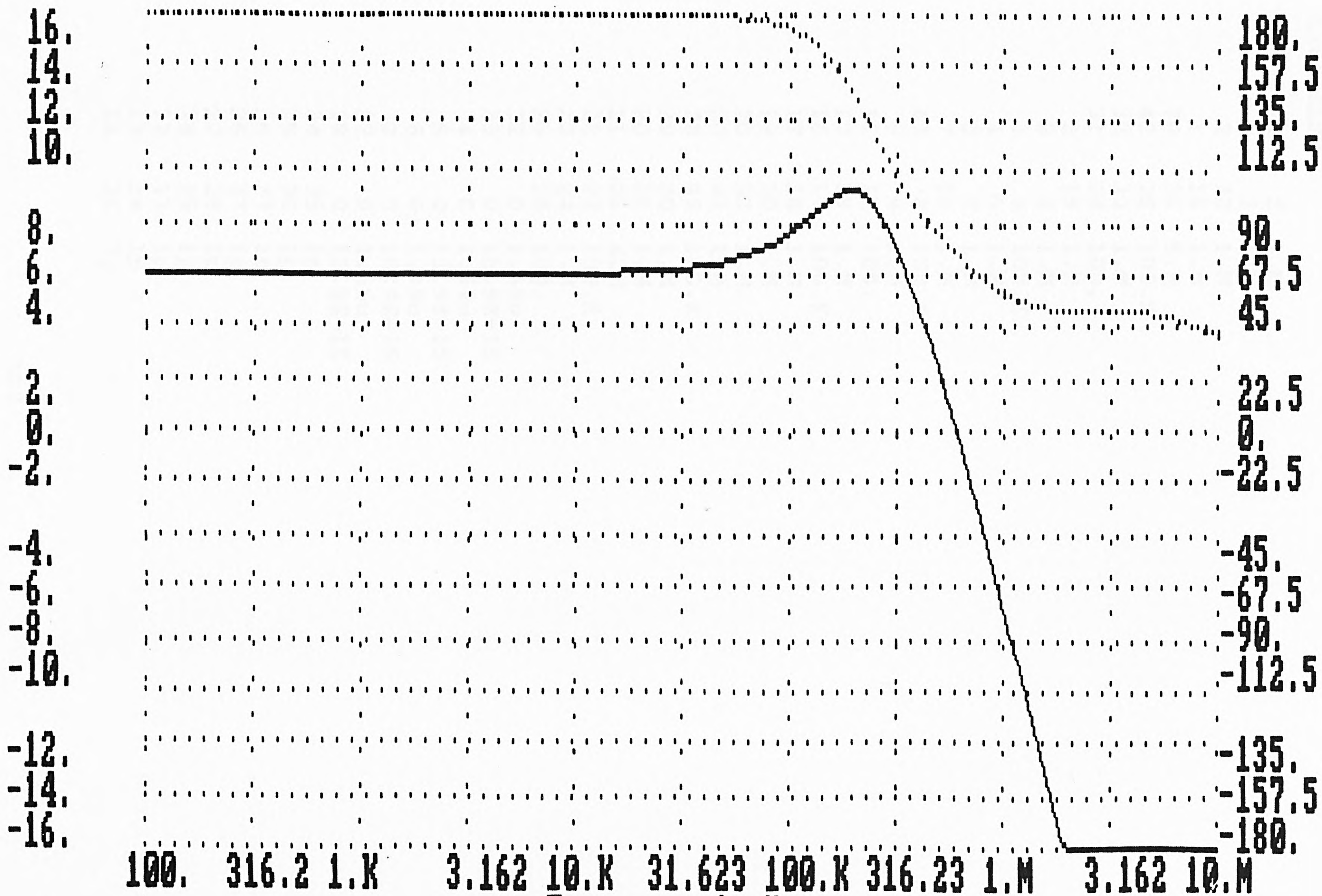
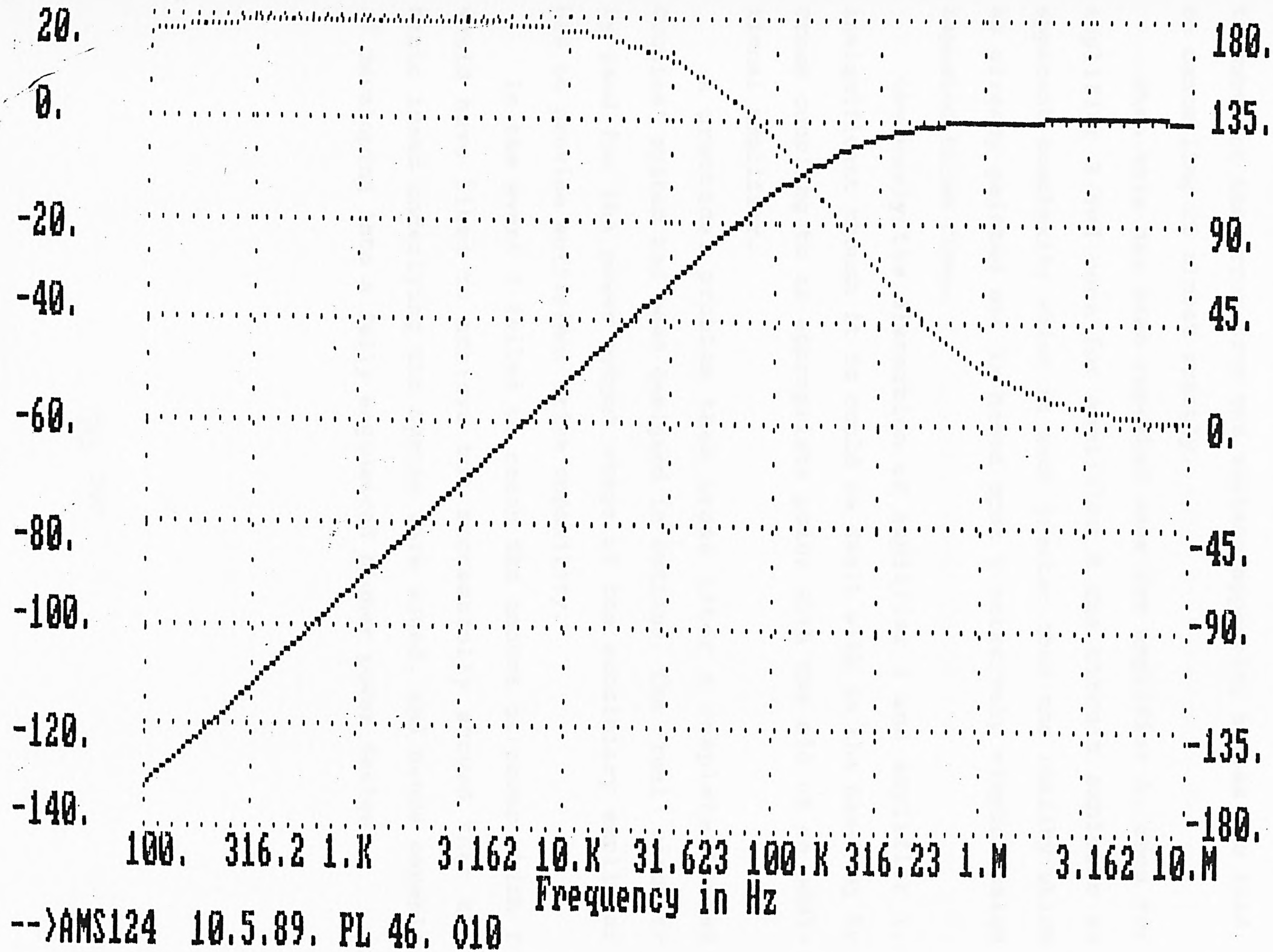

16.
14.
12.
10.
8.
6.
4.
2.
0.
-2.
-4.
-6.
-8.
-10.
-12.
-14.
-16.
180.
157.5
135.
112.5
90.
67.5
45.
22.5
0.
-22.5
-45.
-67.5
-90.
-112.5
-135.
-157.5
-180.
100.
316.2
1.K
3.162
10.K
31.623
100.K
316.23
1.M
3.162
10.M
Frequency in Hz
-->

FIG. 3.8a

Fig. 3.8b

1	R1	29	2	1.M
2	R2	1	2	1.M
3	R3	2	30	1.M
4	R4	1	10	1.M
5	R5	30	31	5.M
6	R6	42	32	100.K
7	R7	32	33	1.M
8	R8	31	34	100.
9	R9	2	15	1.M
10	R10	2	6	1.M
11	R11	3	4	100.K
12	R12	4	5	1.M
13	R13	0	7	1.M
14	R14	7	11	1.M
15	R15	43	8	100.K
16	R16	8	9	1.M
17	R17	11	12	100.
18	R18	10	19	1.M
19	R19	16	17	100.K
20	R20	17	18	1.M
21	R21	19	20	1.M
22	R22	20	37	1.M
23	R23	20	41	1.M
24	R24	38	39	100.K
25	R25	39	40	1.M
26	R26	20	21	1.M
27	R27	21	25	4.M
28	R28	22	23	100.K
29	R29	23	24	1.M
30	R30	25	26	100.
31	C1	32	0	1.59u
32	C2	33	0	15.9E-15
33	C3	4	0	1.59u
34	C4	5	0	15.9E-15
35	C5	8	0	1.59u
36	C6	9	0	15.9E-15
37	C7	17	0	1.59u
38	C8	18	0	15.9E-15
52	L1	34	35	1p
53	L2	35	36	1p
54	L3	36	37	1p
55	L4	26	27	1p
56	L5	27	28	1p
57	L6	28	29	1p
58	L7	12	13	1p
59	L8	13	14	1p
60	L9	14	15	1p

61	VA1	30	0	1.
62	RA1	42	0	1.u
63	VB1	0	33	1.M
64	RB1	31	0	1.u
65	VA2	2	0	1.
66	RA2	3	0	1.u
67	VB2	0	5	1.M
68	RB2	6	0	1.u
69	VA3	7	10	1.
70	RA3	43	0	1.u
71	VB3	0	9	1.M
72	RB3	11	0	1.u
73	VA4	10	0	1.
74	RA4	16	0	1.u
75	VB4	0	18	1.M
76	RB4	19	0	1.u
77	VA5	20	0	1.
78	RA5	38	0	1.u
79	VB5	0	40	1.M
80	RB5	41	0	1.u
81	VA6	21	0	1.
82	RA6	22	0	1.u
83	VB6	0	24	1.M
84	RB6	25	0	1.u
85	R31	6	44	100.M
86	R32	41	45	100.M
87	R33	45	0	100.M
88	R34	44	46	100.M
89	VA7	45	44	1.G
90	RA7	46	0	1.u
91	VI	0	0	1.
92	RI	0	1	1.u



20.
0.
-20.
-40.
-60.
-80.
-100.
-120.
-140.
180.
135.
90.
45.
0.
-45.
-90.
-135.
-180.
100. 316.2 1.K 3.162 10.K 31.623 100.K 316.23 1.M 3.162 10.M
Frequency in Hz
--)AMS124 10.5.89. PL 46. 010

FIG. 3.9A

amplifiers A, 3 and B), then to measure the voltages at their virtual earths, restore, with the aid of amplification, the error which the voltage at the 'virtual earth is an attenuated measure of and if the sign of the gain is right to 'cross couple' it to the other side of the amplifying system when the result will be to subtract the error from the voltage appearing across the load, so cancelling it almost exactly.

When this has been repeated once for amplifier A, once for amplifier 3 and once for amplifier B the circuit exhibits an apparent complexity which is much greater than the reality which as already pointed out is based upon a relatively simple design repeated three times.

Obviously the distortion of amplifier 4 and amplifier 1, insignificant though it is could be dealt with in the same way by cross coupling to an appropriate point with the aid of an additional amplifier.

A practical problem then arose after a complete 'Cross Coupled' system had been designed in outline. The 'real' circuitry used for the power output stage of the subsidiary amplifier has to provide sufficient drive capability.

In the event I failed to reach the amount of power which I would have liked to achieve but successfully showed that the basic ideas underlying the system were sound, and hence capable of developing into a fully engineered higher power design.

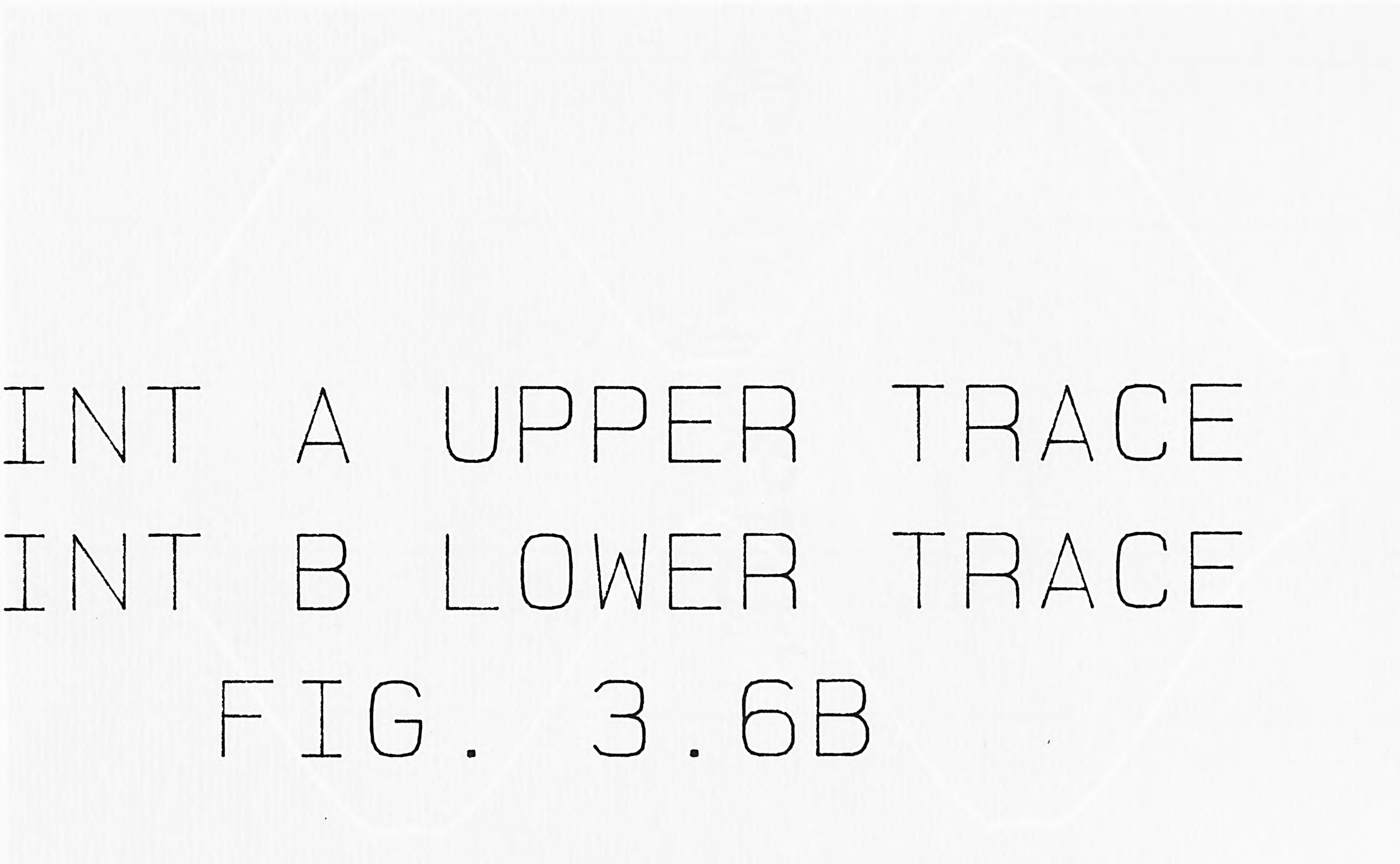
Results using 'Class S' in a 'Cross Coupled' arrangement.

The results clearly show that 'Cross Coupling' works.

Since unlike power feedforward systems such as error take-off, with the bridge circuit the error correcting voltages are not applied at the output where resistors are needed to make then work but earlier in the system, where signal processing operations at low power level are possible.

The photographs, figs.3.6b and 3.6c taken from an oscilloscope screen, are referenced to the circuit diagram (fig.3.6) by alphabetics corresponding to points on the circuit diagram.

