LOW POWER HIGHLY LINEAR CMOS OUTPUT BUFFER

FOR

LOW TEMPERATURE INFRARED FOCAL PLANE ARRAYS

By

David Paul Foley

B.S. Physics, M.I.T. (1980)
B.S. Elec. Eng., M.I.T. (1980)
M.S. App. Physics, Stanford University (1982)

SUBMITTED TO THE DEPARTMENT OF
ELECTRICAL ENGINEERING AND COMPUTER SCIENCE
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF

MASTER OF SCIENCE

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
June, 1992

Signature of Author .....
Department of Electrical Engineering and Computer Science, May, 1992

Certified by .....
Charles G. Sodini, Associate Profesor of Electrical Engineering

Certified by .....
Dr. Neal Butler, Loral Infrared and Imaging Systems

Accepted by .....
Arthur C. Smith, Department Committee on Graduate Students

LOW POWER HIGHLY LINEAR CMOS OUTPUT BUFFER

FOR

LOW TEMPERATURE INFRARED FOCAL PLANE ARRAYS

By

David Paul Foley

ABSTRACT

A design for a low power highly linear CMOS output buffer is presented. The buffer is designed for use on NASA's Atmospheric Infrared Sounder's scanning focal planes, and is intended to operate at 60 K. Given the requirements for power dissipation, load capacitance, and settling time a class A output buffer would be unsuitable. The approach taken uses a class AB amplifier configured as a charge integrator. Thus, it converts a charge packet in the focal plane multiplexer to a voltage which is the output of the focal plane array.

Measured performance meets or exceeds design goals in the majority of devices tested. With a quiescent current of 18 microamperes, and a load capacitance of 100 picofarads, settling to .03 % is obtained in less than 3.5 microseconds. Integral nonlinearity is less than .03% over 5.5 volts with a 6 volt supply. An output current limit imposed by device mismatch is responsible for low slew rate in some units.

Measured performance agrees with predicted performance in most areas. Excess high frequency noise is noted in the amplifier input referred noise voltage, but consistently high levels are also found in device measurements. Excess sampled noise contributions from the switch elements is observed, with RMS noise 3.5 times greater than the theoretical value. Measured maximum positive output current is larger than expected.

A technique for calculation of loop gain, suitable for numerical computation, is presented and shown to give results consistent with traditional approaches.

Thesis Supervisor : Neil R. Butler, PhD

Title: Engineering Fellow, LORAL Infrared and Imaging Systems

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

First and foremost I wish to thank my wife, Maureen. For this undertaking, sacrifices were made. More than any other single factor, her support was responsible for the success of this thesis.

I wish to thank Dr. Jill Wittels for establishing the Masters of Industry program at LORAL. For at least one person this program has provided a path where there was none.

I wish to thank Professor Charles Sodini for his comments and advice regarding this document. Dr. Neal Butler is greatly appreciated for introducing this topic to me and providing many insightful discussions which kept things on track.

I wish to thank the AIRS program team for providing a vehicle for this research. Special thanks go to Jim Stobie for his efforts to maintain an atmosphere in which this research could be conducted as part of an ongoing development program. Jim has also helped with many useful suggestions. I wish to thank David Verville and Jay Baker for help in setting up and maintaining the test set as well as assistance in performing measurements. I wish to thank Jim Stobie and Maureen Griffis for providing much of the device measurements that I have used in this analysis. I wish to acknowledge the tireless efforts of Margene Gustafson at ADS in laying out the test chip and those of Rosanne Tinkler in checking it.

I am grateful to Dr. Sato Iwasa, as head of the Focal Plane Lab at LORAL, for his efforts to provide the right environment for all of the programs and individuals in his group to thrive.

Finally I wish to thank all my co-workers in the the Focal Plane Group at LORAL. Their help and friendship was very important in making this research effort successful. In particular I wish to thank all those who refrained from borrowing my equipment while I was away at school.

This research was supported by JPL under contract 958970, and NASA prime contract NAS 7-918. This support is gratefully acknowledged.

1. INTRODUCTION

1.1 Overview of NASA's Atmospheric Infrared Sounder

The Atmospheric Infrared Sounder (AIRS) is a research instrument that has been selected to fly on the first Earth Observing System (EOS) polar orbiting platform (4th qtr. 1996 launch)[1]. The EOS is designed to improve understanding of the earth as a system through remote observation of our environment with a comprehensive suite of instruments.

AIRS is the chief atmospheric sounder on EOS and will provide data for NASA's earth science research programs as well as NOAA operational programs (i.e. weather prediction). A major objective of AIRS is providing data which can be used to determine atmospheric temperature profiles accurate to 1^{o} K in a vertical resolution of 1km.

To meet these and other requirements AIRS spectral resolution will cover the range from 3.4 um to 15.4 um with more than 3600 spectral measurements at a resolving power of $\lambda/\Delta\lambda$ = 1200. The noise equivalent change in temperature must be 0.2^{o} K.

The AIRS instrument design consists of two multi-aperture, pupil imaging grating spectrometers. One spectrometer covers the spectral region (3.4-8.6 um) while the other (8.6-15.4 um). Each spectrometer contains an echelle grating and utilizes detection in several grating orders. The optical design provides spectral separation by imaging different spectral regions to different locations on the focal plane.

The focal plane/dewar assembly consists of optical filters, IR detectors, multiplexers, and dewar elements with stringent performance and alignment requirements. There are two focal planes within a common dewar both of which operate at 60^{o}K by means of a closed cycle refrigerator. The dewar assembly

isolates the 60°k focal planes from the 155°k optical benches. The focal planes are comprised of linear mercury cadmium telluride (HgCdTe) detector arrays. Each linear array is covered by a different optical bandpass filter to isolate the proper grating order and reject stray radiation.

Every 2.667 seconds AIRS will scan a +/- 49.5 degree swath in a direction perpendicular to the EOS satellite's ground track. The scanner employs a continuosly rotating 45 degree mirror and provides a spectral calibration, and a two point radiometric calibration. The AIRS instantaneous field of view (IFOV) is 1.1 degree, providing a 13.5 km diameter footprint while looking straight down from a 705 km orbital altitude.

After considerable flowdown from the system requirements, a set of design goals for the output buffers of the focal plane multiplexers were determined.

1.2 Output Buffer Characteristics

Before considering the design goals for the AIRS output buffer a general discussion of output buffer characteristics is appropriate. As a circuit block the function of a buffer is to accept a signal from a previous circuit block and present it as input to a following circuit block. In the simplest case the input and output signals are circuit voltages. The signal domain of the buffer may be continuous or discrete. The same distinction can be made with the time domain. The AIRS buffer will deal with a continuous value , discrete time signal.

Figure 1.1 shows an idealized buffer for which Vout=Vin. Real buffers have performance limitations which must be considered when considering there suitability for a particular application.

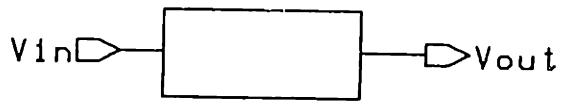


Fig. 1.1

When a signal passes from one stage in a circuit to another it is generally reduced and distorted by the voltage divider that is comprised by the output impedance of one stage and the input impedance of the following stage. The principal reason for using a buffer is to improve the impedance relation that would otherwise exist between two portions of a circuit. An ideal buffer would have infinite input impedance and zero output impedance which would eliminate that portion of signal distortion which arises from this voltage divider. The input and output impedances of real buffers must be designed to interface with adjacent circuit blocks.

Aside from impedance considerations a buffer may introduce additional errors into the signal. For DC signals these limitations may be classified as input range, offset, gain, and nonlinearity. An ideal buffer would have an input range equal to the entire signal range, an offset of zero, a gain of one, and a nonlinearity of zero.

For AC signals these limitations may be classified as bandwidth, supply rejection,noise, and slew rate. An ideal buffer would have an infinite bandwidth, infinite supply rejection, infinite slew rate and zero noise.

Other buffer characteristics would include power consumption, physical size and semiconductor process. To design the AIRS output buffer values for the above characteristics need to be developed.

1.3 Development of Design Requirements for Output Buffer

Some of the design requirements for the AIRS output buffer follow quite naturally from system requirements, others have arisen from baseline budgets which represented a system designer's best estimate of an optimal configuration. With a projected 4th qtr. 1996 launch, the AIRS system will certainly undergo many revisions before that date. The requirements form the starting point for the design. An appreciation of the rational for these requirements is important for a number of reasons. First it indicates the relative importance of the design requirements. This is important because it may be possible to achieve better than spec performance. In this case the designer should improve performance in that area of the design which will give the most system return. A better overall design will result if the specifications convey more information than minimum requirements. Secondly, the designer may judge for himself the validity of the requirements. An inaccurate or overly conservative requirement may render a solution impossible or overly complex. A designer should know when to challenge a requirement in order to achieve overall system improvement.

To minimize heat loss from the 60°K focal plane to the off focal plane electronics which will be at a substantially higher temperature, it is planned to use interconnections which have low thermal, and therefore unfortunately low electrical conductance. For this reason undesirable ground potential differences between the focal plane and the electronics are expected. To combat this signal corruption differential outputs are a requirement. For increased common mode rejection the differential output must be balanced. This is to say, that the impedance of each of the two lead outs must be matched. In this way their response to external stimuli will be as matched as is possible and greater rejection of common mode signals will result.

Due to the physical layout of the AIRS instrument, it is expected that three feet of cable will connect the focal plane to the electronics. The capacitive load due to this cable is projected to be 100pF.

The largest detector array is baselined to have 208 channels. Each channel integrates photocurrent for 1.23 milliseconds. This time results from a dwell time of 22.22 ms and a desire to have 18 samples per dwell in order to perform image smear compensation. The output data rate from the largest FPA multiplexer must therefore be greater than 208/1.23ms = 169 kHz. Because all channels are designed to integrate simultaneously a sample and hold function must be performed prior to readout and the output data rate requirement has been specified at 200 kHz. This leaves 5 microseconds to output each pixel.

Available cooling capacity limits focal plane active power dissipation to 60 milliwatts. 8.5 milliwatts of this amount has been budgeted to outputs. NASA redundancy requirements dictate two output buffers per multiplexer and with 17 multiplexers on the short wave focal plane this leaves 8.5mw/(17x2) = 250 uW available for each output.

A 1.5 micron CMOS process has been selected for the focal plane multiplexers. VDD-VSS is set for 6 volts. As large a signal range as possible is desired as it increases the dynamic range of the system. A minimum signal range of 3 volts has been baselined.

It is certainly desired that the output buffer not degrade the signal to noise ratio of the signal being output. On the other hand as long as the noise introduced by the output buffer is small compared to other noise sources then it is not worth trading off other performance parameters in order to reduce the noise still further. If one considers the channel of the AIRS instrument servicing the wavelength 3.4um one finds that taking into account the scene which is being viewed, the optics of the instrument, and the detector characteristics that the

maximum number of signal photoelectrons that can be collected in the 1.23 ms integration time is $4x10^4$. It is desired to integrate this charge onto a capacitor small enough to generate a full signal range of 3V but the smallest capacitor that can be consistently fabricated is estimated to be around 20 femtofarads. On this size capacitor a signal of .3V results. The input cell on the focal plane multiplexer which couples with the 3.4 um HgCdTe detector has an input referred rms noise of 40 electrons. These numbers indicate that the signal to noise ratio for this channel will in no case be any better than about 1000 to 1. By providing gain to this channel, the output signal could be brought up to 3Volts. In this case the rms noise level of the signal would be 3mV. Analysis of each channel shows that at the output of the focal plane the signal will always have a rms noise greater than 1 millivolt. Based on this the rms noise of the output buffer has been specified at 500 microvolts. Since independent noise source combine in quadrature, output read noise would always contribute less than 12% of the total noise.

Settling time requirements depend on off chip sample and hold characteristics and the maximum signal change from one pixel to another. As neither of these have been fixed, the author has adapted a goal of settling to .03% (1 part in 3000) in 4μsec for a 3V step output change. This leaves 1μsec for the external sample and hold to acquire the output which will have settled to within a millivolt for a full signal range step change.

An initial estimate of available real estate was set at .5x.75 mm^2. As only two output buffers exist on each focal plane this specification is not a very hard limit and somewhat more space could probably be made available without very much difficulty.

Focal plane multiplexer architecture dictated that the output buffer must take charge as its input signal from a bus shared by up to 208 channels. Pixel pitch is set at 50 microns which determines the length of the bus and therefore its

capacitance. Including the capacitance from 208 source/drain diffusions hanging on the bus, the total stray input capacitance is 1.7 picofarads. The sample capacitor feeding the bus from each channel is 3pF (fig.1.2).

Nonlinearity of the multiplexer is targeted to be better than .5%. The output buffer itself hase been budgeted to have .1% nonlinearity. This figure is interpreted as integral nonlinearity over the specified output range. A summary of output buffer requirements is given in Table 1.1

Semiconductor Process	Orbit 1.5u twin well
Operating Temperature	60°K
Data Rate	200 KHz
Settling Time(to .03% for 3V step)	4μs
Output Range	3 volts
Integral Nonlinearity	.1%
Noise	500 μV RMS
Power	250 μW
Capacitive load	100 pF
Area	.375 mm^2

TABLE 1.1

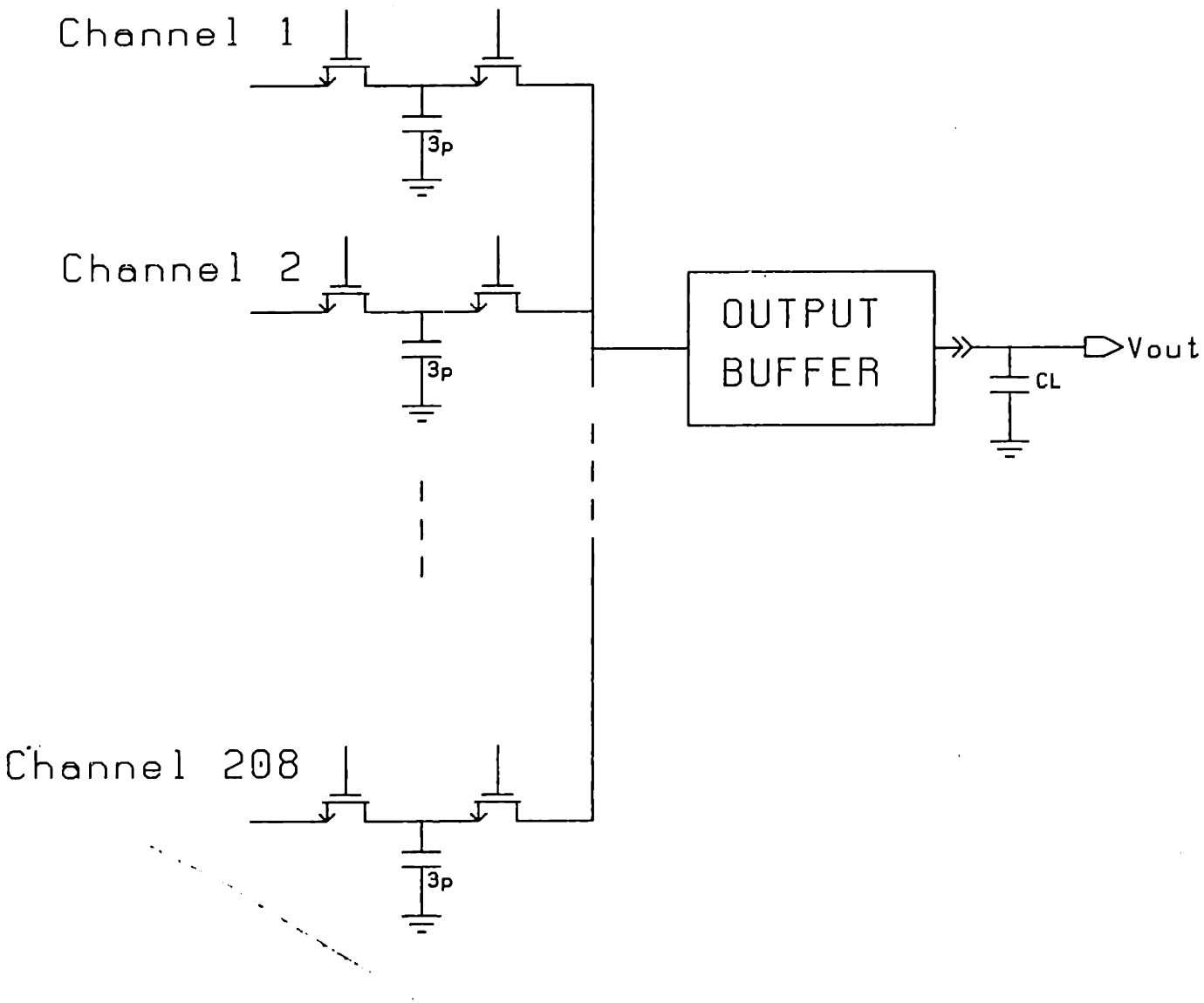

Channel 1
3p
Channel 2
3p
OUTPUT
BUFFER
Vout
CL
Channel 208
3p

Fig. 1.2

1.4 Review of Previous Output Buffer Designs

1.4.1 Hybrid IR Focal Plane Arrays

1.4.1.1 Source follower with voltage input

The simplest output structure is the source follower (fig.1.3).

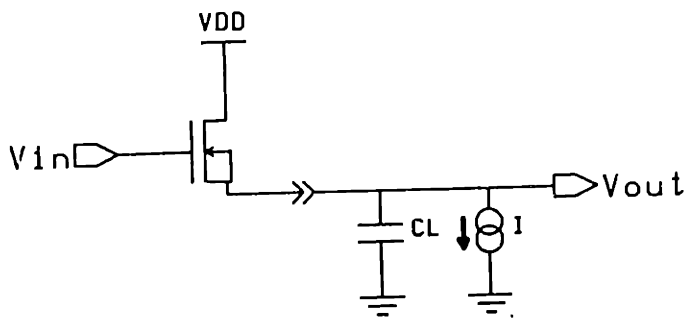


Fig 1.3

Because of it's simplicity, this has been enormously popular. This circuit comes in p and n channel varieties which may have their wells connected to there sources if the fabrication allows. Characteristics of this circuit can be found in introductory circuit texts. The current source is generally located off chip to reduce on chip power dissipation. Because it is the prevalent output buffer a summary of its characteristics follow.

Signal range: The output for the n channel follower is one V_{gs} below the input. This V_{gs} drop is often lost from the overall signal range. Unfortunately, MOS transistor threshold increase by about .4 volts when cooled to LN2 temperatures so zero bias thresholds of up to 1.4 volts may be anticipated. If the output transistor is not source connected a further increase in threshold occurs due to V_{sb}. When operating in strong inversion the transistor V_{gs} exceeds the threshold by an amount Δv which depends on the bias current.

$$\Delta v = \sqrt{((2I/k(W/L))} \qquad (1.1)$$

Equation 1.1 is derived from the square law relation for MOS drain current and is valid for V_{gs} sufficiently above threshold yet small enough so that mobility degradation effects have not yet become apparent. Even at LN2 temperatures (77°K) MOS device equations such as those used in SPICE are generally valid. Certain parameter values do change in a way which is not consistent with the temperature dependance assumed in SPICE. This is because SPICE was not designed for cyrogenic operation. One must extract parameters from device data measured at operating temperature to use for cryogenic circuit design. In order to recover some of the lost signal range the source of an n channel follower can be allowed to go below VSS if it's well is tied to it's source. In this configuration the voltage on the well substrate junction exceeds VDD-VSS. This may be acceptable because the well to substrate breakdown voltage is generally larger than the source/drain diffusion to well breakdown voltage because it is a more lightly doped junction.

Offset, Gain, Linearity: The offset of a voltage follower has just been described. The small signal gain of a source connected follower is given by

$$a_v = g_m r_o/(1+g_m r_o) \qquad (1.2)$$

This figure can be very close to unity at low currents because the product $g_m r_o$ increases with decreasing current until the subthreshold region of operation is entered. However n channel followers can suffer severe degradation of r_o at high V_{ds} due to impact ionization. This degradation results in a drop in small signal gain and a nonlinearity in the overall transfer characteristic.

Impact ionization occurs because the electric field in the pinch off region of the transistor channel becomes quite large as the Vds increases. Despite fabrication steps to keep this field low scaling laws dictate that smaller geometry processes must use higher doping densities and the mechanisms for large electric

field strengths are ever present. When electric field strengths reach a certain value known as E_{sat} the energy acquired by a drift electron between collisions is enough to create an electron hole pair. This energy loss mechanism is sufficient to keep the drift velocity of the electron from increasing as the electric field is increased beyond E_{sat}. This is why E_{sat} is known as the saturation field strength. At large V_{ds} the field in the pinch off region is often large enough for impact ionization to be a major effect. Because the electron created in the electron hole pair is free to itself cause further ionizations avalanche multiplication takes place and the drain current increases rapidly with V_{ds}. In an n channel device the excess electrons created in the pinch off region are attracted to the drain while the holes are attracted to the substrate. For this reason a corresponding increase is seen in substrate current when impact ionization is taking place. Impact ionization is not generally a problem in p channel transistors because holes have lower mobilities and therefore larger saturation field strengths.

A follower which is not source connected will have a nonlinearity due to the back gate effect. The back gate effect increases the threshold of the transistor by an amount which depends nonlinearly on voltage between the source of the transistor and the well of the transistor.

Bandwidth: The small signal bandwidth of a follower is given by

$$\omega_{3db} = gm / CL \qquad (1.3)$$

As it consists of a single transistor, all the power dissipated by a source follower is used in generating g_m and hence bandwidth. In subthreshold operation the g_m of the transistor is given by

$$g_m = I/\eta \cdot V_{th} \qquad (1.4)$$

while in strong inversion it is given by

$$g_m = \sqrt{(2 \cdot I \cdot k \cdot W/L)} \qquad (1.5)$$

The transition between these expressions can be more than a decade broad in drain current (fig 4.14). In order to obtain the most g_m for a given current it is desirable to operate the transistor in the subthreshold region. In the subthreshold region the g_m does not depend on the width and length of the transistor so that if the transistor is too deeply into the subthreshold region it is larger than it need be and has more gate and stray capacitance than it need be to obtain a given g_m. A good tradeoff of these characteristics exists in the transition region between subthreshold and strong inversion. Actual device measurements are required to determine g_m accurately in the transition region.

Slew Rate: A serious drawback of the follower is its slew rate limitation. For the n channel follower shown in fig. 1.3 there is no positive slewing limitation, but

$$SR- = I / C_L \qquad (1.6)$$

This limitation can become severe if the capacitive load is large enough. Operation at fixed power and speed can become impossible.

Noise: The input referred noise model of a MOS transistor has been given by

$$V_{eq}^2/\Delta f = 8kT/3g_m + k_F/WLC_{ox}^2 f \qquad (1.7)$$

Since the source follower has relatively large g_m and large gate area for a given current it has good noise properties.

Input, Output Impedance: The source follower connection has good input impedance since only a fraction of the C_{gs} of the transistor appears as input impedance. Because the follower has near unity gain the voltage of the source follows the voltage at the gate. For this reason Vgs changes by a small fraction of the input voltage change and the effective capacitance presented by C_{gs} is diminished by this fraction. The gate to drain overlap capacitance C_{gd} can be quite small in all but very large devices. The output impedance is given by $1/g_m$ and for a follower designed to give maximum bandwidth for a given power dissipation the impedance is minimized.

1.4.1.2 Source follower with charge input

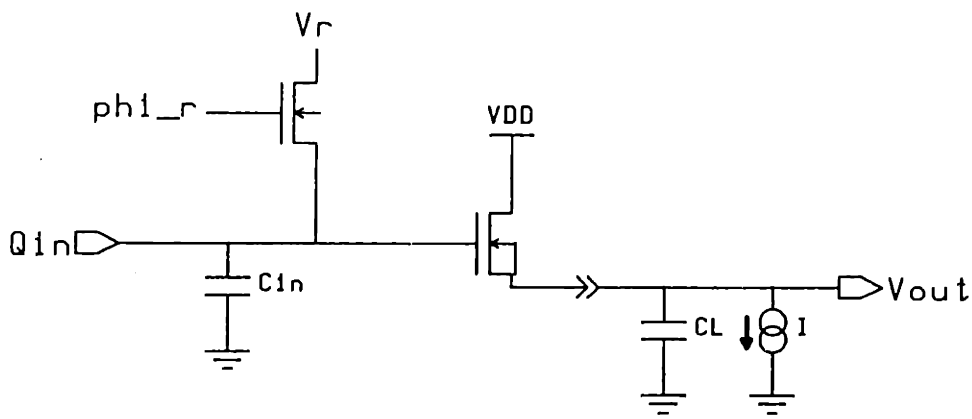


Fig. 1.4

The source follower can also be configured to accept charge as an input. This technique was prevalent in hybrid IR focal plane arrays with CCD multiplexers and can also be used in CMOS multiplexers. It is basically the same as the floating diffusion amplifier used in silicon CCD imagers. All of what was said about a voltage input source follower applies here.

In addition, another noise term is present due to the resetting of the input capacitance. When a capacitor is reset to a given voltage it is found that the voltage on the capacitor after the reset operation differs in a random way from the reset voltage. This phenomenon is due the the thermal noise in the switch element. The spectral voltage noise power in the switch element is proportional to its resistance. The switch resistance and the capacitor form a low pass filter whose noise bandwidth is inversely proportional to the switch resistance. The mean square noise voltage is equal to the product of these terms and is independent of the switch resistance. This fluctuating noise voltage is sampled on the capacitor when the switch is opened. It is interesting to note that the switch opening may be looked on as the switch resistance becoming very large. In this model the noise bandwidth of the the low pass filter becomes very small and the voltage on the capacitor is constrained to fluctuate quite slowly compared to when the switch was in the on state. In this way the fluctuating noise voltage becomes sampled when the switch is opened. The RMS value of this fluctuation is given by

$$V_{RMS}^2 = kT/C \qquad (1.8)$$

Unlike CCD imagers the noise already present in the input charge packet Q is generally larger than the read noise of the output amplifier. Read noise of the output amplifier is seldom a dominant noise source in hybrid IR focal plane arrays so amplifier designs are not driven by low noise requirements.

1.4.1.3 Source follower with bipolar buffer

In order to solve the negative slewing problem of the source follower, the circuit of fig 1.5 has been used.

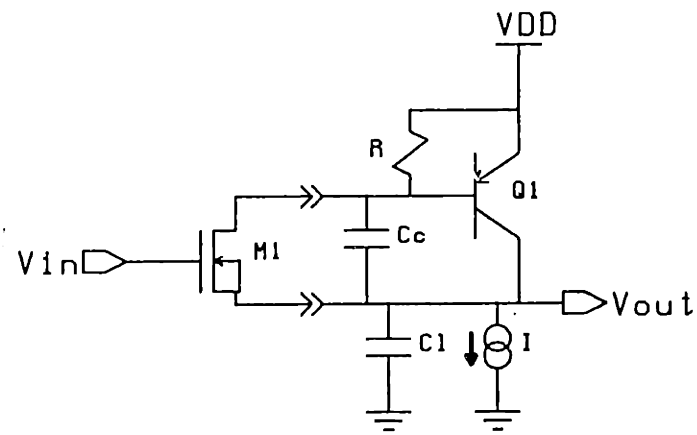


Fig. 1.5

A detailed analysis of this circuit shows similar small signal characteristics as the follower while the slewing characteristics can be determined by a much larger current than the on chip quiescent current. Unfortunately the number of lead brought off chip must be doubled in this configuration. The pair of leads brought off chip for this output buffer do not have matched impedance characteristics. To provide a differential balanced output would require four leadouts which is a doubling of the allotted two. This is not a problem for a focal plane with a single output, but the airs focal planes (long wave and short wave) have a total of 68 outputs with an allotted 132 leadouts. It would be a rather poor trade to double this number.

1.4.2 Monolithic CCD Image Sensor Output Amplifiers

Charge Coupled Devices (CCD's) were invented in 1970 [2]. Although originally conceived as a memory device, the primary use of CCD's over the past two decades has been as an imaging sensor.

Evolution of output amplifiers for CCD operation has been in the direction of lower noise. These systems apparently have not required operation in a domain where speed, power, and capacitive load are incompatible. All outputs have been based on a single transistor and differ only in the way in which the signal charge is coupled to the transistor.

As process improvements in CCD manufacture improved the read noise of the output amplifier became the dominant noise source.[3] This provided a considerable impetus to improve the output amplifier design. This is especially true at low light levels where the shot noise of the signal packet is very small.

The first CCD output amplifier was the floating diffusion amplifier [4,5] which is shown in figure 1.6

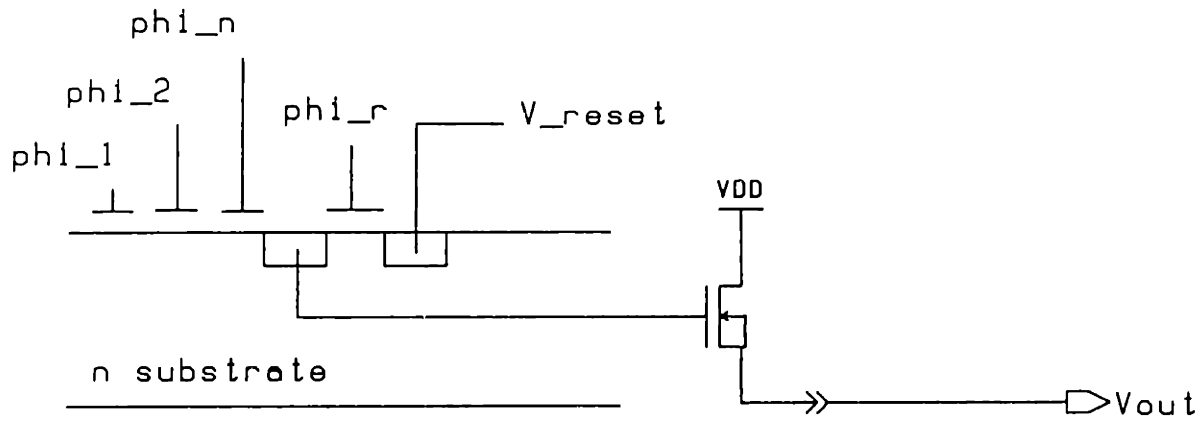


Fig. 1.6

The signal charge is transferred to a floating diffusion where it generates a certain voltage according to the capacitance of the diffusion. This voltage appears at the gate of a MOS transistor where it can be read out with the transistor acting as a source follower. The diffusion is then reset to fixed voltage and the

circuit is ready for the next output. There are three sources of noise in this circuit. They are the kT/C reset noise of the floating diffusion, MOSFET 1/f noise and MOSFET thermal noise. The output noise voltage can be referred to the input by multiplying by the diffusion capacitance.

The reset noise of the diffusion capacitor was eliminated by a technique known as correlated double sampling (CDS) [6] in which two samples were taken for each output, one just after resetting the floating diffusion and another after the signal charge has been transferred to the floating diffusion. By subtracting these two measurements, the reset noise is eliminated from the measurement. This technique also reduces some of the low frequency 1/f noise.

The thermal noise is inversely proportional to the transconductance of the amplifier so to reduce this a large width transistor is desirable. The 1/f noise is inversely proportional to the gate area of a transistor so likewise a large geometry device is desirable. A small diffusion capacitance is desired because this increases the node sensitivity and reduces the input referred noise measured in electrons. No consideration is required of the floating diffusion reset noise since it has been eliminated by using CDS. The considerations to produce minimal noise come into conflict because a larger transistor (lower 1/f and thermal noise) introduces a larger parasitic gate to drain capacitance on the floating diffusion (smaller node sensitivity). Studies have been done to optimize these geometries for minimum input referred noise electrons.[7]

Wen introduced the floating gate amplifier in 1974.[8] Figure 1.7 is a diagram of a floating gate amplifier. In the floating gate amplifier the charge packet is shifted directly under the transistor gate electrode. The signal is capacitively coupled to the transistor gate which can be operated as a follower. The charge is then shifted out from under the gate and the amplifier is ready for another input.

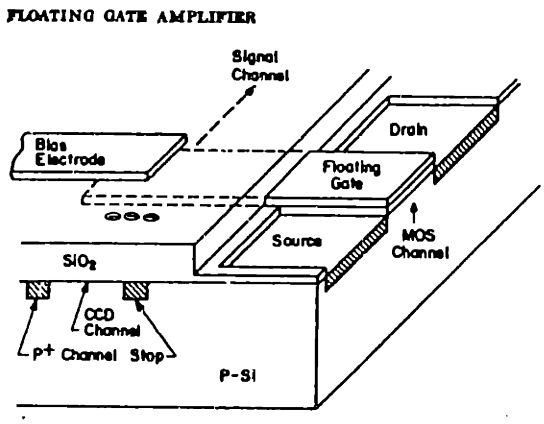


Fig. 1.7

The important thing to note about this technique is that it is nondestructive. The same charge packet can be read out more than once. Reset noise is also eliminated since there is nothing to reset.

More recent amplifier designs[9-13]have utilized novel fabrication steps to achieve a higher node sensitivity and therefore smaller input referred noise electrons. Reducing parasitic capacitance at the transistor gate is one way to reduce the input capacitance and increase node sensitivity. Lightly Doped Drain (LDD) structures reduce the parasitic gate to drain capacitance by reducing the overlap of the gate over the diffusion. The LDD tip is driven to a shallower depth than a conventional source/drain diffusion and therefore diffuses laterally under the gate by a smaller amount. Dynamic range is also an issue with designs with high node sensitivity. Video sensors must operate with a wide bandwidth so achieving low noise is more challenging than with slower readout rate applications.

In astronomical applications readout rates can be much slower and thermal noise can be reduced by using a smaller bandwidth. New floating gate

designs[13-14] are designed to output designated pixels repeatedly and average the results off focal plane for a square root on N noise reduction.

1.5 Review of Class AB CMOS Amplifiers

Class AB CMOS amplifiers may provide potential solutions for the AIRS output buffer because the total current they can deliver to or pull from a load is larger than the dc quiescent current of the amplifier. This feature should allow large output voltage swings without slew rate limitations.

An amplifier used in a feedback application may provide better performance than a source follower output alone because the amplifier may have more than one stage of amplification. The dc gain of the amplifier is an important factor in determining the linearity of the buffer and with two stages of amplification a dc gain of the order of $(g_m r_o)2$ can be obtained. With unity feedback the closed loop gain will be of the form

$$a_v = (g_m r_o)^2 / (1 + (g_m r_o)^2) \qquad (1.8)$$

When compared with the small signal gain of the follower it is apparent that better linearity can be obtained. The linearity of a follower will probably be determined by the amount of impact ionization. This is hard to estimate as it is very process dependant so that measured results should quantify this distinction.

The architecture of figure 1.2 is easily achieved by configuring an amplifier as a charge integrator. With a differential amplifier the quiescent voltage on the bus is easily controlled with a bias voltage on the noninverting input. This concept is shown in figure 1.8

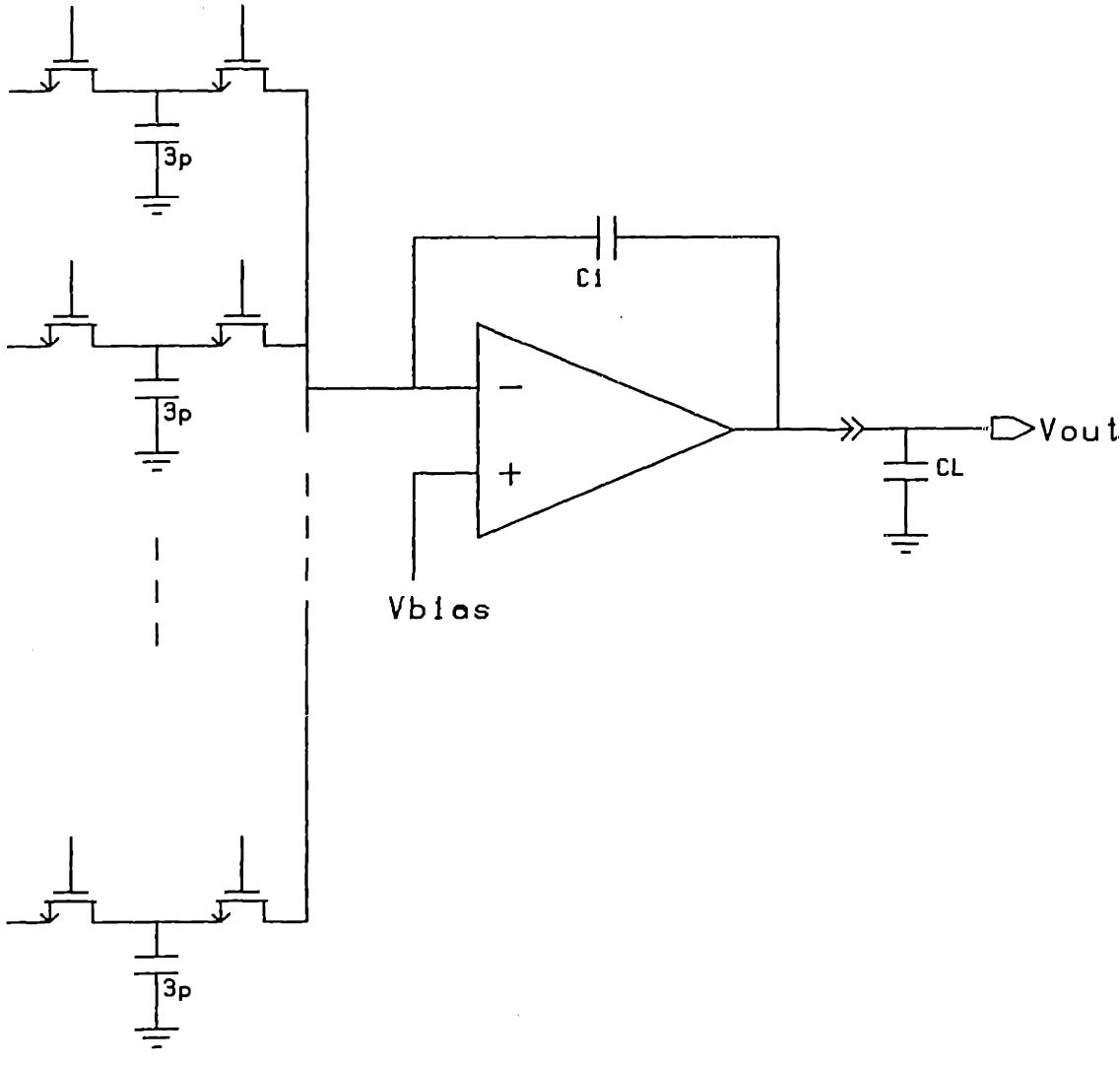


Fig. 1.8

Unfortunately a Class AB CMOS output stage with wide applicability does not exist. In bipolar design the complementary emitter follower stage has received widespread use. The CMOS analog to this circuit is the complementary source follower stage[15]. The drawback which has prevented this from becoming a standard is the reduction in output voltage range which is caused by the gate to source voltage drops of the n and p channel MOS transistors. Given the low supply voltages found in small geometry CMOS processes, threshold voltages in the .5v to 1.0v range, and the fact that one of the thresholds will be increased due to the back gate effect, the resulting output range is often inadequate. Operation of this circuit is even more problematic at cryogenic temperatures due to the fact that threshold voltages increase by .4 volts as the temperature is reduced from 300 ^{o}K to 77 ^{o}K. Special process modifications can be employed to reduce thresholds, but with standard processing this circuit has limited applicability. In this case 3 volts of output range are needed with a 6 volt supply. If each threshold is assigned a magnitude of 1.15 V then the increase of threshold due to back gate affect must be less than .7 volts. This is unlikely because in the complementary source follower connection the back gate voltage is maximum when the limits of output range are encountered. Orbit semiconductor's 1.5 um twin well process has been selected for fabrication. Process options such as low threshold devices or bipolar transistors are not available. Only designs with standard process options are considered here.

1.5.1 Dynamic Amplifiers

Dynamic MOS amplifiers were introduced by Copeland and Rabaey.[16] The idea is to allow the amplifier bias current to change as a function of time in switched capacitor circuits. The clocks used in switching charge from capacitor to capacitor are used within the amplifier itself to control its bias current.

A charge integrator proposed in [16] is shown if figure 1.9 The amplifier is within the dotted line.

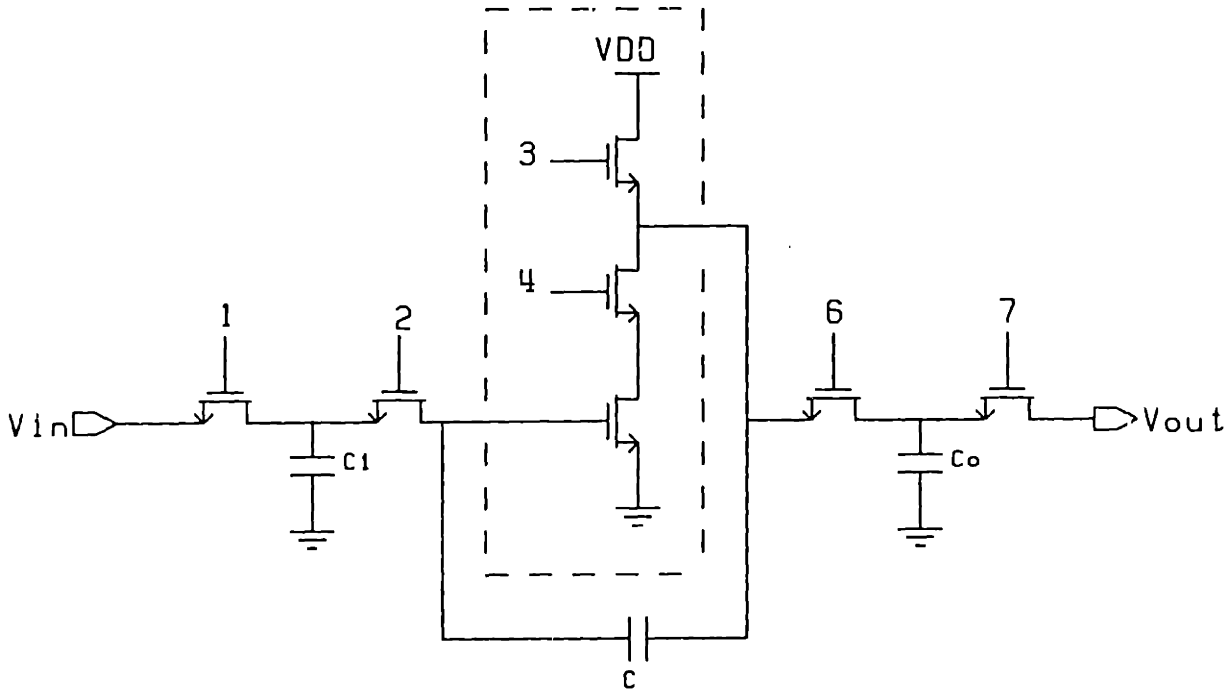


Fig. 1.9

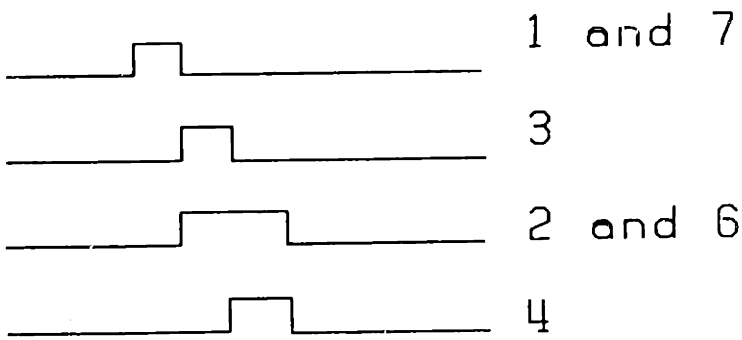


Fig. 1.10

The circuit works as follows: The input voltage is sampled on C_i. Next the output is pulled high and the sampled charge is dumped on the inverting node of the amplifier. Finally the transistor is connected with the output and it draws current from the output capacitor C_o and the feedback capacitor C until the voltage at the inverting node returns to the threshold of the device. The bias current of the device can be quite large initially but will always approach subthreshold values at the end of the clock cycle. The potential for power savings from this circuit is clear since there are no dc bias currents. The dc accuracy is good because the current is at its lowest at the end of the cycle. This particular switching scheme resets the output at each cycle which is not a good feature with a large capacitive load since this can represent a significant power loss. Another modification which is necessary for use as a charge integrating output amplifier is a means of resetting the integration capacitor.

Hosticka[17] showed how classical amplifiers could be converted to dynamically biased amplifiers by replacing current sources by pulsed current sources. He uses pulsed current sources which are nothing more than capacitors that are periodically charged and discharged. Figure 1.11 shows a CMOS differential stage with such a current source.

During the clock period phi the tail capacitor is discharged and during the clock period phi-bar it becomes charged as the amplifier is operating. The size of

Co must be chosen properly since for classical topologies the total charge that the amplifier can put out in one cycle is directly proportional to the charge stored on a current source capacitor in the circuit. Also the size of the capacitor places a limit on how large a frequency the circuit can be run at. It is desired that the tail current drop to a subthreshold value by the end of the cycle so that high gain is obtained. If the tail capacitor is too large, the bias current may still be too large at the end of the cycle.