I had at first planned to use a circuit based on an idea by Blomley (ref.3.2). Although not ultimately used in the final circuit, for reasons already discussed, never the less the idea and results are of some interest.

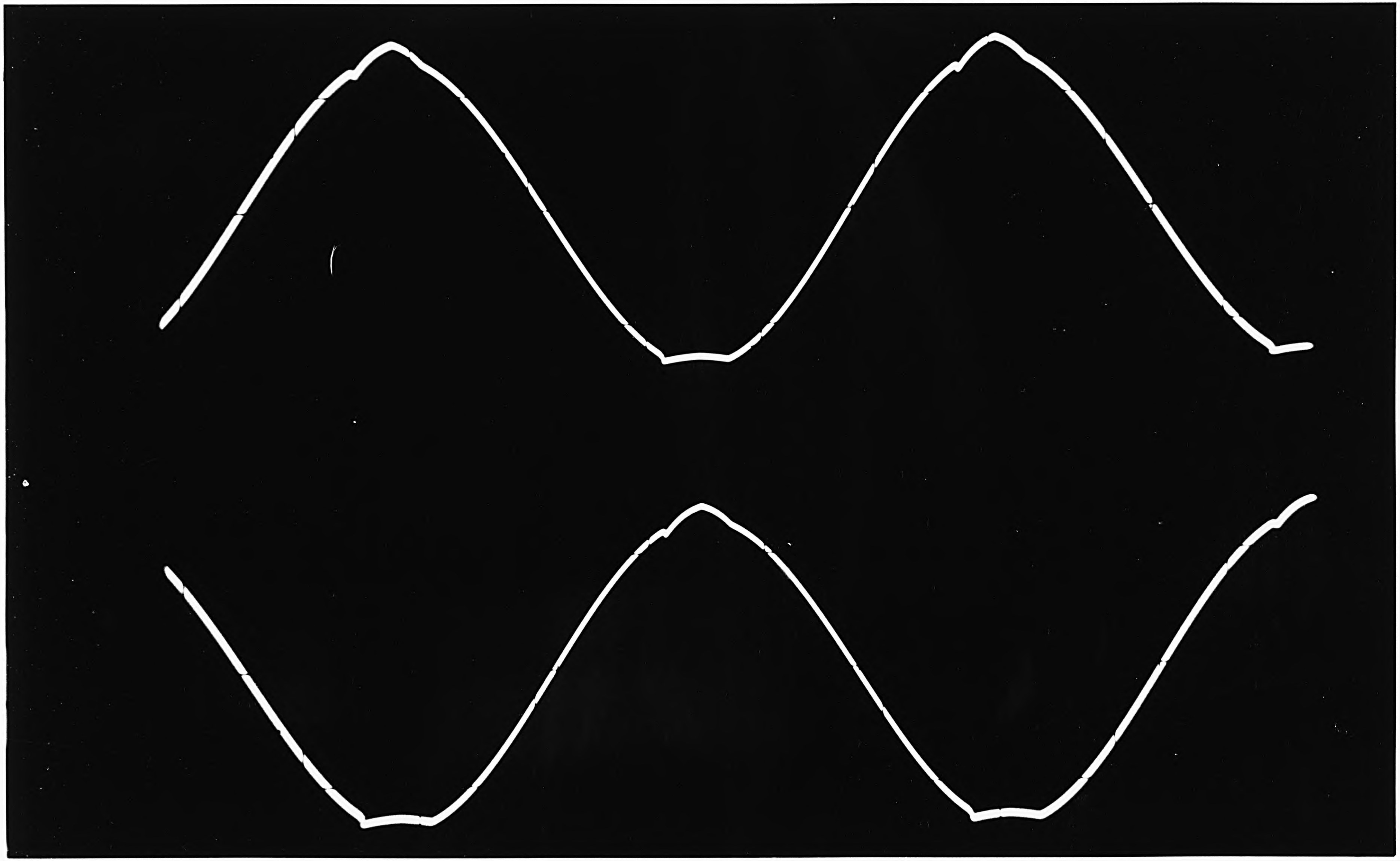


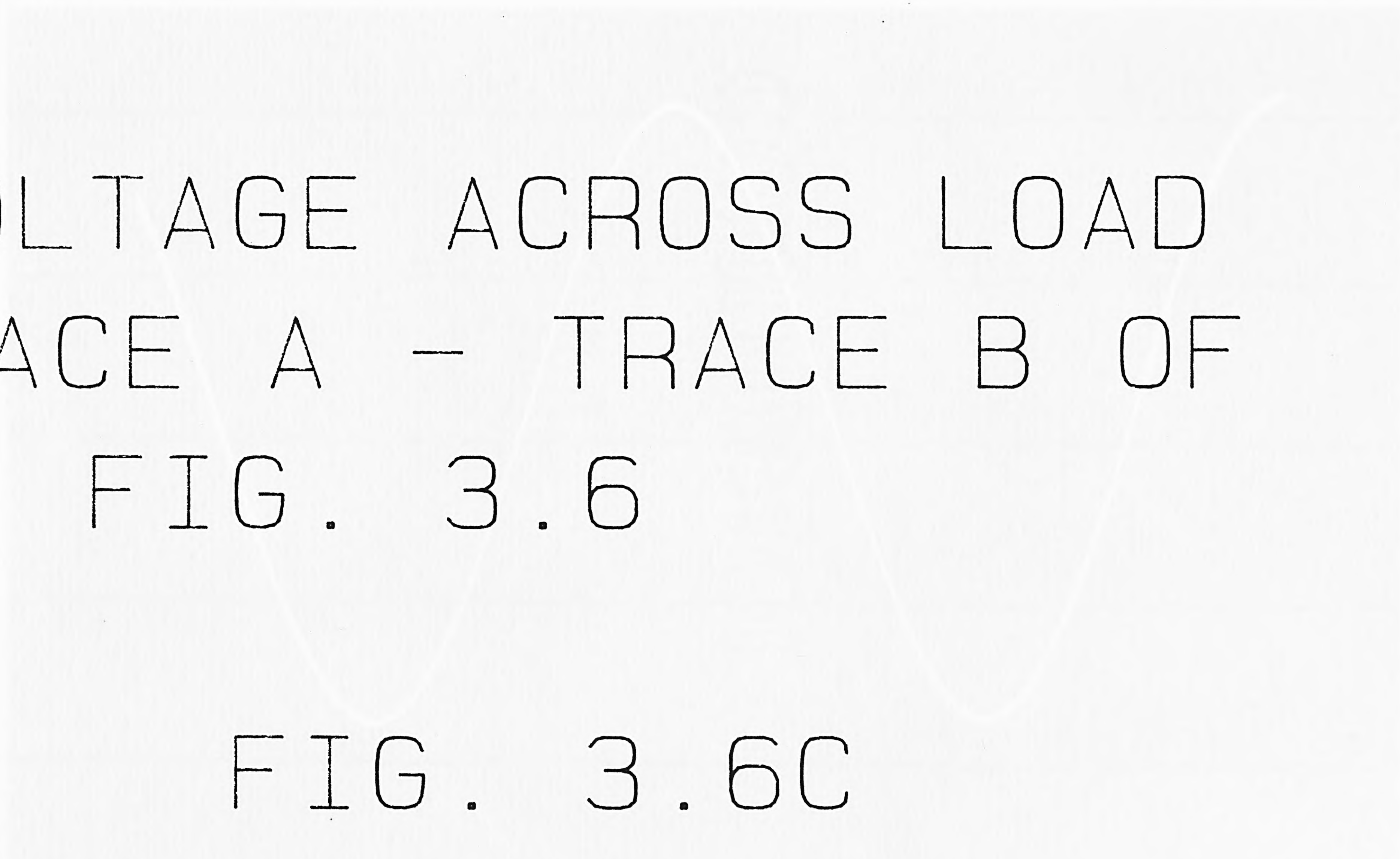
POINT A UPPER TRACE

POINT B LOWER TRACE

FIG. 3.6B

3.6B

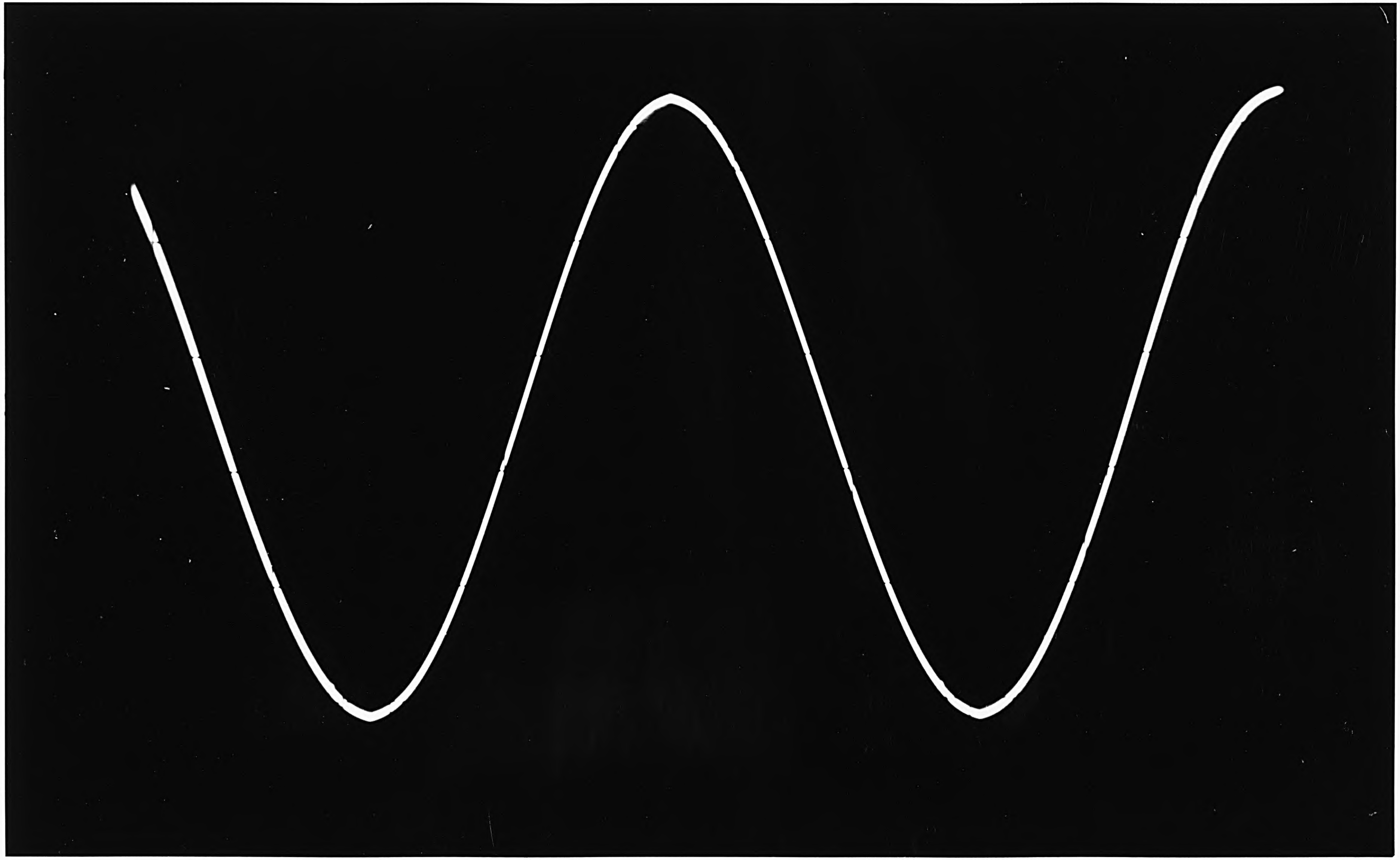




VOLTAGE ACROSS LOAD
TRACE A - TRACE B OF
FIG. 3.6

FIG. 3.6C

3.6C



A modified Blomley circuit: deliberate error generation and correction applied to an amplifier.

Background.

For signals of a widely varying amplitude with time, Class B has great advantages compared to Class A, since over a period of time when the signal to be amplified has only a small amplitude the power dissipated is low compared to the Class A amplifier. Indeed in the Class A amplifier dissipation actually rises with decrease in output power.

As typical real audio signals (such as speech and music) have very 'peaky ' waveforms, as well as periods of silence and low level from the point of view of heat dissipation and thus of keeping the heat sink size low, Class B has great advantages over Class A.

In addition there is the advantage of economy of current consumption in battery equipment as well as the fact that the high temperature associated with high power dissipation much reduces the life of many electronic components.

However even with the great advantage of power efficiency compared to Class A, Class B has been used with reluctance for high quality audio applications because of its inherent problem of cross over distortion.

Current Dumping is a well known attempt to solve the problem which has previously been mentioned but which will not be used in this thesis as part of a 'Cross Coupling scheme'.

What follows is a clear example that a brilliant idea is not enough, but that the details must also be right. In 1971 Blomley (ref.3.2) designed an amplifier in which two current output amplifiers were combined together, one to provide positive output current the other to provide negative output current, with the essential feature that at all times these amplifiers operated in the linear region. In order to obtain proper operation it was necessary that a very well controlled distorted signal waveform was generated for the phase splitter and a very stable pair of power current amplifiers used in the output stage (Blomley, ref.3.2. p. 57, fig. 1.).

On both counts the circuit failed.

By the use of my 'Error Takeoff' concept and a standard constant current circuit these problems have, to the best of my knowledge, been overcome, for the first time.

The first of Blomley's errors was (Blomley p. 59, fig.6) the assumption that by using a constant current source it was possible to have instantaneous switching from one diode to the other in his fig. 6. Although at frequencies below a few Hz it is a good approximation, it is not at higher frequencies since an unavoidable shunt capacitance appears at the junctions of the diodes of his fig. 6b so producing unavoidable distortion as the diodes change state. It is significant that his fig. 2. p. 128 is at 1 kHz and not 10 kHz.

The second error was the need for initial setting up by adjusting R20 and R21 (last paragraph Blomley p. 128). This a serious production problem since setting up costs time and needs either skilled staff or special automatic equipment. In addition such circuits are sensitive to ageing, component tolerances and temperature effects.

Component sensitivity makes maintenance a problem because setting up needs to be carried out every time a component is changed and temperature sensitivity could introduce problems every time a burst of high power output raised the equipment temperature.

The requirement is a circuit which is inherently insensitive to parasitic capacitance, needs no setting up and retains its performance with ageing components.

Such a circuit would depend for its operation upon operational amplifiers because of their good bias voltage input stability as well as upon diodes matched to themselves alone, that is to say no diode would be required to track the Vbe (emitter to base voltage) temperature coefficient of a bipolar transistor,as well as the effect of variations in base current of such a power transistor.

Apart from ageing and component to component variations another reason to avoid trying to get diodes to track transistors with temperature is that it is difficult to ensure that as the power in an output transistor varies (so varying its actual temperature) the temperature of this transistor and a matching diode is the same.

In any case the circuit to be described does not depend on any of these difficult to control effects.

It needs no setting up uses ordinary readily available components and as functions as a good approximation to a true Class B amplifier.

Operation of sub-circuits

The first requirement is an accurate phase splitter of such a design as not to need any setting up.

An essential element of such a circuit is an accurate half wave rectifier capable of high speed, high amplitude (10 V peak) operation. The circuit to be described is my own design, which has been put into commercial production working at up to 20MHz. Unlike the conventional circuit (Fig 3.17.) it is insensitive to stray, parasitic, capacitors.

However in this application the high speed facility is not required but the low d. c. error which it possesses is.

Consider fig.3.10. When Vs is positive (assuming ideal diodes) D1 conducts and D2 is cut off, so Vo=Vs.

When Vs is negative, D1 is cut off and D2 conducts.

The result, for ideal diodes, is that only positive values of Vs appear at Vo.

Now in reality of course, diodes are not ideal and a drop of roughly 0. 5 volts for a Schottky diode may be expected, so that for the waveform shown the output will be from +11. 5 to -0. 5 volts.

This problem is readily overcome using a third diode and resistor (fig.3.11) and a second supply. (In practice both supplies would often already exist). Although the maximum voltage swing will be reduced, in the present case it does not matter. The result is a high accuracy amplifier which is insensitive to parasitic capacitors since even as Vs approaches zero, about 0. 6 mA is still flowing through D1 this current being sufficient to

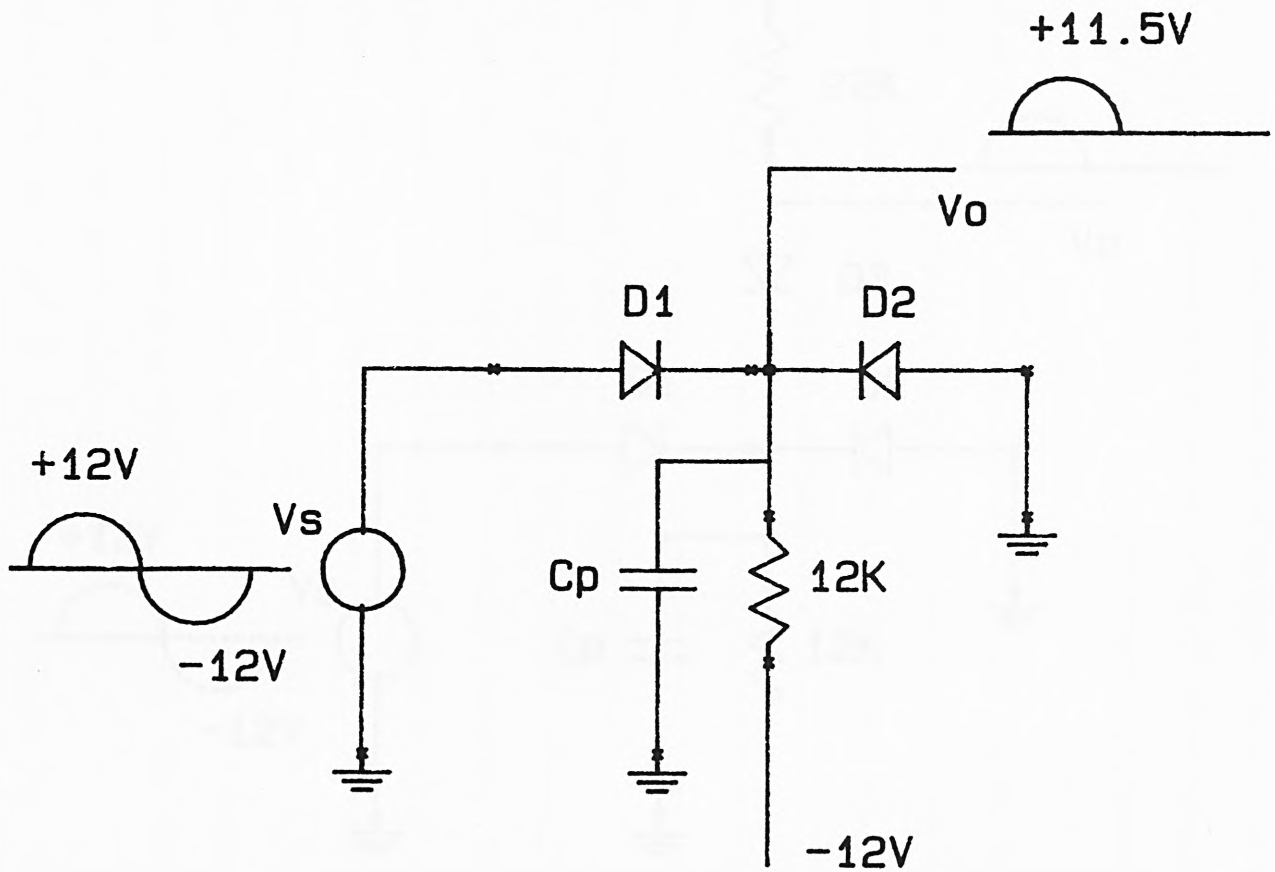


FIG. 3.10

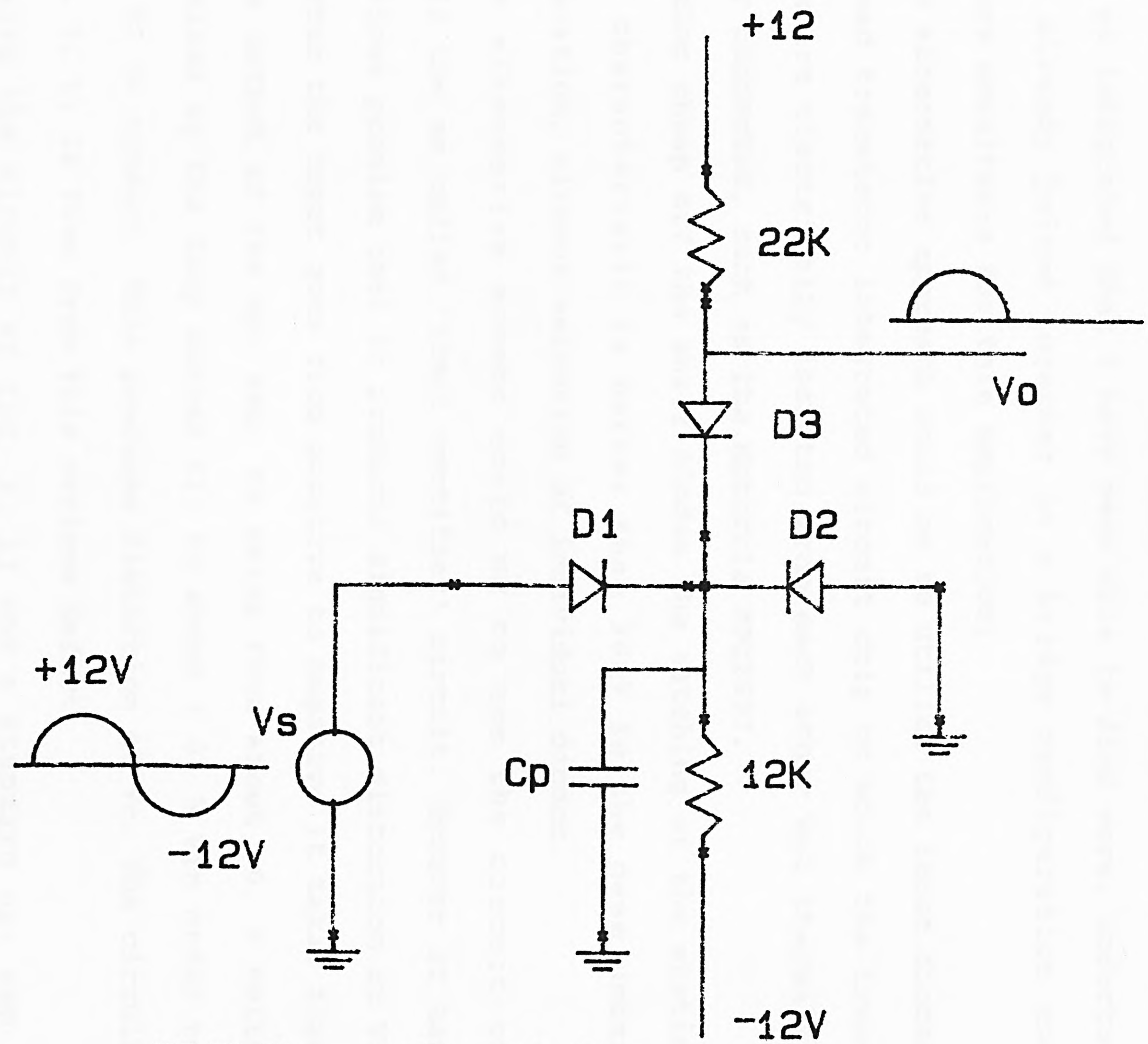


FIG. 3.11

charge or discharge any parasitic capacitors effectively, unlike the close to zero current flowing through D in fig.3.12, which is the conventional circuit.

The power dissipated in D1, D2 and D3 (Fig.3.11) is so small that even though it varies with signal amplitude it is less than 1mW and so may be ignored. As a result it is necessary only they be physically close together to main temperature matching. (This is a recipe for integrating the diodes on a single chip. The only diodes so integrated that I have been able to find were, unfortunately already joined together in a bridge configuration and therefore unsuitable for this application).

An alternative approach would be to utilise the input diodes of a quad transistor integrated circuit chip on which the transistors are electrically isolated from each other but thermally closely connected, such as the Motorola MPQ2907.

Using cheap off the shelf diodes the matching of the static I vs V characteristic is better than 30mV in the case under consideration, without selection of individual diodes.

An alternative scheme would be to use the circuit of fig.3.12 the so called 'ideal rectifier' circuit. However it has the serious problem that it produces significant distortion at Vo since when the input goes from positive to negative it takes time for the output of the op. amp. to swing from about -0. 5 volts (determined by the drop across D1) to about + 0. 5 Vin order to enable D2 to conduct. This produces distortion at Vo. The circuit of fig. 3. 11 is free from this serious defect.

Using the circuit of fig. 3. 11 and a standard op. amp.

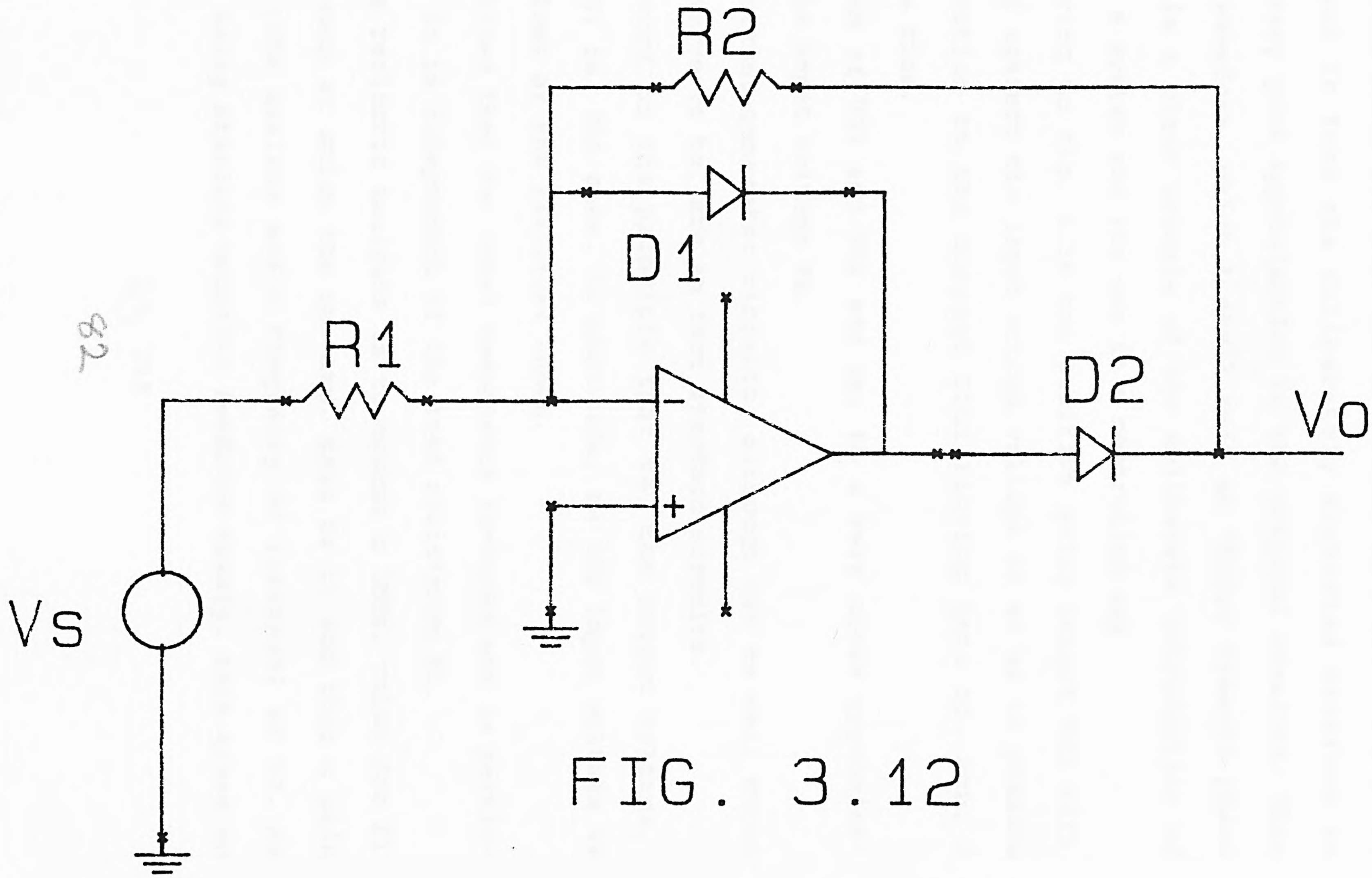


FIG. 3.12

inverter together with a voltage follower produces a circuit which has a low output impedance.

At this point the author's fig.3.13 'Error Takeoff' scheme (ref.2.4) is used in modified form.

Having generated a precisely distorted waveform i. e. that of a half sine wave, all that has to be done is measure the error and subtract it from the deliberately distorted waveform to produce a very good approximation to the original waveform. This distorted waveform, which I shall call an 'Error Takeoff phase splitter' is a clear example of the deliberate introduction of error into a system and its use in a controlled way.

Referring to fig. 3.14 the positive going output VO1 will balance off against the input source voltage Vs so as to produce no contribution to the current Iin3 flowing into op. amp. 3 during this time.

The sum of VO1 and VO2 add up, to a very close approximation, to the input voltage Vs.

The current generator circuits, although not as well known as they deserve to be, are in fact standard circuits.

They work on the principle that Vo, the output voltage, (fig. 3. 15) is the same, in magnitude, as the input voltage Vs for the values of the resistors shown.

It follows that for ideal components Io=Vo/Rs and in particular that Io is independent of the load resistance RL.

A more realistic analysis is to assume a 1MHz. value for f1 (the frequency at which the op. amp. gain is 1) and thus a gain at 20 kHz (the maximum audio frequency of interest) of 50. As calculated using standard negative feedback theory, this gives an