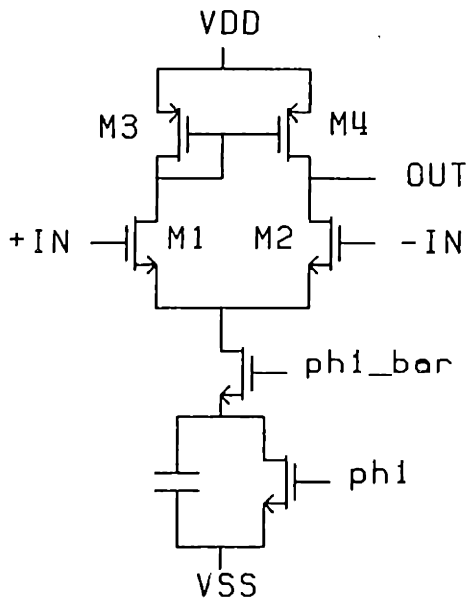


Fig. 1.11

Hosticka also presented a switched capacitor integrator similar to that of Copeland and Rabaey. This is shown in figure 1.12 . Phi and phi-bar are two phase nonoverlapping clocks. This is essentially a complementary version of the Copeland and Rabaey circuit where the M1 and M2 may pull the output up or down depending on the polarity of the input signal. Once again as the input mode is

returning to ground the bias current is dropping into the subthreshold region and high gain is obtained.

If this circuit is considered for use as an output buffer one must consider that while high gain is obtained at the end of the cycle, the subthreshold currents flowing will yield such a low g_m that a problem develops with high output impedance. If the output is capacitively coupled to any unwanted signals,whatever there origin, the output impedance of the output buffer will determine how effectively these signals will be rejected. No requirement has been given for the output impedance of the AIRS buffer. Balanced differential outputs will help reject common mode pickup, but it is hard to quantify the amount of rejection obtainable or the amplitude of the spurious signals present.

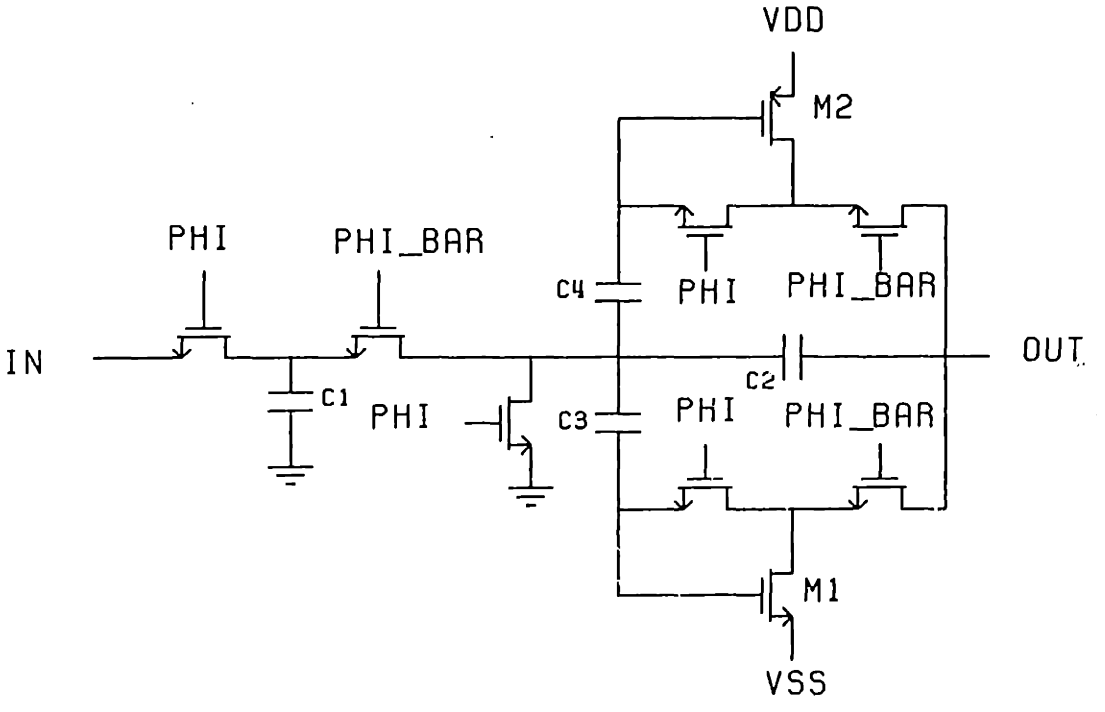


Fig.1.12

In consideration of the output impedance issue it seems that a version of figure 1.12 in which the currents settle to a controlled finite value rather than to zero would be desirable.

1.5.2 Adaptive Biased Amplifiers

Degrauwe et al.[18] introduced a class of amplifiers in which the bias current is dependent on the input signal without requiring internal clocks in the amplifier. These amplifiers were not limited to sampled data analog circuits but are generally applicable.

For an input differential stage the idea is to increase the tail current based on the amount of imbalance in the input stage currents. The imbalance is measured with a current subtractor as shown in figure 1.13

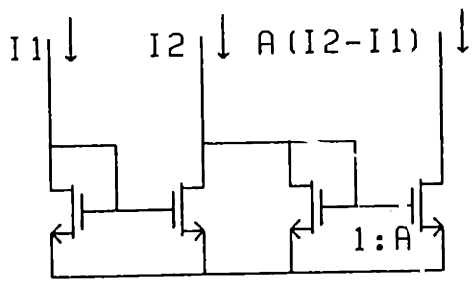


Fig. 1.13

The currents in each leg of the input stage are mirrored into the current subtractor. The output of the current subtractor augments the quiescent tail current bias. Subtractors must be added for both possible polarities of input stage imbalance since the bias current must be increased in either case. An operational transconductance amplifier (OTA) based on this principle is shown in figure 1.14. For simplicity, the current subtractor is only shown for one polarity

of imbalance. The effects of mismatches in the current mirrors as well as the stability of the local feedback are issues that must be considered.

An attractive feature of the adaptive bias concept is that if input signals are not present power dissipation remains at its quiescent value. Pulsed current sources in the dynamic amplifiers introduce large bias currents independent of the input signal. Thus an adaptively biased amplifier may show lower power dissipation than a dynamic amplifier.

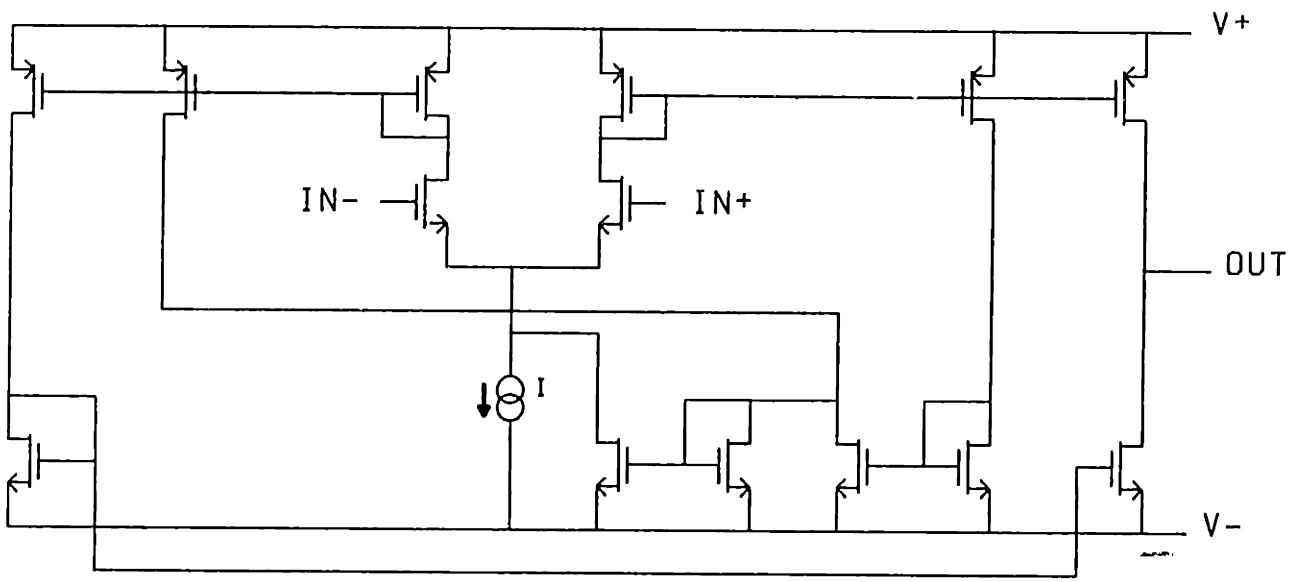


Fig.1.14

Another attractive feature of the adaptive biased amplifier is that as the amplifier settles the bias current approaches a well controlled value. The bias current of the dynamic amplifiers approaches zero as the amplifier settles. Thus an adaptively biased amplifier will show lower output impedance than a dynamic amplifier.

A class AB OTA which utilized n channel - p channel source coupled inputs was used by Costello and Gray.[19]. A simplified version of this is shown in figure. 1.15. Biasing for proper quiescent current and input range are issues with this design. A low voltage process coupled with high transistor thresholds can make this difficult to use for cyrogenic applications. The most severe problem however is the reduction of gm in the input due to the source coupled n/p transistors. The effective transconductance of a source coupled n/p pair is the parallel combination of the individual transconductances.

$$1/g_m = 1/g_{mn} + 1/g_{mp} \qquad (5.1)$$

Bandwidth for a given power is compromised due to this relationship.

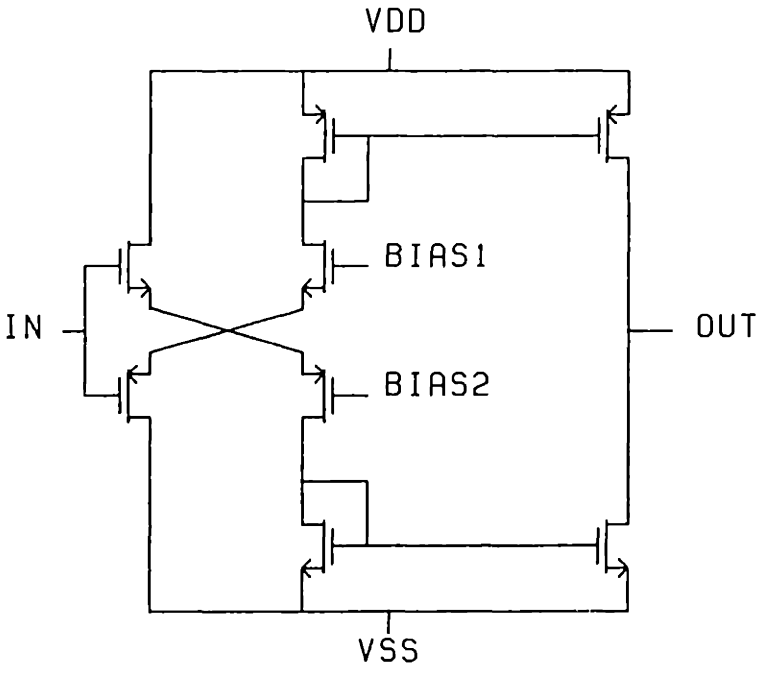


Fig. 1.15

Wong and Salama[20] have modified a conventional OTA so that large gate to source voltages become available to the output transistors when the input stage is saturated. The conventional OTA is shown in figure 1.16. It is possible to look at this stage as a single transconductance stage or as a low gain wide bandwidth differential input stage followed by a high gain low bandwidth output stage. CL creates the dominant pole and if large enough stabilizes the amplifier. Wong and Solama have inserted a low gain wide bandwidth intermediate stage as shown in figure 1.17.

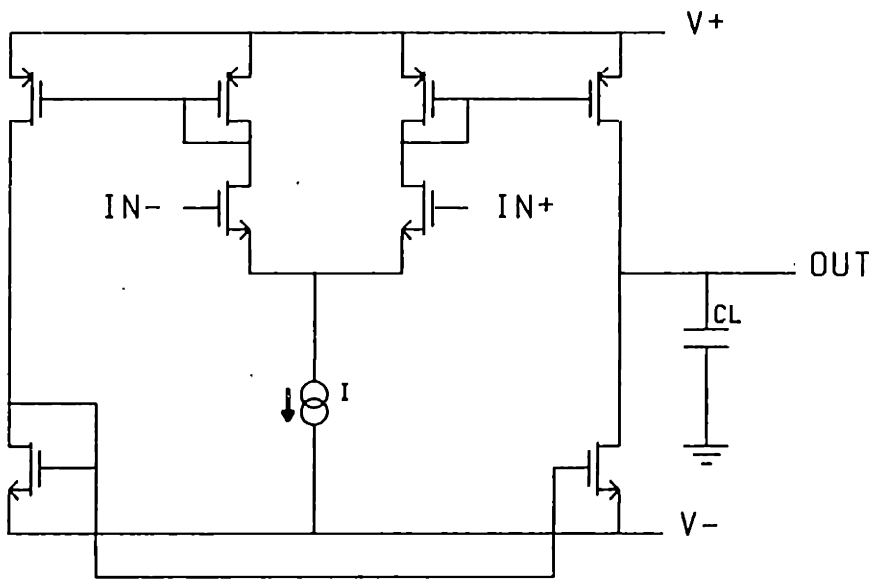


Fig.1.16

A gain of 3 is taken at the intermediate stage by the relative sizes of MX1 and MX2. Note that one extra inversion has taken place and the input is relabeled accordingly. Under slewing conditions with negative input MX1 is on with all available bias current while MX4 is off. The gate of MX8 is pulled up through R and large output currents flow. Under linear conditions little current flows through R and it does not determine the output biases. The major concern here is the proper selection of bias voltages to control the output quiescent current. The bias circuit of Wong and Salama is supply and process dependent and could show quite a large variability for small supply voltages and large device thresholds.

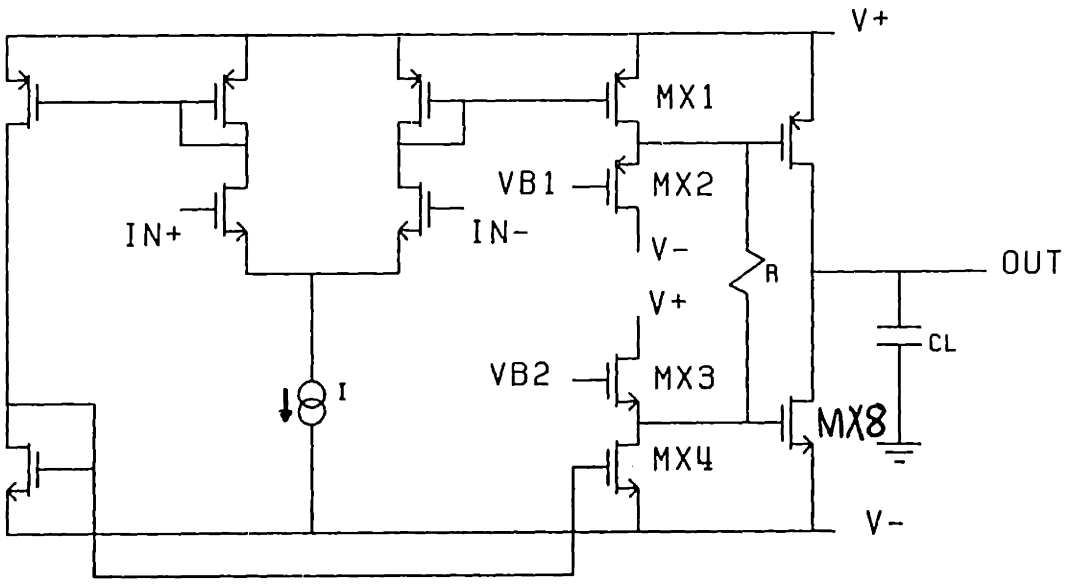


Fig. 1.17

Callewaert and Sansen[21] have developed another class AB circuit which increases input pair tail current as imbalance develops due to input signal. The circuit is shown in figure 1.18. The differential pair output current is routed through a cascode transistor to a current mirror which augments the tail current of the pair itself. In this case the input current becomes a quadratic function of input voltage. The input current is mirrored to the load. A complementary stage is required, but not shown, to drive a p channel output transistor and provide push pull action. As with the current subtractor approach details of current mirror matching is important for good performance. Bias current control is determined by the current sources and is relatively simple.

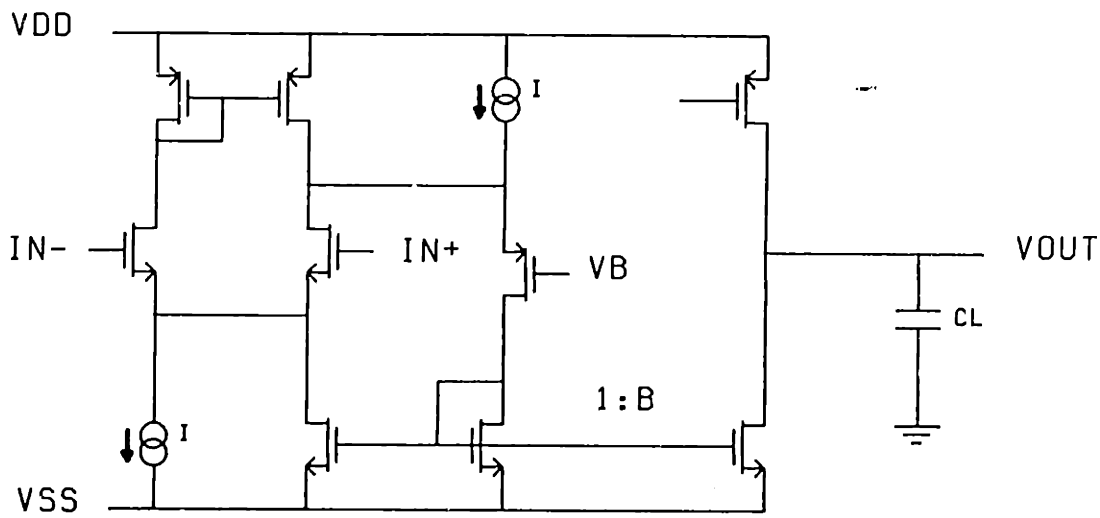


Fig. 1.18

1.5.3 Quasi-Complementary Approaches

Several designs have been of the quasi-complementary type as shown in figure1.19[22,23] These designs have tended to be quite elaborate due to problems associated with properly controlling the quiescent current and insuring stable operation with large capacitive loads. These designs are more suited to delivering a lot of power to a load than they are for low power operation,

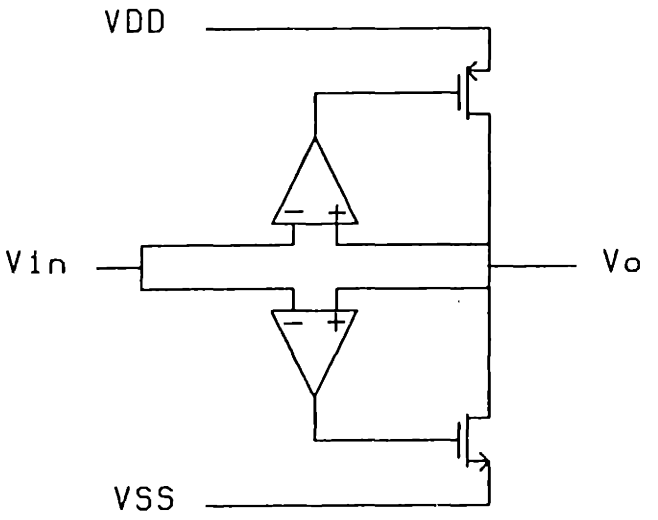


Fig. 1.19

1.6 Design Approach for AIRS buffer

A source follower is the most common type of output buffer for IR hybrid focal plane arrays. It provides design simplicity and performance characteristics which make it a suitable choice for a wide class of output requirements. In this case it is not appropriate. The requirements on power dissipation, capacitive load, and slew rate are incompatible with the characteristics of a source follower. In Class A circuits, such as a source follower, the output current is limited by the bias current, and the slew rate is given by the ratio of the bias current to the load capacitance. Power dissipation limits the total bias current to 40 μa. The requirement for a balenced differential output implies that at most 1/2 of the

available bias current will be available for each output. The maximum slew rate for a Class A output is therefore.

$$SR = I / CL = 20\ \mu a / 100\ pf = .2\ V/\mu s \qquad (1.10)$$

This slew rate is to low to meet the settling time requirement by about an order of magnitude. A Class AB output is therefore called for.

If one assumes that the slewing problem has been solved then the bandwidth problem must be addressed. In order to settle to a given degree of precision in a given time there is a minimum bandwidth which will be required. In this case settling to .03% will requires roughly 9 time constants. This time constant can be no larger than C_L/G_m, where G_m is the transconductance of the amplifier in the buffer or a stage of it. The maximum value of G_m which can be obtained in a single transistor is given as its weak inversion value.

$$Gm = I/ (n \cdot Vth) \qquad (1.11)$$

$$= 20\ \mu a / (2 \cdot .0067) = 1490\ \mu S$$

This gives a time constant of 67 ns if this transconductance is driving 100 pf. Nine such time constants is .6 μs. This implies that it is possible to achieve the degree of settling with the present power budget.

Because of the ability to achieve a well defined output impedance an adaptively biased design was chosen. The design of Costello was attempted first but the obtainable transconductance was too low, primarily due to the source coupled n and p input transistors. The designs of Degrauwe and Callewaert promise higher obtainable transconductances with well controlled bias currents in the output

stage. The topology of Callewaert was selected, although similar results could probably be obtained with that of Degrauwe.

The AIRS output buffer calls for a differential output. In this instance it is proposed to use two output circuits. The first circuit provides a signal output as depicted in figure 1.8. The second circuit contains a duplicate buffer but the input voltage is constant so the output voltage is constant. Another approach would be to employ a fully differential OTA and provide a truly differential output. In this case differential inputs are not available to the output buffer, so a fully differential output would require changes to the focal plane architecture. It is proposed here to meet the requirements without requiring such changes.

2.1 General Principle of Operation

The Class AB Operational Transconductance Amplifier (OTA) used here is that proposed by Callewaert and Sansen[21]. A review of it operation is given here, including some aspects not elaborated in [21]. The basic circuit of the amplifier is shown in figure 2.1.

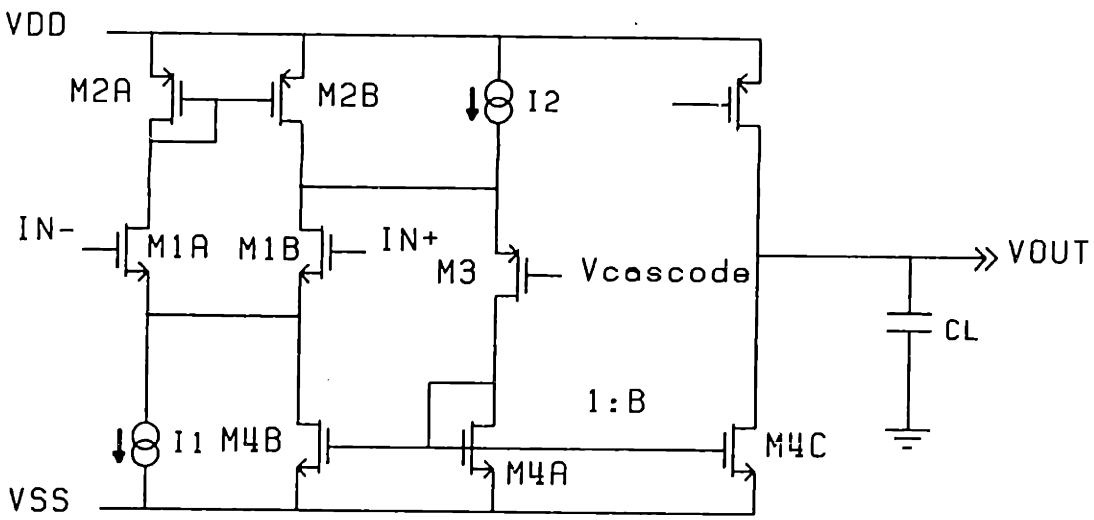


Fig. 2.1 OTA basic circuit

The input is differential and is applied to the gates of the source coupled pair M1A and M1B in the usual manner. A current source on the source node provides the quiescent bias current for the input stage. A current mirror load (M2A,M2B) directs the difference current from the input pair through a cascode transistor (M3). This current is then mirrored to two places. The first is via M4A

and M4B, and creates additional tail current for the input pair. The transistor M4B alters the bias current in the input stage and gives the amplifier its class AB operation. The second is via M4A and M4C, and provides current for the output load. A step up ratio of 1:B is conveniently taken at the M4A,M4C mirror to provide adequate drive for the load.

The input to the p channel drive transistor in the output stage is not shown in figure 2.1. It is driven by a parallel complementary input stage to provide symmetrical push pull drive capability. The differential input signal goes to two input stages. One input stage drives the n channel driver of the output stage, while the other input stage drives the p channel driver of the output stage.

A simple calculation, presented in [21], helps to explain the operation of the circuit. Suppose that the transistor M1A is conducting a dc current I1A. This current is mirrored with a ratio R2 to the source node of the cascode transistor M3. Applying Kirchoff's current law (KCL) to this node yields the equation

$$I3 = Ibias \cdot Rbias + I1A \cdot R2 - I1B . \qquad (2.1)$$

Rbias and R2 are nominally equal to one but are included here for generality. The current I3 is mirrored back to the common source point of M1A and M1B with a ratio of R4. Using KCL at this node yields

$$I1A + I1B = Ibias + I3 \cdot R4 . \qquad (2.2)$$

Equations (2.1) and (2.2) can be used to solve for I1B and I3 in terms of Ibias and I1A.

$$I1B = Ibias \cdot (R4 \cdot Rbias + 1) / (R4 + 1) + I1A \cdot (R2 \cdot R4 - 1) / (R4 + 1) \qquad (2.3)$$

$$I3 = Ibias \cdot (Rbias - 1) / (R4 + 1) + I1A \cdot (R2 + 1) / (R4 + 1) \qquad (2.4)$$

The equation in [1] which corresponds to (2.3) disagrees due to a printing error. If all of the current ratios are replaced by their nominal values of one, these equations are drastically simplified.

$$I1B = Ibias \qquad (2.5)$$

$$I3 = I1A \qquad (2.6)$$

These equations show that the local feedback loop acts in a way to maintain the current in M1B constant and equal to Ibias. The current through M4B increases or decreases as needed to maintain the current through M1B. If it is assumed that V+ is fixed, then it also follows that the source node of M1 is fixed. This follows because fixed current through M1B, and a fixed gate voltage on M1B, imply a fixed V_{gs} for M1B. V+ will be a fixed bias voltage in the output buffer application considered here. Therefore the action of the local feedback loop can also be described as maintaining the source voltage of M1 equal to its quiescent bias value despite variations in the input signal V-. Assuming that M1A is in saturation and strong inversion, its drain current is given by the equation

$$I1A = 1/2 \cdot K \cdot W/L \cdot (V_g - V_s - Vtn)2$$

$$= 1/2 \cdot K \cdot W/L \cdot (\Delta V - V_d)2 . \qquad (2.7)$$

V_d is the differential input voltage (V+ - V-) and ΔV is the quiescent value of the voltage above threshold (V_{gs} - V_{tn}) for transistor M1.

Equation (2.7) shows that for large negative differential input voltages I1A increases quadratically. According to (2.6) I3 is equal to I1A. The current I3 is mirrored to the output via M4A and M4C so that for large negative differential input voltages large currents can be sunk from the load. For large positive differential input voltages the transistor M1A turns off and equation (2.7) is not valid. For large positive differential input voltages the operating current in the input stage which drives the P channel driver increases quadratically. This conditions sources large current to the load.

This amplifier has a single high impedance node at the output. The dominant pole in the frequency response is determined by the load capacitance. If large enough, the load capacitance will stabilize the amplifier, and produce a single pole response. In this case, the amplifier is well modeled as a single transconductance stage. It is also possible to look at the amplifier as having two stages. The first is a low gain wide bandwidth differential input stage. The second is a high gain low bandwidth output stage. The two stage description is more accurate because it models the internal poles of the amplifier. If the design is pushed to produce maximum bandwidth, these internal poles must necessarily come into play as limiting factors.

The inverting output stage has extremely good output range. Let ΔV be the voltage above threshold (V_{gs}-V_{tn}) of the n-channel drive transistors. The output voltage can approach within ΔV of the lower rail before pushing the drive transistor into the triode region of operation. Similar considerations apply to the p-channel drive transistor. Without cascode transistors the output range goes virtually rail to rail.

To achieve highly linear operation, high dc gain is desirable. It is relatively straight forward to put cascodes on the output stage to increase the dc gain. Cascode transistors will decrease the output range of the amplifier, but if biased properly, this reduction can kept to a few tenths of volts off each end of the output range. The characteristics of this amplifier make it well suited for low power operation in applications where high linearity and large output swing are desired, and large load capacitance is encountered.

2.2 Minor Loop Stability

The stability of the local feedback loop is treated in Appendix 1. A new technique is used which does not require opening the loop. The new technique is suited for machine computation but is included in Appendix 1 for illustrative purposes.

2.3 Small signal frequency response

A schematic of the amplifier with a cascoded output stage is shown in figure 2.2. There are two parallel signal paths in the amplifier. One path goes through the n channel driver of the output stage, while the other goes through the p channel driver of the output stage. The step up ratio at the current mirror M4A, M4C is denoted as B. It is assumed that the step up ratios for the n and p channel circuits are the same. All the results presented here could easily be modified to take into account different step up ratios if necessary.

The transconductance of the amplifier has contributions from each input stage stepped up by a factor of B. If only the n channel input stage is considered, and the output stage p channel driver were considered to have a constant bias, the amplifier transconductance would be

$$G_m(n) = B \cdot g_m(1An) \ . \qquad (2.8)$$

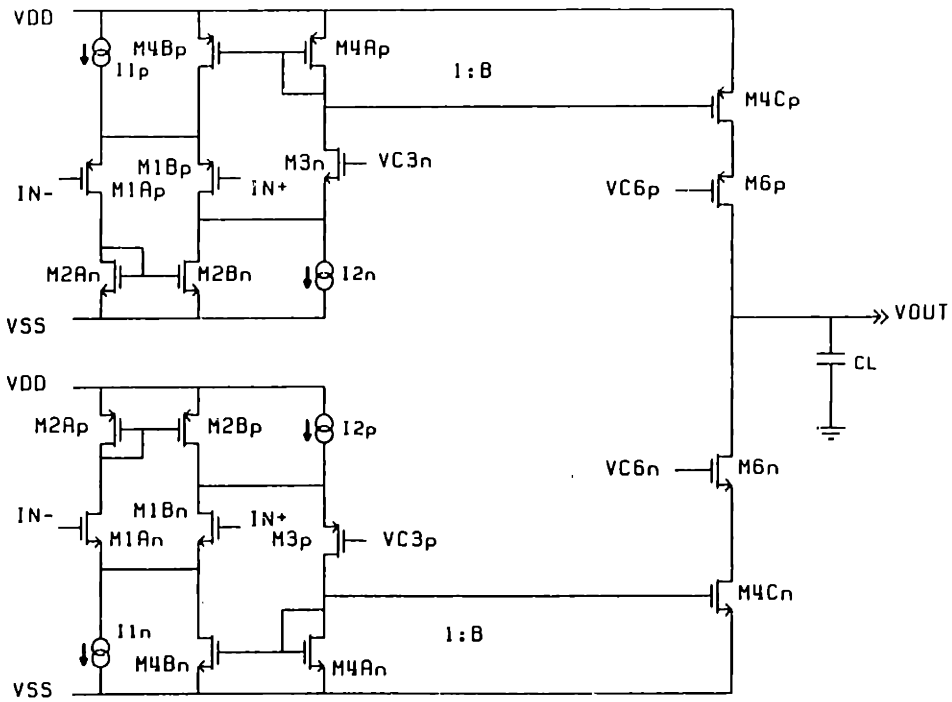


Fig. 2.2 OTA with cascoded output stage

Because the local feedback keeps the source at a fixed voltage there is no correction to account for the bulk to source transconductance. In a conventional source coupled connection, the transconductance is higher if the transistors are in there own well because of this effect. For this circuit there is no such distinction. If only the p channel input stage is considered, and the output stage n channel were considered to have a constant bias, the amplifier transconductance would be

$$G_m(p) = B \cdot g_m(1Ap) \ . \qquad (2.9)$$

The amplifier transconductance is a small signal linear property so that these two contributions may be added together to obtain the total amplifier transconductance.

$$G_m = B \cdot (g_m(1An) + g_m(1Ap)) \qquad (2.10)$$

The amplifier output conductance is the sum of output conductances from the pull up and pull down portions of the output stage. For the case without a cascoded output stage the total output conductance is given by the sum of the device output conductance of the driver transistors.

$$G_o \text{ (no cascodes)} = g_o(4Cn) + g_o(4Cp) \qquad (2.11)$$

The configuration with the cascoded output stage is of primary interest here because it is adopted in this design to yield high dc gain. The cascode configuration yields an output conductance with is smaller than the device output conductance by a factor equal to the intrinsic gain g_m/g_o of the cascode transistor[24]. This is valid as long as the intrinsic gain of the cascode transistor is much greater than one, as it is in this case.

$$G_o = g_o(4Cn) \cdot \frac{g_o(6n)}{g_m(6n)} + g_o(4Cp) \cdot \frac{g_o(6p)}{g_m(6p)} \qquad (2.12)$$

The dc gain of the amplifier is given by the ratio of the amplifier transconductance to the amplifier output conductance.

$$A = G_m/G_o$$

$$= B \cdot \frac{g_m(1An) + g_m(1Ap)}{g_o(4Cn)\cdot g_o(6n)/g_m(6n) + g_o(4Cp)\cdot g_o(6p)/g_m(6p)} \qquad (2.13)$$

This expression could also be expressed as the sum of the gain from each of the two parallel gain paths.

This OTA has a single high impedance node at the output. This is true for both the cascoded and noncascoded configuration. The load capacitance, which is estimated at 100 pf, and is much greater than any internal capacitance, also appears on the output node. The amplifier frequency response has a single dominant pole determined by the ratio of the output conductance to the load capacitance.

$$p1 = \frac{g_o(4Cn)\cdot g_o(6n)/g_m(6n) + g_o(4Cp)\cdot g_o(6p)/g_m(6p)}{C_L} \qquad (2.14)$$

This result can be arrived at from a zero-value time constant analysis, where the time constant from the output node dominates the sum. It can also be arrived at in another way which is discussed now.

When two portions of a circuit have a high degree of isolation from each other, the poles from each portion of the circuit can be calculated independently. This can be argued physically by considering an experiment where one portion of the circuit is placed in an initial state. Its state variables then decay according to the natural frequencies of that portion of the circuit. The isolated portion of the circuit has no influence on these decay frequencies. For two circuits which are completely isolated this argument is rather obvious. The argument yields useful results in a case where unilateral information flow exists between the two

subcircuits. This is to say that, dependant sources can exist in one subcircuit which are driven by variables from the other, but not vice versa, In this case the above argument concerning decay frequencies is still valid. If either of the two subcircuits is placed in an initially excited state, while the other is unexcited, it will decay in a manner independent of the other subcircuit. The same result can be argued mathematically, although somewhat more abstractly, by showing that the determinant of the nodal matrix can be factored into two separate expressions.

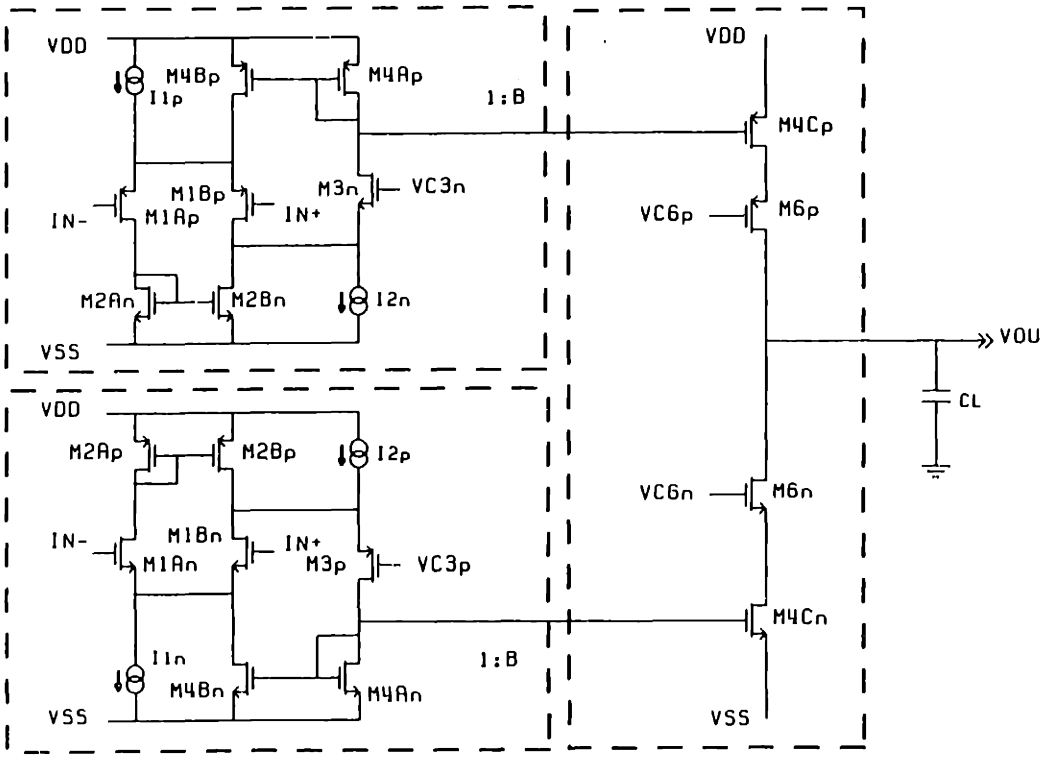


Fig. 2.3 Three isolated subcircuits

Figure 2.3 shows how the amplifier under consideration can be considered to be three isolated subcircuits. This isolation was implicitly assumed in figure A.1 where the output stage was considered immaterial to the local feedback loop in the

input stage. The poles of each section may be considered independently. The pole from the output stage can calculated directly without further approximation. Equation 2.14 results where once again the intrinsic gain of the cascode transistors are assumed greater than one.

An advantage to looking at the circuit as three separate subcircuits is that zero-value time constant analysis may now be performed on the other two subcircuits to determine the position of the nondominant poles [25]. The impedance looking into each of the four nodes in the input stage is straight forward to calculate. The impedance of node 2 is affected by the local feedback loop, while the other three are not.

$$R_1 = 1 / (2 \cdot g_{m1}) \tag{2.15}$$

$$R_2 = 3 / (2 \cdot g_{m2}) \tag{2.16}$$

$$R_3 = 1 / g_{m3} \tag{2.17}$$

$$R_4 = 1 / g_{m4} \tag{2.18}$$

The zero-value time constant analysis predicts a dominant pole location for the input stage subcircuit as

$$p_2 = 1 / \tau_2$$

$$\tau_2 = \frac{C_1}{2 \cdot g_{m1}} + \frac{C_2}{(2/3) \cdot g_{m2}} + \frac{C_3}{g_{m3}} + \frac{C_4}{g_{m4}} . \tag{2.19}$$

Each of the input stage subcircuits will contribute its own set of poles and zeros. The frequency response of the entire amplifier will be the sum of the individual frequency responses of the two parallel signal paths. Each signal path shares the output stage so the dynamics of the output stage is common to both.

If each input stage subcircuit were modeled as having a single pole the open loop frequency response of the amplifier would be given as

$$a(S) = \frac{1}{S - p_1} \left[\frac{A_n}{S - p_n} + \frac{A_p}{S - p_p} \right] . \quad (2.20)$$

A_n and A_p are the dc gains for each of the two parallel signal paths. The pole locations of the n channel and p channel input stages are denoted by p_n and p_p. The dominant pole from the output stage is given as p_1. The dc gain of the amplifier is the sum of A_n and A_p, and has been given in equation (2.13). Equation (2.20) can be rearranged into a more useful form.

$$a(S) = \frac{A_n + A_p}{S - p_1} \cdot \frac{S - z}{(S - p_n)(S - p_p)} \quad (2.21)$$

$$z = \frac{A_n \cdot p_p + A_p \cdot p_n}{A_n + A_p} \quad (2.22)$$

A pole occurs at each of the subcircuit pole locations and a zero appears at a location between the two poles. If the two poles were at approximately the same location, then combination of two poles and one zero would behave approximately as a single pole at that location.

The gain bandwidth product of the amplifier is given by the product of the dc gain and the dominant pole location.

$$GBW = A \cdot p_1 = G_m / C_L$$

$$= B \cdot \frac{g_m(1An) + g_m(1Ap)}{C_L} \qquad (2.23)$$

The input impedance of this amplifier differs somewhat from that of a conventional source coupled pair with current source bias. In a conventional source coupled pair, in the linear region of operation, the source node does not stay at a fixed voltage. As one input goes up in voltage, so does the source. The increase in current through one transistor is offset by a corresponding decrease in the current through the other, so that the sum is constant. To achieve this, the source node moves by 1/2 of the applied differential input voltage. As a result, the differential input capacitance is equal to 1/2 of the C_{gs} of an input transistor. In this amplifier the local feedback keeps the source voltage constant. The entire differential input voltage falls across one transistors gate and source nodes. The resulting differential input capacitance is given by

$$C_{in} = C_{gs}(1n) + C_{gs}(1p) \qquad (2.24)$$

Higher input capacitance is one penalty that is incurred by the Class AB operation.

2.3 Dynamic cascode biasing

Cascode transistors provide a very useful technique in achieving high dc gains. One detail of proper cascode design is the generation of the bias voltage. Figure 2.2 did not show the circuitry needed to generate the cascode bias voltage for both the input stage cascode and the output stage cascode transistors. Figure 2.4 shows the output stage with cascode transistors included.

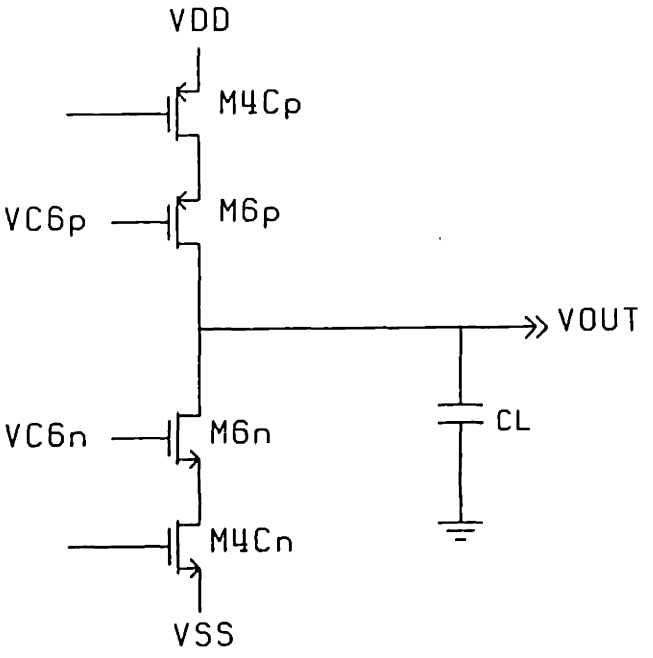


Fig. 2.4 Output stage with cascode transistors

The value of the bias voltage for the n channel cascode transistor, Vbiasn, is important. If it is too small, the drive transistor, M4cn, will be in the triode region and the voltage gain of the output stage will be too low. If it is too large, then it will limit the output range of the output stage. It is generally desirable to have the cascode bias only just large enough to guarantee that the drive transistor will be in the saturation region. Transient conditions may cause the operating current to increase over its quiescent value. The drain voltage on the drive transistor decreases as the output stage operating current increases. This is

because the Vgs voltage drop of the cascode transistor increases. The largest operating value of the output stage current must be used in determining the required bias voltage. Class A amplifier designs have generally used a constant voltage for the cascode bias with due regard to the above considerations.

Class AB operation causes a complication in the biasing of cascode transistors. It has been shown in section 2.1.1 that the amplifier operating current can become much greater than its quiescent value. The drive transistor M4C will fail to operate as a current mirror with M4A if it enters the triode region of operation. Output stage operating current will be limited to the value of current which causes this condition to take place. If the cascode biasing criteria discussed above is applied to the largest obtainable operating current, then the required cascode bias voltage may become so large that output range is severely reduced. To reduce the Vgs drop of the cascode transistors under maximum current conditions may require a prohibitively large cascode device. Dynamic cascode biasing was developed to address this problem [19].

In a dynamically biased cascode, the bias voltage is generated from the amplifier operating current. The bias voltage for the n channel cascode is generated so that it will increase as the amplifier operating current increases. Maximum output range is obtained under quiescent conditions, which is when it is needed. Maximum operating current is obtained under transient conditions, which is when it is needed. Figure 2.5 shows how easily dynamic biasing can be incorporated into this amplifier design. Transistor M5n is sized to produce maximum output range under quiescent conditions. As operating current increases in the n channel input stage, the increase in the bias voltage for M6n is more than adequate to compensate for increases in the M6n Vgs drop. A complementary circuit can be used to provide bias for the p channel output stage cascode transistor.

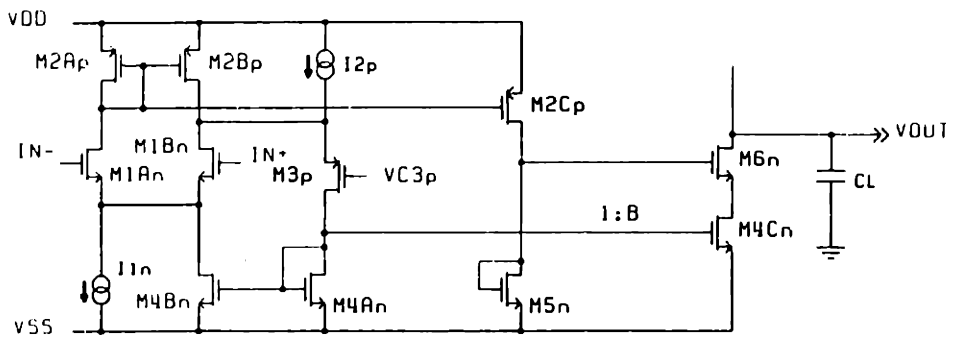


Fig. 2.5 Dynamic output stage cascode biasing

2.4 Large Signal Performance

Although the operating current increases quadratically for large differential input signals, there are at least two mechanisms which limit how large this current can become. The first is caused by limited supply voltage and has been discussed in reference 1. The second is caused by device mismatch and is introduced here for the first time.