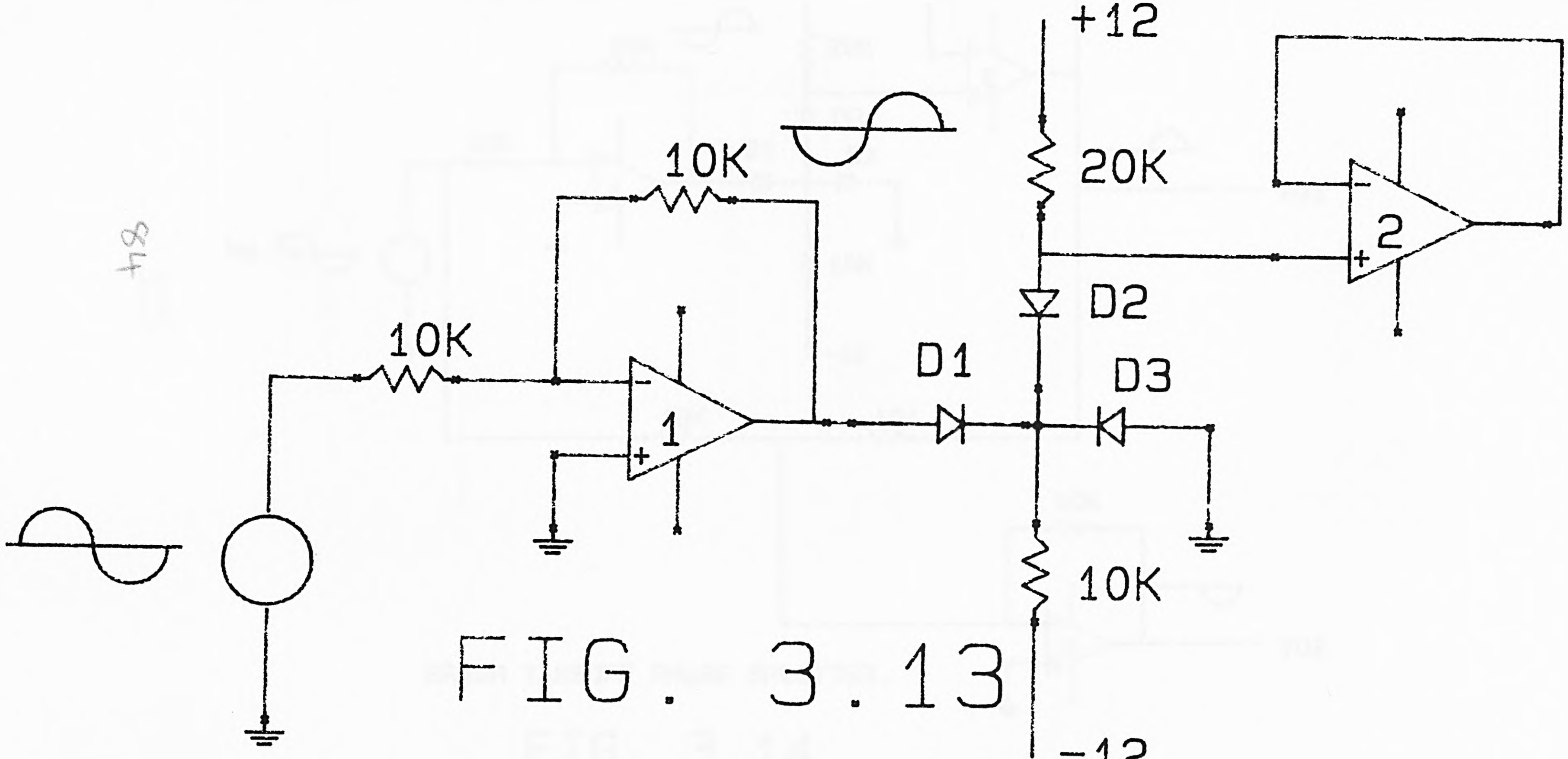
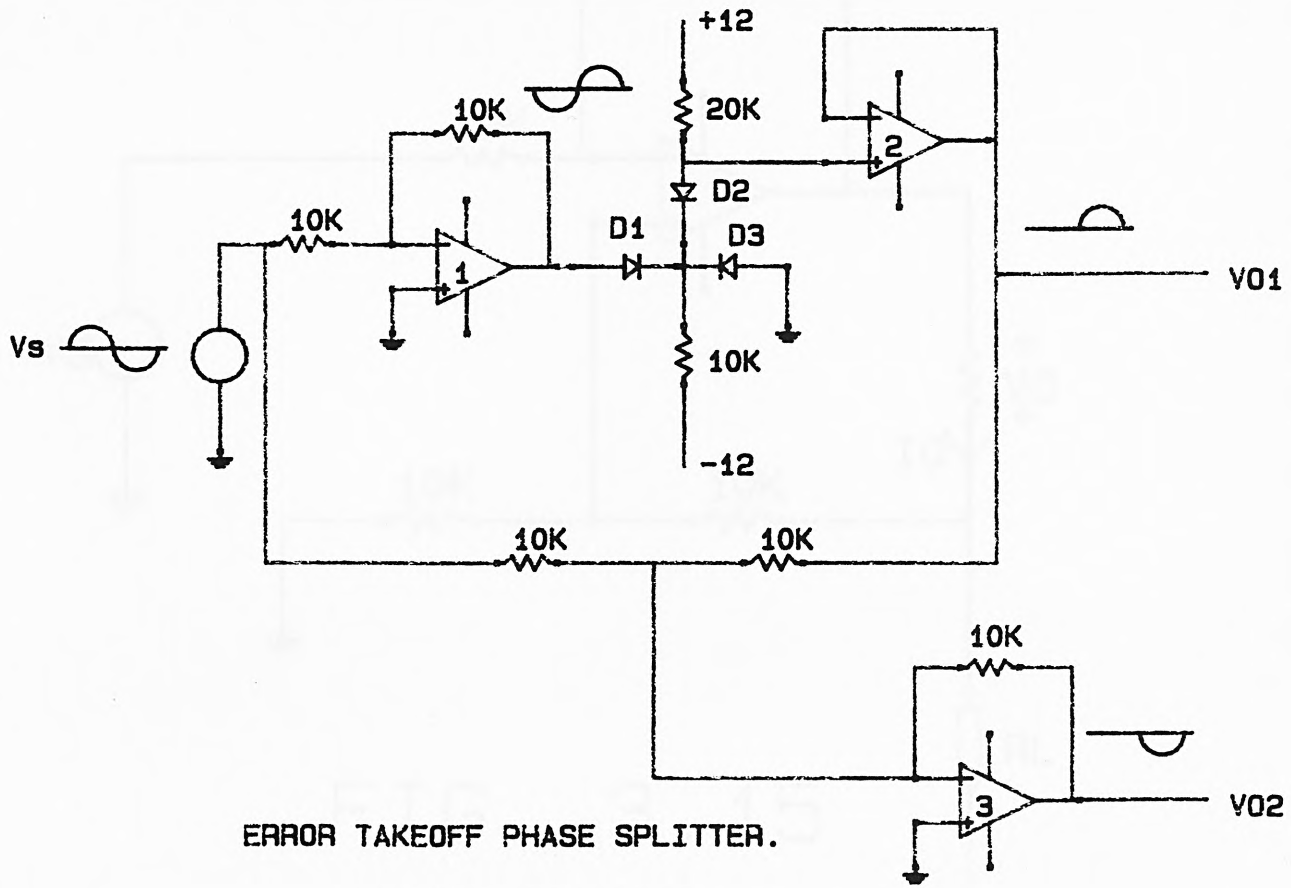


FIG. 3.13



ERROR TAKEOFF PHASE SPLITTER.

FIG. 3.14

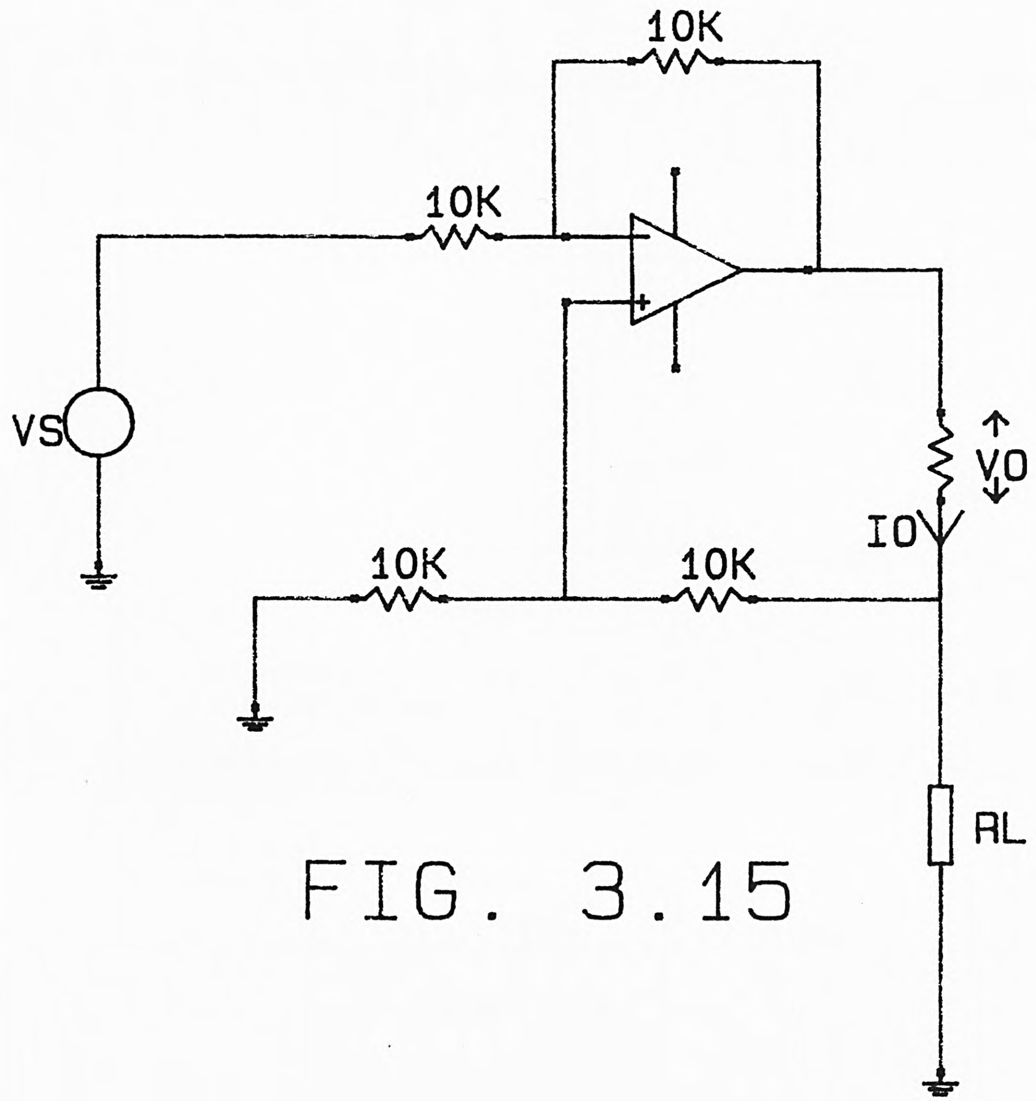
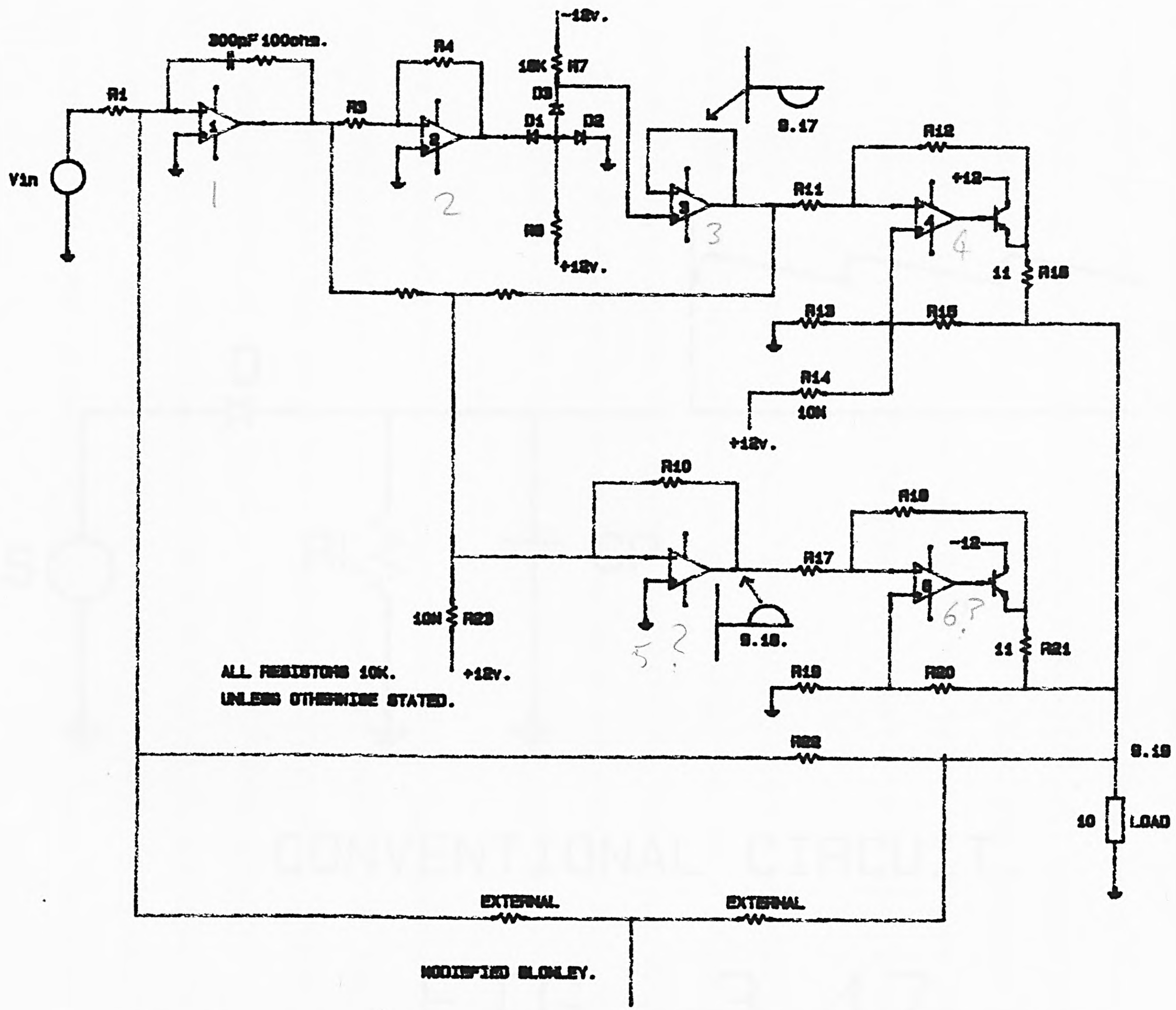


FIG. 3.15

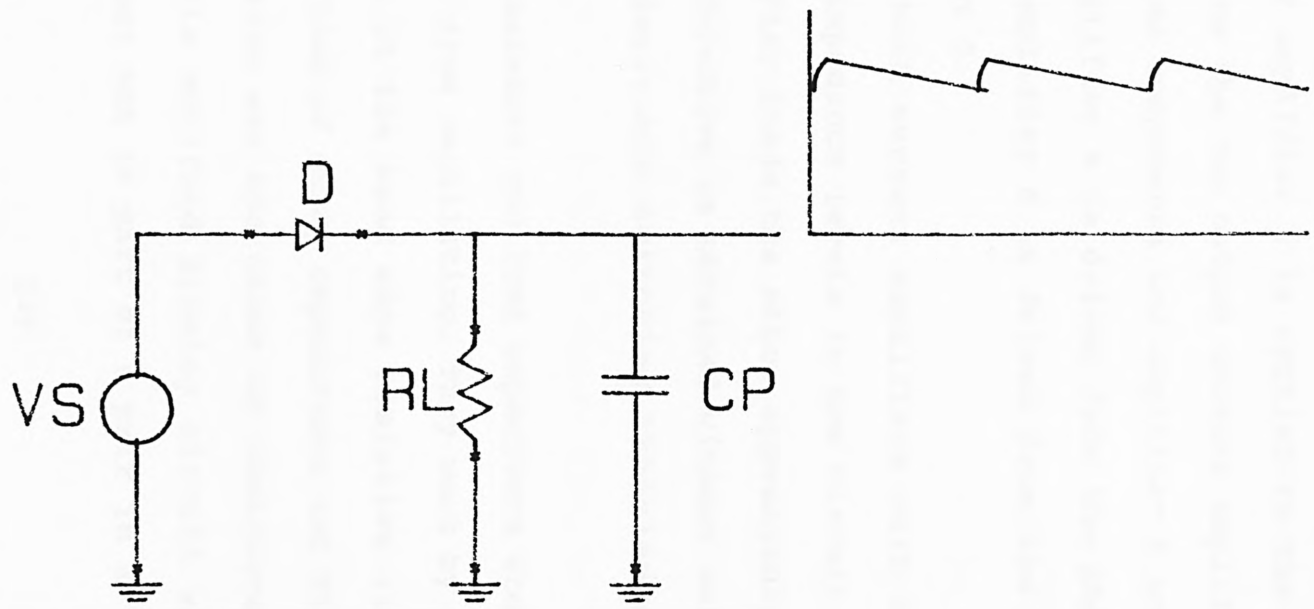
output impedance of -j50 ohms which is so much greater than 1ohm as well as being reactive that the assumption of an infinite output impedance is justified).

As well as the sub-circuits already described another standard circuit, that of an integrator with a resistance in series with the feedback capacitor(fig. 3.16) is used. Its only relevant feature that needs special mention is that it also acts as a summing junction.



Vin
R1
300pF 100ohms.
R3
R4
-12v.
10K R7
D3
D1
D2
R8
+12v.
9.17
R11
R12
+12
R16
11
R13
R15
R14
10K
+12v.
R10
R18
-12
R17
9.18.
10K R23
+12v.
11 R21
ALL RESISTORS 10K.
UNLESS OTHERWISE STATED.
R19
R20
R22
9.19
10
LOAD
EXTERNAL
EXTERNAL
MODIFIED BLOWLEY.

FIG. 3.16



CONVENTIONAL CIRCUIT.

FIG. 3.17

Overall operation of a Blomley circuit.

The combined summing junction and inverter (part of fig. 3.16, the complete circuit) amplifier 1 and its associated components), ensures that the small d. c. offset that could otherwise appear at the output is much reduced because the op. amp. employed has a much smaller offset error than the errors involved in using discrete diodes in the three diode rectifier.

The output of amplifier 1 is applied to the phase splitter which in turn drives the two output current amplifiers, amplifier 4 and its associated components and amplifier 6 and its associated components. Amplifier 4 is driven from the phase splitter at amplifier 3 and amplifier 6 is driven from the phase splitter output at amplifier 5.

As they are both current amplifiers with high impedance (compared to the impedance levels in the circuit at this point) neither sub-amplifier loads the other appreciably and the basic 'Blomley' design objective is obtained without setting up or the production of undesirable switching transients in the phase splitter.

The 100 ohm resistors and 10nF capacitors stop the output of the TL071 op. amp. from oscillating. They work by making the load the op. amp. sees at the band edge resistive (100 ohm) rather than a capacitive load of stray capacitance and Miller effect.

This oscillation was the cause of considerable trouble in the sense that this modified Blomley circuit would only work stably by itself but not as part of a pair in a 'Cross Coupled' arrangement. This difficulty persisted in other designs including an automatically adjusting 'Class B' amplifier (ref.3.3) as well as a close copy of the original primitive 'Class S' circuit.

On changing to 741 amplifiers the primitive circuit became stable.

In other words the more modern TL071 appears to have a much less desirable phase response at its band edge than the old fashioned 741 op. amp.

However as it resulted in an improvement to the automatic 'Class B' amplifier and in order to be able to demonstrate a working implementation, a much simpler circuit (the 'primitive 'Class S' amplifier) the consequences were positive in the sense that both the automatic 'class B' and the Blomley circuit were improved.

Returning to the Blomley circuit of fig.3.16.

R1 and R22 together provide overall feedback as well as being the combined input resistance to the integrator formed by op. amp. 1 and the 300 pf. capacitor and its 100 ohm series resistor. The 100 ohm series resistor so alters the phase as to stop oscillation of the 'Nyquist' type, that is to say oscillation caused by the point -1 being enclosed on the Nyquist diagram. It and the 300 pF capacitor were found experimentally.

The accurately rectified half wave output of op. amp. 3 drives current amplifier 4 and similarly but in opposing sign, amplifier 5 drives its associated current output amplifier 6 the

91

two currents from 4 and 6 combining to provide the total output current. Figs. 3, 18, 19, & 20 are photographs which show the significant circuit waveforms. The two external resistors in fig. 3.20 are adjusted to balance the output to zero at their junction. This is a measure of the error in the output waveform.

It will be observed that no setting up potentiometers are required and no sudden changes in voltages occur either.

Both such faults, as already pointed out, are to be found in Blomley.

Although this circuit certainly worked, and the original 'Wireless World' article contains proof of this in the form of photographs (ref.3.1.) it now appears to me that it must have been only marginally satisfactory, and defects in the design were overlooked.

This was because the problem of measuring the minimum current that flowed in the circuit was not solved in a reliable way.

The problem has now been solved and the circuit now functions in a satisfactory manner.

Postscript

At a very late date a survey of the Field (ref.3.4) was published which shows that 'Current Dumping' is derived from my 'Error Takeoff'.

Of more interest is the information that 'Class S' has been adapted by several Japanese manufacturers. In particular Messrs Technics have, using 'Class S', produced an amplifier with a THD of one part in 500,000.

This is referred to in ref.3.4.

93

FIG. 3.18

3.18

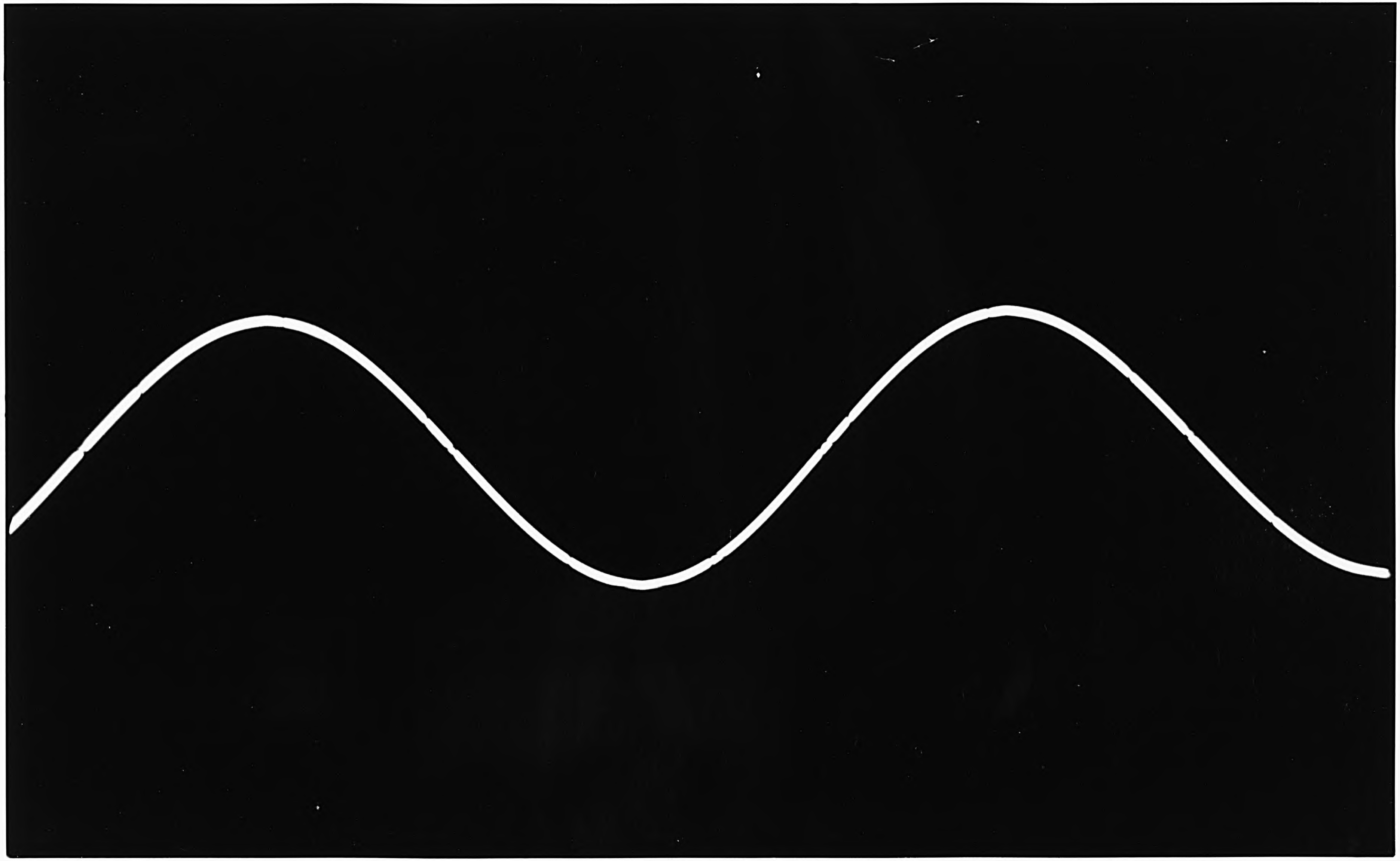


FIG. 3.19

3.19

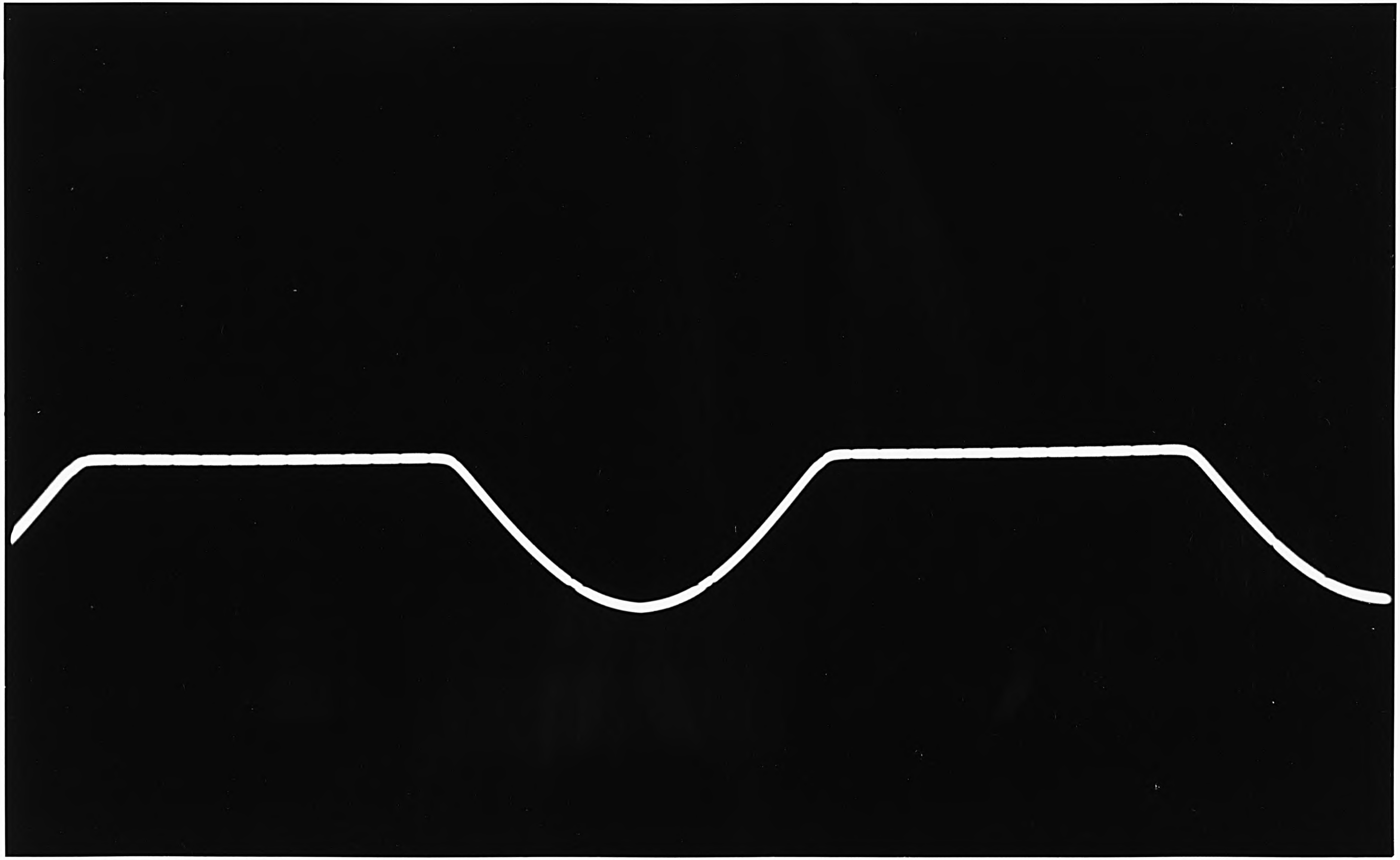


FIG. 3.20

References.

3.1 Class 'S', A. M. Sandman, Wireless World, Sep. 1982, pp 38-39.

3.2. New approach to Class B amplifier design, P. Blomley, Wireless World, Feb. & March 1971.

3.3.Low cross over distortion class B amplifier. Wireless World, July 1971.

3.4.Solid State Audio Power, J.Linsley Hood, Electronics World + Wireless World, Dec.1989.

Introduction.

Before describing the proposed scheme, a discussion of the limitations of the present scheme and the obstacles to changing it, will be made.

This is because, in my view, the technical problems of Railway Signalling are not the real difficulty standing in the way of a much improved system.

In an age when space vehicles can be sent out into the planetary system and years later produce impressive results the Railway System with its drivers and minimum separations of hundreds of yards between trains travelling at speed is not just an anachronism, it is also an example of refusing to change modes of thinking which tolerate this. For this reason I debate the background to my proposals before coming to them.

Towards the end I present as succinctly as I can an outline of a reliable communications system, to pass measured information between vehicles adjacent to each other and for them to use this information to determine the maximum speed of each vehicle, which in my view is what Railway Signalling is about.

It is deliberately an outline since such a task will require a large team to complete the details, in addition it may well be that a better way to achieve the desired result, which is to make it possible for individual vehicles to run at speed under automatic control whilst almost touching, could be found. The real problem is to believe it can be done.

For these reasons I have deliberately allowed myself more space than I might, in order to emphasise that the difficulties with Railway Signalling are of attitude not of technology.

General Discussion

The Engineering Problem is by no means in the top league of difficulty, many more difficult problems have been solved in the past. In a nutshell the problem of Railway signalling is the problem of knowing the distance between a vehicle and the one in front and reacting accordingly.

In other words the problem may be changed to that of designing a very reliable communications system between fixed and moving points.

This problem is in my view one of the most important of all Engineering problems, because we do not have sufficiently safe transport.

Who has not met people whose lives have not been changed for the worse by a road accident either because of an injury to themselves or a close relative or because a relative has been killed on the roads.

In my view something could be done to significantly increase safety. The solution is to make the use of transport both as simple and as safe as using a lift. A modern transport system should involve no more than entering an address, or for frequently used destinations a name into an on board computer and the rest would be done automatically, the vehicle going to its destination without further intervention from the passenger or passengers. Anyone able to dial a telephone number would be able to travel, what ever their age, in speed, safety and comfort.

A modern railway system would be a first step towards such a scheme.

Let us examine why we are not making progress towards achieving this.

There are numerous reasons, amongst which are the long history of the railways with their more modern roots in Victorian England. The slow evolution from a 'see and be seen' system with accidents being by no means unknown, via the well known system of operating signals with the aid of long connecting wires from the signal box, (a system requiring amongst other things some physical strength), and its associated features of 'fail safe 'and block signalling to the present system which has been described as 'modernising a well by attaching an electric motor to the winding handle.'(This phrase first appeared in a letter to the 'Times' twenty years ago, I ask my readers indulgence for the lack of a more precise reference). A phrase which sums up the current railway signalling system, which in my view is Victorian 'block signalling' with a veneer of electronics.(refs.4.1 & 4.2). ref.2 is a manufacturers leaflet which clearly demonstrates that an overlay of electronics does not make a modern efficient system. ref.4.2 clearly shows that history is the root of 'Block Signalling' not considerations of what is possible and desirable at the present time. Although still current the GEC publicity which comprises ref.4.2 was printed in 1979, this says something about Railway Signalling technology.

In addition, there is the innate conservatism which long established ideas generate in some people, an inability to change, to think afresh from basic facts.

A famous example of this is the phlogistic theory of heat, in which heat was assumed to be a substance called phlogiston, similar to say iron or gold. This theory had such a firm hold that even the simplest and most convincing experiments were rejected at first and it took some considerable time for a more realistic and correct theory to be accepted. Even Joseph Priestly the discoverer of oxygen, remained a believer in the Phlogiston theory, evidence that great ability can be combined with a rigidity of thought.

An identical rigidity of thought blocks progress in Railway Signalling.

Furthermore the current system certainly works and works in a safe and certain manner for most, though not all, of the time. The occasional railway accident shows up the weaknesses of a "fail safe" system in which failure, which is not safe, can and does occur.

The enquiry into the Clapham disaster " ref.4.3 states that there were more than 100 " wrongside failures " in the Waterloo area in the four years up to the crash. A wrongside failure is situation in which a signal does not acknowledge a fault and allows a train to proceed. Certainly this pointed up the weakness of fail safe.

Finally, modernising the present system to make it fully automatic, although worthwhile, is only marginally so, since it would not produce a very great improvement in capacity with all the positive benefits that would imply.

No doubt there are more reasons for the lack of major progress in Railway Signalling but before we deal with the real issue, let us consider again reasons for its importance.

The most important reason is that if the railways could be made to operate more efficiently by means of radically improved signalling road transport could be much safer, since a modified version of such an improved railway signalling system could be applied to the roads. Indeed vehicles capable of running on both rail and road could be developed so that at the start of an evolving process the percentage of traffic carried by road would drop which in itself would probably reduce road deaths. At the very beginning when only a small part of the network was automated and Rail/Road vehicles were not available it is impossible to tell what the effect would be because, increasing the capacity of the transport system as a whole might simply increase the amount of usage. This factor is well known today on the roads, one result of increasing road capacity (for example, by building new roads), being that more journeys are then made.

An automatic railway capable of dual rail/road running with the possibility of passengers running their own car sized vehicle on it would of necessity beat the roads, whilst even a system capable only of running on rail alone with, the impracticality of individual passengers owning their own vehicles would certainly offer stiff competition to the roads for many journeys even though it could not offer door to door transport.