As the differential input voltage increases, the operating current increases first linearly and then quadratically. As the operating current increases, the gate to source voltage of all the transistors in the signal path increases. Eventually one of the transistors will enter the triode region. The drain of M1A and M2A is a node where this condition can happen. The source of M3 and the drain of M2B is another. The onset of a transistor entering the linear region may place a limit on the operating current. The maximum obtainable operating current establishes the maximum output current. For a given load capacitance, this determines the slew rate. For acceptable settling time, the slew rate must be adequate. Transistors must be sized so that the maximum current increase is acceptable.

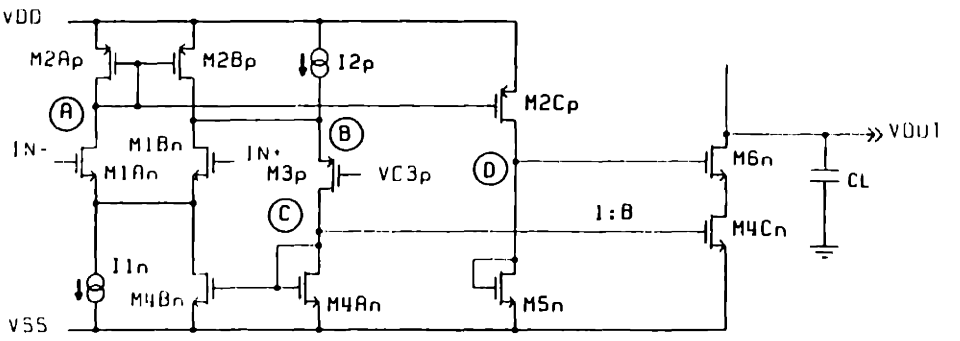


Fig. 2.6 Nodes sensitive to operating current increases

Figure 2.6 shows four nodes at which operating current increases could cause a transistor to enter the triode region of operation. The transistor which enters the triode region at the lowest current will limit the output current of the amplifier. At large operating currents the transistors in question are in strong inversion, so their currents should be accurately represented by the square law relation.

$$I_d = 1/2 \cdot K \cdot W/L \cdot (V_{gs} - V_t)2 \quad \text{for} \quad V_{ds} > V_{gs} - V_t \qquad (2.25)$$

It is assumed, in equation (2.25), that the currents are not so large that mobility degradation effects are significant.

At the node labeled A, operating current increases cause the drain voltage of M1A to decrease, and the gate voltage of M1A to increase. When the gate to drain voltage exceeds the threshold, M1A will enter the triode region of operation. A further increase in V- will only produce a small increase in the operating current. Using equation (2.25), the operating current at which this condition will occur is given as

$$I(An) = (1/2)\cdot K_n\cdot(W/L)_1 \cdot \left[\frac{\Delta VM1B + (V_{DD} - V+) + (V_{tn} - |V_{tp}|)}{(1 + g_{m1A}/g_{m2})} \right]^2 . \quad (2.26)$$

Vtn in (2.26) is the threshold of M1A, which may be increased by the back gate effect depending on its substrate connection. If the back gate effect is present the maximum current will be even larger than (2.26). The gm ratio in the denominator is the ratio when both M1A and M2 are in strong inversion. M1B need not be in strong inversion when this condition arises because it continues to conduct the bias current even while M1A may be conducting the maximum current. If it is assumed to be in strong inversion so that (2.25) is valid, then equation (2.26) may be expressed in a somewhat different form.

$$\frac{I(An)}{I_{BIAS}} = \frac{1}{(1 + g_{m1A}/g_{m2})2} \cdot \left[1 + \frac{(V_{DD} - V+) + (V_{tn} - |V_{tp}|)}{\Delta V_{M1B}} \right]^2 . \quad (2.27)$$

,Equation (2.26) should be more accurate, as it does not make this assumption. A simplified version of (2.27) is given in [21] where V_{tn} has been assumed equal to $|V_{tp}|$ and g_{m1A} has been assumed equal to g_{m2}. I(A) is a potential current limit of the amplifier.

At the node labeled B, operating current increases will increase the drain voltage of M2B, and lower the gate voltage of M2B. When the voltage gate to drain falls below the threshold, transistor M2B will enter the triode region of operation. The current source I2 could conceivably enter the triode region but this is

unlikely because it continues to conduct only the bias current. With the onset of this condition, the operating current will level off. The maximum current is given by

$$I(Bn) = (1/2)\cdot Kp\cdot(W/L)_{2B}\cdot \left[\frac{V_{DD} - V_{CASCODE} - |V_{tp}|}{1 + g_{m2B}/g_{m3}} \right]^2 \quad (2.28)$$

$$\frac{I(Bn)}{I_{BIAS}} = \left[\frac{V_{DD} - V_{CASCODE} - |V_{tp}|}{\Delta V_{M2}\cdot (1 + g_{m2B}/g_{m3})} \right]^2 . \quad (2.29)$$

As before, the ratio of transconductances in the denominator is the value that applies with both M2 and M3 in strong inversion. Vtp is the threshold voltage of M3, which may be increased due to the back gate effect depending on its substrate connection. The back gate effect, if present, at this node will decrease the limiting current. To be conservative, it should be accounted for even though the magnitude of V_{SB} will be small when this current limiting condition arises. Equation (2.29) is valid if M2 is in strong inversion under quiescent conditions. I(B) is a potential current limit for the amplifier.

At the node labeled C, operating current increases will increase the drain voltage of M3. When this voltage is large enough, M3 will enter the triode region of operation. When this happens, M3 will operate as a switch and further increases in current will begin to increase the source voltage of M3. Eventually M2B would enter the triode region of operation with the same effect as described with node B. For any reasonable values of cascode voltage, the current limiting condition at node B will occur first. The current at which M3 enters the triode region is given by

$$I(Cn) = 1/2 \cdot K_n \cdot (W/L)_{4A} \cdot (V_{CASCODE} - V_{SS} + |V_{tp}| - V_{tn})2 \qquad (2.30)$$

$$\frac{I(Cn)}{I_{BIAS}} = \left[\frac{V_{CASCODE} - V_{SS} + |V_{tp}| - V_{tn}}{\Delta V_{M4}} \right]^2 . \quad (2.31)$$

V_{tp} is the threshold of M3, which may be increased by the back gate effect, depending on its substrate connection. Equation (2.31) is valid only if M4A is in strong inversion under quiescent conditions. I(C) does not, however, represent a current limiting value.

At the node labeled D in figure 2.6, operating current increases will raise the drain voltage of M2C and lower the gate of M2C. When the gate to drain voltage falls below the threshold, M2C will enter the triode region of operation. M2C will fail to mirror the operating current and the bias voltage for M6 will stop increasing. This condition is unlikely to limit the output current of the amplifier. The gate voltage on M6 is already quite large when this condition happens. It may even be large enough so that M6 is operating as a switch rather than a cascode transistor. This has no drawback. M6 is needed to operate in the saturation region only under quiescent conditions. The current at which M2C enters the triode region is given by

$$I(Dn) = (1/2) \cdot K_p \cdot (W/L)_{2C} \cdot \left[\frac{V_{DD} - V_{SS} - V_{tn}}{1 + g_{m2C}/g_{m5}} \right]^2 \quad (2.32)$$

$$\frac{I(Dn)}{I_{BIAS}} = \left[\frac{V_{DD} - V_{SS} - V_{tn}}{\Delta V_{M2C} \cdot (1 + g_{m2C}/g_{m5})} \right]^2 . \quad (2.33)$$

As before , equation (2.33) is valid if transistor M2C is in strong inversion under quiescent conditions. As with I(C), I(D) does not represent a current limit for the amplifier.

A set of equations analogous to (2.26)-(2.33) which describe the currents at which transistors in the p channel input stage enter the triode region can easily be arrived at.

$$I(Ap) = (1/2)\cdot K_p\cdot(W/L)1 \cdot \left[\frac{\Delta V_{M1B} + (V+ - V_{SS}) + (|V_{tp}| - V_{tn})}{(1 + g_{m1A}/g_{m2})} \right]^2 . \quad (2.34)$$

$$I(Bp) = (1/2)\cdot K_n\cdot(W/L)_{2B} \cdot \left[\frac{V_{CASCODE} - V_{SS} - V_{tn}}{1 + g_{m2B}/g_{m3}} \right]^2 \quad (2.35)$$

The two most important equations concerning current limits in the p channel input stage are given as equations (2.34) and (2.35). Assuming that the current limiting mechanisms given above were responsible for limiting the output current of the amplifier, the maximum negative output current would be

$$I- = B \cdot MIN (I(An), I(Bn)) . \quad (2.36)$$

B is the step up ratio between the input stage operating current and the output stage operating current. The maximum positive output current would be

$$I+ = B \cdot MIN (I(Ap), I(Bp)) . \quad (2.37)$$

Other mechanisms, however, might limit the current to even lower values.

The input stage of this amplifier has three pairs of matched devices. Care must be taken to determine which properties are affected significantly by mismatch in these pairs. Device mismatch places limits on how large the operating current can be. Equations (2.2) and (2.3) were reduced to (2.4) and (2.5) by setting the current mirror and current source ratios equal to there nominal values of one. If each ratio is assumed to have a small deviation from its nominal value, the following equations result.

$$I1B = I_{BIAS} + I_{BIAS} \cdot \delta R_{BIAS} / 2 + I1A \cdot (\delta R_2/2 + \delta R_4/2) \qquad (2.38)$$

$$I3 = I1A + I1A \cdot (\delta R_2/2 - \delta R_4/2) + I_{BIAS} \cdot \delta R_{BIAS}/2 \qquad (2.39)$$

Higher order terms in the mismatch parameters have been neglected. Under transient conditions, I1A becomes much larger than I_{BIAS}. Inspection of (2.38) reveals that for I1A large enough, I1B may differ significantly from its nominal value of Ibias. In fact, I1B may go to zero if the mismatch terms have the correct polarity. For ($\delta R_2 + \delta R_4$) less than zero, the current through I1B will go to zero at an operating current of

$$I(\text{mismatch-n}) = - 2 \cdot I_{BIAS} / (\delta R_2 + \delta R_4) \qquad (2.40)$$

When the current through M1B goes to zero, the local feedback loop ceases to operate. The source voltage of M1B is no longer pinned and the current levels off. This effect may limit amplifier output current.

If mismatches have the correct polarity, equation (2.39) shows that I3 may be less than its nominal value of I1A. In this case the current limiting condition at the location labeled A will occur at lower than nominal output currents.

In general, current ratio errors may have systematic and random components. When a current mirror pair of transistors are operating at different drain voltages, their drain currents are mismatched by a term proportional to the device output conductance and the difference in drain voltages. This difference in drain currents represents a systematic mismatch component. Differences in device threshold voltage introduces a random mismatch component.

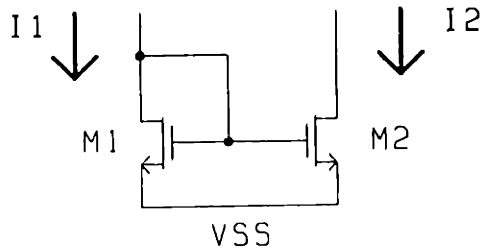


Fig. 2.7 Matched current mirror

The current mirror shown if figure (2.7) is assumed to be operating in strong inversion and with a nominal ratio of 1. The current in M2 can be written as

$$ID = (1/2) \cdot K \cdot (W/L) \cdot (V_{GS} - V_{tn})^2 \cdot (1 + \lambda \cdot V_{DS}) \qquad (2.41)$$

With equal drain voltages, the current ratio is nominally equal to one. The deviation in this ratio due to drain voltage mismatch is given by

$$\delta R(V_{DS}) = \lambda \cdot \delta V_{DS} \qquad (2.42)$$

and is a systematic mismatch component. The deviation in the ratio due to other parameter mismatches are

$$\delta R(V_{tn}) = -2 \cdot \delta V_{tn} / (V_{GS} - V_{tn}) \qquad (2.43)$$

$$\delta R(K) = \delta K/K \qquad (2.44)$$

$$\delta R(W/L) = \delta(W/L)/(W/L) \qquad (2.45)$$

These mismatchs are generally random and uncorrelated [25,26]. Mismatch in the threshold voltage can produce severe ratio mismatchs if ($V_{GS} - V_{tn}$) is small, as shown by equation (2.43). In cases where random mismatch effects dominate over systematic ones, and where mismatch effects dominate over other current limiting mechanisms, the maximum amplifier output current would be quite variable. Factors known to affect matching accuracy are transistor area and layout technique. With mismatch effects taken into account, the maximum output current of the amplifier can be expressed as

$$I- = B \cdot MIN (I(An), I(Bn), I(mismatch\text{-}n)) \qquad (2.46)$$

$$I+ = B \cdot MIN (I(Ap), I(Bp), I(mismatch\text{-}p)) \qquad (2.47)$$

where, once again, B is the step up ratio between input stage and output stage operating current.

2.5 Noise performance

A prescription for good noise performance is a simple input structure, large input devices, and negligible noise contributions from other devices in the

circuit. Unfortunately these consideration are often in conflict with other design objectives, and compromises must be made.

It is most useful to represent the effects of noise sources as being caused by equivalent input noise sources. When dealing with multiple incoherent noise sources, it is most useful to deal with their noise powers because these add linearly. In the treatment given here, noise powers from noise sources internal to the amplifier will be systematically referred to the input of the amplifier.

In order to calculate the input referred noise contributions of each subcircuit consider the MOS small signal noise model shown in figure 2.8 [27].

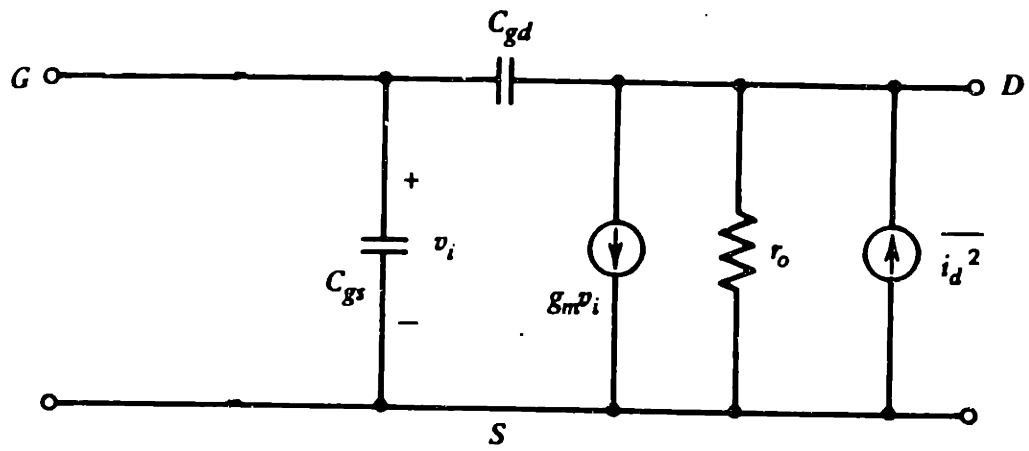


Fig 2.8 MOSFET noise model

The noise current generator id^2 is given as

$$i_d^2 = 4\,k\,T\,(2/3 \cdot g_m)\,\Delta f + K \cdot ID \cdot \Delta f / f \quad . \qquad (2.48)$$

The constant K is here different than previously used. Equivalent gate referred current and voltage noise generators can also be used in the place of the drain current noise generator. The equivalent input referred noise voltage and current source generators can be given as

$$v_i^2 = i_d^2 / g_m^2$$

$$= (8/3)\, k \cdot T \cdot \Delta f / g_m + K_f \cdot \Delta f / (W \cdot L \cdot C_{OX} \cdot f) \qquad (2.49)$$

$$i_i^2 = \omega^2 \cdot C_{gs}^2 \cdot v_i^2 \qquad (2.50)$$

In cases with low source resistance, the input referred noise voltage source contributes all of the output noise current. The current generator makes no contribution. In cases with high source resistance, the current noise generator is entirely responsible for the output noise current. At intermediate impedance values, both noise generators make a contribution. The intermediate case is further complicated, because it is clear that the two noise generators are not independent. They are both generated from the same noise source i_d^2. Because they are correlated, their noise powers do not add linearly. To avoid this complexity, it is useful to investigate the source impedance regimes in which one of the two generators suffices without the other.

The factor which determines whether the source impedance falls in the low regime or the high regime is how it compares with the input impedance of the transistor. If the source impedance is much lower than the input impedance of the transistor, the voltage noise source fluctuation will appear essentially full strength at the gate of the FET. The current noise source on the other hand will go essentially entirely into the source impedance and have no impact on transistor output current noise. If the source impedance is much higher than the input impedance of the FET, then this situation is reversed. The voltage fluctuation appears across the source impedance while the current fluctuation goes into the transistor gate.

The input impedance of the transistor is that of the gate capacitance C_{gs}, and is equal to $1/(S \cdot C_{gs})$. For a given source impedance, it will be smaller than the input impedance at low frequencies, but larger than the source impedance at high frequencies. If the frequency at which the source impedance equals the input impedance is larger than any frequencies of interest, the current generator may be neglected with negligible loss of accuracy. As an example, if frequencies over 1Mhz are not of interest due to bandwidth limitations of the circuit, and the input capacitance is as large a 1pf, then the current noise generator will be have negligible impact for source resistance lower than 150 kΩ.

Equation (2.49) shows that larger transistors have lower input referred 1/f noise. It also shows that input referred thermal noise power is inversely proportional to transconductance. Input transistors should therefore be as large as feasible. The down side of large transistors is real estate and input capacitance. Larger input capacitance will slow down the output buffer as in shown in section 3. A reasonable trade off between speed and noise is to bias input transistors in the transition region between strong inversion and weak inversion. In this region, they have essentially as much transconductance as can be obtained for a given bias current. This is good for speed and thermal noise. The area can be made as large as feasible without unduly impacting input capacitance.

Figure (2.9) shows one input stage subcircuit with incrementally grounded output stage. Inputs are incrementally grounded so that the equivalent input referred noise voltage can be calculated. Source impedances should be small enough so that input referred current sources are not needed. Although it is conventional to make this calculation with input referred noise sources for each of the individual transistors, drain current noise sources are used here.

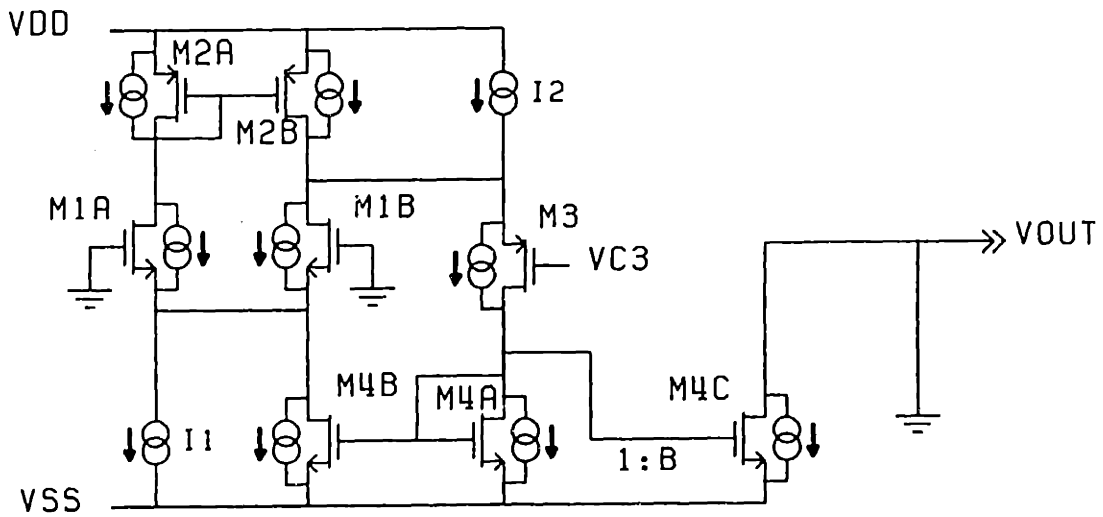


Fig. 2.9 Input stage subcircuit with noise sources

Figures (2.9) must be considered a small signal circuit model, although transistors are shown in their normal large signal representation for simplicity. Grounded inputs and outputs must be considered incrementally grounded. In figure (2.9) the contribution of each current noise source to the output current noise power can be calculated individually. The contribution of each current noise source to the output noise is calculated in order to calculate the total input referred voltage noise of this subcircuit.

The current noise source for the current source I1 has no contribution to the output current. It merely perturbs the tail current of the input pair. The input pair is balanced, however, and none of this current perturbation is routed through M3. The current noise source for M4B falls into the same category. It appears in parallel with the current noise source of I1 and so must behave in the same way.

The current noise source from the cascode transistor M3 also has no contribution. The input pair is balanced, so the current through M3 is set by the current source I2. The source node of M3 fluctuates in a way to keep the total current through it constant.

The noise current source from M2A is injected into the drain node between M2A and M1A. Looking toward the drain of M1A is high impedance, while looking toward M2A the current sees a diode connected transistor. The current heads to the low impedance direction and manifests itself as a fluctuating voltage on the gate of M2A. An equal current is generated by M2B and passes through the low impedance path provided by M3. It is then mirrored to the output through M4A with a step up factor of B.

The noise current source from M2B is injected into the drain node between M2B and M1B. It heads to the low impedance path provided by M3 and is mirrored to the output with a step up ratio of B. The noise current source from I2 appears in parallel with that of M2B and thus behaves in an identical way.

The noise current source from M4A is injected into the drain node between M4A and M3. The low impedance path is to M4A where it appears as a fluctuation on the gate voltage of M4A. This fluctuation is mirrored to the tail of the input pair where it has no affect. It is also mirrored to the output with a step up ratio of B. The noise current source from M4C simply adds directly to the short circuit output current.

The noise current source from M1A injects current into the drain node between M1A and M2A. This current is also injected into the source node of M1A. The injection of current into the source node of M1A is in phase or coherent with the injection of current into the drain node of M1A but has opposite polarity. The current fluctuation into the source node of M1A is a common mode disturbance, as discussed before, and does not make a contribution to output current. The

injection into the drain node behaves in the same way as the current noise source of M2A, which also is injected into the same node. Using the same arguments, the noise current source of M1B behaves in the same way as that of M2B.

Adding all the noise current powers at the output gives the total short circuit output current with grounded inputs.

$$i_o^2 = i_d^2(M4C) +$$
$$B^2 \cdot (i_d^2(M1A)+i_d^2(M1B)+i_d^2(M2A)+i_d^2(M2B)+i_d^2(I2)+i_d^2(M4A)) \quad (2.51)$$

Figure (2.10) shows the same circuit with an equivalent input referred noise source.

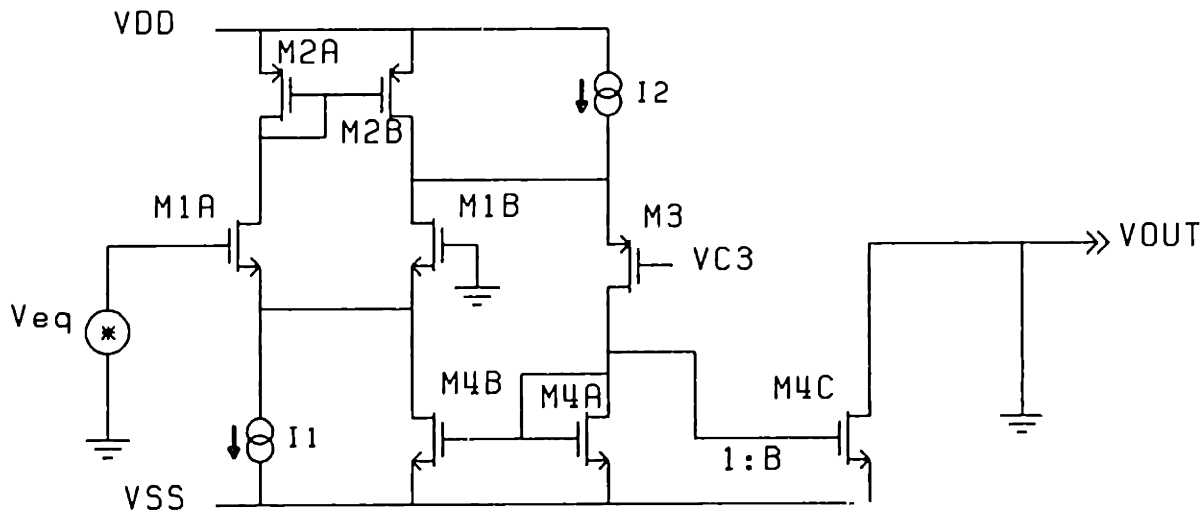


Fig. 2.10 Equivalent input referred noise source

The output short circuit current in Fig. 2.10 is

$$i_o^2 = B^2 \cdot g_{m1}^2 \cdot v_{eq}^2 \qquad (2.52)$$

ve_{q1}^2 is defined so that the output currents in (2.63) and (2.64) are the same.

$$v_{eq1}^2 = i_d^2(M4C) / (B^2 \cdot g_{m1}^2) +$$
$$(i_d^2(M1A)+i_d^2(M1B)+i_d^2(M2A)+i_d^2(M2B)+i_d^2(I2)+i_d^2(M4A)) / g_{m1}^2$$

$$= v_i^2(M1A) + v_i^2(M1B) + (g_{m2}/g_{m1})^2 \cdot (v_i^2(M2A) + v_i^2(M2B))$$
$$+(g_{m4A}/g_{m1})^2 \cdot (v_i^2(M4A) + v_i^2(M4C)) + (g_{mI2}/g_{m1})^2 \cdot v_i^2(I2) \qquad (2.53)$$

The relation that g_{m4C} is B times that of g_{m4A} has been used along with equation (2.51). Equation (2.65) is a familiar result showing that noise sources are referred back to the input according to the ratio of the transconductance of the transistor and the input transistor. It should be noted that the validity of (2.65) relies on the inequality $g_o \ll g_m$ at various locations in the circuit. This inequality determines which way the noise current will flow, and what the output short circuit noise current will be.

For a moderate step up ratio, the contribution from M4C can be negligible by comparison with M4A. For the other sources, equation (2.49) can be used to show that their contribution, relative to an input transistor, to the thermal noise power is in the ratio of their transconductance to that of an input transistor.

$$v_{eq}^2(th) = v_i^2(M1,th) \cdot \{ 2 + 2 \cdot g_{m2}/g_{m1} + (1+1/B) \cdot g_{m4A}/g_{m1} + g_{mI2}/g_{m1} \} \qquad (2.54)$$

If the devices are assumed to be operating in strong inversion, then their contribution, relative to an input transistor, to the 1/f noise power is in the ratio of $K_f \cdot \mu / L^2$ to the same expression for the input transistor. This is also a familiar

result which shows the importance of channel length in the non-input transistors[19]. In the general case

$$v_{eq}^2(1/f) = v_i^2(M1,1/f) \cdot \{ 2 + 2 \cdot (g_{m2}/g_{m1})^2 \cdot (K_{fp}/K_{fn}) \cdot (A_1/A_2) +$$

$$(g_{m4}/g_{m1})^2 \cdot (A_1/A_{4A} + A_1/A_{4C}) + (g_{mI2}/g_{m1})^2 \cdot (K_{fp}/K_{fn}) \cdot (A_1/A_{I2}) \} \quad (2.55)$$

A_1 is the gate area of M1 and the same notation is used for the other transistors.

The two parallel gain paths in this amplifier play an important role in it's noise performance. Noise sources in one of the two input stage subcircuits propagates in one of these gain paths while the signal propagates through both. Noise sources present in only one gain path must be reduced when referred to the input of the amplifier. Figure 2.11 shows the amplifier with noise sources at the inputs of the subcircuits, and the amplifier with an equivalent input referred noise source.

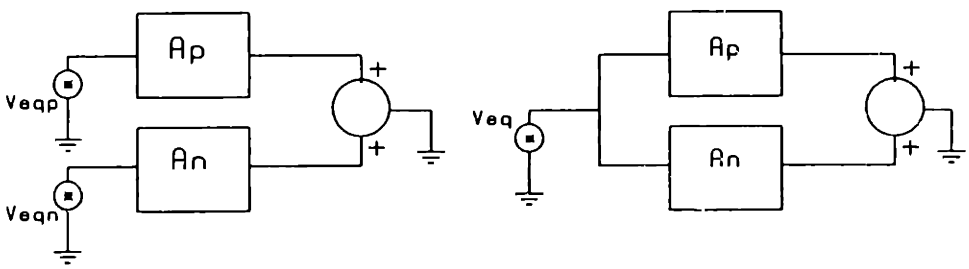


Fig. 2.11 Equivalent input referred noise source

The equivalent input noise source is defined so that the output noise powers are equal. This yields the prescription for referring input noise sources for one subcircuit to the input of the amplifier.

$$V_{eq}^2 = V_{eq1}^2 \cdot \left[\frac{A_n}{A_n + A_p} \right]^2 + V_{eq2}^2 \cdot \left[\frac{A_p}{A_n + A_p} \right]^2 \qquad (2.56)$$

If the gains of each path were equal, then noise powers would be divided by four when referred to the input of the amplifier. If the gain of one path were much larger than the other, then the equivalent input noise power would be equal to the input referred noise power of that subcircuit. In any case, the equivalent input referred noise power of the amplifier is less than or equal to that of either subcircuit. The added complexity of two input stages has not increased the input referred noise, because signal gain is obtained through each subcircuit. In fact, the input referred noise is reduced by providing two gain paths.

3. THEORY OF OPERATION : OUTPUT BUFFER

The concept for this output buffer is to operate as a charge integrating amplifier. Figure 1.8 shows the way it will be configured to multiplex many signal channels into one output. Overall buffer performance will depend on amplifier characteristics as well as those of the feedback elements.

3.1 Small signal frequency response

The feedback network in this case is entirely capacitive and is shown in figure (3.1) . Some series resistance elements due to switches will also be present, but for frequencies of interest here, these resistances play no significant effect. This will be shown later in section 3.2.1.

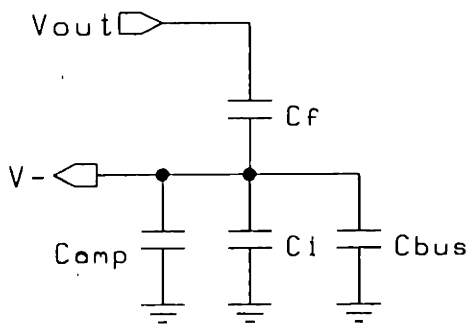


Fig. 3.1 Capacitive feedback network

It is convenient to consider the load capacitor as part of the amplifier for this analysis, so it is not shown in figure (3.1). Ci is the input capacitor which holds the charge from the channel which is being output. Cf is the feedback capacitor of the buffer. Camp is the amplifier differential input capacitance. Cbus is the bus capacitance which includes metal to substrate capacitance as well as switch diffusion to substrate capacitance. The fraction of the output signal that is fedback to the amplifier input is given as

$$f = C_f / (C_f + C_i + C_{amp} + C_{bus}) . \qquad (3.1)$$

The ideal closed loop gain of the circuit is given by the ratio of the feedback impedance to the input impedance.

$$G = - C_i / C_f \qquad (3.2)$$

This is not affected by the capacitive elements C_{amp} and C_{bus}. The closed loop bandwidth, on the other hand is affected by these elements. For frequency independent feedback the closed loop bandwidth is given by the product of the feedback function and the gain bandwidth of the amplifier.

$$\omega_{3db} = f \cdot GBW \qquad (3.3)$$

It is apparent that the presence of Camp and Cbus is detrimental to the bandwidth of the buffer. In this application the maximum number of channels is about 200 and each channel is approximately 60 μm wide. This leads to bus to substrate capacitance of 1.3pf. Two hundred minimum geometry switches gives a total switch to substrate capacitance of about .7pf. The differential input capacitance of the amplifier depends on the size of the input transistors but will be .1pf for every 100 μm2 of input transistor gate area. In order that the denominator of (3.1) not be dominated by the undesirable capacitive elements, Cf and Ci should be several picofarads.

By making C_f less than C_i a gain greater than one can be obtained in the output stage, but this is at the expense of closed loop bandwidth. Having C_f greater than C_i would mean even greater bandwidth, but this would give an output stage

with a gain less than one and the output range would be compromised. In the interest of bandwidth and output range in this application, C_f and C_i have been taken to be equal to yield an inverting gain of one. They have been chosen to be 3 picofarads. Larger values would use more real estate with diminishing returns, while values smaller than 3pf would begin to suffer dramatically from bandwidth reduction.

In this application, f will be approximately 1/3. This is an important parameter to know when evaluating the open loop frequency response of the OTA. If used at a feedback gain of 3, the stability requirements on the amplifier are considerably relaxed. A higher gain bandwidth can be obtained, at the same phase margin, than can be obtained with an amplifier which must be unity gain stable.

One objective in the design of this buffer is to minimize the amount of time that it will take the amplifier to settle to a given degree of precision. A hypothetical output from the buffer is shown in figure (3.2).

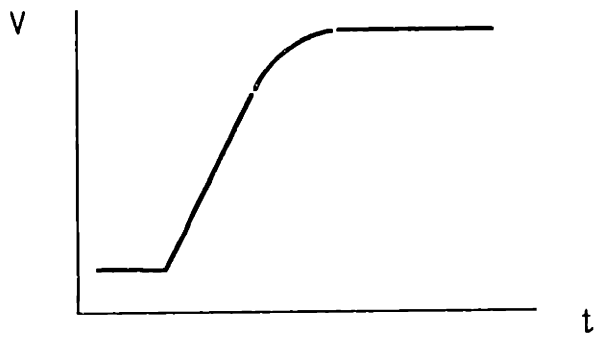


Fig. 3.2 Hypothetical output waveform

In general the output waveform will have two portions. The first, is a nonlinear portion in which the amplifier may be slewing. The second, is a linear portion in which the amplifier is operating in a small signal regime. To minimize the overall

settling time, the slew rate should be made large and the linear region time constant should be made small. Admittedly a trade off may have to be made between both of these objectives. At interest here, is how to minimize the linear region time constant.

If the open loop frequency response of the OTA is modeled as containing a single pole, then the small signal settling portion of the output waveform contains a single exponentially decaying component with a time constant given by

$$\tau = 1 / \omega_{3db} . \qquad (3.4)$$

In this model, increasing the open loop unity gain frequency or the feedback fraction, f, will result in smaller time constants. This model, however, is overly simplistic. Nondominant poles in the frequency response will alter this result.

If the open loop frequency response of the OTA is modeled as containing two poles, then the small signal setting component will contain two frequency components. For large values of the f, the response may be underdamped and the two frequency components are complex. The response is that of a decaying sinusoid. For small values of f, the response is overdamped and there are two decaying exponential components. Figure (3.3) shows a hypothetical open loop frequency response for an amplifier with two poles.

When this amplifier is placed in a feedback configuration, with a feedback fraction of f, the closed loop poles move to different locations in the complex plane. When complex, the location of the poles in the complex plane are customarily given in terms of two parameters, ω_n and ζ . Figure 3.4 shows the way that their locations are related to these parameters.

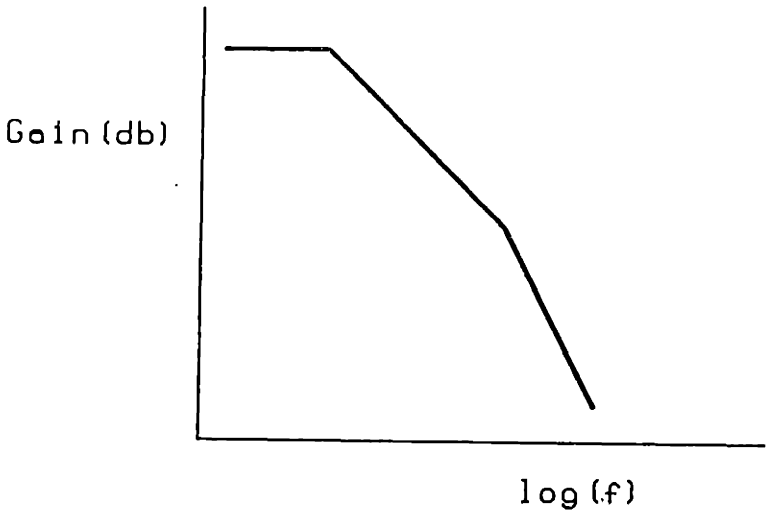


Fig. 3.3 Open loop two pole response

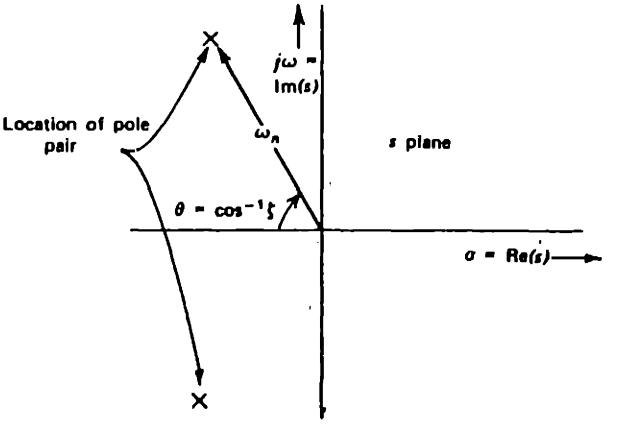


Fig. 3.4 Complex pole locations

The small settling component for a system with poles given by ω_n and ζ will decay with a time constant determined by the real part of the poles location. It will oscillate with a frequency determined by the imaginary part of the poles location.

Figure 3.5 shows a root locus plot of the closed loop pole locations for the two pole amplifier with variable feedback fraction f.

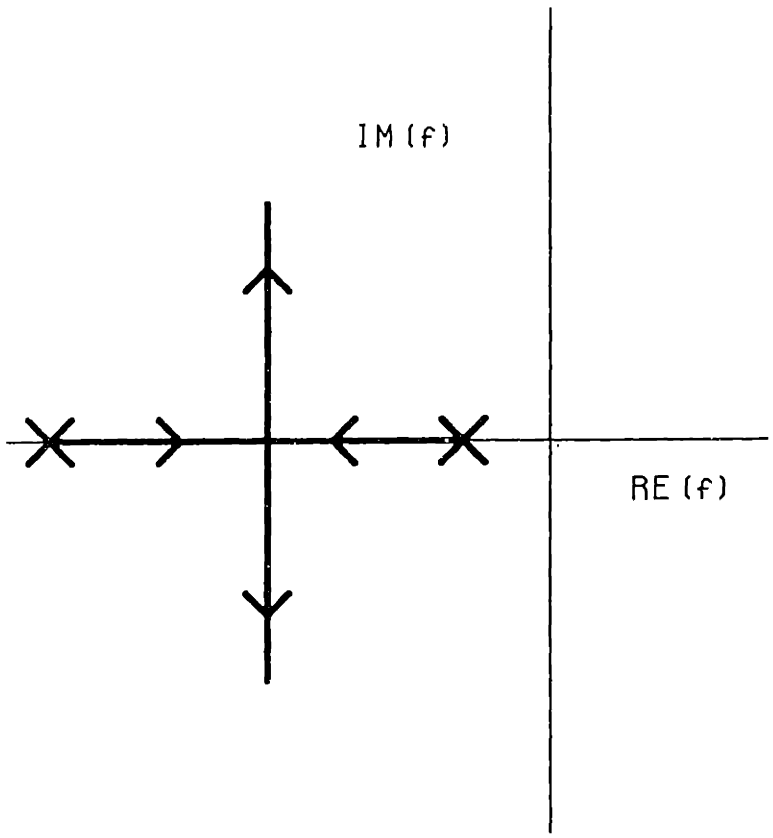


Fig. 3.5 Root locus plot for two pole amplifier

When f = 0, the pole locations are at their open loop positions. As f increases, the poles move toward each other. This is the case of an overdamped system. It is possible for the poles to meet before f=1 and split off into the complex plane. This is the case of an underdamped system. Assuming that that $p_2 \gg p1$, $a \cdot f \gg 1$, and the pole locations are complex, the pole locations can be given by

$$\omega_n = p_2 / 2$$

$$\zeta^2 = p_2 / (4 \cdot a \cdot f \cdot p_1) \quad . \qquad (3.5)$$

Equation (3.5) shows how the feedback fraction, f, determines the complex pole locations for a given amplifier response. Conversely, it shows how the amplifier frequency response determines complex pole locations for a given feedback fraction. The complex pole locations are important because they determine such characteristics as small signal rise time, overshoot, settling time, etc. In this application the most important characteristic is settling time.

The real part of the complex pole location determines the exponential time constant for the small signal settling component. There is no significant dependance of the settling time on the oscillation frequency. The settling time improves as the two poles move toward each other on the real axis but remains essentially constant as they split off into the complex plane. To achieve optimum settling time performance, the network should be in the underdamped regime.

$$f > p_2 / (4 \cdot a \cdot p_1) \qquad (3.6)$$

The two pole model is overly simplistic, however, and it is not advisable to have the network too underdamped.

Figure (3.6) shows the root locus plot for a three pole system. It is clear that one effect of the third pole is to cause the complex pair to move toward the real axis. The settling time performance of this network is optimum at the point where the complex pole pair first leaves the axis. The pair leaves the axis at right angles, so there is no penalty for having a slightly underdamped network. This corresponds to an equal sign for equation (3.6).

$$f = p_2 / (4 \cdot a \cdot p_1) \qquad (3.7)$$

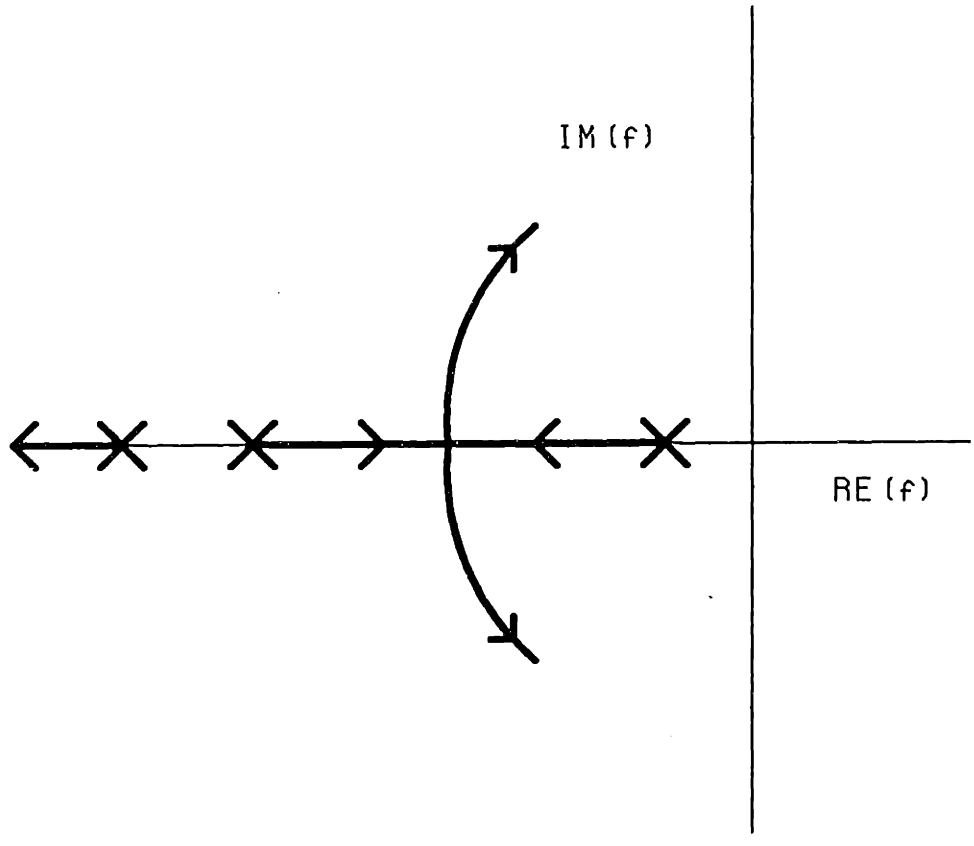


Fig. 3.6 Root locus plot for three pole amplifier

It is understood that it is better to have a slightly underdamped system then a slightly overdamped one. If a value of f = 1/3 is given, then (3.7) indicates that the amplifier should be designed somewhat less than

$$p_2 \approx 4 \cdot f \cdot a \cdot p_1 = 4/3 \cdot a \cdot p_1 = 4/3 \cdot \omega_o \ . \qquad (3.8)$$

Numerical simulation can give accurate assessment of the effects of higher order singularities. Equation (3.8) can be used as a guide, and simulation results can be used for higher order corrections.

The output impedance of the buffer can be arrived at quite easily by modeling the OTA as a single transconductance stage. Figure 3.7 shows such a small signal model suitable for the calculation of the output impedance.

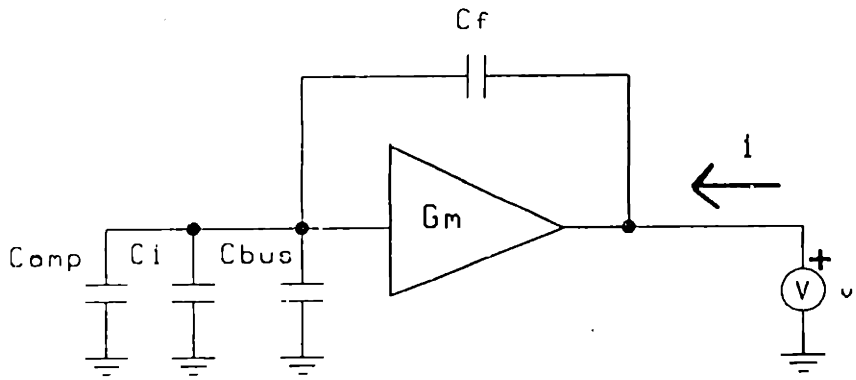


Fig. 3.7 Output impedance model

The result is rather general and in agreement with the well known result that the closed loop output impedance is equal to the open loop output impedance divided by the loop gain.

$$R_o = G_o /(G_o \cdot G_m \cdot f) = 1/ (G_m \cdot f) \qquad (3.9)$$

This is generally valid for feedback connections where the feedback network shunts the output, and the loop gain is much larger than one.

3.2 Large signal considerations

3.2.1 Integration capacitor reset

The feedback capacitor in the output buffer must be reset after each output cycle. Each output should be independent from previous or following samples. The signal charge which has integrated on the feedback capacitor must be removed before another signal can be integrated on that capacitor. Up to this point, no mention has been made as to how this will be accomplished.

One way that this might be done is to place a switch across the feedback capacitor. This switch could be closed between successive outputs when nothing was on the input bus. The capacitor would be discharged and the output would be reset to the bias voltage on the amplifier noninverting input. The switch would then be opened, and the buffer would be ready for the next channel. This techniques has three undesirable consequences.

First, the load capacitance would be reset between each output; and since the load is estimated at 100pf, the excess power dissipated would be a large fraction of the entire power budget. Second, the time taken to perform this function would be lost from the total time allocated to output the signal. The settling associated with this reset would take at least as much time as that associated with the signal itself. Thus the time lost would be a significant fraction of the total time allocated. Third, the amplifier would have to be stable under unity gain operation. This would reduce the obtainable bandwidth.

To address these problems, one might consider switching the load out of the circuit for the reset interval. This is not an option, however, because the load determines the dominant pole of the amplifier. It would generally by unstable without the load capacitance. To design it to be stable both with and without the load capacitance would result in severely degraded performance.

The technique adapted here is to switch the feedback capacitor out of the feedback path, reset it , and switch it back into the feedback path[29]. During the time that the capacitor is out of the feedback path, the amplifier is operating open

loop. By using two feedback capacitors, a previously reset capacitor may be switched into the feedback path immediately after the capacitor with signal charge is switched out. With two capacitors, the period of time that the amplifier is operating open loop may be reduced to a negligible amount. This architecture is shown in figure 3.8.

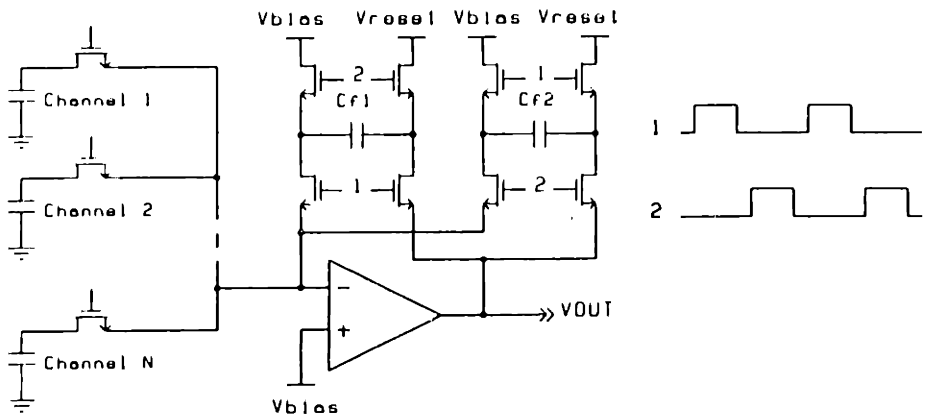


Fig. 3.8 Dual switched capacitors with two phase nonoverlapping clocks

$\phi 1$ and $\phi 2$ are two phase nonoverlapping clocks. Four switches are used to switch each capacitor in and out of the feedback path. This configuration has two advantages. First, it allows for independently resetting the voltage on each plate of the capacitor. Second, it provides a great deal of isolation between adjacent channels. The presence of stray capacitances on each plate of the capacitors provides a mechanism for crosstalk between adjacent channels. As an example of this type of crosstalk, consider the resetting configuration shown in figure 3.9.

Stray capacitances are included in figure 3.9. For the time being, the switch elements are presumed ideal. When a feedback capacitor from figure 3.9 is discharged, the potential difference between the two plates is reduced to zero. The

potential difference between each plate and ac ground is not, however, properly reset. The presence of the stray capacitances cause each plate to assume a potential after reset of

$$V_{reset} = (V_{sig} \cdot C_{s2} + V_{bias} \cdot C_{s1}) / (C_{s1} + C_{s2}) . \quad (3.10)$$

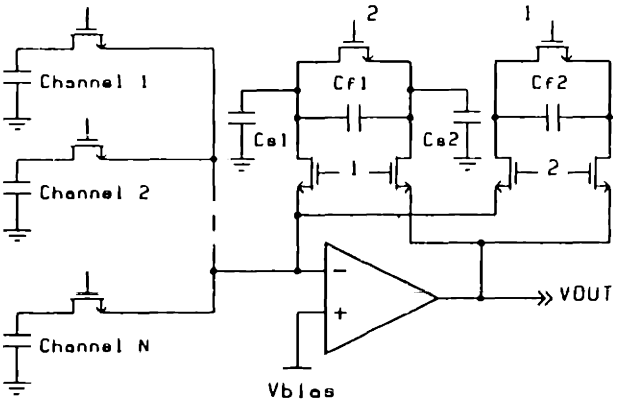


Fig. 3.9 Alternate resetting configuration

This result is very disturbing. The reset voltage depends on the signal. When this capacitor is next used it will carry a contribution from its previous output signal level. This mechanism of crosstalk is eliminated by the configuration of figure 3.8. Other resetting configurations which attempt to reset each feedback capacitor with less than the four switch elements shown in figure 3.8 have similar crosstalk problems.

The clocks, $\phi 1$ and $\phi 2$, must be made nonoverlapping to provide complete isolation between adjacent output signals. The nonoverlapping period should be made as small as feasible. During this time the amplifier is operating open loop and will attempt to drive the output toward one rail of the other. The direction and the amount of output current will depend on the polarity and magnitude of the

differential input to the amplifier. At the end of one integration cycle when a capacitor is about to be switched out of the feedback path, the differential input voltage to the amplifier will be approximately equal to the amplifier offset voltage. If the switch elements were ideal, no charge would be injected onto the input node during the switching process; and the differential input voltage would remain unchanged. Under these circumstances, the amplifier would not drive the output at all. In reality, however, charge will be injected onto the input node, and the resulting change in input voltage will drive the output. The output current will be given by

$$I_{out} = G_m \cdot \Delta V_{in} = - G_m \cdot Q_{in} / C_{in} \ . \qquad (3.11)$$

G_m is the total OTA transconductance, Q_{in} is the charge injected onto the input node, and C_{in} is the total capacitance on the inverting node of the amplifier. The negative sign results from the fact that the charge is injected onto the inverting node of the amplifier. This current slews the output and results in a change in the output voltage.

$$\Delta V_{out} = I_{out} \cdot T_n / C_L = - G_m \cdot Q_{in} \cdot T_n / (C_{in} \cdot C_L) \qquad (3.12)$$

T_n is the length of the nonoverlapping interval. T_n can easily be made as small as 10 nanoseconds. Using conservative estimates for the remaining terms, the change in output voltage is limited to only a few millivolts during the nonoverlapping period.

In figure 3.18 the bias voltage on the noninverting input to the OTA is given as V_{bias}. The reset level for the output plate of the feedback capacitors is given as V_r. To this point, actual values for these voltages have not been

discussed. Two expressions are now developed that relate these voltages to the output voltage range. The OTA itself has a large output range, and with the proper selection of V_{bias} and V_r this may be maintained.

When a channel is switched onto the output bus, charge will either flow to or from the sample capacitor on that channel. If the voltage on the sample capacitor of the channel in question is greater than Vbias, charge will flow from that capacitor to the inverting node of the amplifier and onto the feedback capacitor. If the voltage on the sample capacitor is initially less than Vbias, current will flow in the opposite direction. In either case, current continues to flow until the voltage on the sample capacitor is equal to V_{bias}. As current flows onto the feedback capacitor, the output voltage integrates down below V_r by the same amount that the input voltage was above Vbias.

$$V_{out}(n) = V_r - (V_{in}(n) - V_{bias})$$

$$= V_r + V_{bias} - V_{in}(n) \qquad (3.13)$$

If $V_{in}(n)$ is assumed to have any value between V_{SS} and V_{DD} then equation (3.13) implies that V_r and V_{bias} should obey

$$V_r + V_{bias} = V_{DD} + V_{SS} \qquad (3.14)$$

to maintain maximum output range. If this condition were not valid then $V_{out}(n)$ would want to be higher than V_{DD} or lower than V_{SS} for some values of $V_{in}(n)$. The output would saturate. When equation (3.14) is adhered to, values of $V_{out}(n)$ given by (3.13) will necessarily fall between V_{SS} and V_{DD}.

Another consideration which affects signal range is the charge redistribution which takes place when a new input and feedback capacitor are initially switched into the circuit. This charge redistribution takes place rapidly, before the amplifier has a chance to respond. It is due to the voltage inequality that exists between the input capacitor, feedback capacitor, and the load capacitor. These three capacitors form a loop as shown in figure 3.10. By Kirchoff's Voltage Law (KVL) the sum of the voltage drops around the loop must equal zero.

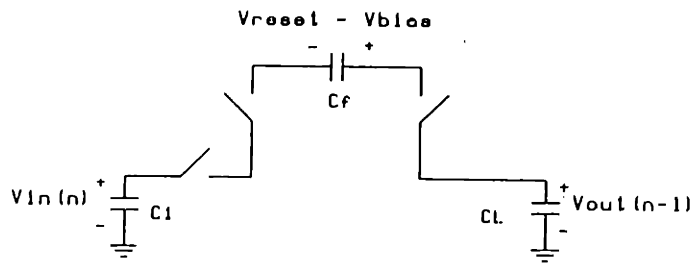


Fig. 3.10 Capacitor voltage loop before insertion

Initially the sum of the voltage drops on the capacitors does not equal zero. There are, therefore, voltage drops across the switch elements. When the switches are closed, there is a rapid flow of current until the voltages equilibrate around the loop. This happens with a time constant determined by the switch resistances and the capacitances, in the loop. For a loop containing a series combination of three resistances and three capacitances the time constant is given by

$$\tau = R \cdot C$$

$$R = R_1 + R_2 + R_3$$

$$1/C = 1/C_f + 1/C_i + 1/C_L \qquad (3.15)$$

With switch resistances less than 10 kΩ, this time constant is less than 50 ns, which is must faster than the amplifier.

Because the load capacitance is much larger than the other two, the output voltage will remain essentially unchanged during this redistribution. Because the other capacitors are equal in size, their voltages change by equal amounts when charge is transferred from one to the other. The approximate capacitor voltages after redistribution has taken place can be determined from these two observations.

$$V_{in}(n) \rightarrow V_{in}(n) - 1/2 \bullet (V_{in}(n) + V_r - V_{bias} - V_{out}(n-1))$$

$$V_r - V_{bias} \rightarrow V_r - V_{bias} - 1/2 \bullet (V_{in}(n) + V_r - V_{bias} - V_{out}(n-1))$$

$$V_{out}(n-1) \rightarrow V_{out}(n-1) \qquad (3.16)$$

The capacitor voltages after charge redistribution are shown in figure 3.11.

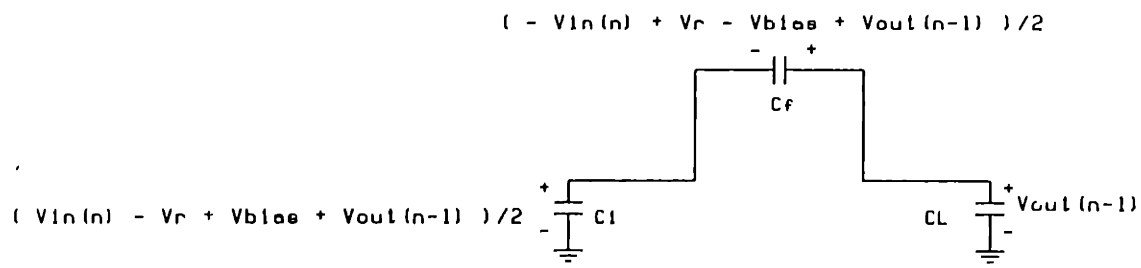


Fig. 3.11 Capacitor voltage loop after charge redistribution

Investigation of the voltage on the input capacitor shows that, if V_{bias} is greater than V_r, this node voltage will be pushed higher than V_{DD} when $V_{in}(n)$ and $V_{out}(n-1)$ are sufficiently high. If V_{bias} is less than V_r, this node voltage will be pulled below V_{SS} when $V_{in}(n)$ and $V_{out}(n-1)$ are sufficiently low. It is certainly undesirable for either of these conditions to occur; because junctions at this node could become forward biased, and signal charge could become lost resulting in signal distortion. Charge redistribution problems are minimized when

$$V_r = V_{bias} \tag{3.17}$$

When equation (3.14) is combined with (3.17), the following prescription is obtained.

$$V_r = (V_{DD} + V_{SS}) / 2$$

$$V_{bias} = (V_{DD} + V_{SS}) / 2 \tag{3.18}$$

With this prescription, V_{bias} acts as a signal ground with V_{DD} and V_{SS} acting as positive and negative supply rails. The output buffer with complete elaboration of feedback elements is shown in figure 3.12.

With Vbias midrail, either n or p channel FETs could be used for the majority of the switch elements. N channels can be used to give lower switch resistance. The switches from the output to the output plate of the capacitors must be complementary to provide low resistance with the output near either rail.

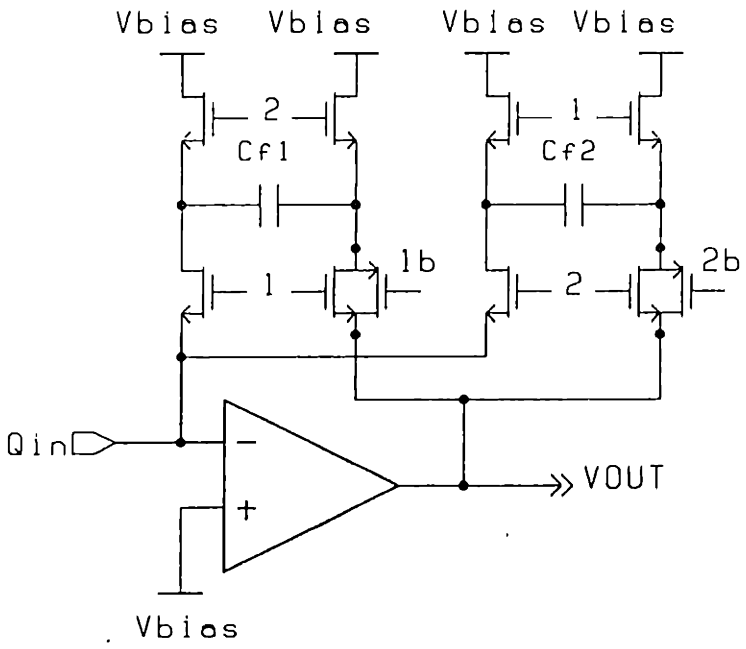


Fig. 3.12 Output buffer with complete feedback description

3.3 Noise Performance

The impact of two noise sources on the noise characteristics of the output buffer will be addressed. Each source is independent of the others so that their effects may be considered individually and their results combined incoherently. The sources are input referred amplifier voltage noise, and switch thermal noise.

The sampled data nature of the output buffer necessitates the use of results from sampling theory. Previous studies of sampled data circuits [30,31] have made the important observation that output noise contributions are of two types. The first type is a direct contribution which results from a source which has a direct path to the output. The second type is an indirect contribution which results from

the sampling of a noise source on a previous clock cycle. The first is continuous and changes continuously over the clock cycle under investigation. The latter is a remnant from a previous cycle which remains constant over the cycle under investigation.

Assume that the output buffer has a grounded or fixed input. Under ideal noiseless conditions, the output voltage would unchanging. In reality this output buffer would show switching transients at clock transitions. These transients are unchanging from cycle to cycle and are not considered noise. This buffer will also show an asymmetry between even and odd outputs due to mismatch in the dual feedback capacitors. Once again, this is a fixed pattern and does not represent noise. In the present discussion, the buffer is assumed to have no fixed patterns; so that the only output with grounded input must be due to noise sources in the buffer. Figure 3.13 shows a hypothetical output waveform under these conditions.

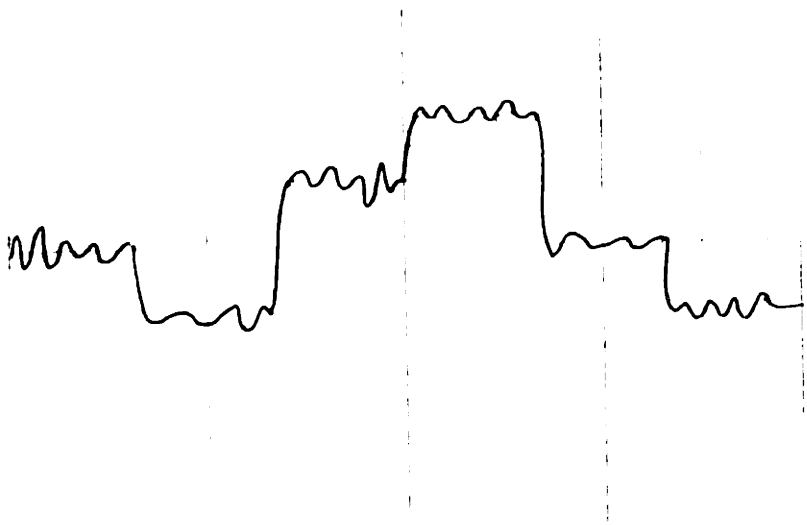


Fig. 3.13 Hypothetical noise output

The noise output shows a continuously varying component, as well as discrete transitions from one cycle to the next. The continuous component represents the direct noise contribution, while the discrete output represents the

sampled component. In an actual example, it will be clear which noise terms are form each category.

3.3.1 Input referred noise voltage

To calculate the noise effects of the input referred noise voltage, it is convenient to use a model in which the noise generator has been brought out of the OTA at the noninverting node. A small signal model can be developed which applies for any particular set of clock states. A noise model which applies for the state in which a signal is being output is shown in figure 3.14.

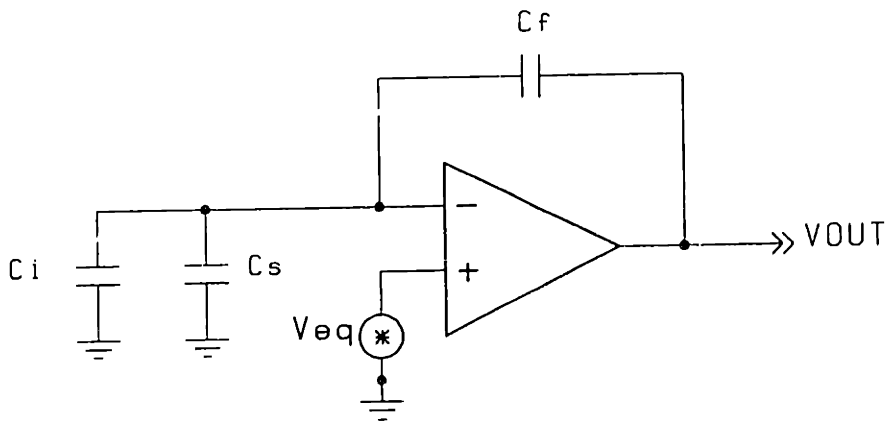


Fig. 3.14 Input referred noise voltage noise source

This noise source provides a direct contribution which is easy to calculate. For frequencies less than the closed loop bandwidth of the circuit, the noise source is amplified by the noise gain.

$$NG = (C_f + C_s + C_i) / C_f \qquad (3.19)$$

At frequencies greater than the bandwidth of the circuit, the gain for the noise source decreases. An equivalent noise bandwidth (NBW) can be defined [30]. Frequency components less than the NBW are accounted for as being amplified by the noise gain, while frequencies above the NBW are accounted for as having no contribution. The NBW is generally close to, but some what larger than, the three db bandwidth. For a single pole response the NBW exceeds f_{3db} by a factor of $\pi/2$. In more realistic cases the factor is closer to one.

The input referred noise voltage also contributes a sampled noise component. This component is more subtle, and arises because the charge state of the stray input capacitance is a function of what the input referred noise voltage was on the previous cycle. This charge state constitutes an initial condition for the current output cycle. The way in which this happens can be quantified by writing equations in the time domain for the output voltage for any given output cycle. Figure 3.15 shows the circuit in question with node voltages and branch currents identified.

$$V_{out}(t) = e_n(t) + 1/C_i \cdot \int i \, dt$$

$$= e_n(t) + \frac{C_i + C_s}{C_f} \cdot \int \frac{d}{dt} e_n(t) \cdot dt$$

$$= (C_f + C_s + C_i)/C_f \cdot e_n(t) - (C_s + C_i)/C_f \cdot e_n(0+) \quad (3.20)$$

The zero of time is taken at the instant that the input and feedback capacitors are inserted into the feedback path. The voltage, $e_n(t)$, is the voltage at the inverting node and is a low passed version of $v_n(t)$. The nonoverlapping clock phase is

assumed to be negligibly small. The first term in this expression is the direct noise contribution, while the second is the sampled contribution.

A limiting case may help to clarify the distinction between these terms. Consider the case where $C_s = 0$. In this case, the inverting node is completely floating during the nonoverlapping clock period. When a freshly discharged set of capacitors are inserted into the loop, the inverting node is set to zero volts. In equation (3.20) the sampled noise contribution must be zero. This is because, with C_s equal to zero, the circuit has no capacity to remember it's charge state at the end of the last cycle. The charge state of the amplifier inverting node is critical in determining the output voltage. The sampled noise contribution is essentially a memory term in the noise expression.

In the general case $e_n(0+)$ is not zero, however, and it makes a contribution to the output noise. At the end of the previous cycle, $e_n(t)$ is given as a low passed version of $v_n(t)$, so that $e_n(0-)$ is a low passed version of $v_n(0-)$. It tracks the inband fluctuations of $v_n(t)$. When affects of stray capacitance are included, the value of en, after new capacitors are switched in place is related to the value which it had before the old capacitors were switched out.

$$e_n(0+) = e_n(0-) \cdot C_s / (C_s + C_i) \qquad (3.21)$$

In the more normal case where closed loop operation is maintained throughout all clock phases, one obtains $e_n(0+) = e_n(0-)$. Equation 3.20 can now be used to calculate the mean square output voltage.

$$< V_{out}(t)^2 > = (C_f + C_s + C_i)^2 / C_f^2 \cdot < e_n(t)^2 > + C_s^2/C_f^2 \cdot < e_n(0-)^2 > -$$
$$(C_f + C_s + C_i)\cdot C_s / C_f^2 \cdot Re_n(t,0-) \qquad (3.22)$$

$R_{e_n}(t,0-)$ is the autocorrelation function for $e_n(t)$. The voltage $e_n(t)$ is a bandlimited version of $v_n(t)$. The bandwidth is the 3 db frequency for the circuit. If the 1/f term in $v_n(t)$ is neglected then the power spectra of $e_n(t)$ will be flat out to ω_{3db}. It's autocorrelation function is equal to the Fourier transform of this power spectra. The autocorrelation function will therefore decay with a time constant $1/\omega_{3db}$. This is also the small signal time constant for the circuit, so that if t is large enough for complete charge transfer to take place in the circuit, then it holds that $e_n(t)$ and $e_n(0-)$ will be essentially uncorrelated. Under these conditions, the third term in (3.22) will be negligible by comparison with the first two. If the power spectra of $v_n(t)$ were dominated by the 1/f component, then a large degree of correlation would exist between $e_n(t)$ and $e_n(0-)$. In the limit where only very low frequency noise components were present, $e_n(0-)$ and $e_n(t)$ would be the same (highly correlated). The relative magnitudes of the thermal and 1/f contributions to v_n^2 will determine which of these cases is more nearly accurate. To be conservative, maximum decorrelation is assumed here, and the third term in (3.22) is neglected. As a practical matter, C_i and C_f are equal, and C_s is approximately equal to that same value. The first term in (3.22) will dominate the second term and is at least three times the third.

3.3.2 Switch Elements

A noise model for a switch element includes the switch resistance and a thermal noise source [28]. Figure 3.15 shows this model where the thermal power spectra is given by

$$v_n^2(f) = 4 \cdot k \cdot T \cdot r \ . \qquad (3.23)$$

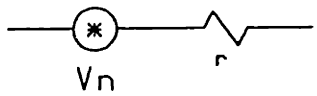


Fig. 3.15 Switch noise model

The incremental switch resistance is given as r, and it should be noted that this is generally a function of the source and drain voltage of the switch.

A form of noise which is commonly referred to as kT/C noise is a sampled noise component of switch thermal noise. As previously discussed in section 1.4.1.2, a capacitor which is shorted by switches will hold a sample of the thermal noise of those switches when the switches are opened. In this buffer, the feedback capacitors hold this sampled noise when they are placed into the feedback loop. Equation (3.20) can be augmented to include this noise voltage.

$$v_{out}(t) = e_n(t) + 1/C_i \cdot \int i\,dt + v_c \qquad (3.24)$$

The sampled noise voltage on the capacitor is identified as v_c. This noise contribution is independent of those previously mentioned, and makes a contribution to the RMS output voltage of

$$< v_{out}(t)^2 > = k \cdot T/C_f \qquad (3.25)$$

fluctuations in the reset value of the input capacitor also makes a contribution to the output voltage noise given by (3.26).

$$< v_{out}(t)^2 > = k \cdot T/C_i \cdot C_f/C_i \qquad (3.26)$$

This is nominally equal to the contribution from (3.25).

The switches which are closed during the output phase make a direct contribution to the output noise. Figure 3.16 is a model for the direct contribution of the switch thermal noise.

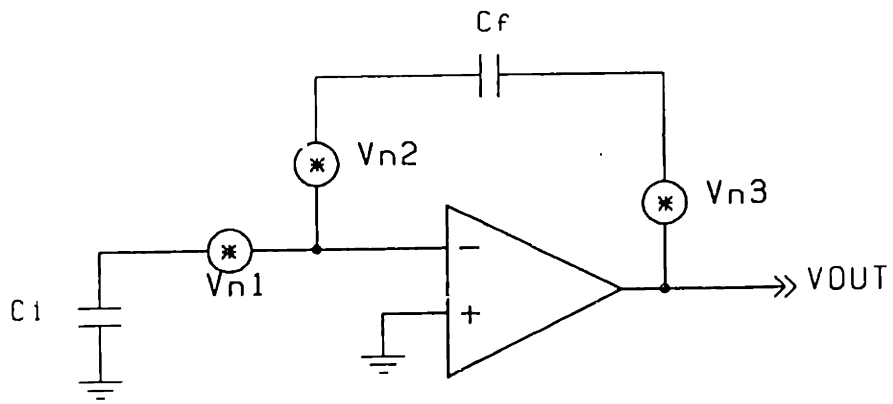


Fig. 3.16 Switch direct thermal noise contribution model

The contribution from each source can be calculated individually and combined incoherently . The result is shown in equation (3.27). Note that the resistors play no role in this expression, because the poles introduced by the resistive elements are at frequencies much higher than the bandwidth determined by the amplifier. If this were not the case, the settling time of the buffer would be affected by the RC time constant established by the switches and the capacitors.

$$< v_{out}^2(t) > = C_f/C_i \cdot v_{n1}^2(f) \cdot NBW + v_{n2}^2(f) \cdot NBW + v_{n3}^2(f) \cdot NBW \; . \qquad (3.27)$$

Only the mean square value of the noise contributions have been calculated here. The spectral shape of all the contributions here identified fall into one of three classes. The first class is the direct contribution of the 1/f term in the amplifier input referred voltage noise. The spectral shaped is as 1/f, as the name implies. It is rolled off by the system bandwidth, but this is of little impact for a function which is already dropping like 1/f. The second class is the direct contribution of the thermal noise sources. This power spectra is flat, but is rolled off by the system bandwidth. The third class is the sampled noise contributions. In the time domain, these contributions look like a series of discrete levels spaced in time at they sampling period Ts. Because the sampled contribution from the 1/f noise source is not included, each level is uncorrelated from the previous or following sample. The autocorrelation function is triangular, and the one sided power spectrum is given as [28]

$$S(\omega) = < v^2 > 2 \cdot T_s \cdot \mathrm{sinc}^2(\omega \cdot T_s / 2) \ . \qquad (3.28)$$

The distinction between a single sided and two sided power spectrum is important, because a one sided power spectrum is what is actually measured. The spectrum is characterized by one parameter, the mean square voltage of the contribution in question. With the inclusion of these spectral shapes, the output noise is completely characterized.

3.4 Power dissipation

Power dissipation can be broken into two components, quiescent and transient. The quiescent power dissipation is the total power dissipated, assuming

that none of the clocks are running. It is the power dissipated by the total bias current of the amplifier. This power can be expressed as

$$P_q = (V_{DD} - V_{SS}) \cdot I_{bias} \cdot (B + 8) . \qquad (3.29)$$

It has been assumed that the bias currents in the N and P channel input stages are the same and that the cascode bias networks are running of that same current.

The transient power is that averaged power which is used charge and discharge capacitances in the circuit. The dominant capacitance in the circuit is the output capacitance. If the output voltage level changes from one cycle to the next the output capacitance must be charged to a different voltage. In supplying this current, power will be dissipated by the driver transistor and it's accompanying cascode transistor. Assume that the output voltage changes from V1 on one output cycle to V2 on the next. Assuming that V2 is greater than V1, the buffer must draw current from VDD and deliver it to the load as it charges up. The instantaneous power dissipated in the buffer is

$$P = (V_{DD} - V_{OUT}) \cdot I = (V_{DD} - V_{OUT}) \cdot C_L \cdot dV_{OUT}/dt . \quad (3.30)$$

The total energy dissipated in the transition from V1 to V2 is

$$E = \int P \, dt = C_L \cdot [V_{DD} \cdot (V2 - V1) - 1/2 \cdot (V2^2 - V1^2)] \quad (3.31)$$

It is clear that the actual energy dissipated is a statistical quantity, which depends on the degree of correlation between adjacent channels. In this case adjacent channels represent radiometric measurements within adjacent narrow spectral

bands. The average energy dissipated in a positive output swing can be expressed as

$$< E_p > = C_L \cdot [\, V_{DD} \cdot < \Delta V_p > - 1/2 \cdot < \Delta V^2_p >] \quad . \tag{3.32}$$

The corresponding expression for a negative transition is

$$< E_n > = C_L \cdot [\, -V_{SS} \cdot < \Delta V_n > + 1/2 \cdot < \Delta V^2_n >] \tag{3.33}$$

It is necessary for the average positive signal transition to be equal to the average negative signal transition or the signal would not stay within the output range of the buffer. It is reasonable to assume that the average difference of the square of the signal for positive transitions is equal to the corresponding quantity for negative transitions. In this case the average energy dissipated in a transition is

$$< E > = 1/2 \cdot C_L \cdot (V_{DD} - V_{SS}) \cdot < \Delta V > \tag{3.34}$$

The average transient power dissipation is dominated by this energy dissipated in the output sample period T_s.

$$P_t = 1/(2 \cdot T_s) \cdot C_L \cdot (V_{DD} - V_{SS}) \cdot < \Delta V > \tag{3.35}$$

This term could conceivably be as large as 360 μW if the signal made a 6 volt transition on each output. This of course is much larger than the entire power budget. Fortunately the system design insures that the average signal swing is much less than this. Due to the selection of spectral bands the signal change from

channel to channel is expected to be no more than approximately .3 V. This would correspond to an average transient power dissipation of 18 μW. It's average change will be significantly less than it maximum change, so it is conservative to say that the transient power dissipated in charging and discharging the load will be less than 20 μW .

3.5 Crosstalk

Crosstalk has not been mentioned in the goals for this design. In the application of this output buffer it is undesirable for the signal from any one channel to corrupt other channels. There are mechanisms in the output buffer which could cause the signals from adjacent channels to influence each other. To be consistent with the other requirements the crosstalk should be lower than -60 db. Two crosstalk mechanisms are considered here.

The first crosstalk mechanism is caused by variations in the settled bus voltage. Under ideal conditions, the bus voltage will return to the OTA noninverting bias voltage of 3 volts at the end of each output cycle. Due to finite OTA gain and incomplete settling time, the bus voltage does not return precisely to the same voltage after each cycle. Because of this effect, the starting voltage of the bus is a function of the signal in the previous cycle. The output buffer operates as a charge integrator, so the charge needed to bring the bus back to its nominal value is integrated as signal in the present cycle.

Imagine that one channel to the buffer is grounded, while the previous channel had the value V. Because of the finite gain, a, of the OTA, the inverting node is settling to a value which is 1/a of the output voltage -V. It starts out with approximately V/2 as discussed in section 3.2.1. Assume that it is settling with a

time constant τ. The noninverting node at the end of the previous cycle will have the value

$$e- = V(1/2 \cdot e^{-T/\tau} + 1/a) , \qquad (3.36)$$

where T is the output sample time. The inverting node for the cycle with the grounded input will settle to zero, but in doing so, the charge

$$Q = C_s \cdot e- = V \cdot C_{bus} (1/2 \cdot e^{-T/\tau} + 1/a) \qquad (3.37)$$

is integrated on the feedback capacitor and an output voltage of

$$V(crosstalk) = -V \cdot C_{bus}/C_f \cdot (1/2 \cdot e^{-T/\tau} + 1/a) \qquad (3.38)$$

results where there should be none. C_{bus} is the bus capacitance. Crosstalk is defined as the ratio of the V(crosstalk) and V and is given as

$$Crosstalk(\text{ bus settling error}) = - C_{bus}/C_f \cdot (1/2 \cdot e^{-T/\tau} + 1/a) \qquad (3.39)$$

This form of crosstalk should be quite small, but can be eliminated altogether by resetting the bus voltage at the beginning of each cycle. This can be done during the nonoverlapping interval of the clock phases. A provision to do this has been made in this design with an additional switch and an additional clock phase, $\phi 3$. It is not anticipated that this should be necessary.

The other form of crosstalk which can arise is due to a stray capacitive coupling between the signal in a channel and the signal bus. The bus is the inverting node of the OTA. When the buffer is integrating, the bus node is at high

impedance and any signal coupling into it will be integrated onto the feedback capacitor. The crosstalk for this mechanism is given by

$$\text{Crosstalk(stray capacitance)} = - C_{stray}/C_f \qquad (3.40)$$

4. PREDICTED PERFORMANCE

Using results from the previous sections, as well as SPICE simulations, a buffer design was performed. Initial device characteristics and parameters were compiled from available data. Two test chips were designed . The first contained many AIRS test circuits, including the output buffer under discussion here. The second was a device test chip containing transistors, resistors, and capacitors, for use in characterizing the Orbit 1.5 μm p-well process. This process is a double poly, double metal process with 250 Å gate oxide. Data from the device test chip is here used to predict buffer performance. It should be pointed out that this data was not available at the time of the output buffer design. The data available at the time of the output buffer design did not include many of the device sizes, and operating currents used in the output buffer design, and is therefore unsuitable to provide quantitative comparison with measured results.

4.1 Orbit 1.5 μm p-well device parameters

AIRS buffer operation is specified for operation at 60 °K. Equipment to operate at this temperature was not available at the time of this research. All device parameter extraction as well as circuit testing was performed at approximately 78 °K. This is the temperature that results with the use of liquid nitrogen as a coolant in the open well of a test dewar. The actual operating temperature depends on the particular dewar being used, the vacuum achieved in the dewar, the thermal contact of the package in the dewar, and the power being dissipated in the device under test. Generally operating temperatures exceed the one atmosphere nitrogen boiling temperature (77.3) by a degree or so. The temperature is reported here as 78 °K as a convenience. Temperature sensor data was not taken with the bulk of the data. There are not expected to be measurable

differences in circuit operation over this range. Full compliance with specifications will require 60°K testing at a later date.

The methodology adapted for device characterization was to extract parameters that would be meaningful in the determination of the buffer performance using the expressions developed in chapters 2 and 3. To make these determinations by hand it is desirable to use a model which is not unnecessarily complex. Here parameters for use in SPICE level 1 models are extracted. This model is sometimes referred to as the Shichman-Hodges model. The most important characterization of this model is that of the FET drain current in saturation.

$$ID = 1/2 \cdot W/L \cdot k' \cdot (V_{GS} - V_T)2 \cdot (1 + \lambda \cdot V_{DS}) \qquad (4.1)$$

$$k' = \mu_n \cdot C_{ox}' \qquad (4.2)$$

$$V_T = V_{TO} + \gamma \cdot (\sqrt{} (2 \cdot \phi_p - V_{BS}) - \sqrt{} (2 \cdot \phi_p)) \qquad (4.3)$$

The expressions are written for an N channel FET, μn is the electron mobility in the channel, C_{ox}' is the strong inversion gate to substrate capacitance per unit area, V_{TO} is the zero bias threshold, γ is the backgate coefficient, and ϕ_p is the Fermi potential of the substrate. The channel length modulation effect is modeled with λ . The channel length modulation effect is often small and the term $\lambda \cdot V_{DS}$ is often assumed much less than one. For long channel lengths L can be considered as the drawn channel length, but for smaller lengths the lateral diffusion of the source and drain implants must be taken into account. SPICE models provided by the foundry are higher level models but are not suitable at cryogenic temperatures. Parameter extraction software was not available for this research, so that use of higher level models would be unworkable. Also, higher level models

have complex or semi-emperical expressions that are unsuitable for hand calculations.

Device sizes and operating currents for the extraction were chosen as close as possible to actual circuit values. In this way, the extracted data is used in its region of maximum validity and errors due to model extrapolation are minimized. Furthermore, simple models are far more useful in performing design tradeoffs because the functional dependance of conflicting requirements are apparent. High level models, if accurate, are useful for final refinement or tweaking. In this case they are not available.

Some device characteristics are known to change significantly at liquid nitrogen temperatures. Others have been reported as having very little change[32]. Junction capacitances are assumed to retain their room temperature values. Junction reverse saturation currents are assumed to be dominated by the conventional exponential temperature dependance $e^{-q \cdot Eg/k \cdot T}$. In this case, they become an issue only when they are the only discharge path for capacitive nodes left floating for a sufficiently long period of time. Because of their importance to circuit performance and their expected or uncertain dependance on temperature threshold voltage, back gate effect, transconductance , output conductance , and noise have been parameterized from actual device measurements. The analysis of these measurements is included as a part of this thesis research. The bulk of the measurements, particularly the noise measurements, have been taken by others.

4.1.1 Threshold Voltage

Threshold voltages were extracted by finding the x intercept in a plot of the square root of drain current versus the gate to source voltage with fixed drain voltage. This follows directly from equation (4.1). A threshold voltage extracted

in this way will necessarily be the correct value to use if the objective is to parameterize the current voltage characteristic of a device operating in the strong inversion region. Threshold voltages are expected to increase as the temperature is lowered according to the relation

$$V_{TO} = \phi_{ms} - Q_{ss}'/C_{ox}' + 2 \cdot \phi_p + \gamma \cdot \sqrt{(2 \cdot \phi_p)} \ . \qquad (4.4)$$

$$\gamma = 1/C_{ox}' \cdot \sqrt{(2 \cdot q \cdot \varepsilon \cdot N_A)} \qquad (4.5)$$

$$\phi_{ms} = \pm E_g/2 - \phi_p \qquad (4.6)$$

$$\phi_p = k \cdot T/q \cdot \ln(N_A/n_i) \qquad (4.7)$$

Equation 4.1 is written for a N channel device. The difference in the gate and substrate work function is given by ϕms. The plus sign in the equation for ϕms is used if the polysilicon gate is doped of the opposite type as the substrate. The negative sign is used if they are doped the same. The effective interfacial areal charge density is denoted by Q_{ss}'. N_A is the concentration of acceptors (assumed uniform) in the p type substrate . E_g is the silicon bandgap, n_i is the intrinsic carrier concentration, and ε is the dielectric constant for silicon. The temperature dependance of the threshold is contained in the explicit dependance in the equation for ϕ_p, and the implicit dependance of n_i. The temperature dependance of n_i is given by[31]

$$n_i \propto T^{3/2} \cdot e^{-q \cdot Eg/2 \cdot k \cdot T} \ . \qquad (4.8)$$

Using the above expressions the expected change in threshold voltage with temperature can be expressed.

$$V_{To}(78) - V_{To}(300) = \phi_p(78) - \phi_p(300) + \gamma\cdot(\sqrt{2\phi_p(78)} - \sqrt{2\phi_p(300)}) \qquad (4.9)$$

$$\phi_p(78) = \phi_p(300)\cdot .26 + .45 \text{ V} \qquad (4.10)$$

Using the value of N_A = 6E16 cm^{-3} provided by Orbit in their SPICE models, a value of $\phi_p(300)$ = .39 V, γ = 1.0 is obtained. With these values equation (4.6) predicts a threshold increase of .28 volts. For P channel transistors the corresponding prediction is .26 volts. Observed increases are .34 V for N channel and .39 V for P channel. This is consistent with previously published results[30] where a partial freeze out of the boron threshold adjust implant can explain larger P channel thresholds. In the N channels, however, the conduction band is bent toward the Fermi level so the acceptor like implant remains fully ionized. Discrepancies such as this show the importance of cryogenic device testing. Table 4.1 shows the extracted thresholds for transistors from the device test chip.

Device Dimension	N	P
30 x 5	1.15	-1.41
30 x 10	1.15	-1.38
30 x 30	1.17	-1.36
5 x 30	1.22	-1.40

Table 4.1 Extracted V_{TO} (V) (78°K)

4.1.2 Transconductance

The transconductance parameter k' was extracted from the slope of a plot of the square root of drain current versus the gate to source voltage with fixed drain voltage. For a device operating in strong inversion the transconductance parameter is given by

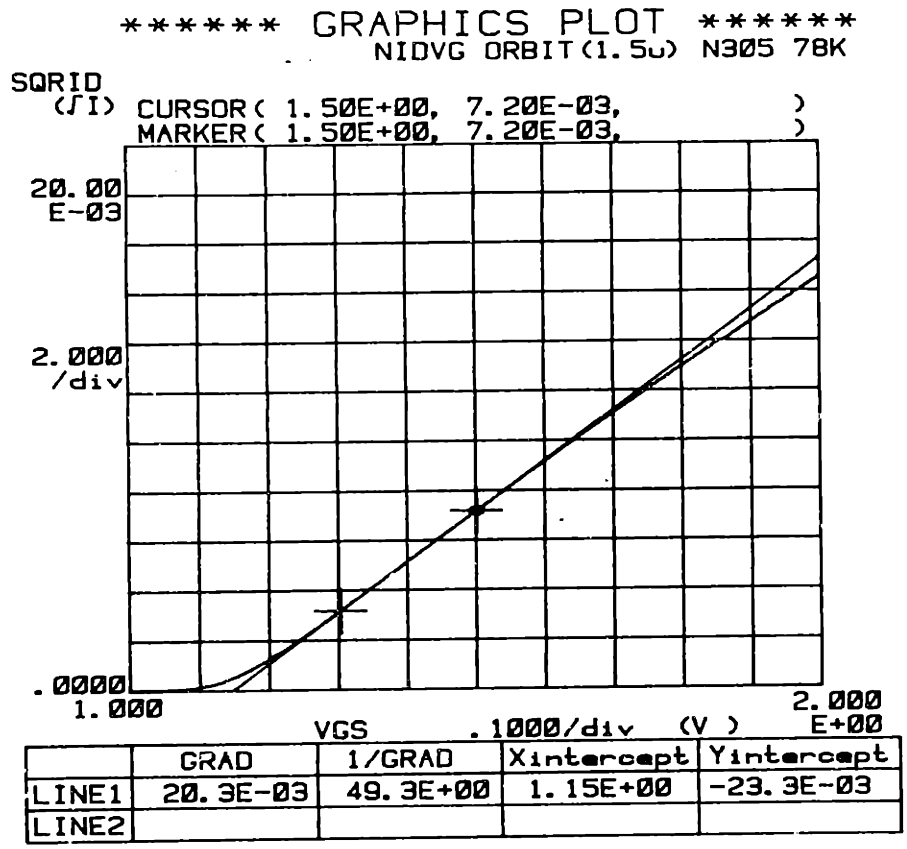
$$k' = 2 \cdot L/W \cdot (\partial \sqrt{I} / \partial V_{gs})^2 \quad . \qquad (4.11)$$

This follows directly from equation (4.1) where the channel length modulation term is assumed small . The maximum value of the slope is used in the extraction so that weak inversion and mobility degraded regions of operation are excluded. Figure 4.1 shows such a plot for an N channel 30/5 transistor. The weak inversion and mobility degraded regions depart visibly from the linear fit. This curve indicates, however, that for currents between 4 μA and 25 μA, the fit is excellent. The threshold voltage of 1.15V is clearly visible as the x intercept.

Table 4.2 show the extracted transconductance parameters values for transistors from the device test chip.