Obviously a dual track vehicle should have a high development priority.

However a rail only vehicle system using small automatically controlled vehicles would give several advantages.

They are the large economic saving in not having a driver, especially for small vehicles, the 'on demand' nature of the service by which I mean being able to travel when and where a passenger pleases (because with a small vehicle one passenger is economic). In addition the number of stations or stopping points could be drastically increased, since, unlike the present system a vehicle would only start and stop at the beginning and end of a journey ignoring stops in which it had no interest. This in itself would much shorten journey times.

Anyone who doubts that a competition, unconscious though it may be, does occur between road and rail should ask themselves why not all car owners drive to work. Certainly many train users own cars. The main reason is that driving to work would, for them, be slower than going by train.

The ideal solution, as already mentioned, would be to develop vehicles capable of running on both rail and road. Such vehicles have been designed and British Rail once had goods trailers which could have their wheels changed by mechanically switching them from rubber to steel.

This would enable the railways to then become high speed, automatically controlled transport arteries, with the most important property of providing a system capable of going between any two points, as roads now do but rail cannot.

In order to realise such an objective, it is essential that a railway signalling scheme be incorporated which provides safety in the presence of high density traffic.

This brings us to the most glaring of all the defects of the current system. This is the state of emptiness on even the most busy lines, lines on which there would be great advantages in being able to run more trains. Even if the present system were made truly safe, which it is not, (and for safety reasons alone an improved signalling system is needed), this inefficiency of low utilisation would by itself call for an improved system.

The reason for this emptiness lies in the history of railway signalling and arises from a basic assumption necessitated by Victorian engineering when velocity measurement and control was impractical. This assumption is that any train in front of the train being controlled by signals is stationary even when it is going at 125 miles an hour. This is a truly erroneous assumption of the first magnitude and its very grossness may be the reason for its acceptance since it suggests some profundity of reasoning which actually does not exist. Whatever criticisms can be made of what I propose, they will be as nothing compared to this error.

I ask my readers to ponder on this fundamental error inherent in the current system before proceeding.

The following basic assumptions should underpin any modern railway signalling system:-

1. Constant self checking which should immediately detect any error as it arises and in addition make faulty installation impossible.

2. The system must have its complexity confined to electronics embedded in integrated circuits. On no account should there be complex wiring outside the immediate vicinity of the circuitry. Ideally all connections involving connections beyond a few centimetres should be made by optoelectronic coupling. A need for

decentralised control with all important decisions being made in the vicinity or on the vehicle, follows from this. In other words there should be no elaborate medium-distance telecommunications links of the kind required for the present centralised 'signal box' system.

3. It follows from this that redundancy and self checking must be an essential requirement at the heart of any modern system.

4. The cost needs to be low and therefore extremely reliable results should be obtained using components of normal reliability.

5. For economy no special 'electronic' skills should be required to install it.(Ideally it should be capable of installation by unskilled workers).

6. Repairs should be easy and cheap involving very simple fault detection of a sub-circuit, followed by its removal and replacement (eventually this could be done automatically by a passing vehicle). From 1. quick diagnosis and correction of any faults is possible.

7. The performance of the actual control system must in no way be inferior to the human operated system employed on the roads. The 'Victorian' Signalling system of block signalling whether based upon fixed blocks or modern moving blocks is absolutely unacceptable because it can never provide a high enough vehicle density. Rail vehicles must be able to run at high speeds whilst almost touching as is done on the roads, without the dangers of the risks taken by road vehicle drivers.

8. It is essential that the control of vehicles is fully automatic. Driver error is a common cause of such crashes as do occur on the railways and the main cause on the roads. Drivers must no longer play any part in the control of moving vehicles.

In addition the economic advantages of not having drivers would be very great since it would make car-sized, individual, vehicles economically feasible something which partly because of track capacity and partly because of the cost of employing a driver, is not really practicable at the moment.

9. Breakdowns should be very rare events indeed, since sufficient redundancy should be built into the system to enable repairs to be done before the system is effected (basic assumption 1.). But if they do occur rapid system recovery is essential.

Docklands Light Railway

The Docklands Railway is a light railway, that is to say that it is not designed to carry the sheer weight of a normal railway, and some of its gradients and curves are much more severe than those of a normal railway which made it cheaper to construct and run although when the cost of tunnelling at the Bank extension is taken into account the saving as a percentage in not having main line standards is probably not all that much, especially as much of it was originally main line track to start with.

As part of its economy of construction it was originally built with short platforms capable of taking only two carriages which formed the total train.

Very soon it was decided that this was insufficient and that so many passengers would be using it that the capacity needed increasing.

A solution was required.

Under the present system of signalling it is impossible to run more trains even though the usual conditions of 'emptiness' on the track hold as well as of course emptiness as far as a train standing at a platform was concerned in the sense that even under busy conditions station platforms are empty.

As the only practical solution with the present system the length of the platforms is being increased to permit longer trains.

This has had two results.

The first is to increase the amount of capital invested and the second to cause the railway to be shut down in the evenings and at weekends for some months in 1989 for the modifications to be carried out.

However with an improved, modern signalling system a very different approach could have been taken.

The platforms could have been left as they were and the number of two carriage trains doubled with the only consequences that the platforms would have had trains standing in them more frequently and the tracks between stations would have had more trains on them. This would have provided the required extra passenger capacity.

The word automatic is normally used for something in which human beings play a very small part. In the Docklands railway they play an essential part both in normal running and in recovery from the too frequent breakdowns.

A reason often given for having a driver on board is that the public's 'nerves' would not stand fully automatic, driverless operation although its 'nerves' seem quite able to endure a lift with no lift attendant during the journey even though there is a tendency to have spells of getting stuck in mid journey in some cases (for example at the Angel Underground station in London)

I suspect other reasons.

The first has nothing to do with the kind of topics normally dealt with in most Engineering PhD. theses. I strongly suspect that Trade Unions would not tolerate a truly driverless automatic system and that the management prefers to take the easy way out by not confronting the Unions.

The second is a technical problem which arises directly from the inability of trains to approach each other closely during normal operation, so that if a breakdown in an automatic control system were to occur it is not possible to get a Railway Official on board quickly.

So in the case of the Docklands Light Railway, the ticket collector takes over the role of driver, who with the aid of a two way radio to communicate with the control centre, takes over from the automatic control and actually drives the train.

A number of questions arise in connection with such a system. For example, what are the reasons for system failures and for those failures which do occur, why cannot recovery be automated with the aid of on board computers or using remote control with the help of the Control Centre.

It may well be argued that requiring corrective information to be taken from the Control Centre is too difficult to achieve technically and is therefore impractical.

In my opinion this is not the case.

The interplanetary space probes launched by the United States (for example, Voyager) have a history of technical breakdowns which have been partially or completely remedied by means of corrective action being taken from earth.

It could be said that a railway official on board is essential because he also looks at tickets and opens and shuts doors. He even pushes a button to start the train. However in many other countries passengers are expected to have the right tickets and if they do not the chances of encountering a travelling ticket inspector are not very small indeed, as they are on the London Underground.

As for the doors the necessity for this operation to be supervised may be questioned as normal passenger lifts work perfectly well under passenger control. I am unable to accept that a sideways movement is so very different from a vertical movement.

The Docklands Light railway thus represents a missed opportunity to provide a fully automatic system.

I will conclude this section by making the claim that the Managing Director of the Docklands railway is close to agreeing with me as an interview with him in the 'Times', entitled " A railway in the dock " (ref.4.4) clearly shows. In particular I enjoyed the paragraph starting 'On each vehicle, the "Train Captain"-'.

Main line operation.

The present system run by British Rail and most other railway systems is a very clear example of the result of evolution of a kind which fails to jettison obsolete methods and ideas.

True automation using the current block signalling system as a base rather than a system which takes into account the velocities of the vehicles involved would have some advantages over the present system in its non automated form even if not as many as the system to be described below, if it were to be used as the basis of automatic control. There would be savings in crew costs, and because of this, it would become economical to run much shorter trains, (many long trains consist of separate, functionally complete short trains coupled together to form a longer train), outside busy periods. In other words, two or three short trains, instead of one long one, would provide a much better service.

Technical Requirements

The previous discussion has clearly suggested that the current system has severe disadvantages both in the number of vehicles or trains which can use a rail route in a given period, and in the actual flexibility and reliability of operation.

I propose, in outline, a scheme, which, in principle, does not have these faults, and which in addition uses modern electronics in a modern fashion, that is to say in a manner other than to duplicate in electronic form the original, basically mechanical Block signalling system.

It is important to recognise at the outset that a complex high-capacity telecommunications network is not required. (unlike the present system, where all the information is taken to a central point, processed, and then used to send instructions back to the trains concerned), since by using a properly designed decentralised system the requirement reduces to communicating over the relatively short distance from a vehicle to the vehicle directly in front and no more.

Although the presence of large numbers of vehicles in close proximity, suggests the possibility of mutual interference arising with the aid of opto-electronic transmission such problems should be capable of being solved.

Basic equations to determine permitted vehicle separation

In order to assess the possible vehicle density that is achievable, it is sufficient to investigate two basic equations. The first determines the maximum speed of a following vehicle, given the speed of the vehicle in front, the distance between the two vehicles and the system deceleration. The maximum deceleration is specified by the system specification and as a realistic example will be taken as one tenth of the acceleration due to gravity, 3.2ft.per sec.^{2} maximum deceleration. Obviously different values of system deceleration will not affect the basic arguments on which these proposals are based. British Rail take 2. 5ft. per sec. ^{2}. The second basic equation is the one which determines the improvement in capacity possible over a stretch of line, assuming that the two vehicles under consideration are almost touching, when compared to a conventional system. Of course the system is not limited to two vehicles as the procedure may be repeated for as many vehicles as there is room for on the railway line.

The proposed system will from now on be called the DAVE (Distance, Acceleration, VElocity) system.

Imperial units are commonly still used in Railway Engineering circles, and to conform with this practice, such units are used in the following sections.

First let the variables used in the equations be defined.

X is the distance, V0 is the initial velocity, T is the time and A is the acceleration, whilst Vtr1 and Vtr2 are the velocities of the respective vehicles and LM is the distance between them.

Now referring to fig.4.1.

Now $X= V0.T + \frac{1}{2}.A.T^2$ (from Newtons Laws) (1)

Thus $dX/dT = V0 + A.t = 0$ (for stationary vehicle) (2)

Whence $T=-V0/A$ and $T^2 = V0^2/A^2$ (3)

sub. 3 in 1 $X=V0.-V0 + \frac{1}{2}.A.V0^2/A^2 = -V0^2/A+\frac{1}{2}.V0^2/A =$

$-V0^2/2.A$ (4)

succinctly $X=-V0^2/2.A$ (5)

So that $X1=-\frac{1}{2}\ Vtr1^2 \div A$ (6)

and $X2=-\frac{1}{2}\ Vtr2^2 \div A$ (7)

now $LM= X2-X1 = -Vtr2^2/2.A + Vtr1^2/2.A$ (8)

$=\frac{1}{2}.A(-Vtr2^2+Vtr1^2)$ (9)

Therefore $-Vtr2^2 + Vtr1^2 = 2.a.LM$ (10)

whence $Vtr2=\sqrt{(Vtr1^2-2.A.LM)}$ (11)

(11) is the first basic equation. This equation is plotted out in fig.4.4. and the program is given in fig.4.3.

Under braking conditions A is negative so that Vtr2 will always be positive. Vtr2 is the maximum speed allowed for Tr2 if it is to end by just touching the vehicle in front. In other words given the instantaneous velocity of the first or leading vehicle the maximum instantaneous velocity of the second vehicle can be calculated.

Now it is of interest to compare the efficiency of the two systems the conventional and the DAVE system.

The track utilisation of the DAVE system for a given length of track with a constant speed limit is almost 100%. Indeed if all the vehicles were actually touching it would be 100% effi-

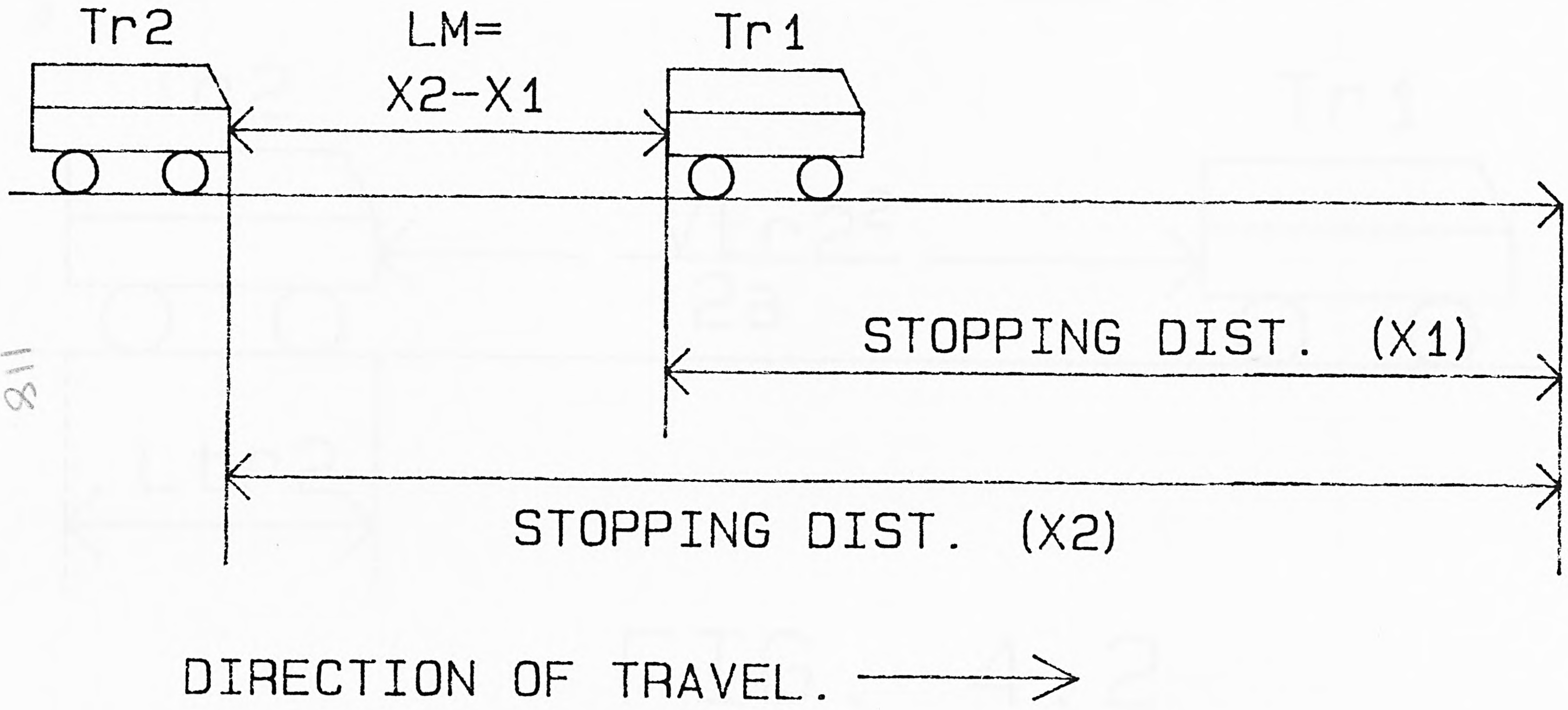

Tr2
LM=
X2-X1
Tr1
STOPPING DIST. (X1)
STOPPING DIST. (X2)
DIRECTION OF TRAVEL.