Dimensions (WxL)	N	P
30 x 5	1.37e-4	8.21e-5
30 x 10	1.44e-4	8.45e-5
30 x 30	1.43e-4	9.42e-5
5 x 30	1.21e-4	7.64e-5

Table 4.2 Extracted k' (A/V^2) (78^oK)



	GRAD	1/GRAD	Xintercept	Yintercept
LINE1	20.3E-03	49.3E+00	1.15E+00	-23.3E-03
LINE2				

Fig. 4.1 Square root I vs. V_{gs}, N channel 30x5(78°K)

Corrections for lateral diffusion and oxide encroachment have been made in the data of Table 4.2. When these effects are taken into account, L and W must be replaced by Leff and Weff , where

$$Leff = L - 2 \cdot LD$$

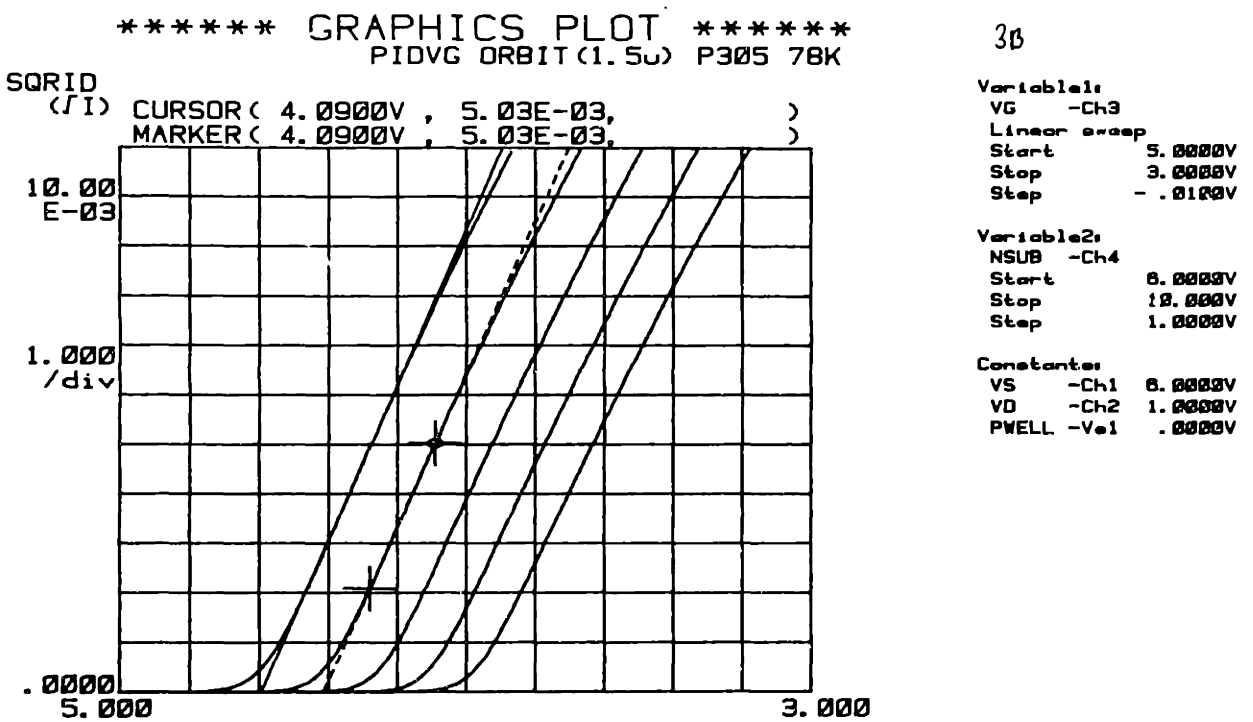
$$Weff = W - 2 \cdot WD \quad . \qquad (4.12)$$

Values for LD and WD provided by Orbit at room temperature are used here. LD and WD models any biasing between drawn and effective electrical dimensions. These dimension corrections are primarily physical in nature and not expected to vary with temperature. LD = .04μ and WD = .3μ . Even with these corrections, the data in Table 4.2 shows some spread. Results from other di concur with the results shown here. It appears that the WD for this lot is probably somewhat larger than .3 μm, although there is not another smaller width transistor on the device test chip to verify this. The effective electrical width of the 5x30 devices, according to Table 4.2 , is only 3.8 μm

The values in Table 4.1 are significantly higher than room temperature values. As shown in equation (4.2), the transconductance parameter is the product of the carrier channel mobility and the gate capacitance. The gate capacitance is expected to have negligible temperature dependance. SPICE models the mobility temperature dependance as being proportional to $T^{-3/2}$. This is because the dominant scattering mechanism at room temperature is with acoustic phonons. The scattering rate for this mechanism has a $T^{3/2}$ dependance [31], and the mobility is inversely proportional to this scattering rate. At 78 ^{o}K, however, the scattering rate from acoustic phonons has dropped by a factor of 7.7 and is no longer the dominant term in the total scattering rate. Scattering is now determined by ionized impurity scattering and surface scattering [30]. The observed mobility increase is therefore less than the factor of 7.7 predicted by SPICE. In this study, mobilities increased by a factor of 4.6 for the N channel and 6.4 for the P channel transistors. This is in agreement with previous results and presumably due to lower substrate doping for the P channel devices.

4.1.3 Back Gate Effect

The threshold voltage of a MOS transistor increases as the voltage between source and bulk increases. This effect is known as the back gate effect and is characterized by equation (4.3). The increase is parameterized by γ and ϕ. Because of the logarithmic dependance of ϕ (4.7) it is confined to a narrow range of values. It is conventional to assign ϕ a nominal value, and use measured values of V_T vs V_{SB} to determine γ. It is also possible to use iterative procedures to find values for both parameters which will give a best fit to the data. Figure 4.2 shows a plot similar to figure 4.1 for a 30x5 P channel transistor with five values of substrate voltage.



	GRAD	1/GRAD	Xintercept	Yintercept
LINE1	-15.7E-03	-63.7E+00	4.59E+00	72.1E-03
LINE2	-15.5E-03	-64.7E+00	4.42E+00	68.3E-03

Fig. 4.2 Square root I vs. V_{gs} for different V_{SUB}, P 30x5(78°K)

Gamma can be assigned as the slope of plot of V_T versus the quantity $\sqrt{(-2\cdot\phi_n+V_{BS})} - \sqrt{(-2\cdot\phi_n)}$. Such a plot is shown in figure 4.3, where ϕ_n has been assigned a nominal value of .55 V corresponding to N_D = 3.3e16 cm^{-3}.

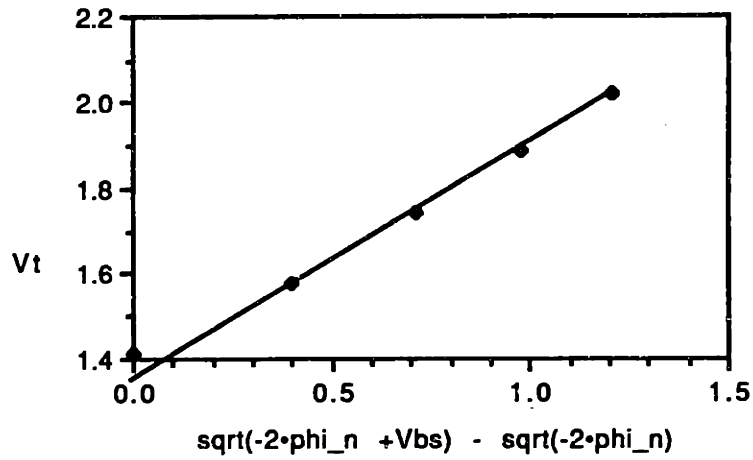


Fig. 4.3 Plot for extraction of γ (78°K)

All points, with the exception of the first, give a good linear fit. This is typical of transistors with implants[37]. The smaller initial slope in this P channel FET is consistent with a threshold adjust implant of acceptors. Room temperature data for the same P channel transistor is shown in figure 4.4.

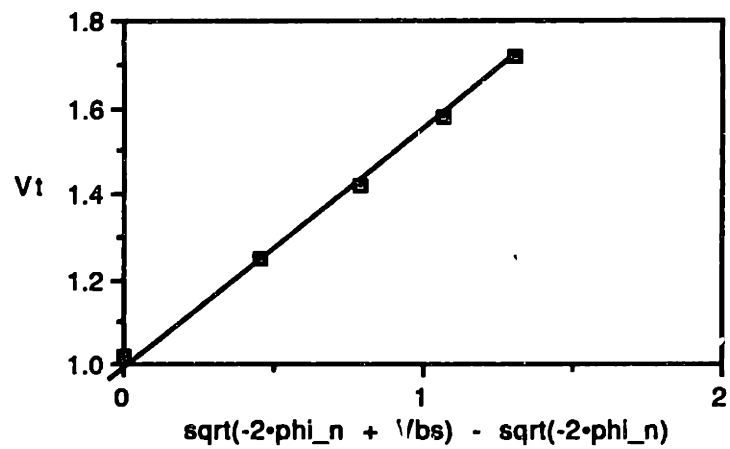


Fig. 4.4 Plot for extraction of γ (300°K)

The slope for V_{SB} greater than 1 V is essentially the same indicating no change in the bulk donor concentration with temperature. The slope in the region with VSB less than 1 V is different, indicating a potential temperature dependance of doping density in the surface region.

Due to a layout error in the device test chip, N channel data for V_T vs V_{SB} was not reliable. A review of existing data was used to provide a value. As with the P channel FETs, the γ as indicated by substrate to bulk voltages greater than 1 volt was temperature independent. A different slope is observed, however, for smaller values. Table 4.3 shows extracted γ's for N and P channel transistors.

Dimensions (WxL)	N	P
30 x 5	1.0*	.54
30 x 10	*	.54
30 x 30	*	.54
5 x 30	*	.57

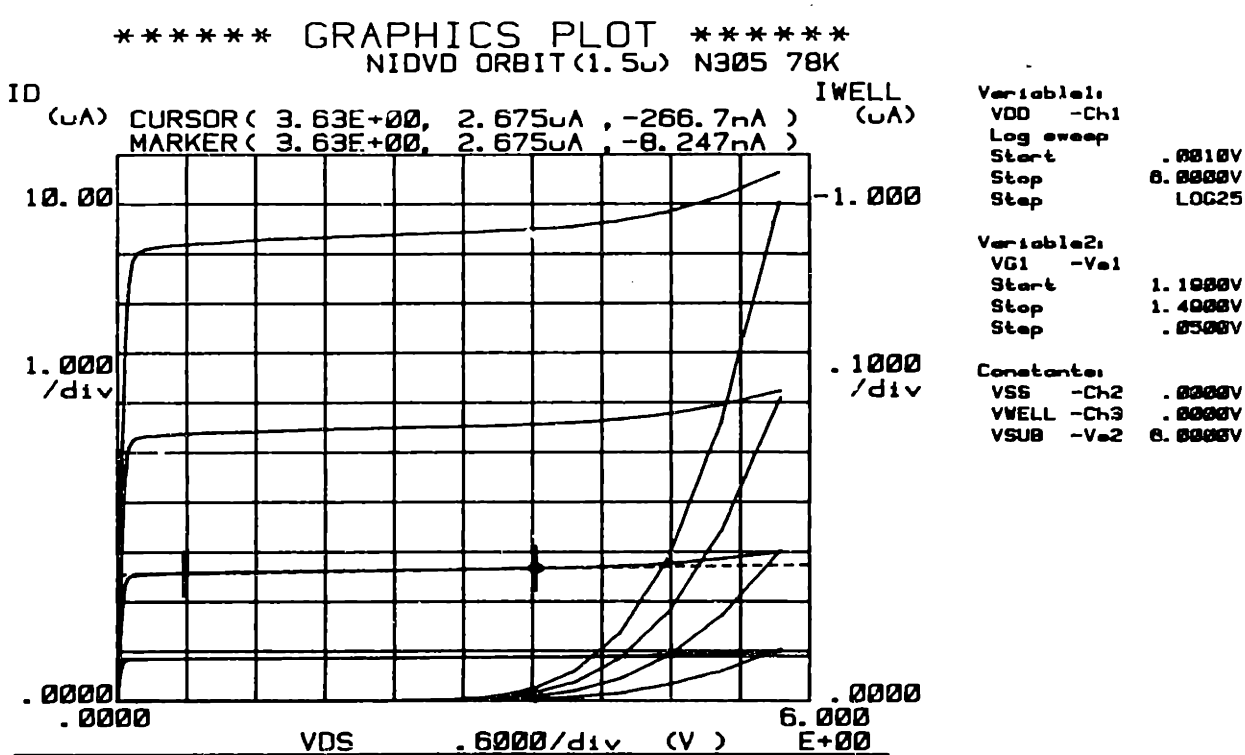
Table 4.3 Extracted γ (78°K)

Values given are for $|V_{SB}|$ greater than 1 volt. Asterisks given for the N channel transistors indicate the lack of data caused by the layout error.

4.1.4 Output Conductance

Output conductance is characterized by the parameter λ, which is analogous to 1 over the Early voltage defined for bipolar devices. Output conductance is defined as the slope of a line fit to the saturation region of an ID

vs VDS plot. Lambda is given as reciprocal of the x intercept of such a line (measured relative to the y axis). Lambda is essentially an empirical parameter to describe the output conductance of the device. Complex models are required to calculate the output conductance of a real device from first principles. Figure 4.5 shows the output characteristics of a 30x5 N channel transistor.

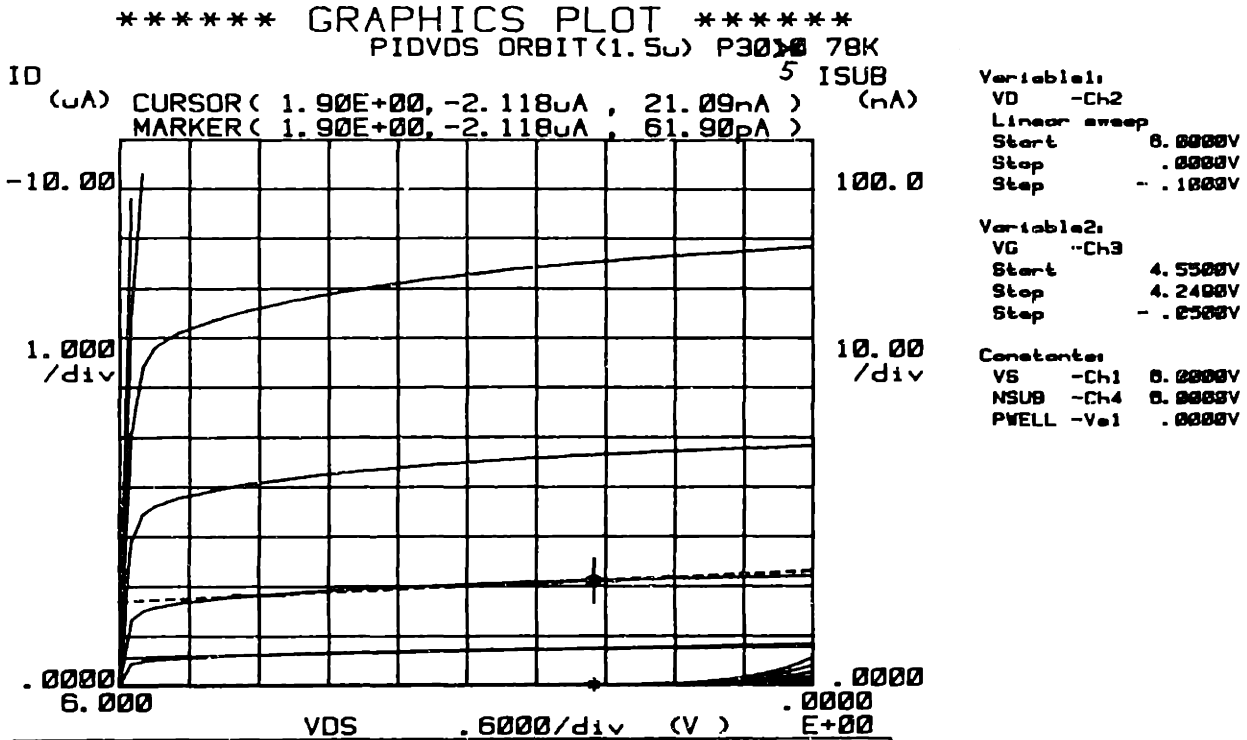


	GRAD	1/GRAD	Xintercept	Yintercept
LINE1	9.12E-09	110E+06	-93.1E+00	849E-09
LINE2	32.7E-09	30.6E+06	-78.1E+00	2.56E-06

Fig. 4.5 Output characteristic N channel 30x5 (78 ^{o}K)

Figure 4.5 also shows the well current with a scale on the right hand side. The output characteristic is quite straight over much of the output range but picks up noticeably at high values of V_{DS}. In this region the output conductance is dominated by impact ionization, as mentioned in Section 1.4.1.1. It must be

appreciated that this is a serious breakdown in the Shichman-Hodges model. Output conductance can easily increase by an order of magnitude over its midrange value.



	GRAD	1/GRAD	Xintercept	Yintercept
LINE1	44.0E-09	22.7E+06	19.0E+00	-836E-09
LINE2	101E-09	9.95E+06	23.0E+00	-2.31E-06

Fig. 4.6 Output characteristic P channel 30x5 (78 ^{o}K)

Figure 4.6 shows a similar plot for a 30x5 P channel transistor. Impact ionization has a negligible effect on the output conductance. Because of the lower mobility, the transition between the triode and saturation region is at a greater voltage. It is also clear that the output conductance over the mid range is about 4 times higher than that of the N channel transistor. The N channel characteristic is relatively linear over the midrange while the P channel characteristic is

bowed over the entire saturation region. The larger output conductance of the P channel is due to lighter doping in the channel region of the device, as compared with the N channel. Table 4.4 show the extracted channel length modulation factor for devices from the device test chip.

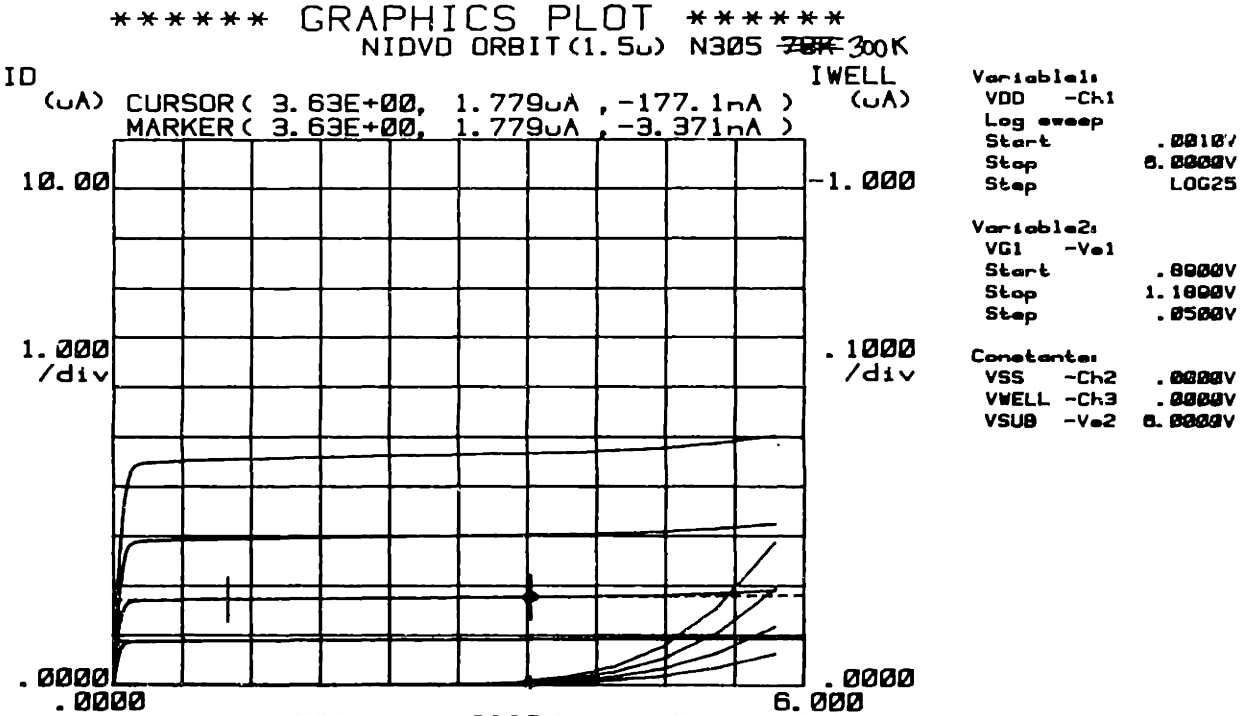
Dimensions (WxL)	N	P
30 x 5	.0120	.068
30 x 10	.0056	.036
30 x 30	.0020	.012
5 x 30	.0023	.012

Table 4.4 Extracted λ (78°K)

The data represent an average of values extracted at approximately 1 μa and 2 μa of drain current. As can be seen from Fig. 4.5 and 4.6, the linear fit to the saturation region is done to avoid both the triode region and the impact ionization region if present. Because of the bow in the P channel data, the value must be interpreted as an average value. Lower output conductances will be obtained at high drain to source voltages, and higher values at lower drain to source voltages. This model must be used carefully, the P channel output conductance can differ from the model prediction by about a factor of two at each extreme of drain to source voltage. Performance predictions such as dc gain which depend on output conductance are therefore expected to show discrepancies.

The data in Table 4.4 does show an expected dependance of λ on device length. Lambda is roughly inversely proportional to L and independent of W. Corresponding room temperature data is quite similar, showing all of the same

features as the 78 ^{o}K data. Figure 4.7 shows the same 30x5 N channel FET from figure 4.5 at room temperature.



	GRAD	1/GRAD	Xintercept	Yintercept
LINE1	10.0E-09	99.7E+06	-87.8E+00	881E-09
LINE2	19.2E-09	52.1E+06	-89.1E+00	1.71E-06

Fig. 4.7 Output characteristic N channel 30x5 (300 ^{o}K)

The conductance values are essentially the same, in agreement with previous findings[32]. The impact ionization current is less by about about 30 per cent. This is in qualitative agreement with the well established fact that impact ionization rates increase with decreasing temperature[33], and with previously published transistor results[35].

4.1.5 Noise

Input referred voltage noise was measured for the P channel devices on the device test chip. Unfortunately , the previously mentioned layout error rendered noise measurements on the N channel devices unusable. Fortunately, an N channel transistor on the circuit test chip was suitable for noise measurements.

In the noise measurements for the P-channel devices the drain, gate, and bulk potentials were specified by batteries and resistive voltage dividers. The source terminal was attached to the virtual ground input of a PAR 181 current preamplifier. The output of the PAR was used as input for an HP 3561A Dynamic Signal Analyzer. The drain current noise was referred to the gate by dividing by the device transconductance, measured on the same setup by dithering the gate voltage. Figure 4.8 shows the input referred noise voltage for a P channel 30x5 FET at 78°K, with a drain current of .1μA. The same result at 1 μA is shown in Figure 4.9.

Both spectra show a low frequency component which varies approximately as $1/\sqrt{f}$. A component in the voltage power spectra(V^2/Hz) with a frequency dependance of 1/f will appear as $1/\sqrt{f}$ in the voltage spectra ($V/\sqrt{Hz}$). In the 1μA measurement, the noise spectra flattens out below 10 Hz. Both spectra flatten out at high frequency well before they obtain the theoretical thermal noise limit, which is marked on each plot. Measurement limitations prevent measuring data above 10kHz. The 1/f component on each plot is relatively independent of drain current, in agreement with the model presented in section 2.1.5. The relatively flat section of each spectrum between 1kHz and 10kHz does drop considerably with increasing current. It is not clear, however, if the spectra will remain flat or drop again in the decades beyond 10kHz. To be conservative, it would be prudent to model this data with a thermal component about 3 times theoretical and a 1/f component with about $1\mu V/\sqrt{Hz}$ at 1Hz. Data for the other P channel

transistors and other current levels roughly agree with this interpretation. Previous data[36] has shown high frequency noise greater than the theoretical thermal noise limit. It also shows that it exceeds the limit by an increasing amount as the temperature is reduced to 77 K. The origin of this excess noise may be a hot electron effect, which would explain why it became worse at low temperatures, but it is not clear.

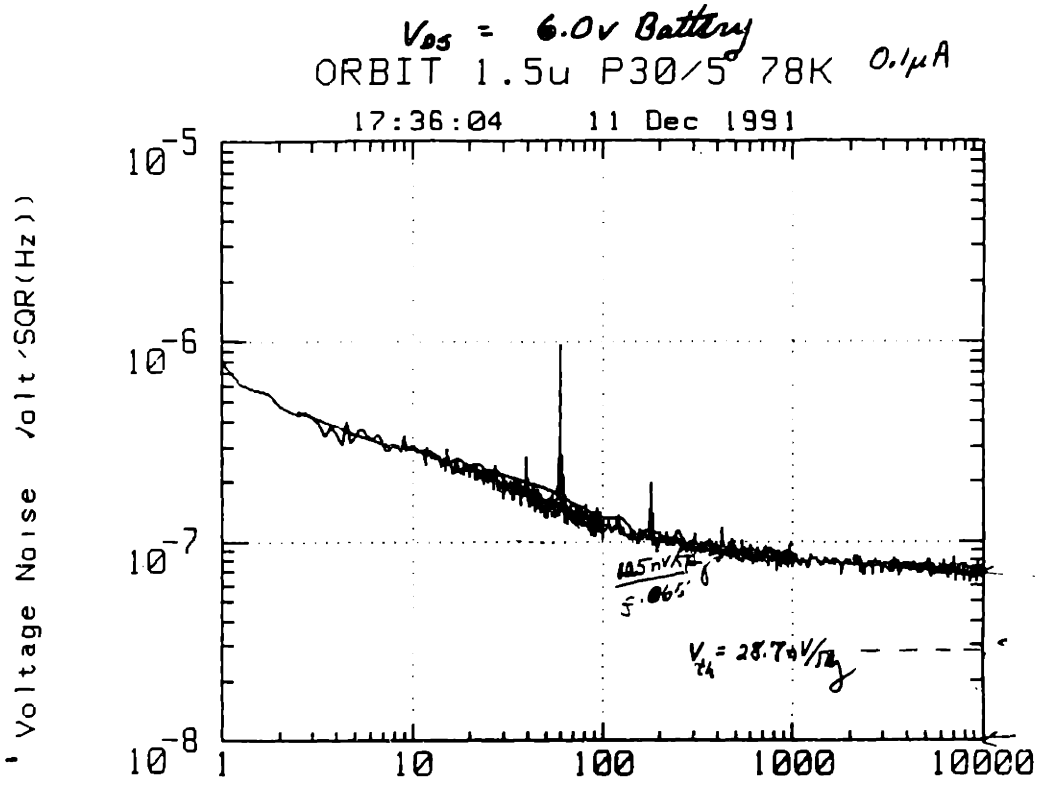


Fig. 4.8 Input referred voltage noise, Id = .1μa P 30x5 (78oK)

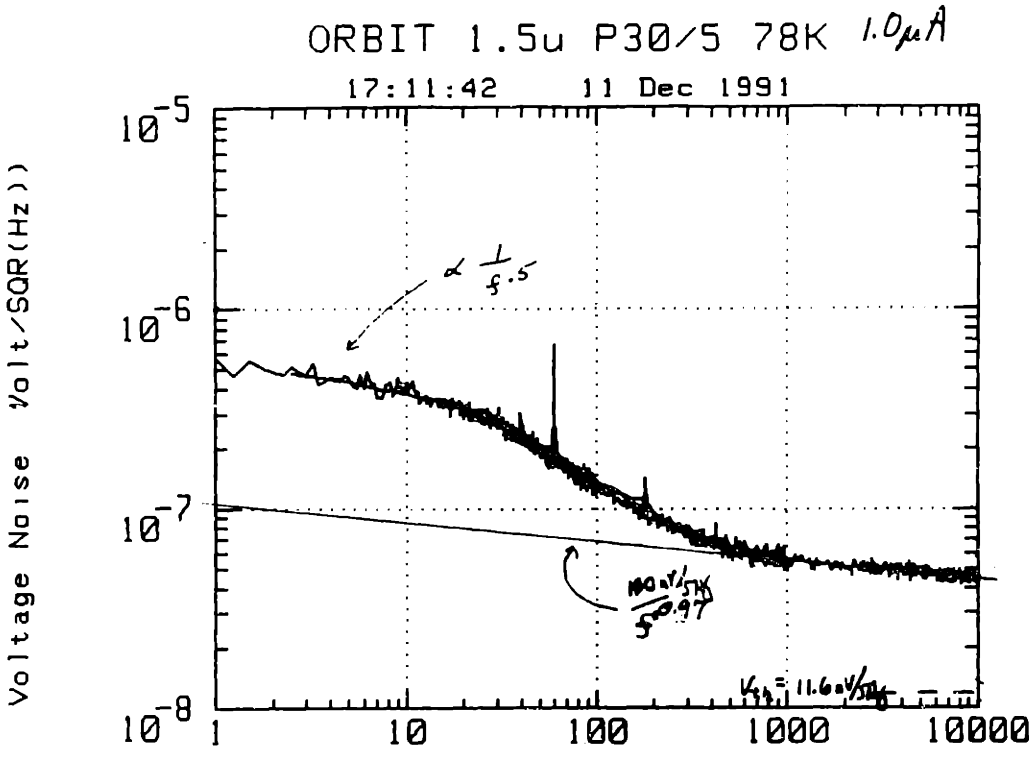


Fig. 4.9 Input referred voltage noise, Id = 1μA, P 30x5 (78oK)

The noise data for the 150 x 5 N channel transistor on the circuit test chip was measured in a different way. The drain and gate were tied to voltage references, the bulk was tied to the source, and the source tied to a low noise current sink. The noise voltage was measured directly at the transistor source. Figure 4.10 shows the voltage noise spectrum at a current of 5 μA.

This spectrum shows a 1/f component and a white component. The 1/f component has a value of .5 V/√Hz at 1 Hz. The white component is a measurement noise floor and is well above the theoretical thermal noise for this device. It is unclear if the N channel devices reach their theoretical thermal noise limits. Table 4.5 shows measured 1/f noise components for the transistors on the device test chip.

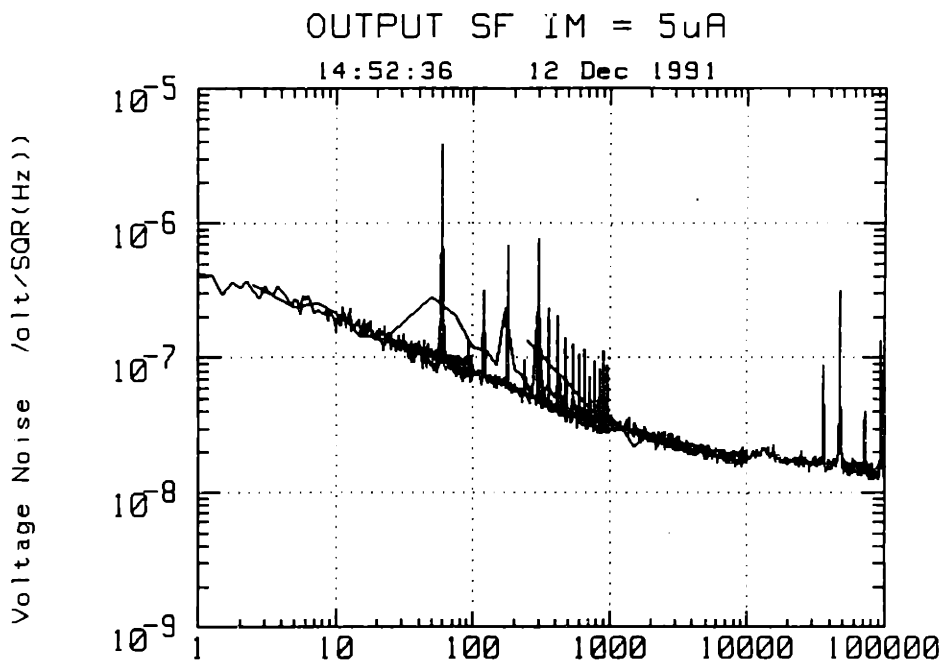


Fig. 4.10 Input referred voltage noise, Id = 5μA, N 150x5 (78oK)

Dimensions (WxL)	N	P
30 x 5	*	1.0
30 x 10	*	.3
30 x 30	*	.2
5 x 30	*	.5
150 x 5	.5	*

Table 4.5 Voltage noise at 1 Hz (μV/√Hz) (78°K)

The data reported in Table 4.5 is for drain currents of approximately 1 μA, with the exception of the 150 x 5 as previously mentioned. There is, however, no strong dependance on drain current. The data shows that the 1/f component does decrease with increasing gate area as predicted by equation (2.61). The P channel data, when averaged, give an input referred noise voltage of .7 $\mu V/\sqrt{Hz}$ for a 100 μm^2 device. The single N channel measurement suggests an input referred noise voltage of 1.4 $\mu V/\sqrt{Hz}$ for a 100 μm^2 device. The P channel devices are quieter than the N channel, which is in agreement with previous results. It will be assumed here that the input referred noise power scales inversely with area as indicated in equation (2.61).

4.2 Test Chip Output Circuits

Complete test chip schematics are included as figures 4.11, 4.12, 4.13. The complete OTA schematic is figure 4.11. Transistor sizes for the OTA design are shown in Table 4.6. Device designations correspond to figure 2.7. A step up ratio of ten was taken between the input stage and the output stage to achieve a large transconductance to bias current ratio. With a small step up ratio, more of the total current is lost in input stage and cascode biasing circuitry. With a step up ratio of ten, essentially all the bias current contributes to the transconductance. Input devices were biased in the moderate inversion region. Dynamically biased cascodes were used, biased so that the driver transistor had a nominal .3V drain to source voltage under quiescent conditions.

The option exists for placing three N channel transistors in there own well. The present design places the output stage N channel cascode in its own well for improved cascode bias accuracy. The N channel input differential pair was not placed in a separate well. The increased threshold voltage improves the output

current limit. No penalty arises with transconductance because the minor loop feedback keeps the source voltage fixed. The input stage N channel cascode transistor were placed in the common p well with the bulk tied to VSS.

It is clear from the device sizes in Table 4.6 that this design is pushed more to high speed rather than to low noise. This is intentional, and is due to the fact that the most difficult requirement to meet is the settling time requirement. Even as the design stands there is less margin in the area of settling time than for noise. Overall system performance may improve with changes in the direction of more bandwidth at the expense of more noise but is unlikely to improve with changes which decrease noise at the expense of speed.

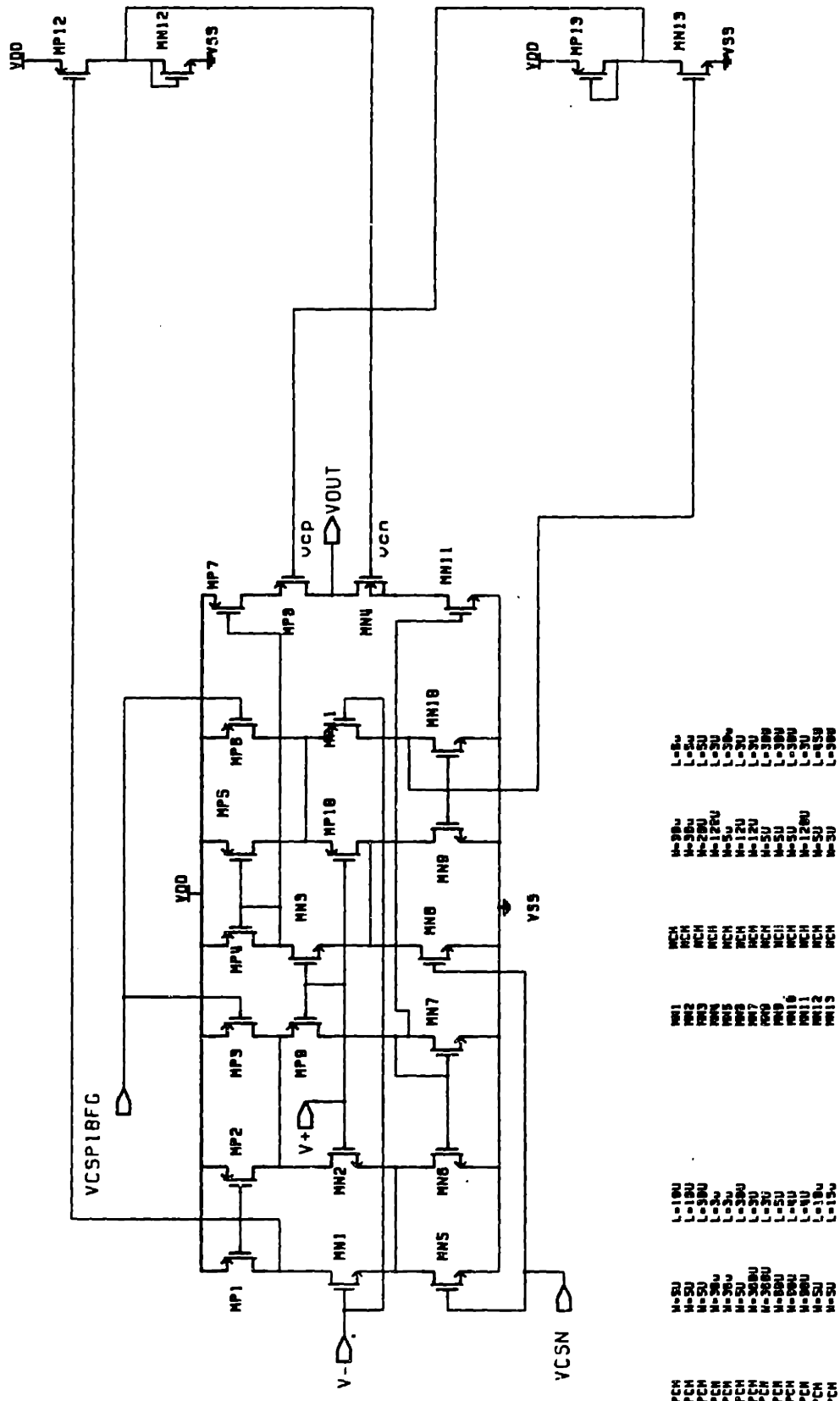


Fig 4.11 Test chip OTA schematic

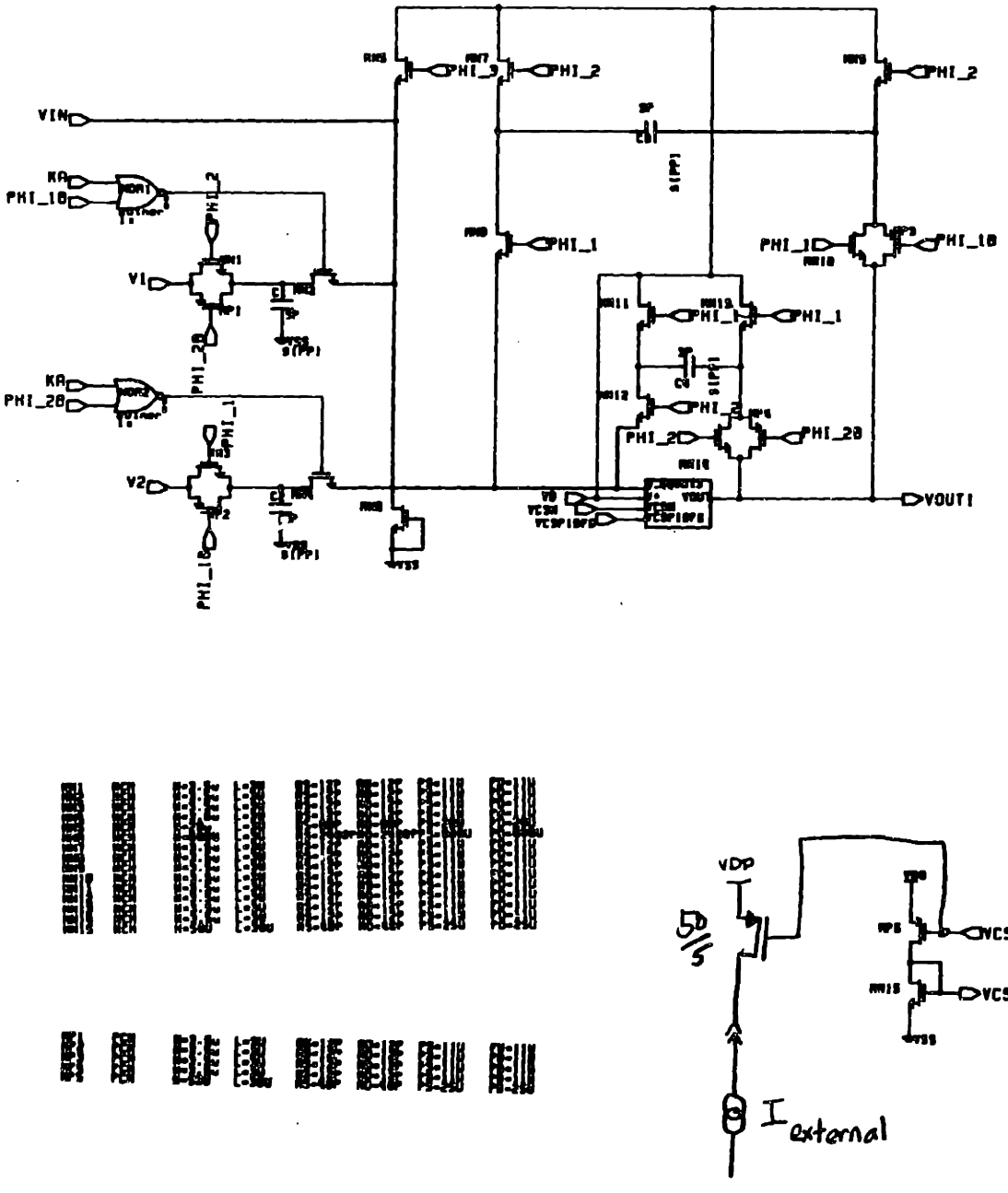


Fig 4.12 Test chip output buffer schematic

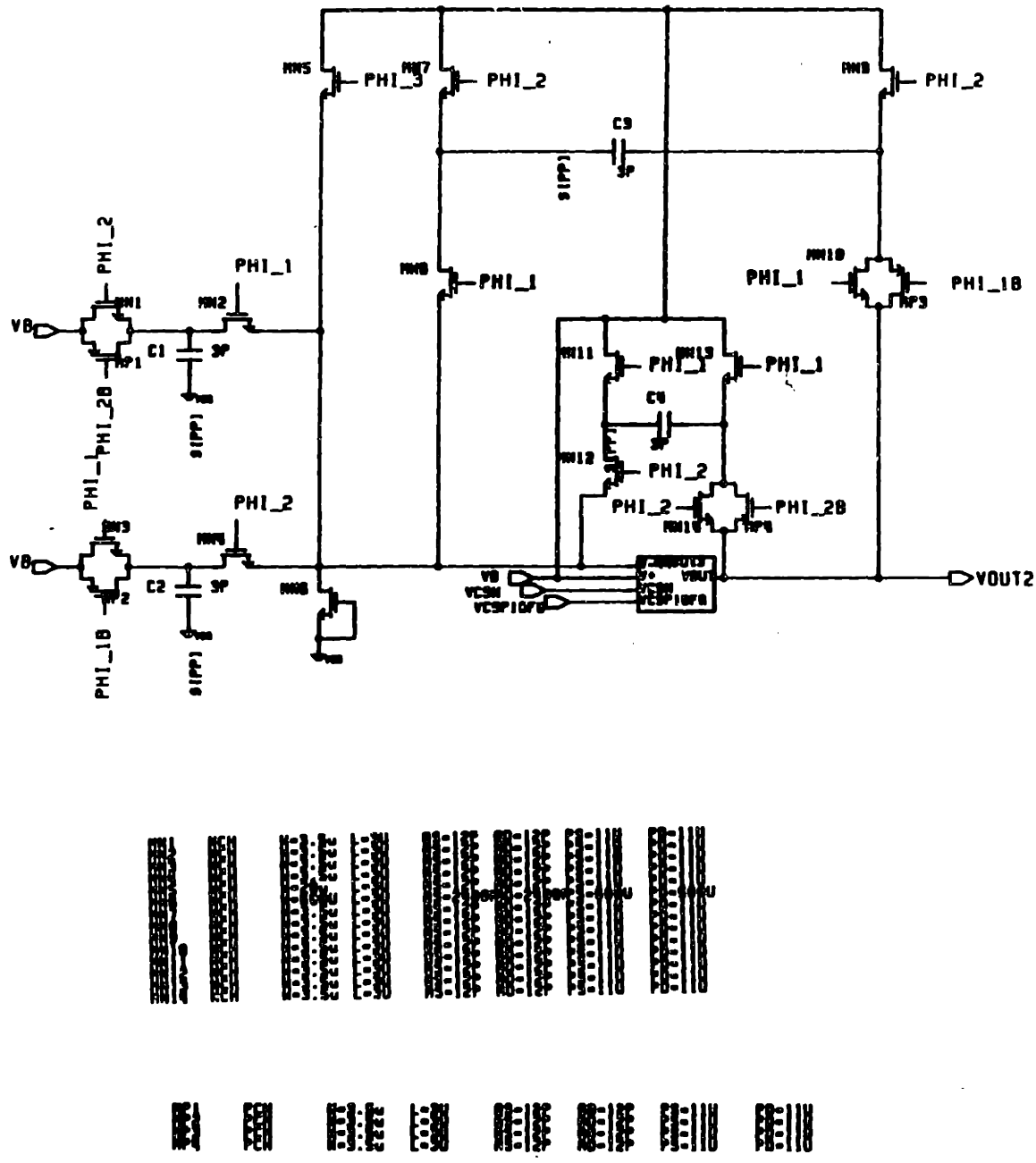


Fig. 4.13 Test chip reference output buffer schematic

Device	W(μm)	L(μm)	Device	W(μm)	L(μm)
M1n	30	5	M1p	90	4
M2p	5	10	M2n	5	30
M3p	60	5	M3n	20	5
M4n A,B	12	3	M4p A,B	36	3
M4n C	120	3	M4p C	360	3
M5n	5	45	M5p	5	15
M6n	120	3	M6p	360	3
Ibias n,p	5	30	Ibias n,p	5	30
Ibias n,p	1 μa		Ibias n,p	1 μa	

Table 4.6 Amplifier device sizes and bias currents

The output buffer is shown in figure 4.12. It has two test inputs, labeled V1 and V2 which simulate adjacent channels of the multiplexer. All switches are 3.5x3.0 with the exception of MN5 and MN6. MN5 is switch which can be used to reset the input bus. It is provided as an option to improve crosstalk in the output buffer. It is sized 20x3 so the bus can be reset in the nonoverlapping clock period. MN6 is provided to simulate the source/drain capacitance of 200 switches on the bus. The nor gates on figure 4.12 are provided so this circuit can interface with other test circuits on the chip. For all testing considered here KA is low and the nor gates operate as inverters. The bias current is adjusted with the input labeled VCSP1BFG. It is generated from a P channel 50x30 current mirror transistor biased with an external current source.

A balanced differential output is required for the AIRS program. The differential output is provided by having a duplicate charge integrating amplifier which provides a reference output. The reference output circuit is shown in figure 4.13. To provide identical output characteristics, the reference circuit is a mirror of the signal output circuit. The OTA noninverting bias voltage, labeled VB, is used to provide the output level of the reference circuit.

4.3 Predicted amplifier performance

4.3.1 Transconductance

The amplifier transconductance is given in equation (2.22). The N channel input transistors are 30 x 5 biased at 1 μA. Figure 4.1 shows that this is not in strong inversion. To get maximum transconductance for a given bias current, and achieve a large gate area, these transistors are biased in the transition region between strong and weak inversion. Figure 4.19 shows the measured transconductance of a 30 x 5 N channel FET from the device test chip.

The weak inversion region of operation is visible at low drain currents where the slope of the curve is seen to be very close to one. The strong inversion region of operation is visible at high drain currents where the slope of the curve is seen to be very close to 1/2. A drain current of 1 μA is in a region of moderate inversion. The measured transconductance is 30 μS, while the strong inversion parameters would predict 40 μS.

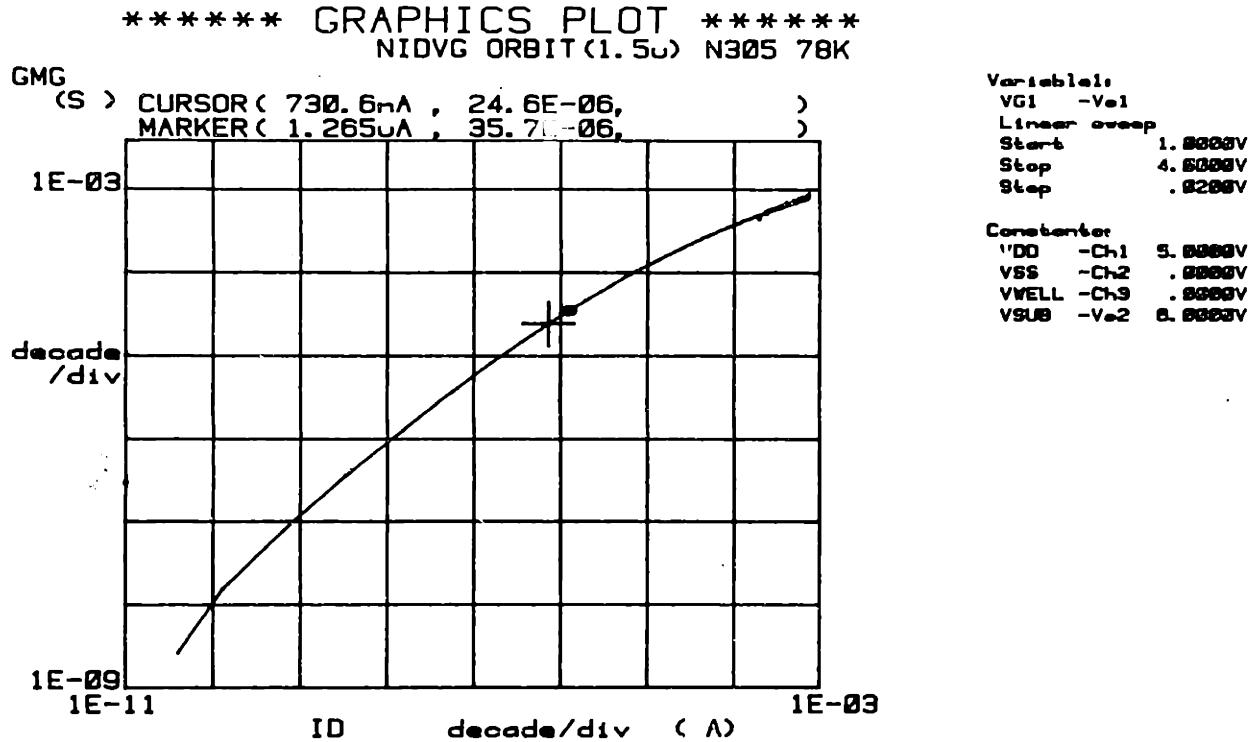


Fig. 4.14 Transconductance plot N 30x5 (78 ^{o}K)

Figure 4.15 shows the corresponding plot for the 90 x 4 P channel FET. In this case the device is operating more into the weak inversion region and its measured transconductance at 1 μA is approximately 29 μS. Strong inversion parameters would yield a value of 60 μS.

With these values the amplifier transconductance is predicted to be

$$G_m = B \cdot (g_m(1An) + g_m(1Ap))$$

$$= 10 \cdot (29 + 30) \mu S = 590 \ \mu S \qquad (4.13)$$

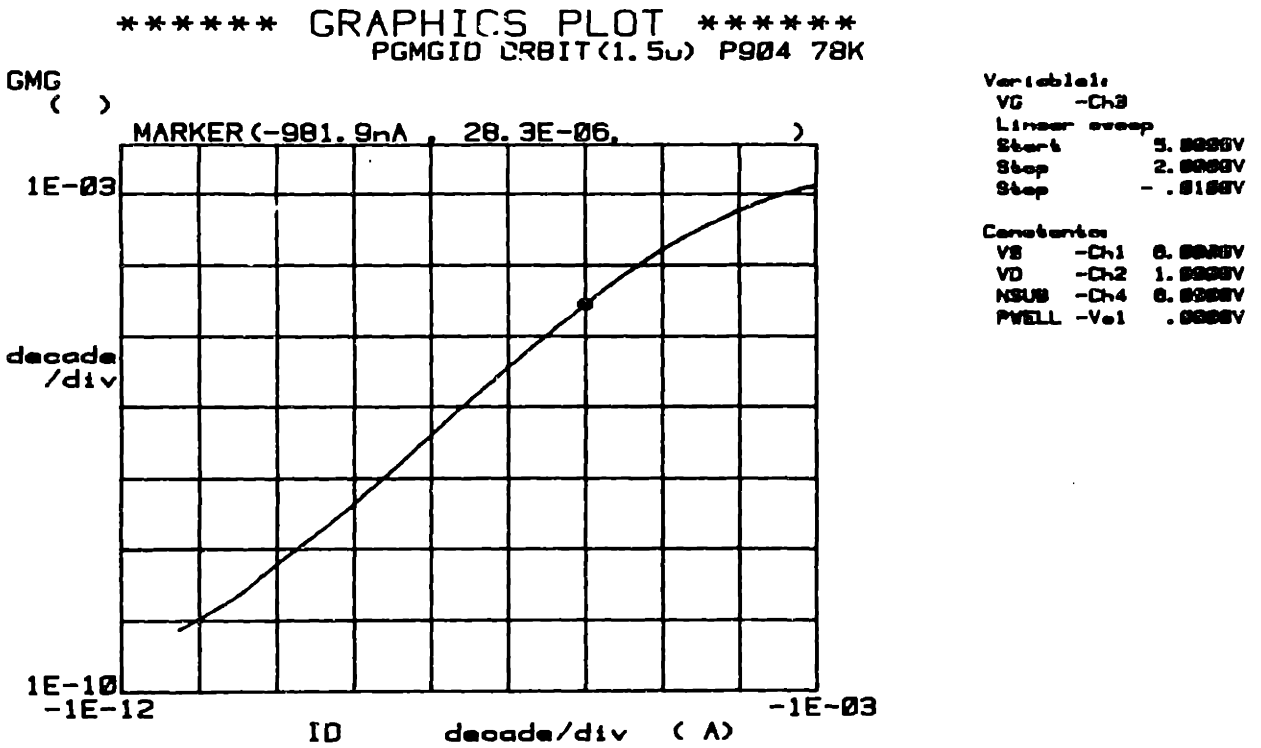
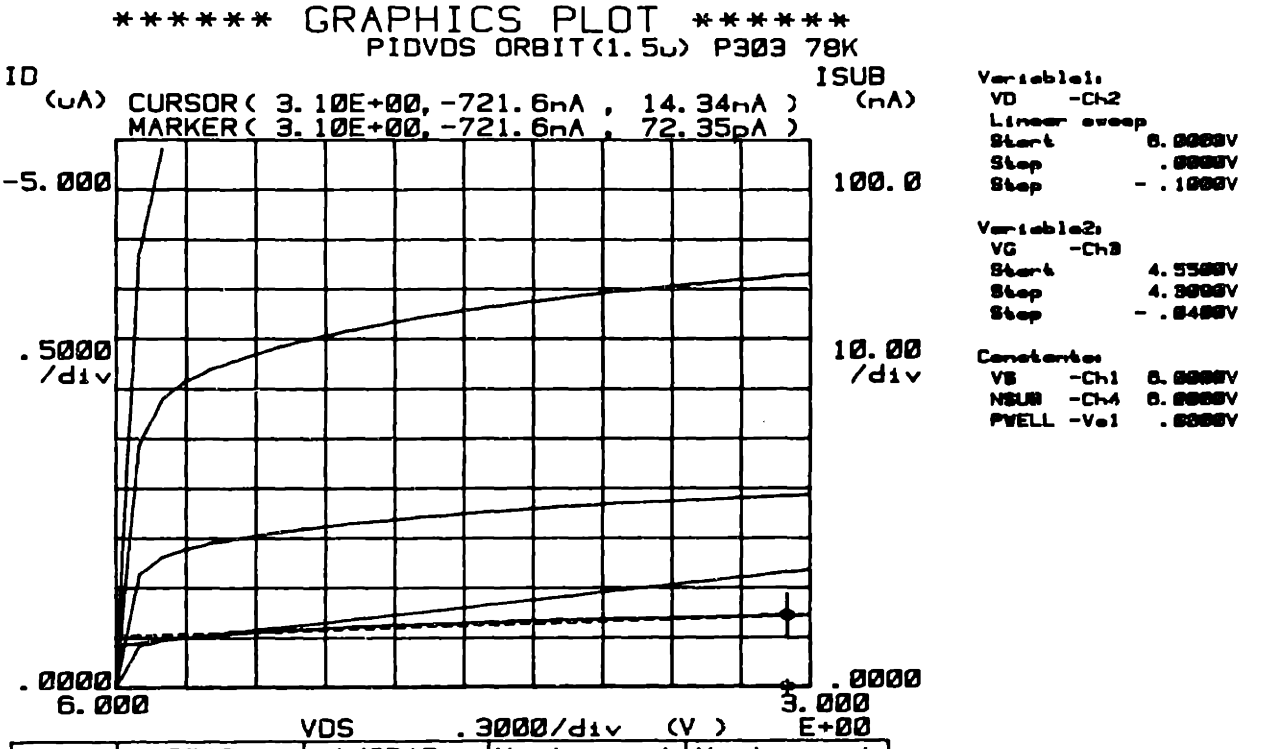


Fig. 4.15 Transconductance plot P 90x4 (78°K)

4.3.2 Output conductance

The amplifier output conductance is given in equation (2.24). One issue that should be taken into account in this expression is that the output conductance of the driver transistors could be increased due to low drain to source voltage. As mentioned in section 4.1.4 this is a problem with the p channel transistors. The drain of the 360x3 P channel output transistor will be approximately .29 volts below V_{DD}. A significant increase in output conductance

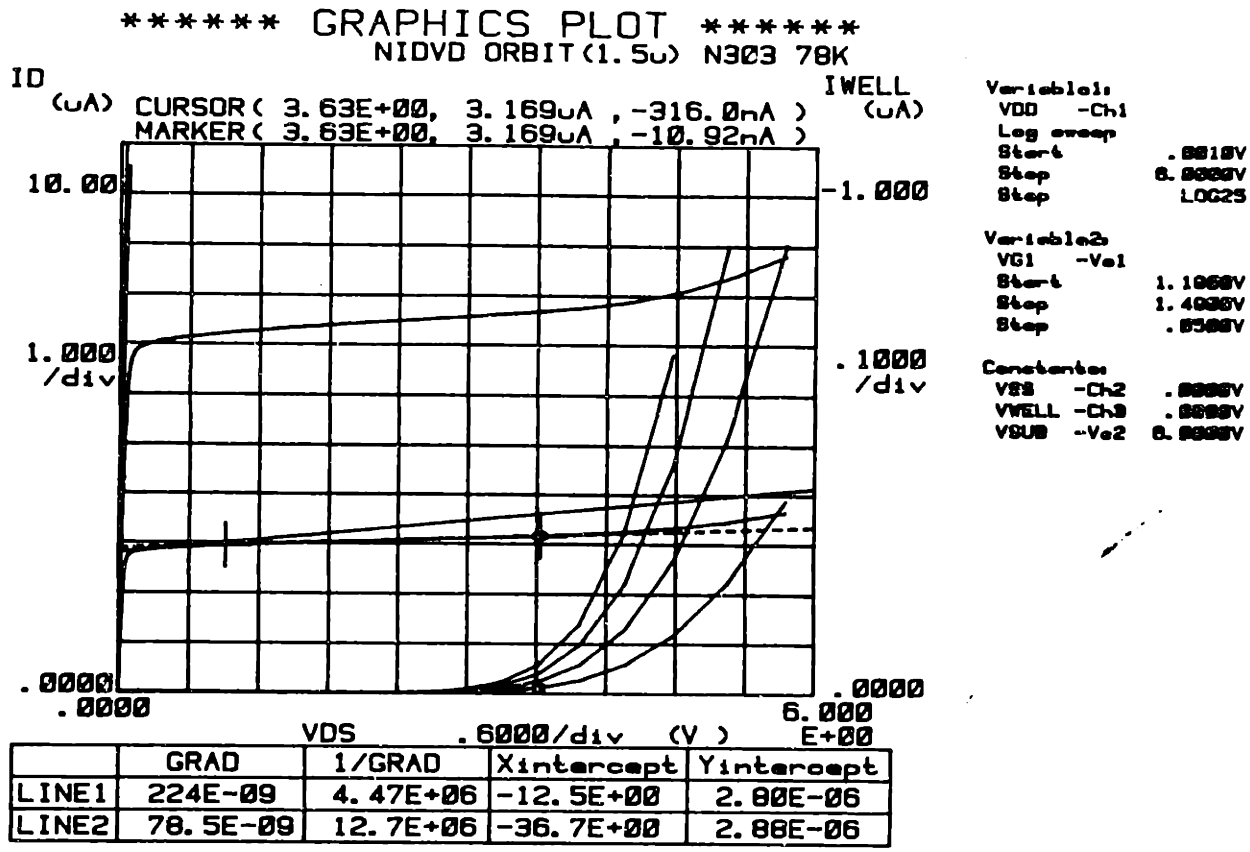
over the mid range value could result. Figure 4.16 shows the output characteristic of a 30x3 P channel FET from the device test chip.



	GRAD	1/GRAD	Xintercept	Yintercept
LINE1	253E-09	3.95E+06	7.69E+00	-1.94E-06
LINE2	70.2E-09	14.2E+06	13.4E+00	-939E-09

Fig 4.16 Output characteristic 30x3 P (78°K)

Using current scaling with device width, the output characteristic of this 30x3 at .8 μA will be a scaled version of a 360x3 at 10 μA . The measured output conductance at Vds = -.29 V is seen to be about 4 times larger than the mid range value. Figure 4.17 shows the output characteristic of a 30x3 N channel FET from the device test chip.



	GRAD	1/GRAD	Xintercept	Yintercept
LINE1	224E-09	4.47E+06	-12.5E+00	2.80E-06
LINE2	78.5E-09	12.7E+06	-36.7E+00	2.88E-06

Fig 4.17 Output characteristic 30x3 N(78°K)

Using current scaling with device width, the output characteristic of this 30x3 at 2.5 μA will be a scaled version of a 120x3 at 10 μA . The drain of the 120x3 N channel output transistor will be approximately .3 volts above VSS. The measured output conductance at Vds = .3V is seen to be about 3 times larger than the mid range value. The current scaling relation used here can be expressed as

$$go(K \cdot W, L, K \cdot I) = K \cdot go(W, L, I) \qquad (4.14)$$

where K is a scale factor. This expression is very general, and relies only on the total current through devices connected in parallel being equal to the sum of the

currents through each device. Edge effects are neglected, so that both devices being compared should be wide enough so that narrow width effects are negligible. Equation (4.14) applies equally well to transconductance, and the 30x3 N and P channel devices transconductance curves can be used to show that transistors M4A, M4B, M4C, and M6 are all operating in the moderate inversion region.

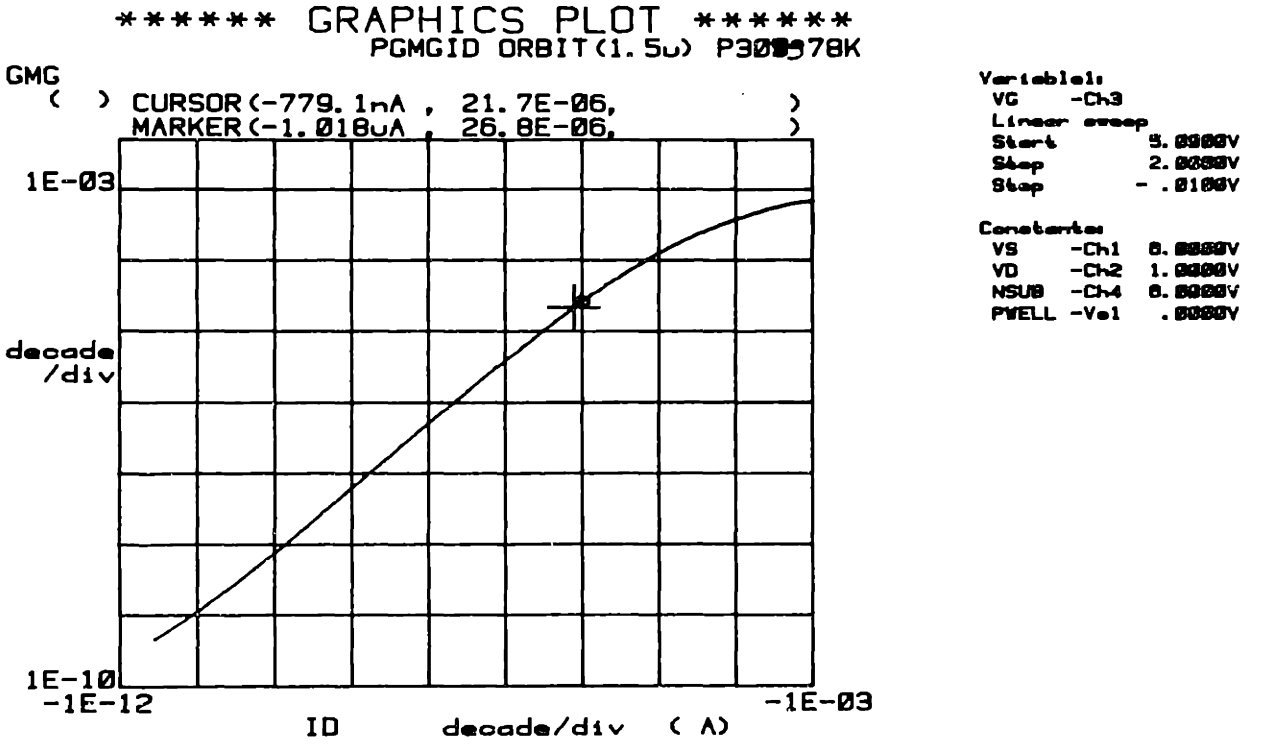


Fig. 4.18 Transconductance P 30x3 (78°K)

Figure 4.18 shows the 30x3 P channel transconductance at .83 μA. This value is .83 times a 36x3 transconductance. The P channel transconductances (M4Ap, M4Bp,

M4Cp, M6p) are only 63% of their strong inversion predictions. The corresponding reduction for the corresponding N channel devices is 83%.

With these corrections the output conductance is given for mid range output voltages as

$$G_o = g_o(4Cn)\cdot g_o(6n)/g_m(6n) + g_o(4Cp)\cdot g_o(6p)/g_m(6p)$$

$$= 7.6e\text{-}10S + 1.5e\text{-}8S = 1.6e\text{-}8S = .016\ \mu S \qquad (4.15)$$

The output conductance is dominated by the P channel devices. The dc gain of the amplifier is given by equation (2.25)

$$A = G_m/G_o$$

$$= 590\ \mu S / .016\ \mu S = 3.7e4 = 91\ dB \qquad (4.16)$$

4.3.3 Small signal frequency response

The dominant pole of the amplifier is given by equation (2.26).

$$p_1 = G_o/C_L$$

$$= .016\ \mu S/ 100\ pF = 160\ rad/sec = 25\ Hz \qquad (4.17)$$

The zero-value time constant for the input stages is given by equation (2.31)

$$\tau_2 = C_1/(2\cdot g_{m1}) + C_2/(g_{m2}\cdot 2/3) + C_3/g_{m3} + C_4/g_{m4}$$

These values for the two input stages are given by

$$\tau_{2n} = 4.9 \text{ ns} + 27.1 \text{ ns} + 6.7 \text{ ns} + 14.3 \text{ ns} = 53 \text{ ns} \quad (4.18)$$

$$p_{2n} = 19e6 \text{ rad/sec} = 3.0 \text{ MHz} \quad (4.19)$$

$$\tau_{2p} = 11.0 \text{ ns} + 104 \text{ ns} + 2.7 \text{ ns} + 41.5 \text{ ns} = 159 \text{ ns} \quad (4.20)$$

$$p_{2p} = 6.3e6 \text{ rad/sec} = 1.0 \text{ MHz} \quad (4.21)$$

In both subcircuits the time constant is dominated by the current mirror load. The P channel input stage is slower by a factor of three than the N channel input stage and it sets the nondominant pole at approximately 1.0 MHz. The gain from the N and P channel input stages are given by

$$A_n = B \cdot g_m(1An) / G_o$$

$$= 18 e3 \quad (4.22)$$

$$A_p = B \cdot g_m(1Ap) / G_o$$

$$= 19 e3 \quad (4.23)$$

Equation (2.33) indicates that the pole locations are at 25 Hz, 1.0MHz, and 3.3 MHz. The zero is given by equation (2.34) and is given by

$$z = (A_n \cdot p_p + A_p \cdot p_n) / (A_n + A_p)$$

$$= 2.0 \text{ MHz} \qquad (4.24)$$

The gain bandwidth of the amplifier is given by equation (2.35).

$$GBW = G_m / C_L$$

$$= 590 \mu S / 100 \text{ pF} = 5.9e6 \text{ rad/sec} = 940 \text{ kHz} \qquad (4.25)$$

The differential input capacitance is given by equation (2.35b).

$$C_{in} = C_{gs}(1n) + C_{gs}(1p)$$

$$= .14 \text{ pF} + .41 \text{ pF} = .55 \text{ pF} \qquad (4.26)$$

4.3.4 Large signal performance

Equations (2.37) and (2.39) give the operating current at which current limiting conditions arise in the N channel input stage. Using the parameters developed above these currents are

$$I(An) = 104 \mu A \text{ (no back gate correction)} \qquad (4.27)$$

$$I(Bn) = 27 \mu A \text{ (no back gate correction)}$$

$$= 21 \mu A \text{ (back gate correction)} \qquad (4.28)$$

According to equation (2.47) the maximum negative output current is determined by the smaller of these two.

$$I- = B \cdot MIN (I(An), I(Bn))$$

$$= 210\ \mu A \qquad (4.29)$$

The current limiting conditions for the P channel input stage are given by equations (2.45) and (2.46).

$$I(Ap) = 99\ \mu A \text{ (no back gate correction)} \qquad (4.30)$$

$$I(Bp) = 20\ \mu A \text{ (no back gate correction)}$$

$$= 14\ \mu A \text{ (estimated back gate correction)} \qquad (4.31)$$

The back gate correction is estimated in (4.31) because this data could not be measured on the device test chip. According to equation (2.48) the maximum positive output current is determined by the smaller of these two.

$$I+ = B \cdot MIN (I(Ap), I(Bp))$$

$$= 140\ \mu A \qquad (4.32)$$

4.3.5 Noise performance

The expressions (2.65b) and (2.65c) give the input referred thermal and 1/f noise powers for the input stages. For the N channel input stage these are

$$v_{eq}^2(th,n) = v_i^2(M1n,th)\cdot(2 + 2\cdot 8\mu S/29\mu S + 1.1\cdot 28\mu S/29\mu S + 5\mu S/29\mu S)$$

$$= v_i^2(M1n,th) \cdot (2 + .6 + 1.1 + .2)$$

$$= v_i^2(M1n,th) \cdot 3.9 \qquad (4.33)$$

$$v_{eq}^2(1/f,n) = v_i^2(M1n,1/f)\cdot(2 + 2\cdot(8\mu S/29\mu S)^2\cdot(1/2)\cdot(150/50) +$$

$$(28\mu s/29\mu s)^2\cdot(150/36 + 150/360) + (5\mu S/29\mu S)^2\cdot(1/2)\cdot(150/150)$$

$$= v_i^2(M1n,1/f) \cdot (2 + .2 + 4.3 + .01)$$

$$= v_i^2(M1n,1/f) \cdot 6.5 \qquad (4.34)$$

The input referred thermal noise has a significant contribution from M4 and the input referred 1/f noise is dominated by M4. The corresponding expressions for the P channel input stage are

$$v_{eq}^2(th,p) = v_i^2(M1p,th)\cdot(2 + 2\cdot 6\mu S/30\mu S + 1.1\cdot 27\mu S/30\mu S + 7\mu S/30\mu S)$$

$$= v_i^2(M1p,th) \cdot (2 + .4 + 1.0 + .2)$$

$$= v_i^2(M1p,th) \cdot 3.6 \qquad (4.35)$$

$$v_{eq}^2(1/f,p) = v_i^2(M1p,1/f)\cdot(2 + 2\cdot(6\mu S/30\mu S)^2\cdot(2/1)\cdot(360/150) +$$

$$(27\mu S/30\mu S)^2\cdot(360/108 + 360/1080) + (7\mu S/30\mu S)^2\cdot(2/1)\cdot(360/150)$$

$$= v_i^2(M1p,1/f) \cdot (2 + .4 + 3.0 + .3)$$

$$= v_i^2(M1p,1/f) \cdot 5.7 \qquad (4.36)$$

As with the N channel input stage M4 makes a significant contribution to the input referred thermal noise and dominates the input referred 1/f noise. When these noise sources are referred to the input of the OTA, they must be reduced by approximately 1/4 as indicated in equation (2.66). The input referred thermal noise power is given by

$$v_{eq}^2(th) = .24 \cdot v_{eq}^2(th,n) + .26 \cdot v_{eq}^2(th,p)$$

$$= .94 \cdot v_i^2(M1n,th) + .94 \cdot v_i^2(M1p,th)$$

$$= 840\ nV^2/hz + 790\ nV^2/hz = 1630\ nV^2/Hz$$

$$veq(th)_{RMS} = 40\ nV/\sqrt{Hz} \qquad (4.37)$$

The surprisingly simple result is approximately the sum of the thermal noise power one input transistor from each input stage. The thermal noise contributions of each input stage are approximately equal, due to their approximately equal transconductance . The input referred 1/f noise power is given by

$$v_{eq}^2(1/f) = .24 \cdot v_{eq}^2(1/f,n) + .26 \cdot v_{eq}^2(1/f,p)$$

$$= 1.6 \cdot v_i^2(M1n,1/f) + 1.5 \cdot v_i^2(M1p,1/f)$$

$$= 2.1/f\ \mu V^2 + .2/f\ \mu V^2$$

$$= 2.3/f\ \mu V^2$$

$$veq(1/f)_{RMS} = 1.5/\sqrt{f}\ \mu V \qquad (4.38)$$

This expression is dominated by the 1/f noise contributions of the n channel input stage.

4.4 Buffer Performance

4.4.1 Frequency Response

On the circuit test chip output buffer, the input bus to the amplifier is reduced to just two elements for simplicity. Thus only two channels out of the potential 200 channels on the AIRS focal plane are simulated. Each channel has a 3 pf capacitor whose voltage can be independently set externally. Figure 4.13 shows how these channels are configured. An additional source diffusion of size equal to 200 switch diffusions was placed on the bus to simulate the bus capacitance of 200 switches. A capacitance to simulate the metal bus run capacitance of 1.7 pf was overlooked. The test circuit output buffer will then operate with a feedback fraction, given in equation 2.60, of

$$f = 3 / (3 + 3 + .50 + .55) = .43 \qquad (4.39)$$

rather than .34, which will be approximately the value in the AIRS focal plane. Equation (2.66) gives the feedback fraction at which the system will be approximately critically damped. Using the results derived above this value for f is .29. On the AIRS focal plane the buffer is expected to be slightly underdamped as expected. The test chip circuit should be somewhat more underdamped. Equation (2.62) gives a prediction for the closed loop pole location for a single pole model.

$$p = GBW \cdot f = 404 \text{ kHz} \qquad (4.40)$$

This model predicts a first order response with time constant of 390 ns. Small signal settling time to .03% is predicted to be 3.2 μs with a one pole model. Equation (2.64) gives predictions for the closed loop pole location for a two pole model.

$$\omega_n = p_2/2 = 550 \text{ kHz}$$

$$\zeta = \sqrt{ (p_2 / (4 \cdot a \cdot p_1 \cdot f)) } = .82 \qquad (4.41)$$

These parameters predict a second order response with time constant of 290 ns and essentially no overshoot . Small signal settling time to .03% is predicted to be approximately 3.2 μs with a two pole model. The effect of higher order singularities will slow things down somewhat.

The buffer output impedance at low frequencies is given by equation (3.9).

$$R_o = 1/(G_m \cdot f)$$

$$= 1/(590\ \mu S \cdot .43) = 3.9\ k\Omega \qquad (4.42)$$

4.4.2 Large signal performance

Output voltage slew rates are determined by the maximum output currents given in equations (4.28) and (4.31).

$$SR+ = I+ / C_L = 140\ \mu A / 100\ pF = 1.4\ V/\mu s$$

$$SR- = I- / C_L = 210\ \mu A / 100\ pF = 2.1\ V / \mu s \qquad (4.43)$$

The slew rate, if large enough, will not have an effect on the large signal settling time. In this case, an estimate of the effect of the slew rate may be made by considering the initial slope in a first order step response. It is

$$(dV/dt)\ max = \Delta V/ \tau \ . \qquad (4.44)$$

Using a step size of 3V and a time constant of 300 ns, this initial slope would be 10 V/ μs. It is fair to say that the slew rates obtained here will retard the large signal settling time. The Class AB nature of this amplifier makes the response nonlinear even in the absence of any slew limitations. Estimates of large signal settling times are best left to numerical simulations.

4.4.2 Noise Performance

4.4.2.1 Amplifier input referred voltage noise

Equation (2.80) gives the mean square output voltage contribution from the input referred voltage noise of the amplifier. The direct noise contribution to the output voltage power spectra can be approximated as

$$V_{out}^2(direct,f) = NG^2 \cdot V_{eq}^2(f) \qquad f < NBW$$

$$= 0 \qquad f > NBW \qquad (4.45)$$

The direct noise contribution to the mean square output voltage is obtained by integrating equation (4.41) over frequency.

$$< V_{out}^2(direct,t) > = NG^2 \cdot \{ NBW \cdot V_{eq}^2(f,thermal) + \ln(f_H/f_L) \cdot V_{eq}^2(f,1/f@1hz) \}$$

$$= 5.6 \cdot \{ 550\ kHz \cdot 1630\ nV^2/Hz + \ln(550\ kHz/.3\ Hz) \cdot 2.3\ \mu V^2/Hz \}$$

$$= 5.6 \cdot \{ 897\ \mu V^2 + 33\ \mu V^2 \} = 5.6 \cdot 900\ \mu V^2$$

$$= 5040\ \mu V^2 \qquad (4.46)$$

The thermal noise contribution is the dominant contributor. The sampled noise contribution to the mean square output voltage is given from the second term in equation (2.80).

$$< V_{out}^2(sampled,t) > = Cs^2/Cf^2 \cdot \{ NBW \cdot V_{eq}^2(f,thermal) + \ln(f_H/f_L) \cdot V_{eq}^2(f,1/f@1Hz) \}$$

$$= (1.05pf/3pf)^2 \cdot 900\ \mu V^2 = 110\ \mu V^2 \qquad (4.47)$$

The sampled noise contribution is seen to be negligible by comparison with the direct contribution.

4.4.2.2 Switch Elements

The direct contribution of the switch elements to the mean square output voltage is given by equation (2.85).

$$< v_{out}^2(direct,t) > = NBW \cdot \{ Cf^2/Ci^2 \cdot 4 \cdot k \cdot T \cdot R_1 + 4 \cdot k \cdot T \cdot R_2 + 4 \cdot k \cdot T \cdot R_3 \}$$

$$= 550\ kHz \cdot \{ 30\ nV^2/Hz + 30\ nV^2/Hz + 30\ nV^2/Hz \}$$

$$= 50\ \mu V^2 \qquad (4.48)$$

A switch resistance of 7 kΩ has been used, which is consistent with the switch dimensions of 3.5x3 and extracted device parameters. The resistance of switch three changes with output voltage, but this is a minor effect because a complementary switch is used at that point. The sampled noise contribution of the switch elements to the mean square output voltage is given by equations (2.83) and (2.84).

$$< v_{out}^2(sampled,t) > = k \cdot T / C_i \cdot C_f/C_i + k \cdot T / C_f$$

$$= 360\ \mu V^2 + 360\ \mu V^2 = 720\ \mu V^2 \qquad (4.49)$$

The total mean square output voltage from all sources is projected to be

$$< vout^2(t) > = (5040 + 110 + 50 + 720) \ \mu V^2 = 5920 \ \mu V^2$$

$$vout_{RMS} = 77 \ \mu V \qquad (4.50)$$

4.4.3 Linearity

The linearity of the output buffer is expected to be quite good because of the feedback configuration. One source of nonlinearity which is not improved by the feedback connection is the nonlinearity of the feedback capacitor itself. In this case a poly-poly capacitor is used. No testing has been done for this process to quantify the linearity of the poly-poly capacitor. If 100 ppm/V were assumed, then this contribution would only be .01% of differential nonlinearity. The specification is listed as .1% integral nonlinearity. Since a measurement of integral nonlinearity is always less than differential nonlinearity[32], it is unlikely that the poly -poly capacitor will cause significant nonlinearity.

On this test chip there is another source of nonlinearity which is likely to be larger. The inputs to the output buffer is charge which has been stored on input capacitors. These capacitors, however, have been charged to fixed voltage by external voltage sources for testing convenience. The nonlinear capacitance associated with the switch elements to these input capacitors will produce a nonlinearity at the output. There are three such switch elements. Two are 3.5x3 N channel switches and one is a 3.5x3 P channel switch. Using diffusion capacitance data provided in Orbit SPICE models, the total switch capacitance is estimated at about 13 ff @ 6 volts, 20 ff @ 3 volts and 27 ff @ 0 volts. This suggests

that the differential gain will differ by about ± .2 % over the entire output range. The integral nonlinearity will be at least a factor of two less than this but no simple relationship exists for its value. This should be the dominant nonlinearity it the test measurements.

4.4.4 Crosstalk

The crosstalk due to incomplete bus settling is given by equation (3.39) as is approximately as

$$\text{Crosstalk(incomplete bus settling)} = \text{Cbus/Cf} \cdot (1/2 \cdot e^{-T/\tau} + 1/a)$$

$$= 1.05/3.0 \cdot (2e\text{-}6 + 3e\text{-}5)$$

$$= -99 \text{ dB} \qquad (4.51)$$

400 ns has been assumed as the time constant. The crosstalk is dominated by finite OTA gain but is much better than is probably needed.

5. EXPERIMENTAL RESULTS

The output buffer was laid out as part of a circuits test chip for the AIRS program. Layout for the circuits test chip was performed by Advance Design Services (ADS) of Minneapolis, MN. Fabrication was performed by Orbit Semiconductor of Santa Clara, CA. Circuit testing was performed at LORAL's Infrared and Imaging Systems Focal Plane Array Laboratory.

5.1 Layout

The current mirror M4A, M4B, M4C was identified as a critical layout issue because, as discussed in section 2.1.4, these transistors are operating with a small (Vgs - VT) and will be sensitive to threshold mismatchs. The circuit which provides the output voltage reference is laid out in an identical fashion as the circuit which provides the output signal so that good common mode rejection is achieved. A plot of the OTA layout is shown in figure 5.1.

5.2 Test Station and support electronics

Figure 5.2 shows the major components of the test station used for testing the output buffer. The test chip was located in a cryogenic dewar to maintain its temperature. All testing reported here was performed under open well liquid nitrogen conditions. All signals to and from the test chip are brought through a bulkhead in the dewar through coaxial cable. The shields of these cable are attached to a ground ring inside the dewar which is held at system ground. Approximately 70 pf of capacitance to ground is introduced by the coaxial shields. The major pieces of equipment used to make the measurements reported here are

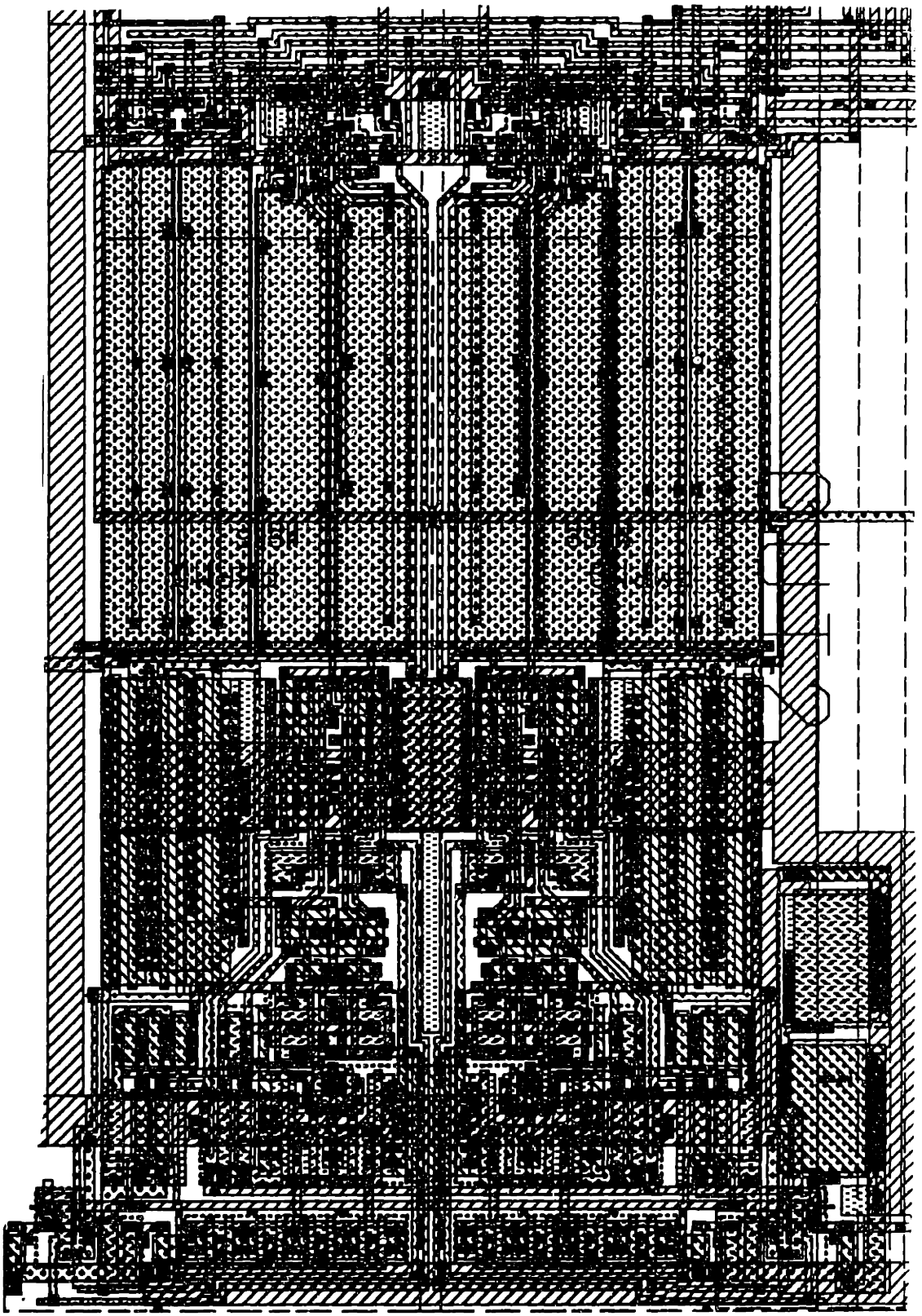


Fig 5.1 OTA layout

an oscilloscope, a network analyzer, a semiconductor parameter analyzer, a 16 bit data acquisition system , and a dynamic signal analyzer. The test chip is wirebonded in a 120 pin grid array package mounted in the test dewar.

To control the load capacitance and to eliminate resistive loading an external buffer box was designed. A schematic of this box is included as figure 5.3.

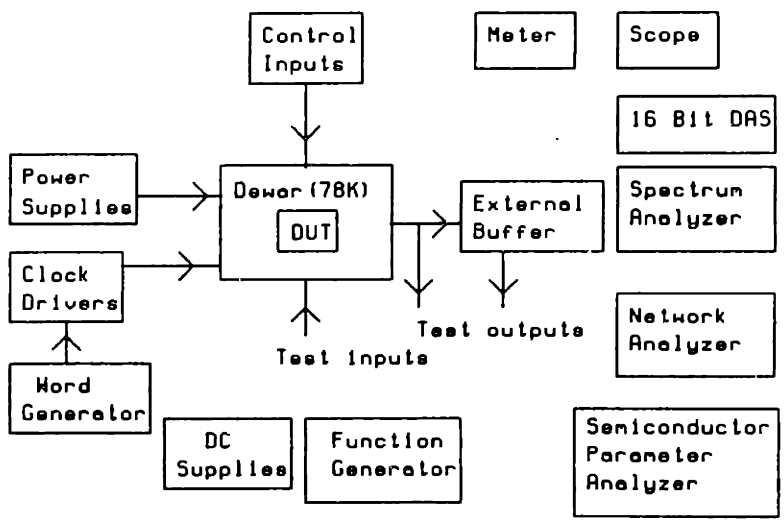


Fig 5.2 Test setup

The LF351 is used because it's low input currents (50 pA), and high input impedance of $10^{12}\ \Omega$ ensure negligible loading on the DUT. The gain bandwidth of 4 MHz is about ten times larger than the closed loop bandwidth of the DUT so frequency response measurements should be undistorted. The slew rate (13 V/μs) is much larger than that expected from the DUT. The noise characteristics (en = 16 nV/$\sqrt{}$ Hz, in = .01 pA/ $\sqrt{}$ Hz) will allow measurement of the expected noise level from the DUT.