FIG. 4.1

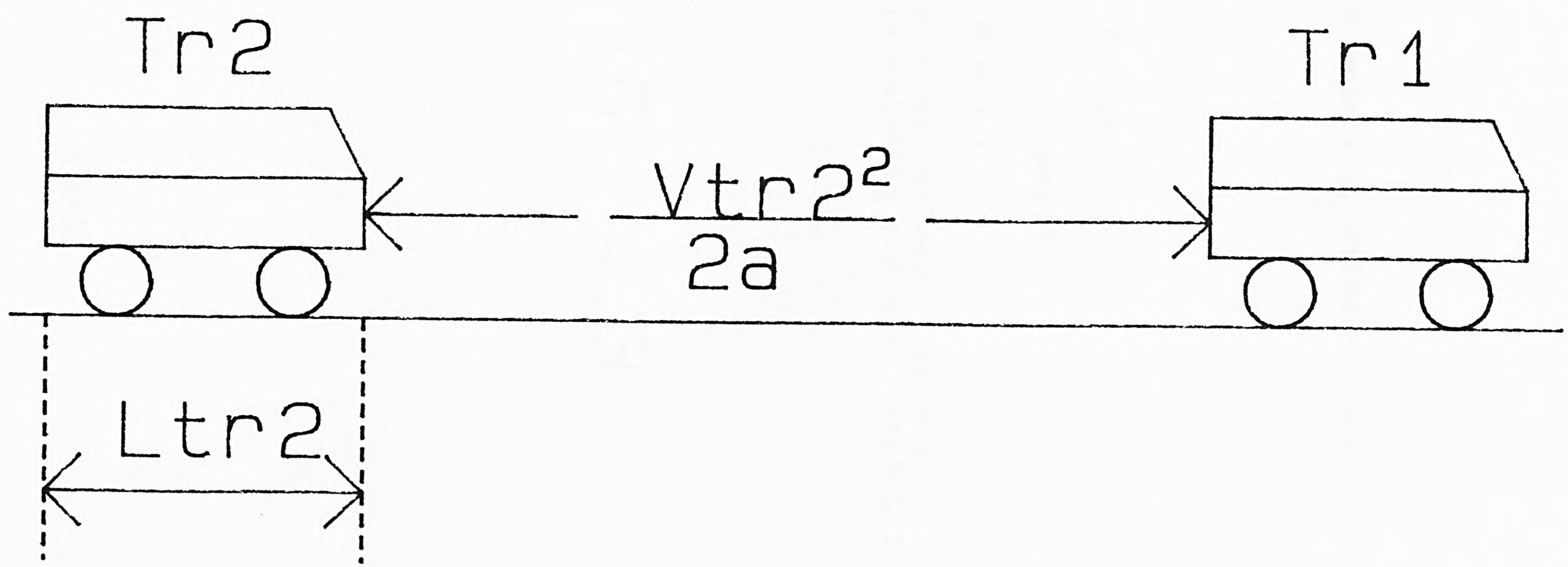
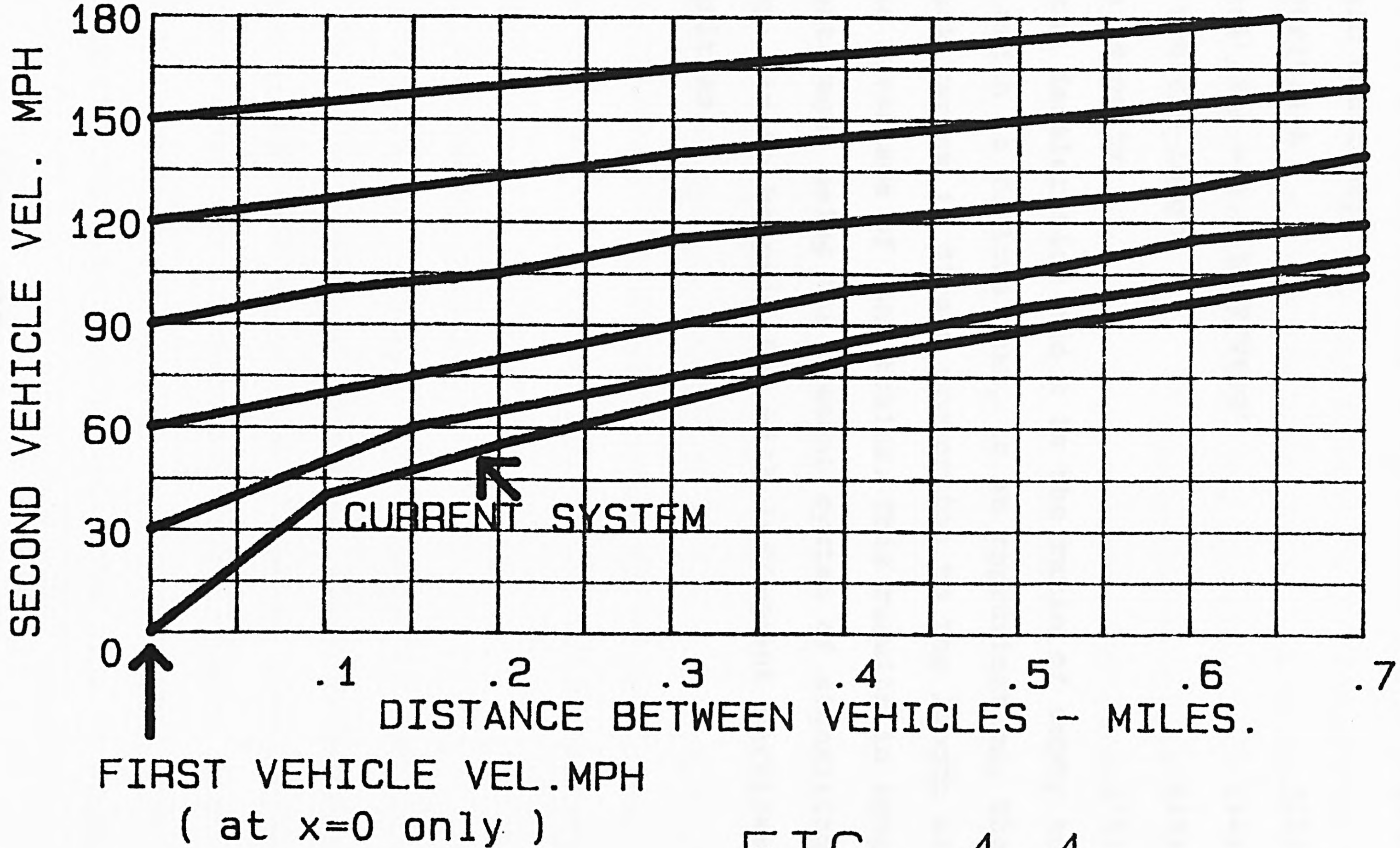


FIG. 4.2

```
program DAVE4;
        const
             a=3.2; { ft.per .sec.^2 }
             vtrl=210; { m.p.h. }
        var
             Vtr2:real;
             xl:real;
             x2:real;
             vl:real;
             v2:real;
             Lf:real;
             LM:real;
             n:integer;
begin
     n:=-1;  { Counter }

     writeln('First vehicle velocity, miles.= ',vtrl);
     repeat
     n:=n+1;

     Lf:=n*100; { ft. };

     LM:=Lf/5280; { Miles };

     vl:=vtrl*5280/3600; { ft./sec.}

     v2:=sqrt(2*a*Lf+vl*vl); {ft./sec.}

     vtr2:=v2*3600/5280; { m.p.h.}

     write('Dist.between vehicles, miles.= ',LM:2:3,'    ');
     writeln('Second vehicle velocity m.p.h.= ',vtr2:2:2);{ m.p.h.}
     until (n=40);

end.
```

FIG. 4.3



180
150
120
90
60
30
0
SECOND VEHICLE VEL. MPH
.1
.2
.3
.4
.5
.6
.7
DISTANCE BETWEEN VEHICLES - MILES.
CURRENT SYSTEM
FIRST VEHICLE VEL.MPH
(at x=0 only)

FIG. 4.4

cient. In other words there would be no empty space on the track.

The conventional system (Fig.4.2) must always have a minimum distance of $Vtr2^2/2.A$, between vehicles, because of its false basic assumption. It follows that the ratio R of occupied to empty track is $R=Ltr2\div(Vtr2^2/2.A+Ltr2)$ (12)

(the second basic equation).

for $Ltr2 \ll Vtr2^2/2*A$ (13)

$R\approx Ltr2 \div Vtr2^2/2.A = 2.A.Ltr2/Vtr2^2$ (14)

Where Ltr2= vehicle length (15)

Vtr2=vehicle velocity (16)

Where A is the deceleration and R is the ratio of empty to full track, from which it follows that, to an approximation, the capacity of a track varies in direct proportion to the length of a train. Hence the advantage of long trains. This results in long trains being advantageous using the present system of signalling but this advantage has to be paid for with infrequent services amongst other penalties.

A suggested implementation.

This scheme hinges on the combined use of three engineering approaches.

The first is to minimise to the largest practicable degree the actual occurrence of errors the second is to immediately detect them and act upon them, and the third is to ensure immediate recovery from a fault in the event of one occurring.

This is why the scheme is part of the positive approach to errors which underlies all the work in this thesis. It should be compared to the approach adopted by the 'Block' system.

The task of securing reliable communication between two vehicles and those vehicles only, can obviously be solved in many ways.

I propose using opto-electronic infrared free-space transmission, using digital transmission and exploiting the most fundamental feature of pulse code modulation to the full, that is that after the initial encoding errors virtually unlimited numbers of repeaters to regenerate the signal can be used, without further distortion.

This is necessary for two reasons.

Firstly it is convenient to have a large number of intelligent terminals which I shall refer to as posts, which both relay signals and compute at the track side spaced closely together as I shall show, no matter how the signals were being relayed, along side the track whilst secondly, with opto-electronics, the praccatical problems rapidly increase with distance.The line of sight a cutting with curves but this is not the only difficulty.In addition increased distance calling on more power than it is convenient to use as well as much increased susceptibly to 'noise' generated by sources such as the Sun, street lamps etc.

Opto-electronics is far more suited to short distance transmission of this type than radio transmission, which has the disadvantages of much increased susceptibly to interference from other users of the system these disadvantages not being relatively easily overcome by directional transmission as is easily possible with opto-electronics, and in addition radio frequency is much more prone to electromagnetic interference,

For these reasons a large number of 'posts' must be used if any appreciable distance is to be covered. Each one will make a low power demand. If sufficiently low, batteries needing replacement only every few months or so could be used,and so provide independence of an electrical mains distribution system. It is possible to use a trickle-charging system with indefinite battery life by using solar power which for very low power operation is perfectly feasible.

In addition the low power requirement leads to the possibility of using cheap mass produced opto-electronic components.

Let us assume that immediately after a sample the vehicle in front, from a steady velocity starts to brake at 3.2 ft per sec.^{2}, the maximum system braking, instead of increasing steadily with distance (fig. 4. 6) we have a shortfall in distance of $\frac{1}{2}at^2$ by the time the next sample of distance is taken, i.e. the front vehicle is $\frac{1}{2}at^2$ where t is the sample rate time.

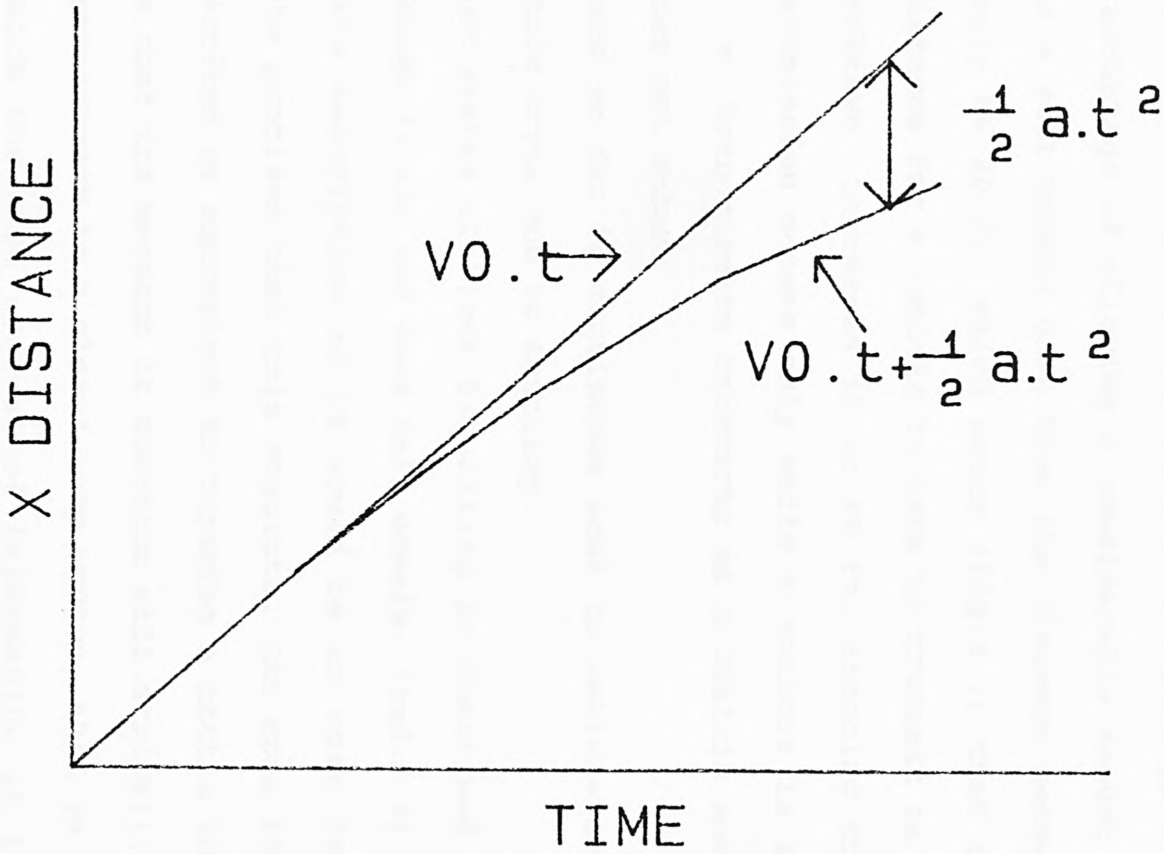


FIG. 4.6

For a 1 second sampling rate $\frac{1}{2}at^2$ gives 1.6ft. for a=3.ft.per sec. ^{2} and for 10 samples a second 0.016 ft.which is negligible, so let us chose 10 samples a second so that the effect of sampling an error is negligible.

I propose a spacing of the order of 10ft. between posts (it could vary according to circumstances), since such a close spacing means that curves in the railway line will never be a problem and that the optoelectronic link should not need too much power.

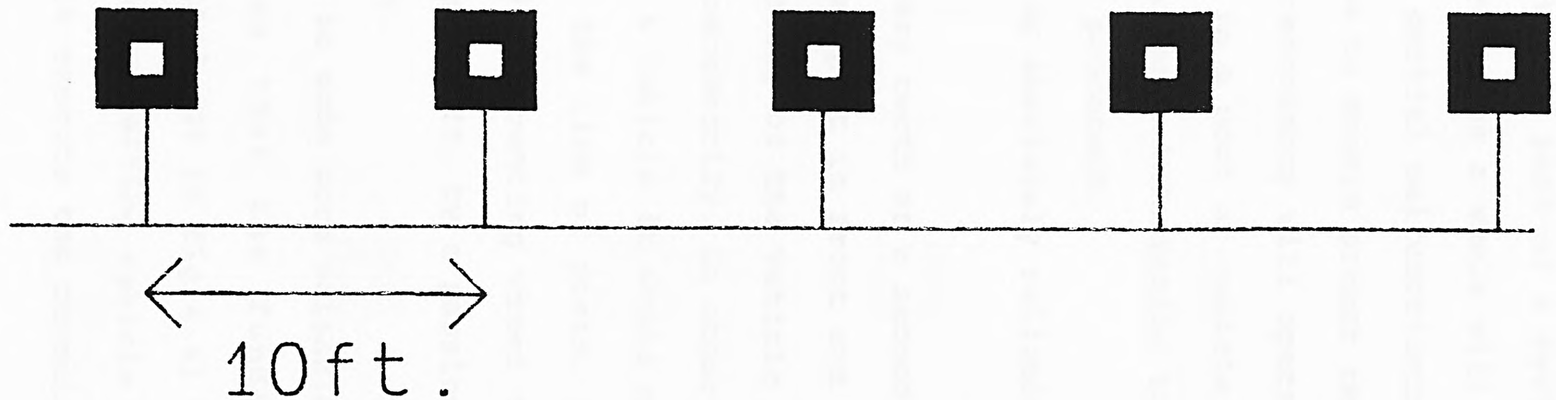
This has the advantage of allowing a considerable amount of redundancy since if a post breaks down then the distance between posts will still only be 20 ft. which means (fig.4.7) that although the usual distance for a vehicle to have to transmit is 10 ft. to a post ,breakdown increases it to 20 ft. assuming that dual direction transmission occurs only while a vehicle is approaching a post i. e. transmission backwards as a vehicle moved away from a post does not occur.

What is proposed so far is continuous post to vehicle contact with a negligible error due to sampling.

Now the present system of Block Signalling is described as 'fail safe' even though it can and does fail unsafe. (ref.4.3).

A more accurate description of it would be an open loop system, with all the problems that this suggests. (An open loop system has been described as equivalent to throwing a bottle into the sea in the hope that the message it contains will arrive!).

So the first requirement is a closed loop system, that is to say a system in which the safe arrival of information at its destination is conveyed by its arrival back, with the original



TRACKSIDE POSTS.

FIG. 4.7

values at the point of origin, with, additionally a label attached indicating that the returned information has actually originated from the target at which it had been aimed.

Another requirement is the generous use of redundancy and majority voting so that even if part of a system or a part of a vehicle breaks down the system as a whole will go on working. In addition an indication of partial malfunctioning must be sent to appropriate personnel so as to enable prompt repairs to be made.

This requirement of redundancy will operate at two levels, that is to say internally in a post or vehicle and externally as when a post and vehicle exchange information to see if it is the same and has been properly processed.

Now a DAVE system using absolutely reliable parts would look like fig. 4.8.

Each vehicle would every tenth of a second send its position and velocity to the nearest post in front and receive from this post the position and velocity of the vehicle in front. No post would relay information unnecessarily. In other words once a post received information from a vehicle it would no longer continue to relay information down the line of posts. The position of a vehicle would be measured by counting wheel revolutions. This counter would be reset every 10 ft. by a passing post. Wheel slip could be taken into account.

How is this system to be made more reliable?

Firstly by checking that the fundamental formula $Vtr2=\sqrt{(2.a.Lm+Vtr1^2)}$ (plotted out in fig.4.4) has been executed correctly. This can be done by getting vehicle Tr1 and its nearest post (fig. 4.9) to both execute the formula and for each to

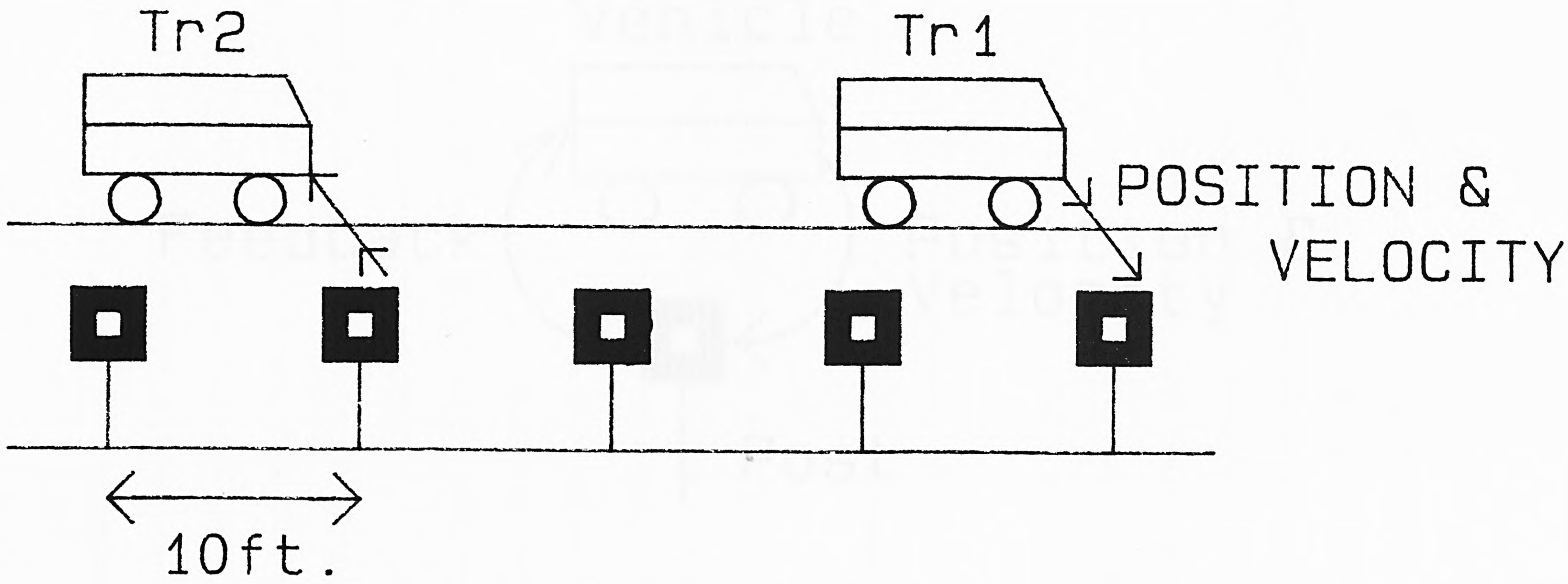


FIG. 4.8

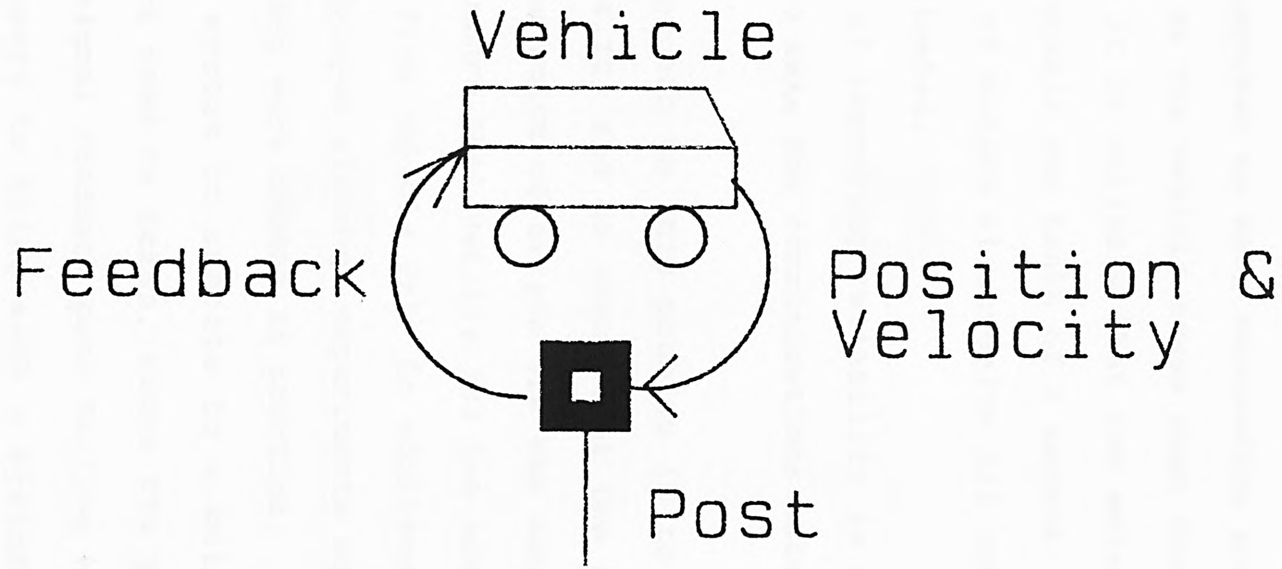


FIG. 4.9

check that the other has obtained the same results.

If they are different an error has been detected and the brakes are applied.

The advantages of using both the vehicles computer and the posts computer are the obvious one of having desirable duplication and the less obvious but equally desirable advantage of having a different computer on each succeeding post to check on the vehicle computer as the vehicle moves down the line, this is an advantage because it is unlikely that two adjacent computers will develop a fault within one tenth of a second. In addition in view of the low cost of modern electronics all computers in the system could be triplicated.

A second method of improving reliability is more difficult to achieve. This is to make the communications between posts more reliable.

The simplest approach to this problem is to have feedback between posts (fig. 4.10) and to check that the signal and the feedback signal at the point of origin are the same, using point of origin labels to ensure that vehicle Tr2 has actually received the information sent from vehicle Tr1. In addition error detecting codes could be employed although experiments would need to be conducted to see if they were needed in practice.

If this were a system to operate in a building further precautions would not need be taken, since the probability of both a signal and a signal feedback path failing together within the short time necessary to bring such a system to a halt is remote especially if parity bits and triple redundancy voting

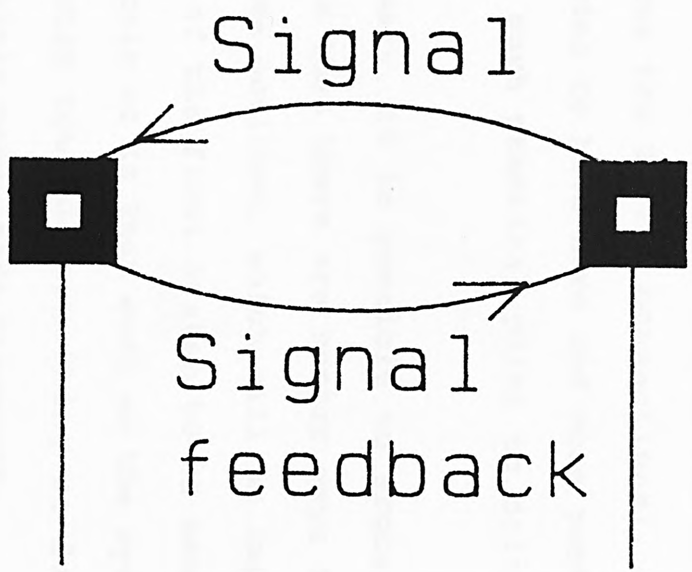


FIG. 4.10

were employed. But this is nor the case and the possibility of a post being physically destroyed must be considered.

The result would be that the system would fail to operate since, obviously, the necessary interchange of information would cease. However provided this lack of information was immediately acted upon, and the brakes applied so as to decelerate at the specified system value all would be well.

In order to overcome this problem a capability for post to post transmission may be added, which takes advantage of the fact that transmission over not just ten but twenty ft. is a system requirement (fig.4.11) in which connections are shown for 10 ft. and 20 ft., the shorter connections are the 10 ft. connections, the longer connections the 20 ft. connections). Obviously the scheme could be expanded to have more and more posts leaped over but whether there is much practical point to doing this is debatable.

Finally, in summary, it is possible to argue over the details and I am certain that there are better ways to implement a DAVE system than I have outlined, which will be implemented as a result of experience of the first system to be manufactured but what I am equally certain of is that even as the system stands it represents a large step towards something we do not have at present and most certainly need, safe transport.

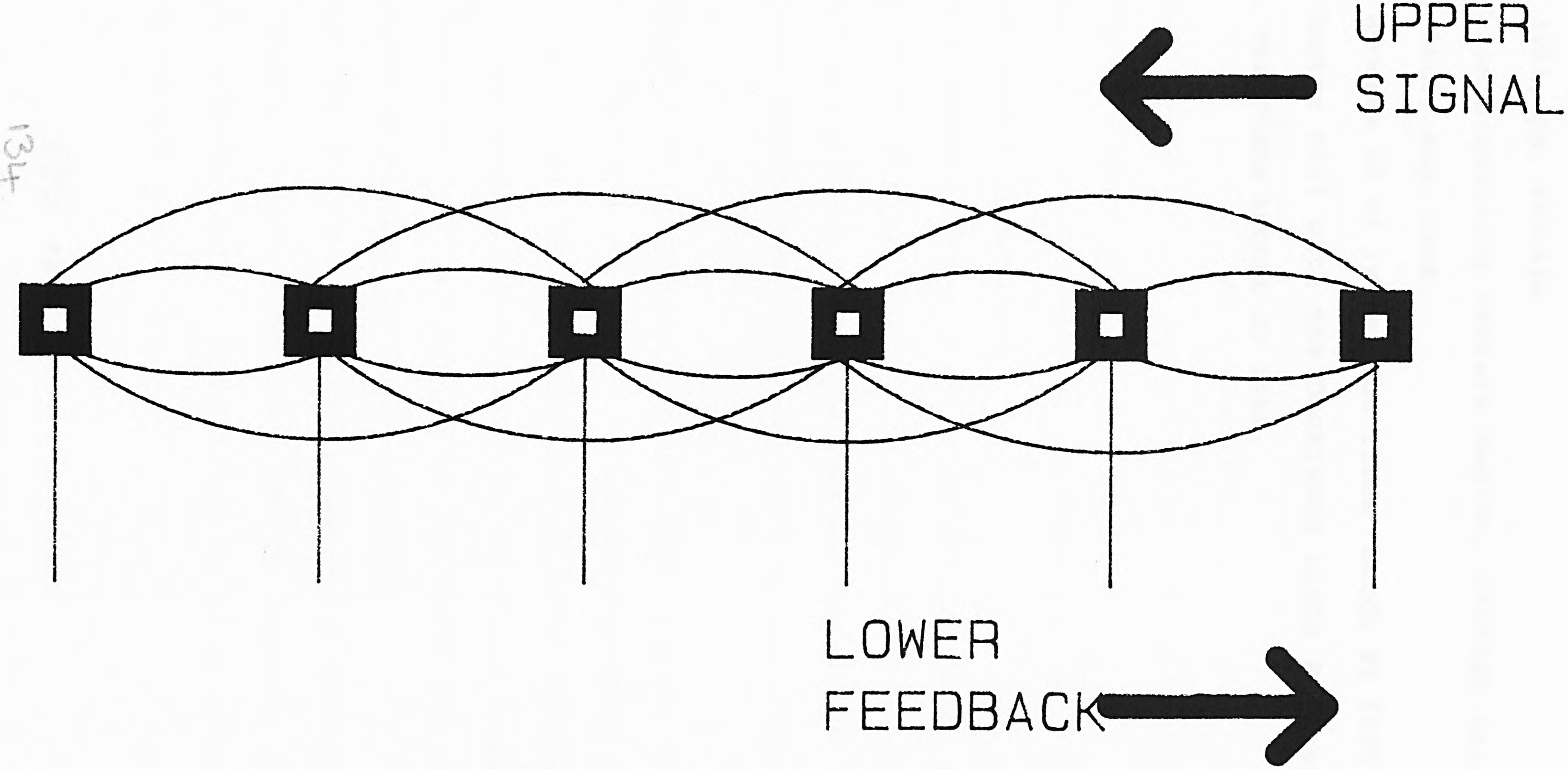


FIG. 4.11

References

4.1 J.Johnson & R.A.Long, British Railways Engineering 1948-80, Mech.Eng.Pub.Ltd. 1981, pp. 389-436.

4.2 Main line & Urban Signalling Eastern Region, British Rail, GEC-General Signal Ltd., Aug. 1979.

4.3.Drivers Union accuses BR of lying. The Times March 29 1989

4.4.Interview by George Hill with the Docklands Light Railway, Managing Director. The Times August 23 1989.

CONCLUSIONS

The work on 'Cross Coupling' shows the power of a method which amplifies errors in order to subtract them from the distorted output.

I claim that 'Cross Coupled Error Takeoff' is a basically different method, when compared to Negative Feedback or Feedforward, of reducing error in a circuit, combining as it does the advantages of Feedforward whilst at the same time minimising the disadvantages of Negative Feedback.

In other words much reduced error is possible without significantly increased stability problems.

It enables power efficient amplifiers to be designed with an improvement factor of 40 dBs per decade, as compared to 20 dBs per decade for negative feedback, as well as making the design of power amplifiers much easier since problems with output stages are much eased.

The railway signalling proposals are included to show the generality of ideas based upon error and are illustrated by an outline design for the signalling work. The use of error enables an outline design to be proposed which uses an efficient algorithm to control speed as well as making it possible for an arbitrarily high degree of reliability to be achieved.

The work on the Third order polynomial has shown that an error controlled 'spline' function has much better fidelity of visual reproduction compared to the well established zero and first order data compression algorithms.

FUTURE WORK

The amplifiers are the subject of a serious effort on my part to get Industry interested as there are obvious applications in both signal processing, for example A-D convertors, as well as in power amplifiers.

Whilst, although I intend to publicise, by means of the technical press, my views on Railway Signalling, I do not think that those in a position to do something about it will.

In my view the Third Order Polynomial is ripe for commercial application which as it has been published in 'Signal Processing' may well occur. From the point of view of performance, this can probably be improved since the basic algorithm and its application are both amenable to further development