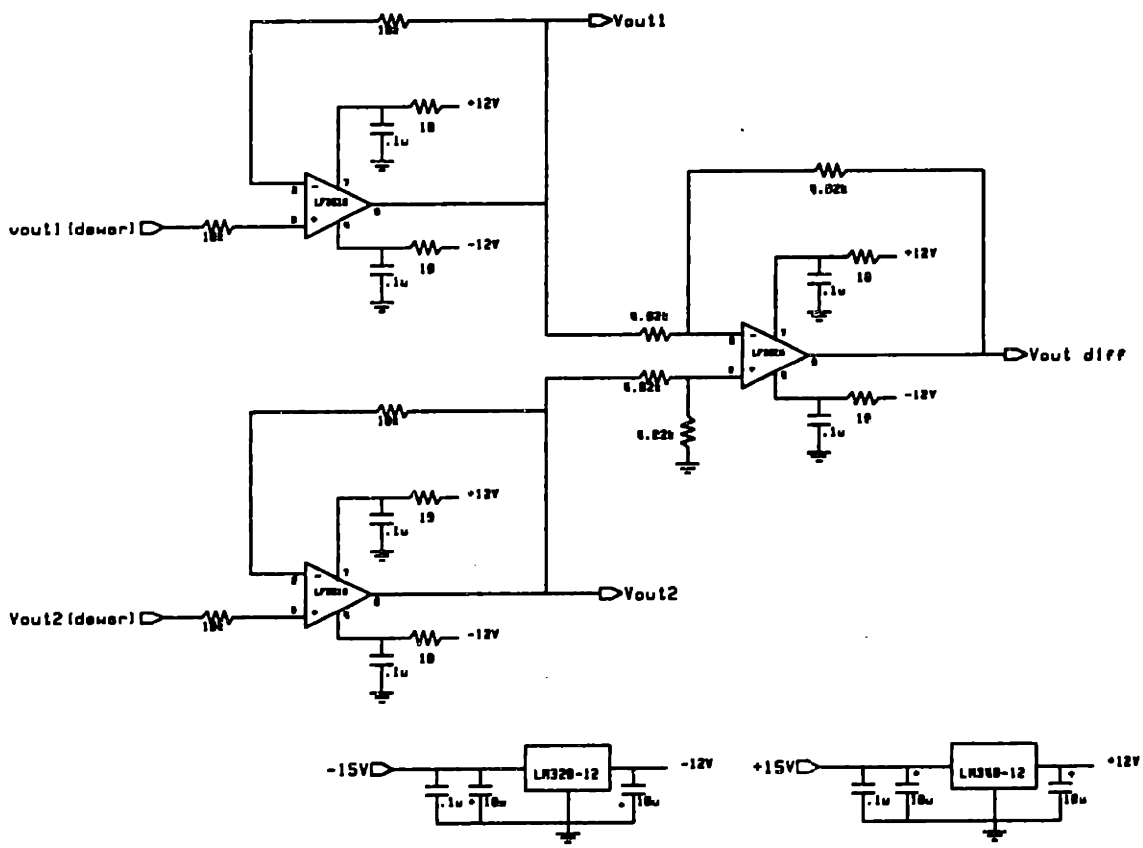


Fig. 5.3 External Buffer

5.3 Amplifier Measurements

Amplifier bias currents were controlled by an external current source which was mirrored and scaled to provide the nominal 1 μA in the input stage current sources. The bias circuitry is shown in figure 5.4. Independent measurement of supply current was not possible on the AIRS circuits test chip, so bias currents are reported as their nominal values. Unless otherwise specified all data presented here is at a nominal bias of 1 μA for the input stage current sources. The temperature, unless otherwise specified, is a nominal value of 78 ^{o}K. Twelve test die were received, ten of which were operational. On the data that follows individual die are identified as part 1, part 2,..., part 12. Tests were performed with VSS = 0 volts and VDD = 6 volts unless otherwise specified.

5.3.1 Transconductance

The test chip is configured so that with proper selection of switches each input of the OTA is brought out of the dewar. It is also possible to switch all feedback elements out of the feedback path. In this way, unrestricted testing of the amplifier is possible. The OTA transconductance was measured using an HP 4145B Semiconductor Parameter Analyzer. The setup is shown in figure 5.4. The external buffer was not used for this measurement. The OTA output was held incrementally grounded at 3 volts. The OTA noninverting input was held at 3 volts. The output current was measured as a function of voltage at the inverting input.

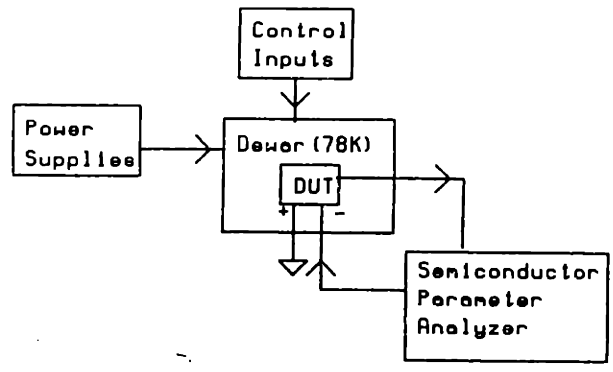
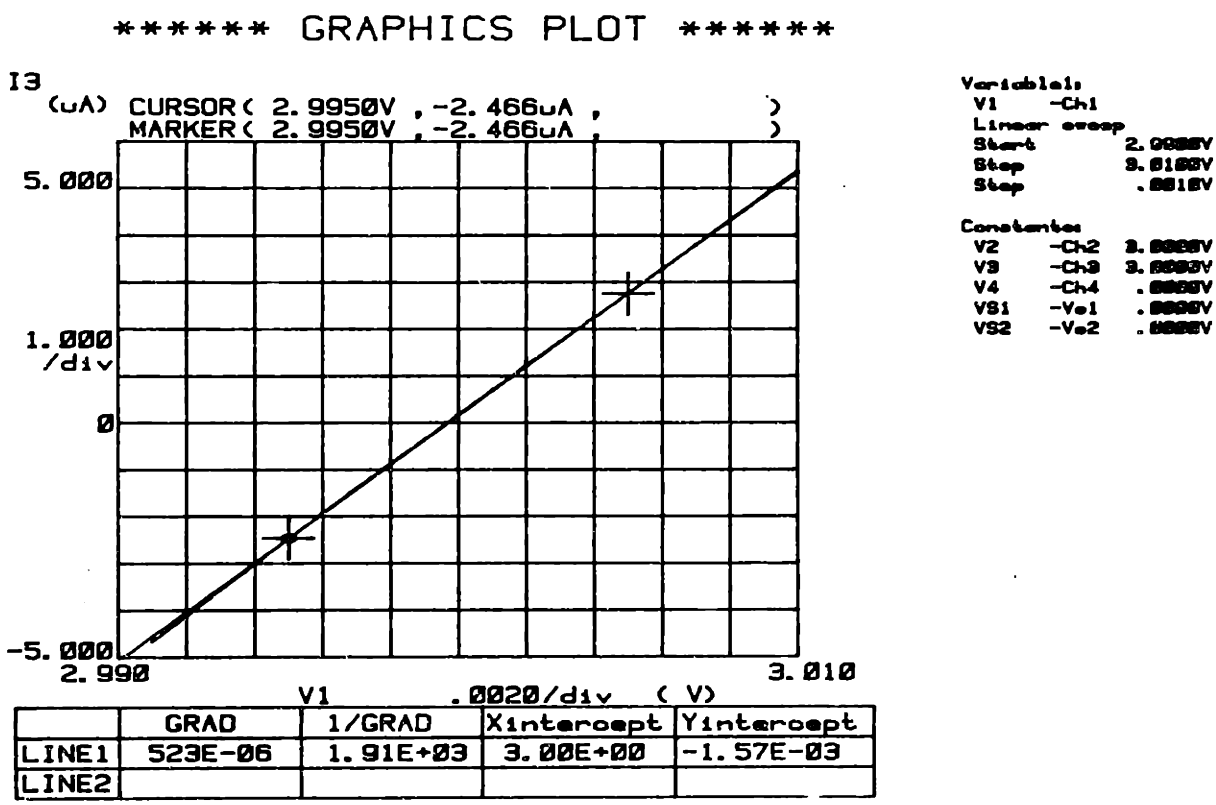


Fig 5.4 Transconductance measurement setup



	GRAD	1/GRAD	Xintercept	Yintercept
LINE1	523E-06	1.91E+03	3.00E+00	-1.57E-03
LINE2				

Fig. 5.5 OTA transconductance Part 8 (78oK)

Figure 5.5 shows a transconductance measurement for part 8. The data is quite linear over the 20 mV region displayed. The measured transconductance of 523 μS

is visible as the grad of line 1. Table 5.1 shows measured values on four different parts

Part	Transconductance
1	313 μS *
2	461 μS
7	505 μS
8	523 μS

Table 5.1 Measured OTA transconductance (78°K)

Subsequent measurements of other properties indicate that the measurement for Part 1 is probably in error. The other results are, on average, 16 % less than the predicted value of 590 μS. The most likely mechanism for the systematically low values would be an error in the bias current generation. This is indeed the case. The current sources in the input stages of the OTA are 5x30. Their bias current of 1 μA is established with a current mirror with a 50x30 device biased at 10 μA. As discussed in Section 4.1.2 the 5x30 devices on the device test chip behaved with an effective electrical width of 3.8 μA. This 24% reduction in bias current could easily account for the 16 % drop in measured transconductance. Possible mechanisms for the variation in measured transconductances include input stage bias current variations, current mirror step up ratio variations, and device variation.

5.3.2 Small signal frequency response

The OTA open loop frequency response was measured with the use of an HP 3577A Network Analyzer. The measurement setup is shown if figure 5.6.

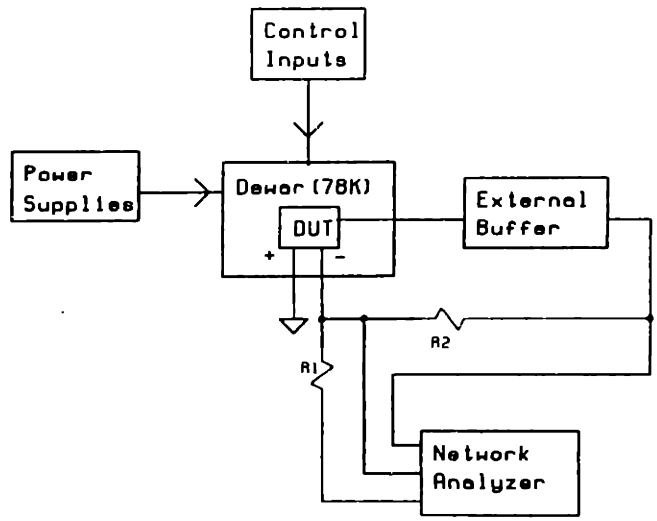


Fig 5.6 OTA open loop frequency response setup

The external buffer is used, and the OTA is configured with resistive feedback using discrete components external to the dewar. The inverting node of the OTA and the output of the external buffer are inputs to the network analyzer. Their ratio define the open loop frequency response of the OTA. Because of the high low frequency gain, an external attenuator is needed for low frequency measurements. Figure 5.7 shows a frequency response measurement for part 8. The output capacitance for this measurement was measured at 97 pF. The dc gain of 91 dB is visible from the marker location on the low frequency section. The three dB frequency is approximately 27 Hz and the unity gain frequency is 905 kHz. The single pole roll off is evident over the majority of the range with a droop of about 1db at the unity gain crossover. These results are in good agreement with predicted values of 91 dB, 25 Hz , and 940 kHz. Measurement of the unity gain frequency of Part 1 with load capacitance of 115 pF was 761 kHz. When the different load capacitance is taken into account, this measurement indicates equivalent transconductance as part 7. This is much higher than would be

predicted from the measured transconductance from Table 5.1 and suggests that the measurement in Table 5.1 was probably in error.

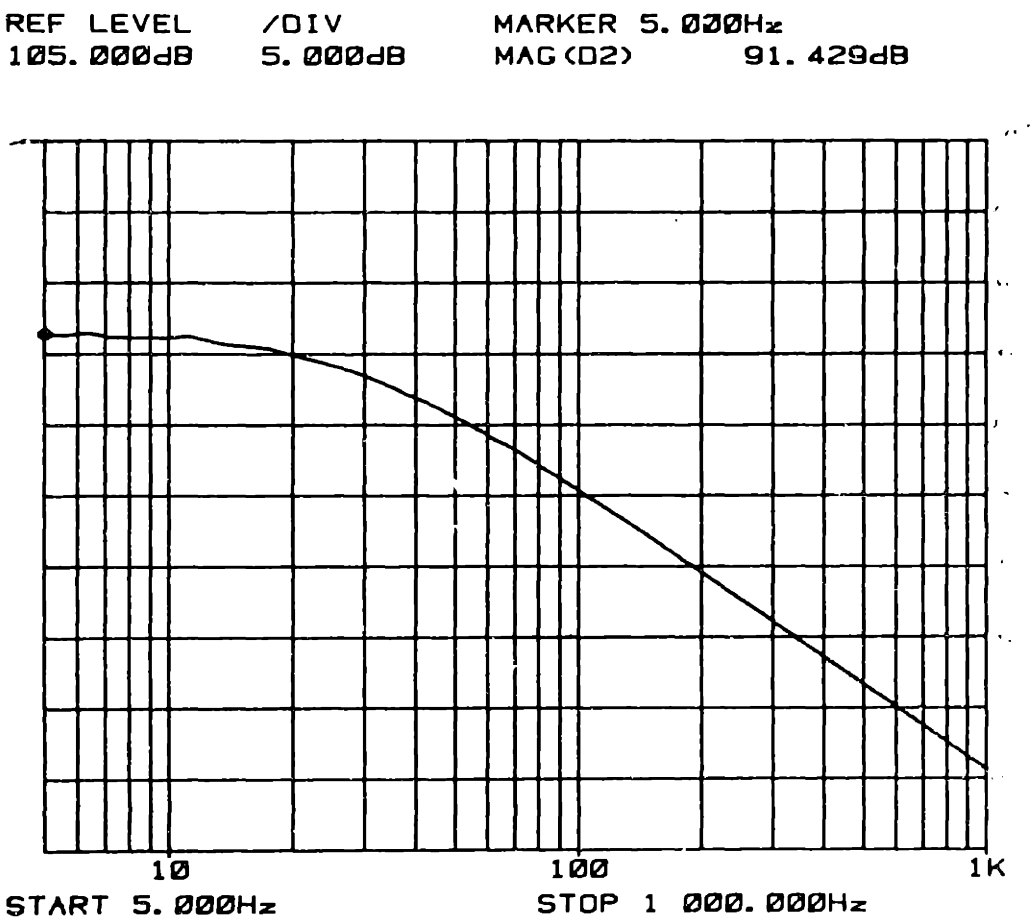


Fig. 5.7a Low frequency open loop frequency response, Part7 (78oK)

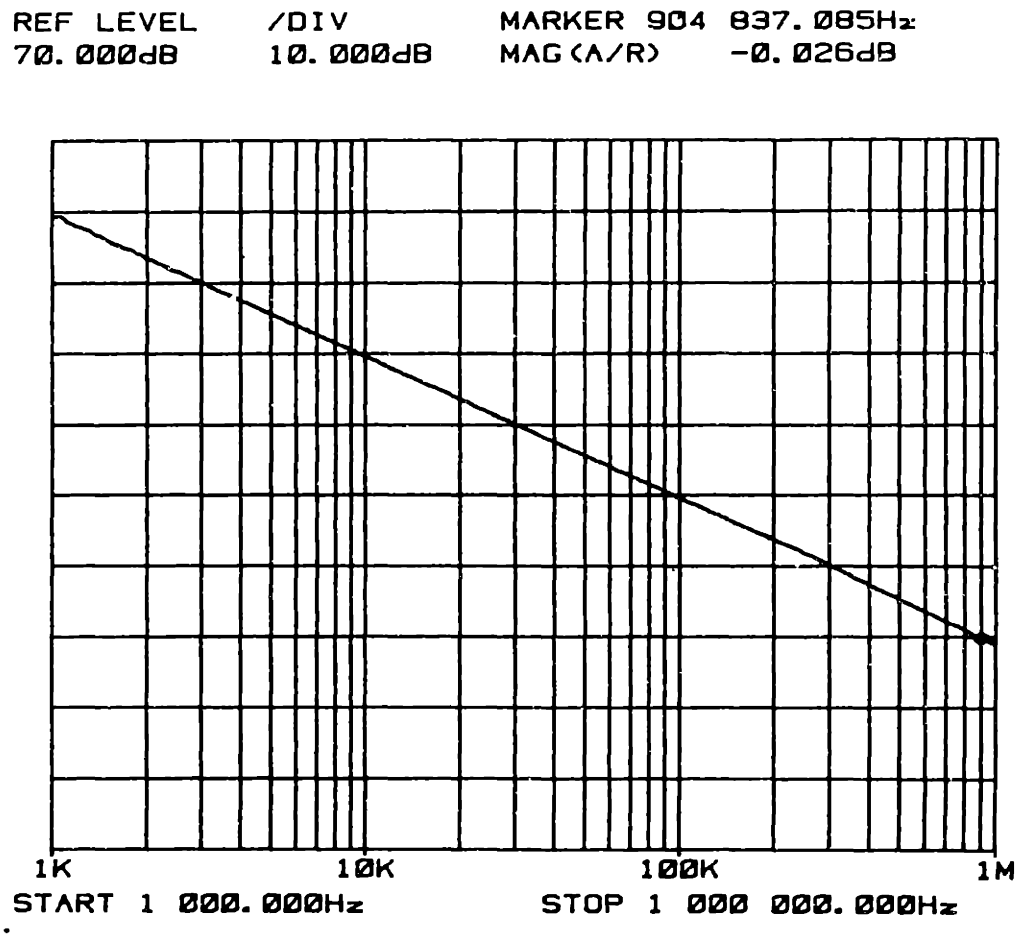


Fig. 5.7b High frequency open loop frequency response, Part7 (78oK)

A variation on the frequency response measurement was made to measure the low frequency gain as a function of the output voltage. The setup for this measurement is shown in figure 5.8. The inverting input to the OTA is here amplified with a voltage divider. With this setup, the gain can be measured with a much smaller output voltage swing [33]. It is important to reduce the output voltage swing as the output voltage approaches the rail.

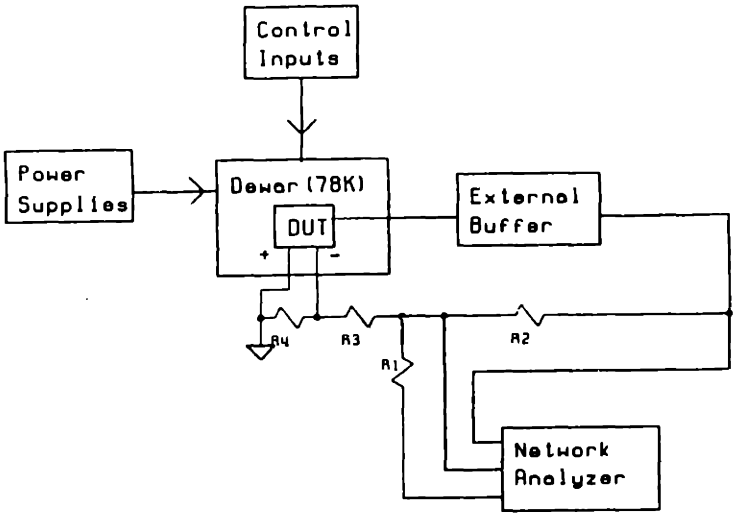


Fig. 5.8 Low frequency gain versus output voltage test setup

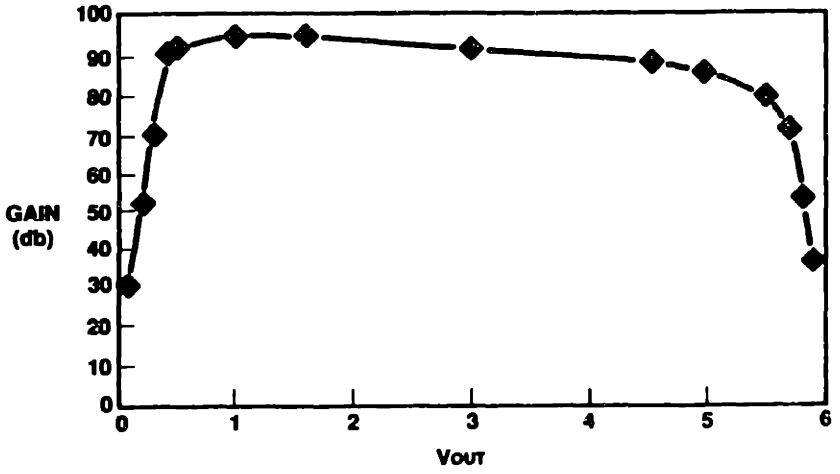


Fig. 5.9 Low frequency gain versus output voltage, Part 7 (78 ^{o}K)

A plot of low frequency gain versus output voltage for Part 7 is shown as figure 5.9. This data shows two interesting features. First, the gain drops rapidly as the output voltage approaches within 3 tenths of a volt of either rail. In this region the cascode transistor goes into the triode region, and the gain drops dramatically. The location of this transition is determined by the cascode bias

voltages, and in this case agrees well with the .3 volt nominal value of drain to source voltage on the driver transistor established in this design. Secondly, the dc gain decreases with increasing output voltage over the output range of the amplifier. This is due to the fact that the output conductance of the n channel cascode increases with increasing drain to source voltage while the output conductance of the p channel cascode transistor decreases with increasing drain to source voltage. As the output voltage increases, the drain to source voltage of the n channel cascode increases, while that of the p channel cascode decreases. This effect was discussed in section 4.1.4

5.3.3 Large Signal Performance

The large signal output current characteristics of the OTA was characterized using the same setup as was used for the transconductance measurements. In this case, however, the input voltage was swept over a much larger voltage range. Figure 5.10 shows the output current of Part 3 for input differential voltages from -200 mV to 200 mV. The second order nature of the current, predicted by equation (2.7) is clearly visible. Due to the polarity of the measurements, negative current in figure 5.10 corresponds to positive output current from the OTA. In this case the P channel input stage shows a larger increase in current for a 200 mV input than the N channel. This is consistent with the larger measured transconductance of the 90x4 P channel FET than the 30x5 N channel. The actual bias currents of the input transistors will show a spread due to matching tolerances of the the M4 current mirror. This spread is large enough so that Part 1 shows larger current in the N channel input stage for a 200 mV input than in the P channel input stage.

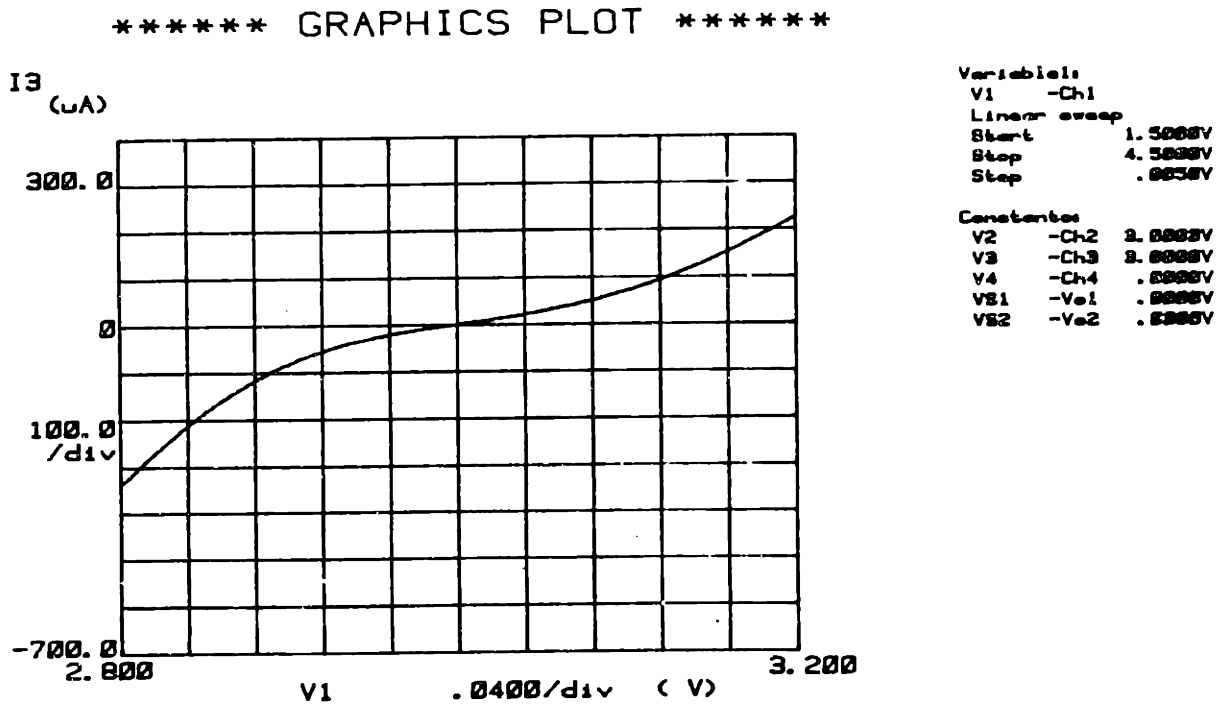


Fig. 5.10 Large signal output current, Part 3 (78 ^{o}K)

At larger still differential input voltages the current saturates. Figure 5.11 shows the output current limits of Part 3. The quadratic increase rolls over into a shallow linear region as the input differential voltage approaches 300 mV. It is clear that the current stops its quadratic increase well before it flattens out.

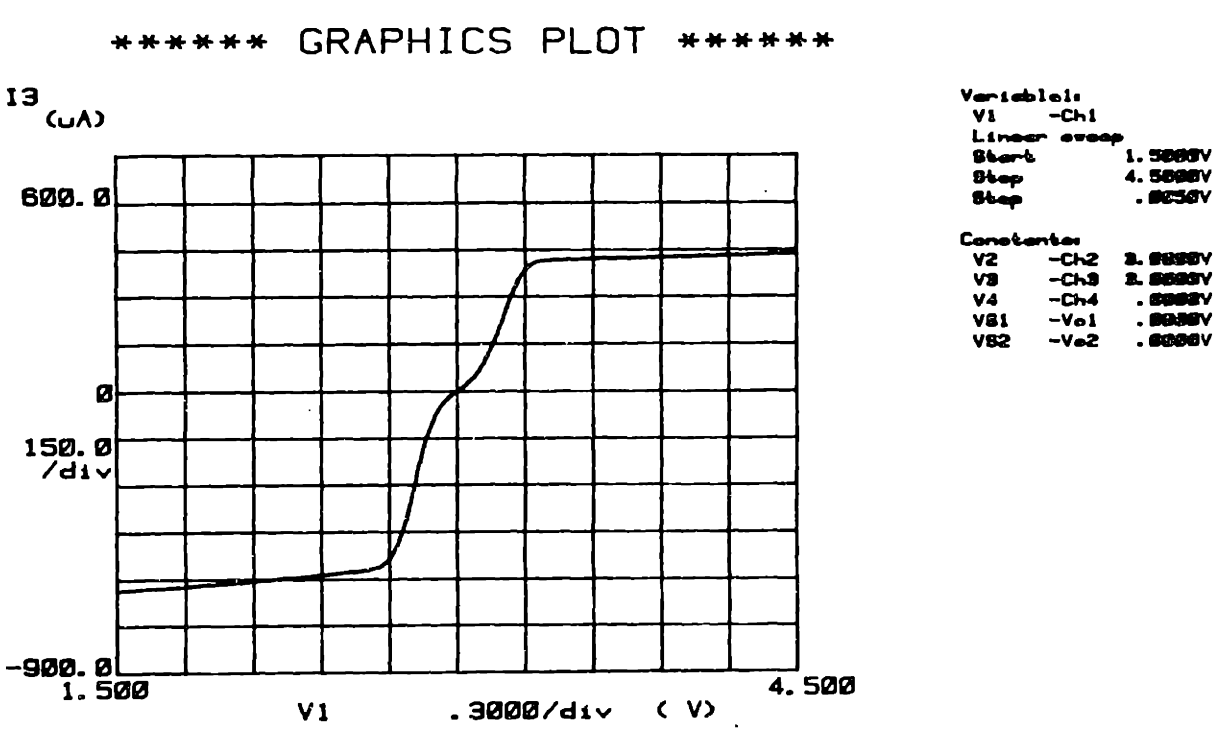
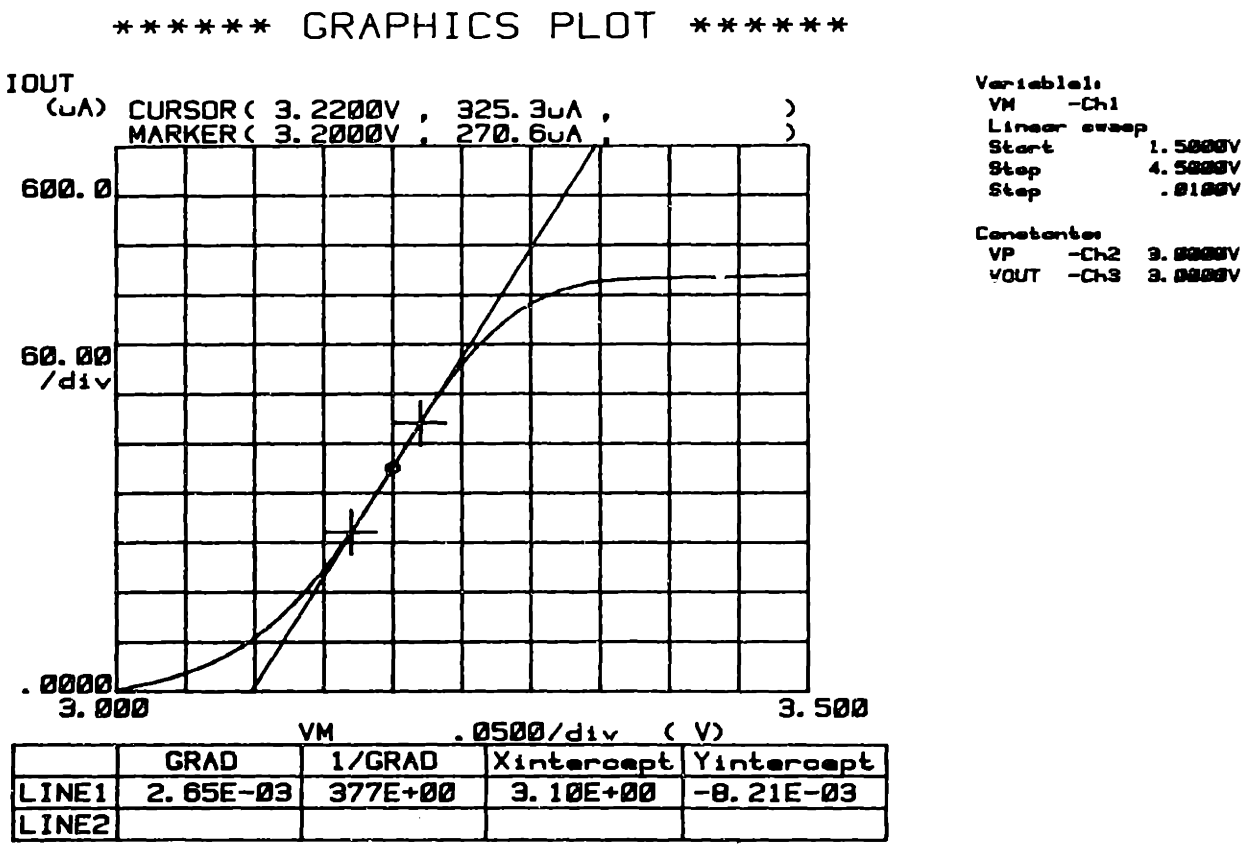


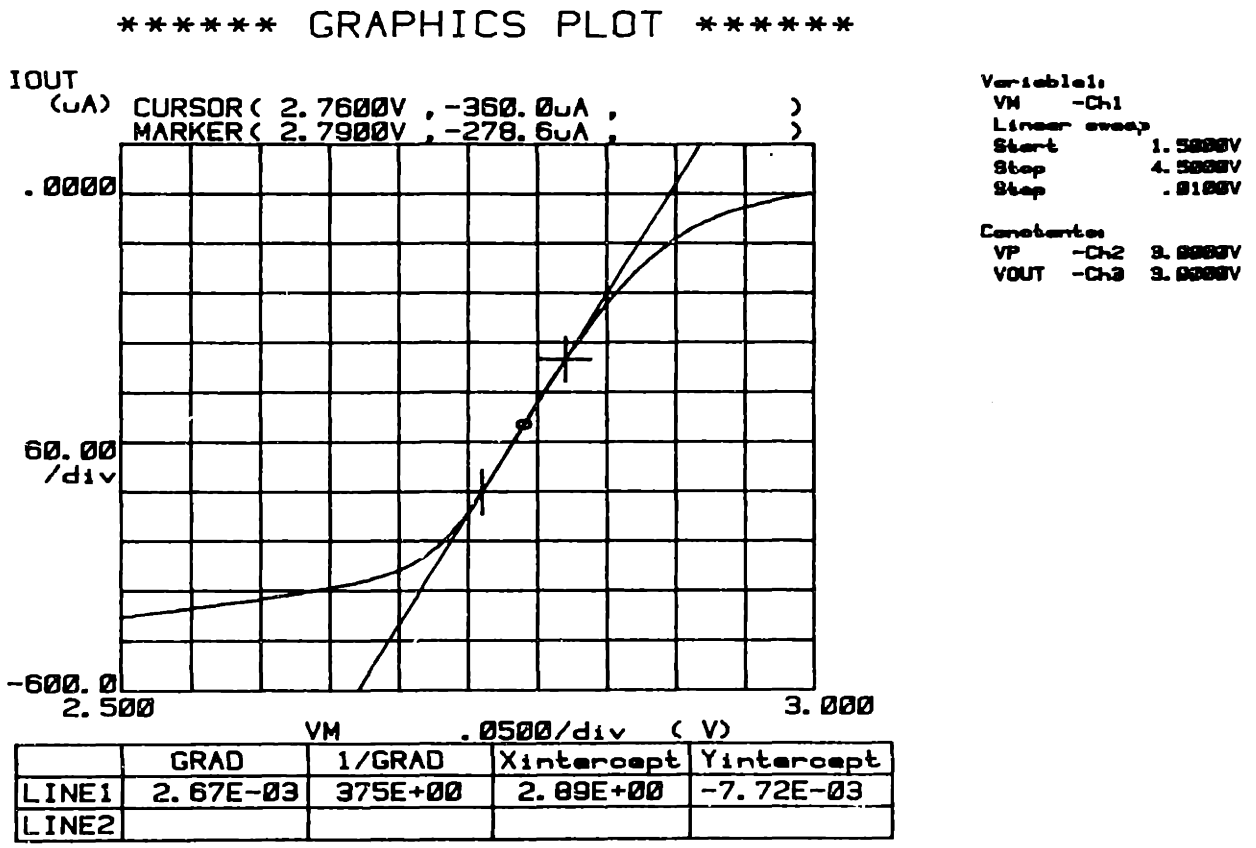
Fig. 5.11 Output current limits, Part 3 (78°K)



	GRAD	1/GRAD	Xintercept	Yintercept
LINE1	2.65E-03	377E+00	3.10E+00	-8.21E-03
LINE2				

Fig. 5.12 Negative output current , Part 1 (78°K)

Figure 5.12 shows the N channel input stage portion of the output current for Part 1. The curve becomes linear at about 200 μA , which is less than 1/2 of the 500 μA that it finally achieves. Equation (4.28) estimated the quadratic increase would stop at a negative output current of 210 μA. Figure 5.13 shows the same result for positive output current.



	GRAD	1/GRAD	Xintercept	Yintercept
LINE1	2.67E-03	375E+00	2.89E+00	-7.72E-03
LINE2				

Fig. 5.13 Positive output current, Part 1 (78oK)

In this case the current becomes linear at approximately 200 μA. Equation (4.31) predicts that the positive current will stop its quadratic increase at about 140 μA. The agreement is good for the negative output current, but a discrepancy exists in the case of the positive output current.

In both cases the maximum output current was much greater than the value at which the current ceases its quadratic increase. The actual maximum output current will depend on the characteristics of the transistors in the triode region. This is best left for numerical simulation but is is clear that accurate modeling in this region of operation would be important to obtain an accurate answer. Table 5.2 shows maximum output current levels for several of the test parts.

Part	IMAX +	IMAX -
1	580 μA	500 μA
2	750 μA	230 μA *
3	600 μA	440 μA
7	800 μA	160 μA *
8	850 μA	420 μA

Table 5.2 Maximum output currents (78°K)

The currents in Table 5.2 has been measured with 1 volt of differential input voltage. The maximum positive output currents are systematically greater than the maximum negative currents and two of the parts show anomalously low negative current. It was observed that the maximum negative output current of parts 2 and 7 was only moderately affected by supply voltage. Current limiting due to transistors entering the triode region should show a strong dependance on supply voltage. It was anomalous behavior of parts 2 and 7 which suggested the mismatch current limit mechanism explained in section 2.1.4. Figure 5.14 shows the large signal output current characteristic of Part 2 with 8 bias currents ranging from nominal values of 1.0 μA to 1.7 μA.

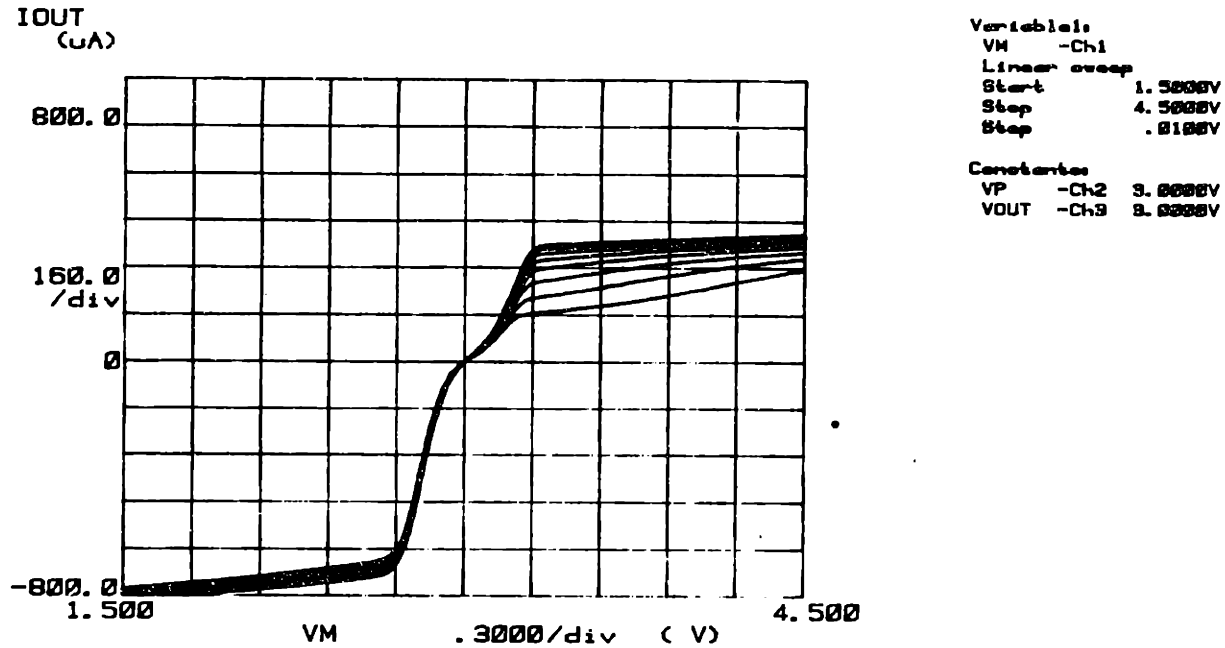


Fig. 5.14 Effect of bias current on mismatch limited output current, Part 2 (78°K)

Bias currents are spaced by .1 μA. Because the output current is limited by mismatchs, the maximum output current increases as the bias current increases. Increases in bias current over 1.5 μA produce negligible increases in maximum negative output current because the output current is no longer limited by mismatch. It settles at a value of 410 μA which is consistent with the other measurements in Table 5.2. Using equation(2.51), the data from figure 5.14 suggests a current mirror mismatch of approximately 10 %. This is consistent with a threshold mismatch of approximately 3 mV, assuming that the dominant mismatch term is the threshold mismatch of the M4A,M4B current mirror. Although the author is unaware of mismatch data for this process, the value of 3 mV for a 12x3 n channel transistor is in agreement with published data.[27]. The p channel input stage does not suffer from this problem because the transistors

M4Ap, M4Bp are 36x3 and have less threshold variation. Only somewhat less than half of the n channel transistors have this problem because it is only manifested when the threshold mismatch has the proper polarity and exceeds approximately 3 mv. In this case, the problem was magnified because the bias current was systematically 25 % low. At the proper bias it would be much less severe, although not altogether absent.

What remains unclear however, is why the maximum output currents are as large as they are and why the maximum positive output current is larger than the maximum negative output current.

5.3.4 Noise Performance

The OTA noise was measured by configuring it as a unity gain follower with an incrementally grounded input. The external buffer was used within the feedback loop so its noise contributions within the bandwidth of the circuit would be negligible when referred to the input. The test setup is shown in figure 5.15.

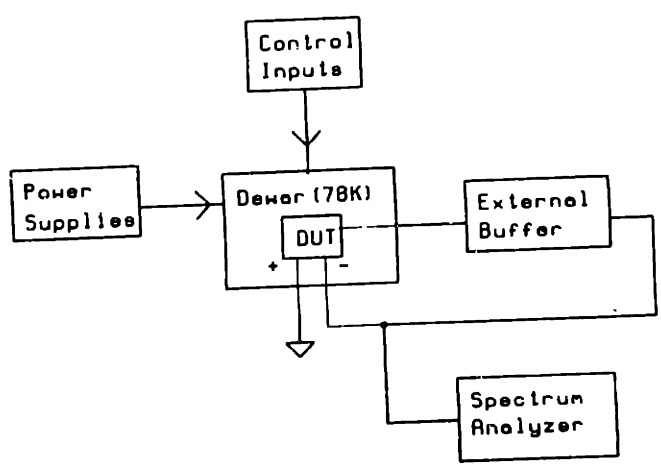


Fig. 5.15 Amplifier noise test setup

The noise was measured with an HP 3561A Dynamic Signal Analyzer. The results from 250 Hz to 100kHz is shown in figure 5.16.

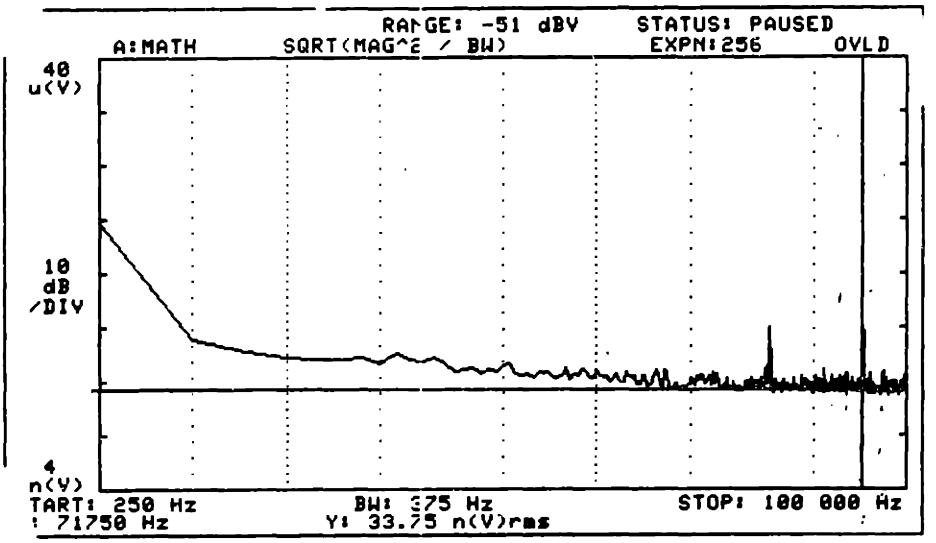


Fig. 5. 16 Input referred OTA voltage noise, Part 3 (78 ^{o}K)

The noise voltage spectra shows a white component of 34 nV/$\sqrt{}$ Hz. The predicted value of equation (4.37) was 40 nV/$\sqrt{}$ Hz. It must be noted, however, that the theoretical thermal noise is only 13 nV/$\sqrt{}$ Hz. The thermal noise estimates were increases by a factor of three over theoretical because P channel noise spectra showed white noise components at a level approximately 3 time the thermal noise value. The n channel noise characteristics were taken from a single transistor. The OTA input referred thermal noise is in agreement with the observed excess high frequency noise. Because of the limited data for N channel noise characterization, agreement with predicted noise values for the OTA cannot be expected to be very accurate. Figure 5.17 shows the low frequency region from .25 Hz to 100 Hz.

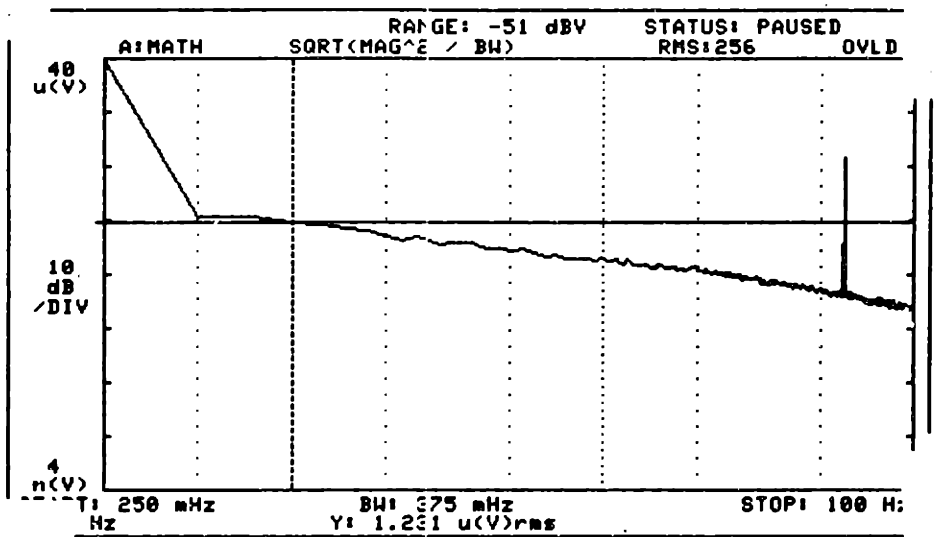


Fig. 5.17 low frequency input referred voltage noise, Part 3 (78oK)

The low frequency noise rolls off roughly as $1/\sqrt{f}$ as expected, and has a value of 1.2 $\mu V/\sqrt{Hz}$ @1Hz. The predicted value of equation (4.38) was 1.5 $\mu V/\sqrt{Hz}$. The difference being that the n channel devices were not characterized for noise due to a layout error. The agreement is reasonable, considering these circumstances.

5.4 Closed Loop Buffer Performance

In the closed loop connection the output buffer is normally tested with the clocks $\phi 1$ and $\phi 2$ running. The default timing pattern is shown in figure 5.18. Each clock is active for 5 μs and the nonoverlapping interval is nominally set for 100ns. The clocks are provided from a Programmable Word Generator (PWG) external to the dewar. The clocks are buffered with a set of inverters on the test chip. Unless otherwise specified all data presented is generated with the default

timing pattern. The external buffer box is used is the following data unless otherwise specified.

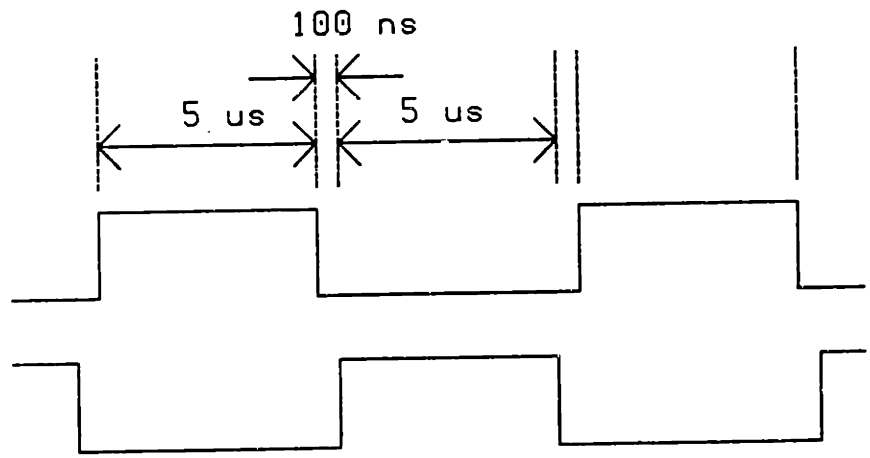


Fig. 5.18 Two phase nonoverlapping clocks

5.4.1 Output Impedance

The closed loop output impedance of the output buffer was measured by coupling a sinusoidal input into the output with a 1 MΩ resistor. By comparing the amplitude of the sinusoid with the amplitude of the modulation the output impedance of the DUT can be inferred. Fig 5.19 shows the setup for the output impedance measurement. The measurement was made with an HP3577A Network Analyzer. The source from the network analyzer drove one side of the 1 MΩ resistor, while the signal input for the network analyzer was taken from the output of the external buffer box.

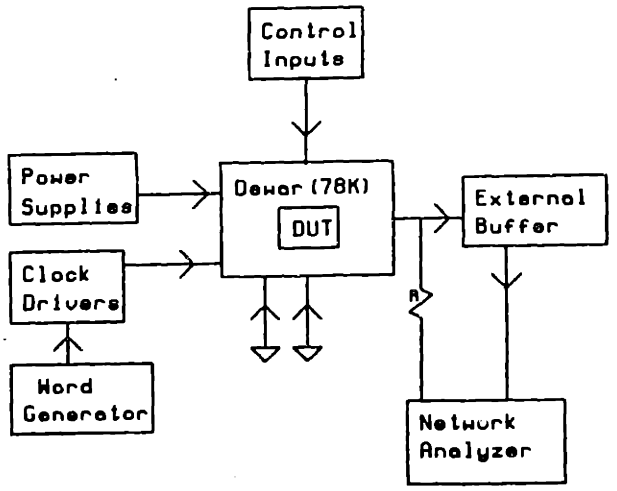


Fig. 5.19 Output impedance measurement test setup

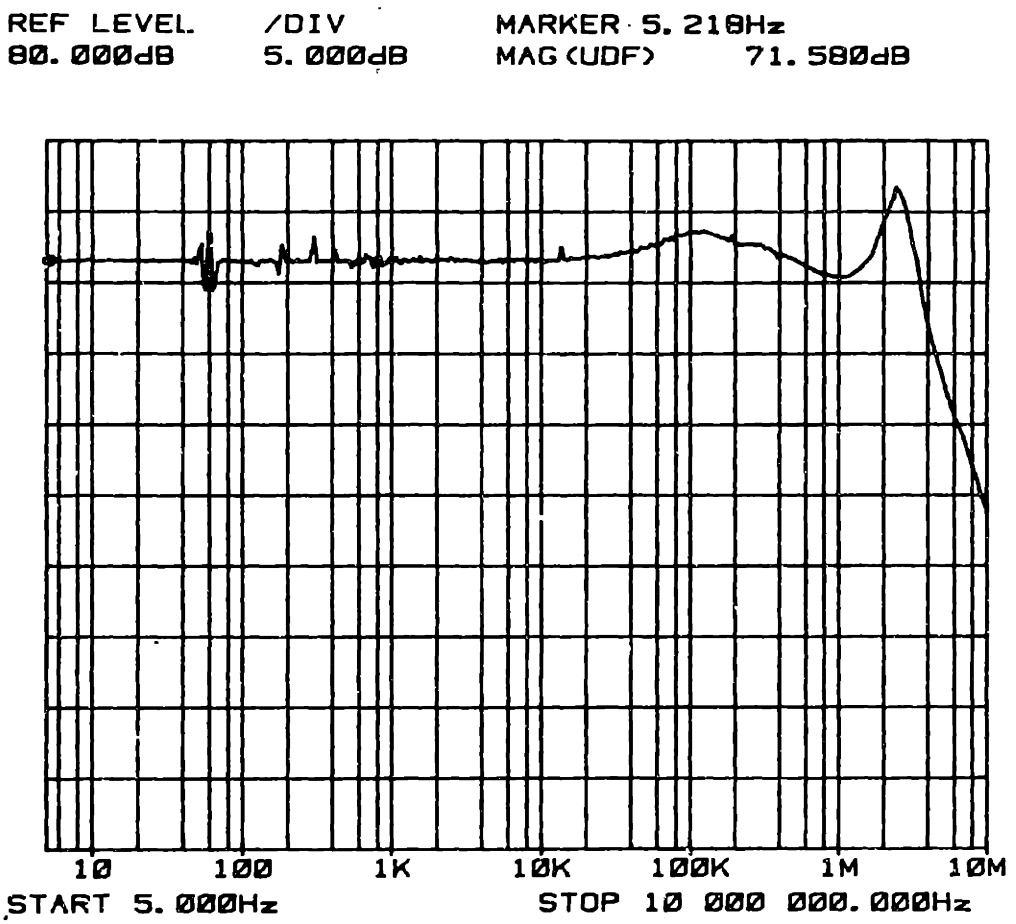


Fig. 5.20 Output impedance, Part 2 (78oK)

The measured output impedance is shown in figure 5.20. The low frequency value of 72 dB is equivalent to 3.8 kΩ. This is in good agreement with the predicted value of 3.9 kΩ in equation 4.4.2. The plot is not accurate for frequencies above 100kHz. Limitations in the external buffer, discussed in section 5.4.1, have distorted the data.

5.4.2 Slew Rate

With V1 and V2, the input voltages for the buffer, set to 1.5 V and 4.5 V the large signal performance of the buffer could be viewed on an oscilloscope. Figure 5.21 is a double exposure oscilloscope photograph.

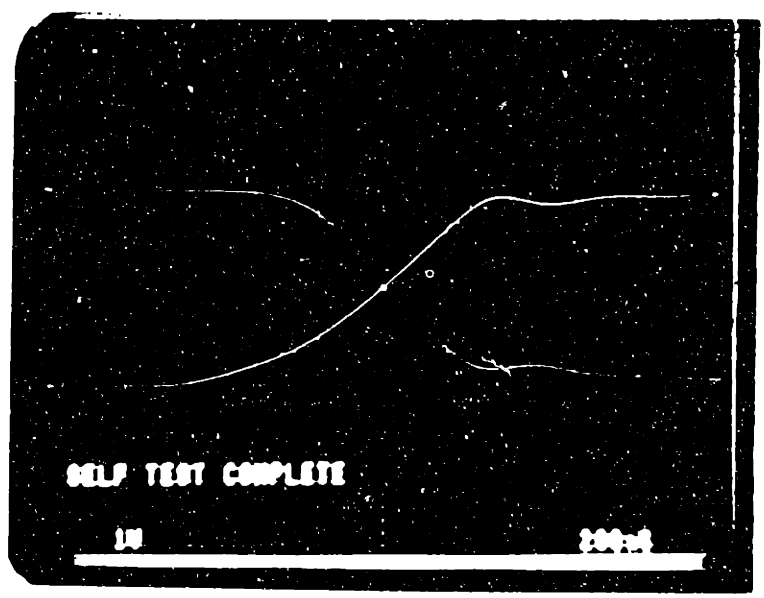


Fig 5.21 Large signal Step response, Part 1 (78°K)

Both a positive and negative 3 volts transition is displayed. The horizontal scale is 200 ns/div and the vertical scale is 1V/div. The load capacitance is 115 pF. From

the measured slew rates the maximum slewing currents from the amplifier can be calculated. Table 5.3 shows the measured slewing currents for three parts.

Part	I+	I-
1	460 μA	610 μA
2	620 μA	150 μA
7	-	130 μA

Table 5.3 Measured slewing currents (78 °K)

By comparison with the steady state maximum output currents reported in Table 5.2, the slewing values are seen to be approximately 20 % lower. This is to be expected because the slewing currents represent an average over a region of the steady state output current characteristic. The values reported in Table 5.2 are not sustained over any period of time because the output current drops as the transition is taking place.

5.4.3 Large Signal Settling Time

Measurement of settling time to a small percent for a large step can pose special measurement requirements. In this case a 16 bit data acquisition system was used to sample the waveform at intervals of 400 ns. The waveform from the output buffer is switched from 1.5 volts to 4.5 volts with a period of approximately 10 μs. The waveform is sampled at a particular time for 4096 cycles. At each cycle the sample is digitized and recorded. The average value of the 4096 samples is saved as the value of the waveform at that particular time. The sample time is then

moved by 400 ns and process repeats. The experimental setup for the large signal settling time measurement is shown in figure 5.22.

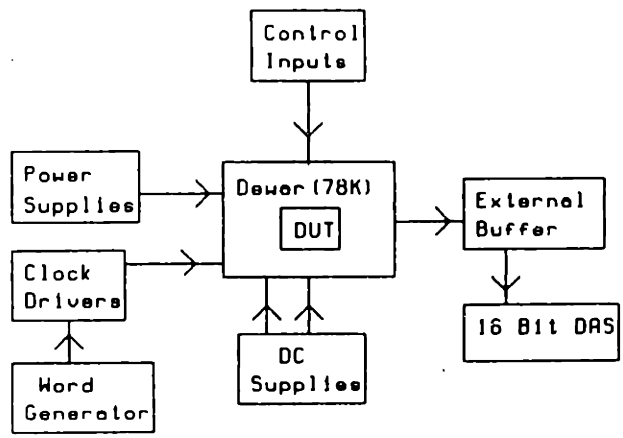


Fig. 5.22 Large signal settling time setup

A potential problem with this technique is that the DAS may corrupt the measurement. If the large signal settling time of the signal chain in the DAS is itself longer than the DUT, then the settling time as measured will be longer than the actual settling time of DUT. It takes about 1 second to take the 4096 samples at one time and about 30 seconds to sample the entire waveform. Any drift in the system over this period of time will appear as an erroneous slow settling component. In either case the settling time will be at least as good as the measurement indicates, so the measurement can be regarded as an upper bound. Any drift in the DAS was reduced by sampling in one direction along the waveform and then reversing direction and sampling in the other.

Figure 5.23 is an oscilloscope photograph of the 3 volt step square wave from the DUT.

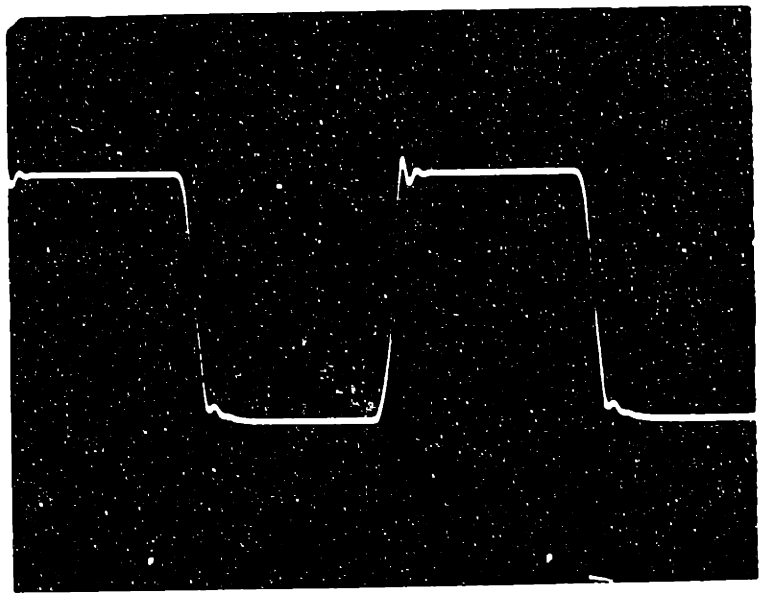


Fig. 5.23 Large signal step response, Part 3 (78°K)

The vertical scale is 1V/div and the horizontal is 2 μs/div. The waveform is switching from 1.5 volts to 4.5 volts. The ringing is the waveform is due to the external LF351 buffer. Unfortunately this was not discovered until after the measurements were complete. Figure 5.24 shows the large signal step response of the external buffer box alone.

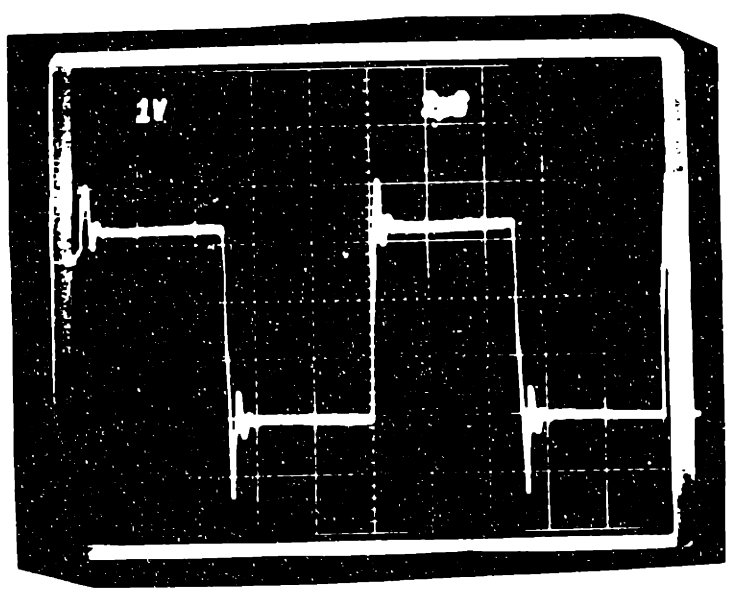


Fig. 5.24 Large signal step response, buffer box

The scales are identical as those for figure 5.22. A measurement of the time constant for this yields approximately 400 ns, which is about equal to the expected time constant for the DUT. It is expected, therefore, that the settling time measurements will be degraded by the external buffer. The problem in the external buffer is due to an overly large feedback resistor (10 k). In fact, the resistor is not necessary at all.

For the large signal settling time measurements the period of the output waveform was doubled so that samples out to 10 μs after the transition could be recorded. Figure 5.25 show the settling time results for the falling edge.

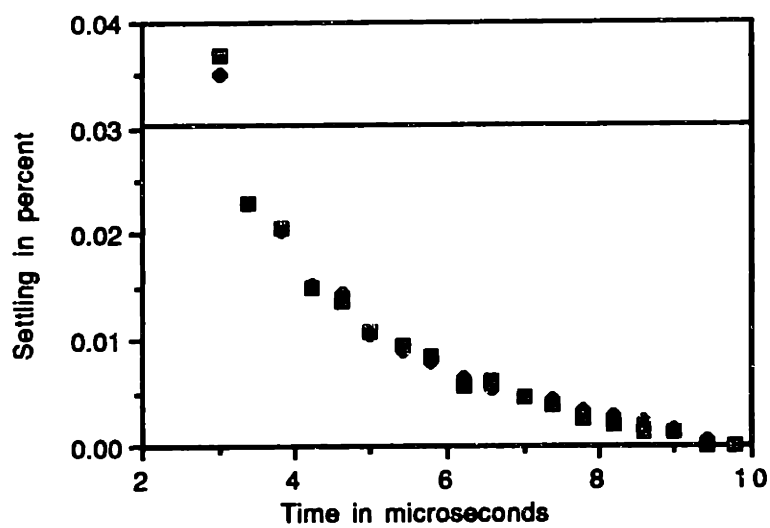


Fig. 5.25 Large signal settling time, falling edge Part 3, (78°K)

Both sweeps over the sampling time are shown. No appreciable drift is visible in the data. Settling to .03% is seen to fall between 3.0 μs and 3.4 μs after the transition. This is in good agreement with settling time estimates made in section 4.3.1. Figure 5.26 shows the data for the rising edge.

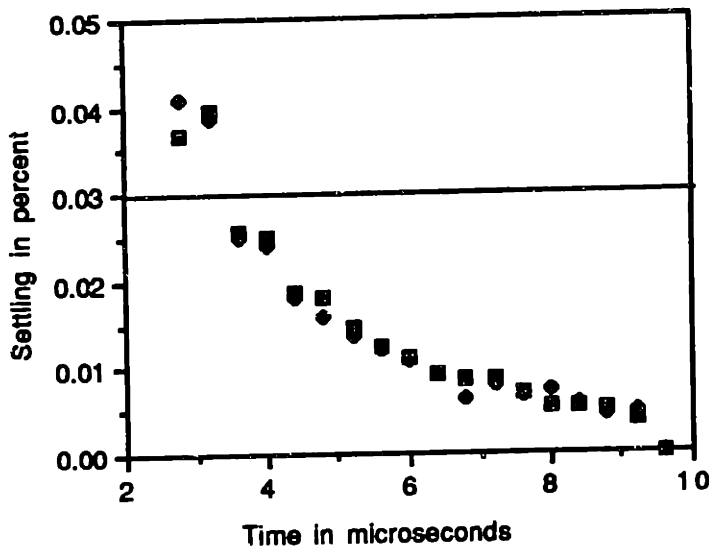


Fig. 5.26 Large signal settling time, rising edge Part 3 (78°K)

In this case the settling time falls between 3.2 μs and 3.4 μs. This agrees well with the falling edge results. Part 2 , which had a degraded maximum negative output current , shows a settling time for a falling edge of 5.0 μs. With the nominal bias increased to 1.5 μA, this settling time is reduced to 3.9 μs.

5.4.4 Noise Performance

To make noise measurements on the closed loop output buffer, the inputs were set to the noninverting bias voltage of 3 volts. This voltage as well as V_{DD} was provided with batteries. The measurement was also made directly with the reference output since the bus inputs for the reference are set to the noninverting bias voltage internally. The setup for the closed loop noise measurements is shown in figure 5.27.

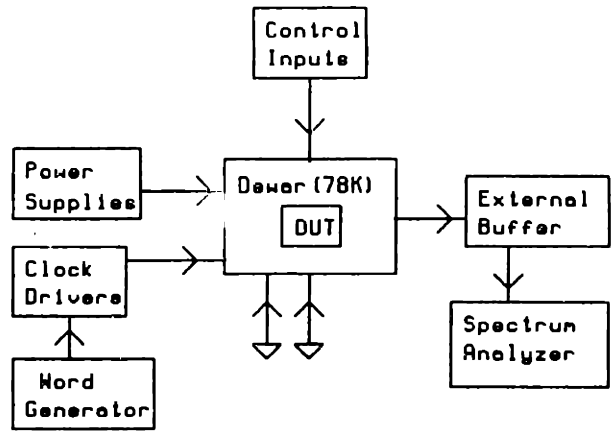


Fig. 5.27 Test setup for noise measurements

5.4.4.1 Without clocks running

In an effort to isolate the various noise contributions, the noise was first measured without the clocks $\phi 1$ and $\phi 2$. It should be mentioned that without the clocks running the inverting node of the amplifier is isolated. Due to leakage currents, it will eventually charge to one of the rails. When at cryogenic temperatures, this can take a considerable period of time. During this time the output buffer can be tested in a closed loop but nonclocking state. To achieve operation in the linear range, the clocks are run for a period of time and then stopped. In this configuration only the direct noise terms will contribute. There are no sampled noise terms. Figure 5.28 shows the noise voltage spectrum for Part 3 in this mode.

Figure 5.28 shows a white noise component of 78 nV/$\sqrt{}$ Hz. It was concluded in Section 4.3.2.2 that the direct contribution of the switch elements was negligible by comparison with the direct contribution of the OTA. The white component of the OTA input referred noise voltage was measured at 34 nV/$\sqrt{}$ Hz

and with the estimated noise gain of 2.4 agrees well with the measured value of 78 nV/√Hz.

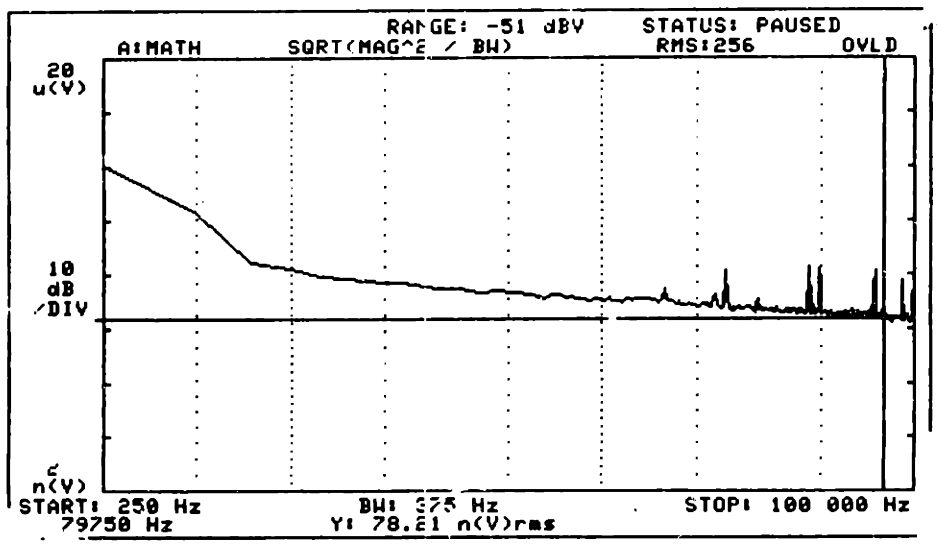


Fig. 5.28 Closed loop noise, no clocks Part 3 (78 °K)

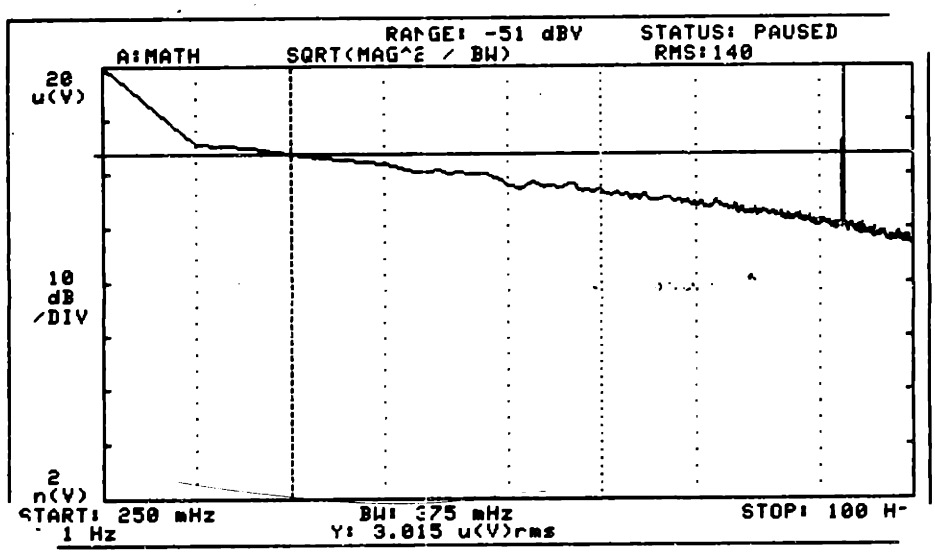


Fig. 5.29 Low frequency closed loop noise, no clocks Part 3 (78 °K)

Figure 5.29 shows the low frequency region of the spectrum from figure 5.28. The measured noise at 1 Hz is 3.0 μV/ √Hz. This is in good agreement with

the measured input referred OTA voltage noise of 1.2 μV/ √Hz and the noise gain of 2.4.

5.4.4.2 With clocks running

To measure the sampled noise effects the clocks were run in their normal mode of operation. Figure 5.30 shows the resulting spectrum.

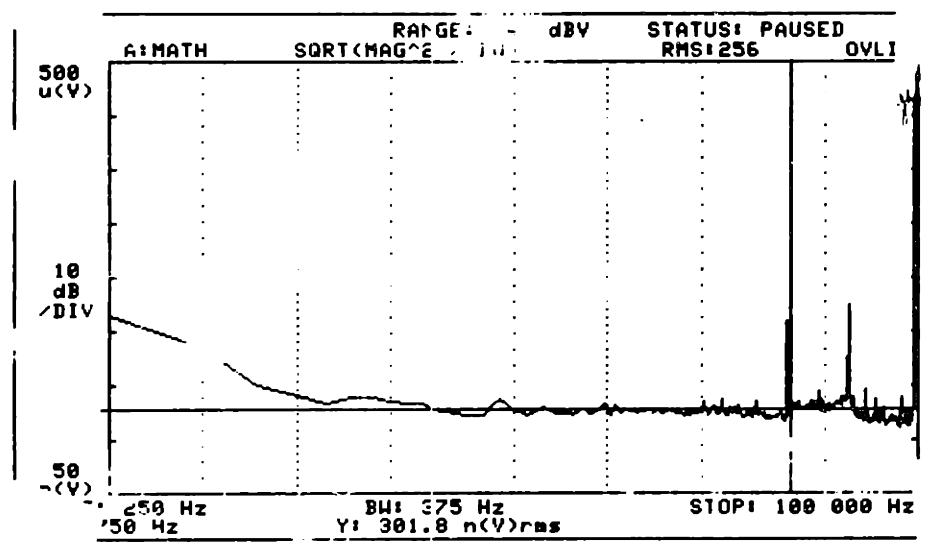


Fig. 5.30 Closed loop noise Part 3 (78 °K)

The first thing to recognize is that the output is a periodic waveform with period of approximately 10 μs. Its spectrum therefore has frequency components at 100 kHz and its harmonics. The fundamental component at 100 kHz is visible at the end of the spectrum. This is not noise, but the Fourier transform of a deterministic periodic signal. In this spectrum we expect to see sampled noise components with a sampling period of approximately 5 μs. These components should have a sinx/x spectrum with the first null at 200 kHz. Unfortunately this is beyond the range of the 3561A Dynamic Signal Analyzer. The degree of rolloff at 100 kHz, however, is

consistent with the expected shape. The sampled noise components were expected to be dominated by the switch elements. The expected mean square voltage from the switch elements was given in (4.49) as 720 μV^2. According to equation (2.86), its spectral form is given as 7.2e-15 V^2/Hz $sinc^2(\omega \cdot Ts/2)$. This would correspond to 84 $nV/\sqrt{Hz}$ [$sinc(\omega \cdot Ts/2)$] on a voltage spectral density plot such as figure 5.30. The 302 $nV/\sqrt{hz}$ reported in figure 5.30 indicates that the sampled switch noise is 3.5 times larger than its theoretical value. Figure 5.31 shows the low frequency region of figure 5.30.

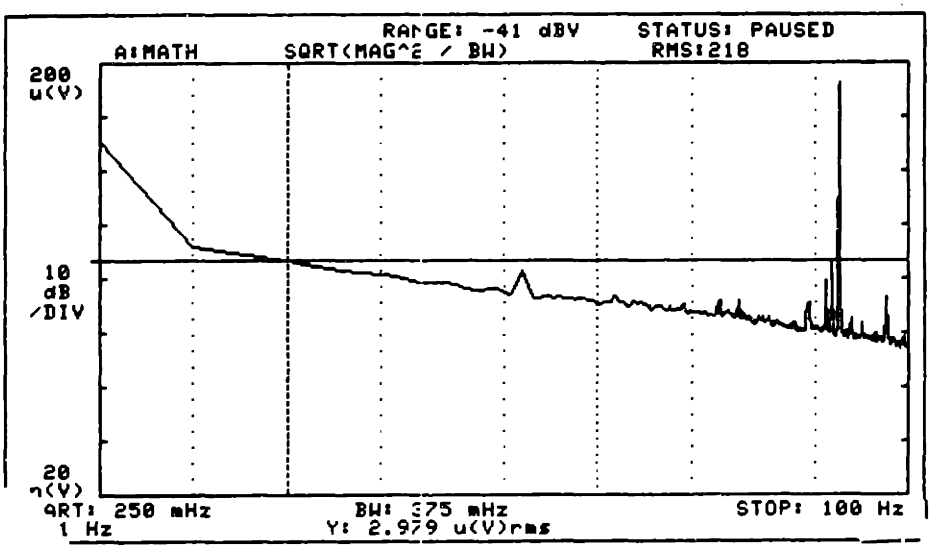


Fig. 5.31 Low frequency closed loop noise Part 3 (78 oK)

The reported noise at 1 Hz is 3.0 $\mu V/\sqrt{Hz}$. This is equal to the value without sampling and in good agreement with the expectation that sampled noise components are broadband. The excessive noise shown in figure 5.24 seems to suggest that the sampled noise components of the switch elements is much greater than theoretical kT/C value. This has been observed before[39], and has been attributed to a statistical component to the splitting of channel charge

which takes place when the switch elements open. In the previous study, the effect is found to be significant only at reduced temperatures.

The spectral data presented above indicate a mean square output voltage of

$$< V_{out}^2 > = (78\ nV/\sqrt{Hz})^2 \cdot 550kHz + (290\ nV/\sqrt{Hz})^2 \cdot 100kHz$$

$$= 3.3\ e3\ \mu V^2 + 8.4e3\ \mu V^2 \ . \qquad (5.1)$$

The first term is the white noise component of the OTA while the second is sampled noise component of the switch elements. The 1/f component of the OTA is negligible. The rms output voltage is indicated to be

$$\text{Vout RMS} = 108\ \mu V \qquad (5.2)$$

The output was digitized by the 16 bit DAS. The RMS voltage of 4096 samples was 210 μV. The excess noise was found to be due primarily to 60 Hz harmonics as well as a spurious component at 38 kHz. The mean square voltage of these single frequency components were individually measured and subtracted from the mean square voltage of the output, yielding an RMS voltage of 141 μV. The difference between this result and equation 5.2 is presumably due to other stray components. The range of the 3561A is only up to 100kHz while the DAS input is bandlimited at 2MHz. Thus, additional stray single frequency components may well have added into the RMS measurement.

The external buffer box also has a differential receiver configuration. In this mode, the signal from the reference circuit and the signal circuit are differenced. In this mode, the RMS output noise is expected to increase by $\sqrt{2}$.

Spurious common mode signals will be attenuated. The expected differential RMS output voltage is

$$Vout(diff)RMS = 108\ \mu V \cdot \sqrt{2} = 153\ \mu V \qquad (5.3)$$

The measured value is 170 μV. In this case the measurement is closer to the expected value because common mode signals have been rejected.

5.4.5 Linearity

The output buffer linearity was measured using a programmable voltage source and the the 16 bit DAS. Both inputs to the output buffer were connected to an HP3314A Function Generator, which is used simply as a programmable voltage supply. The output of the 3314A is monitored by an HP3478A multimeter. The output from the DUT goes to the external buffer and then to the 16 bit DAS. Figure 5.32 shows the test setup.

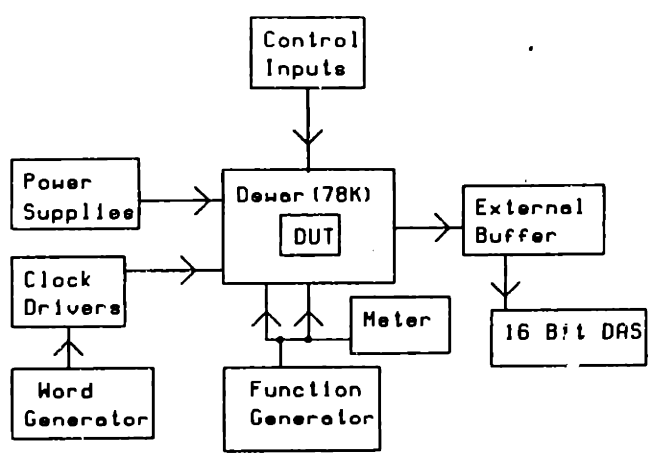


Fig. 5.32 Linearity test setup

The transfer characteristic of the DAS was measured by bypassing the DUT and feeding the voltage supply directly to the DAS. The input voltage was swept from 0 to 6 volts. Figure 5.33 shows the transfer characteristic of the DAS.

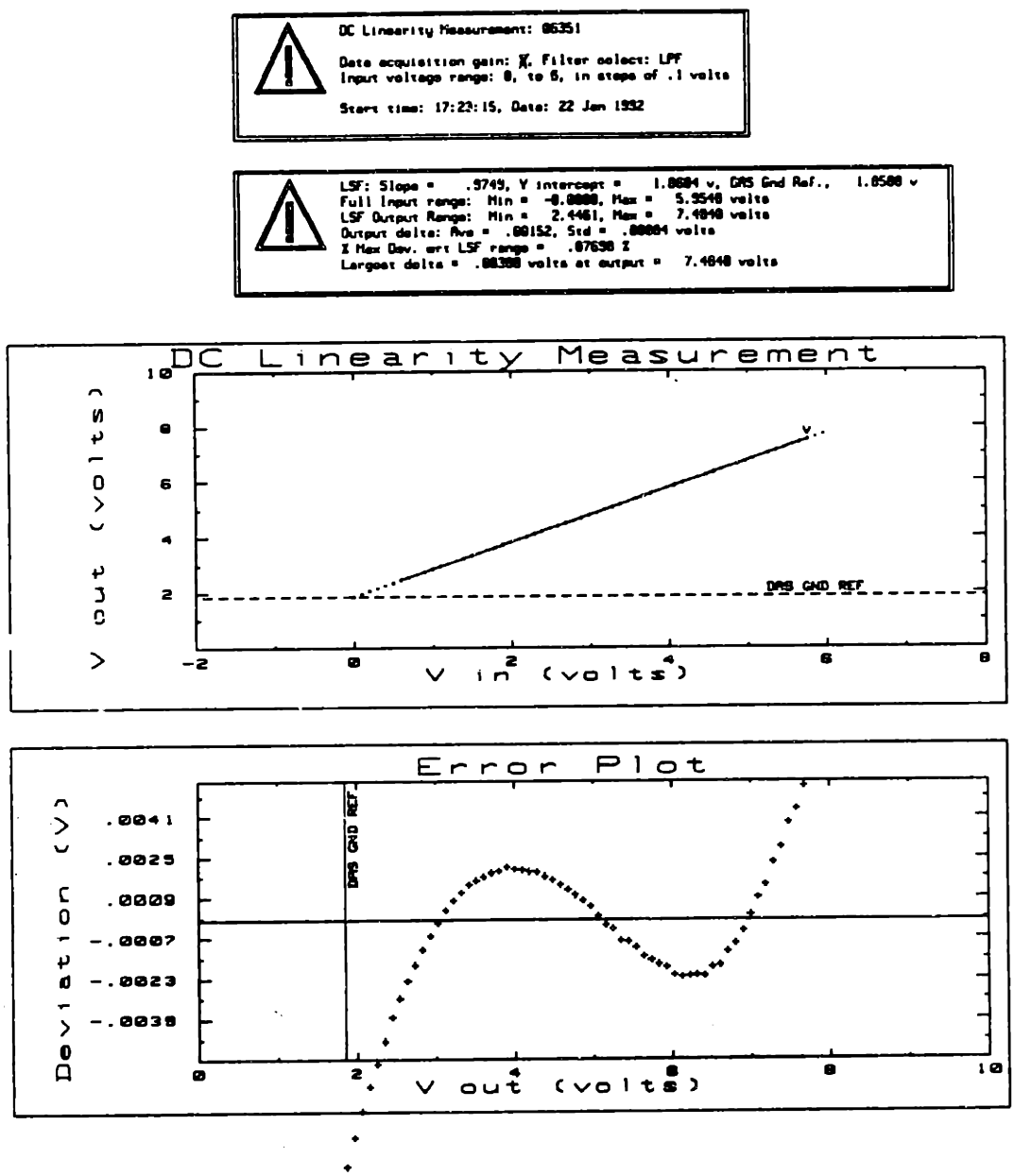


Fig 5.33 Linearity of 16 bit DAS

The input voltage is swept five times over the desired range and the results averaged. The upper plot shows the transfer characteristic of the DAS. The lower

plot shows the error of the data from a least squares line fit over the central five volts of the characteristic. The transfer characteristic of the DAS was recorded in memory so that it could be used to process subsequent data and remove the distortion introduced by the nonlinearity of the DAS. Figure 5.34 is a measurement of the DAS linearity in which the nonlinearity has been corrected.

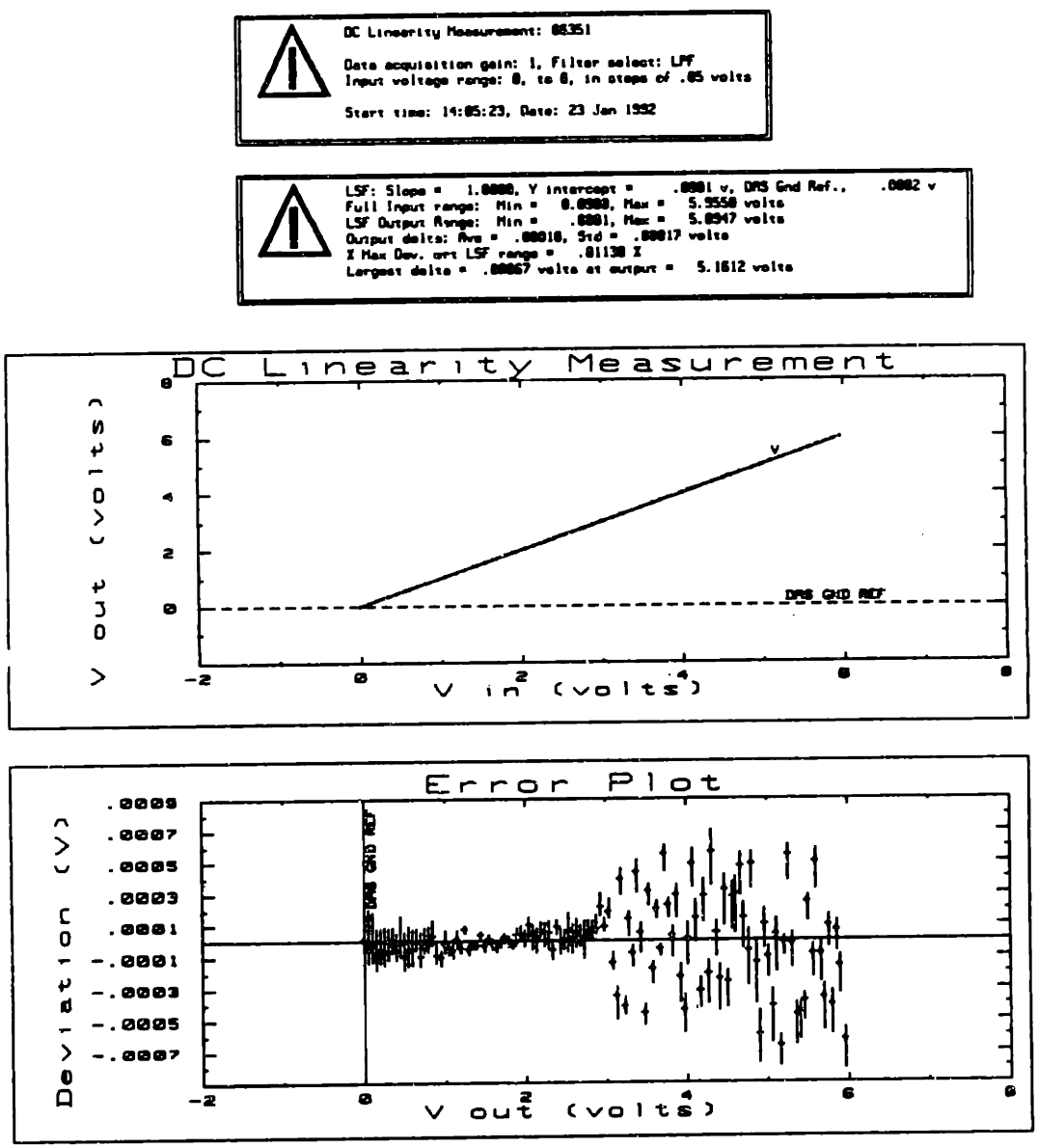


Fig. 5.34 Corrected DAS linearity

Figure 5.34 shows the limits of this technique. Over the lower 3 volts data is corrected to within .1 mv, but over the upper 3 volts the correction is limited to .7 mv. This is apparently due to the fact that the 3478A meter which is used to make the correction switches ranges at 3 volts. The corrected linearity of part 2 is shown in figure 5.35.

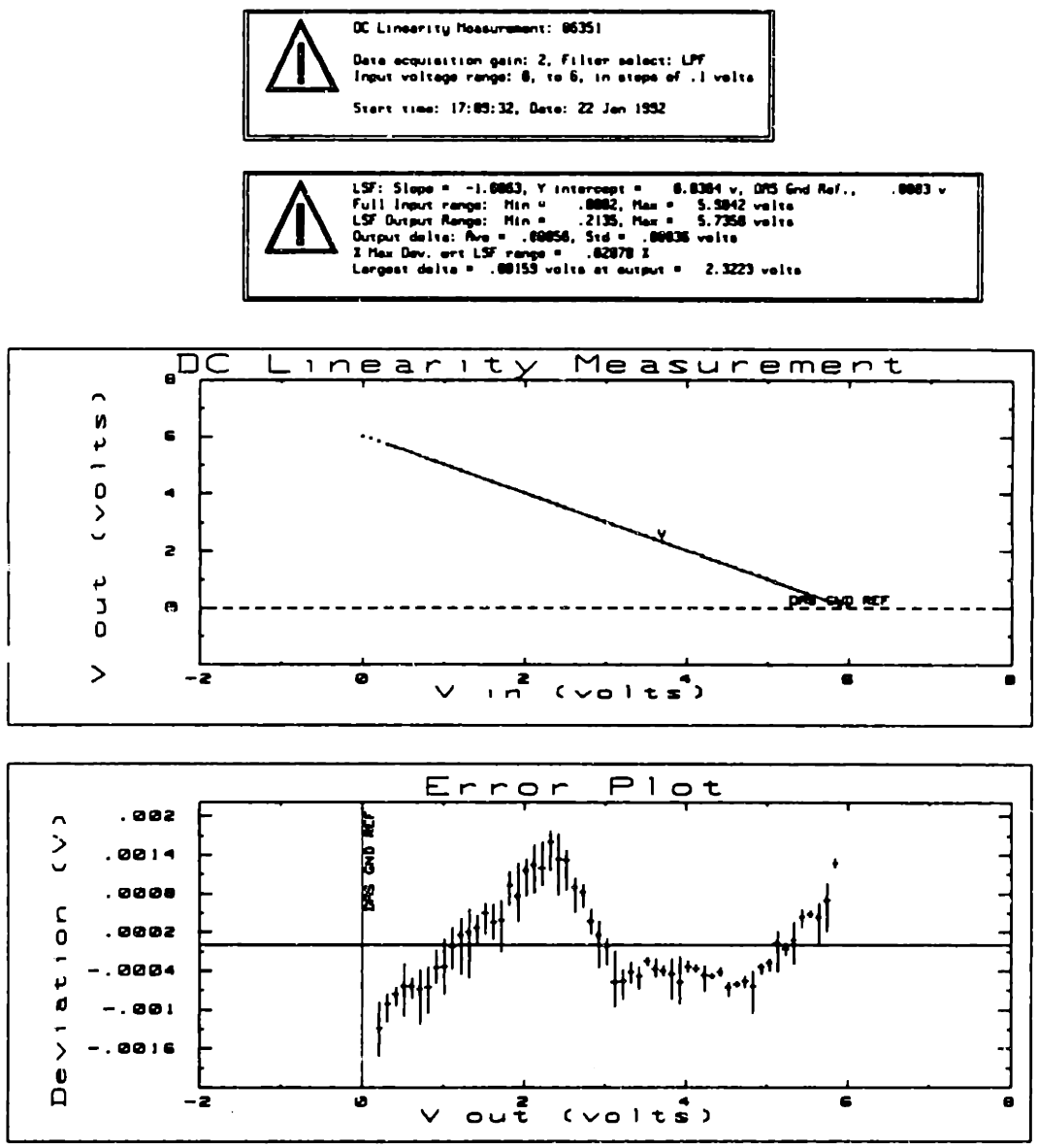


Fig. 5.35 Corrected linearity, Part 2 (78°K)

The gain of -1 from the output buffer is evident in the transfer characteristic. The maximum error from a least squares fit line over the central 5.5 volts is 1.6 mV. If this is represented as a percentage of the total output range, it yields an integral nonlinearity of .029 %. This is well below the specification, but it is not easy to relate this value to any predicted value.

5.4.6 Crosstalk

Crosstalk was measured with an HP3577A Network Analyzer. The source from the 3577A was used as an input for one channel of the output buffer while the other channel was grounded. The channel corresponding to the grounded input was sampled and held on each output cycle and was used as the signal input to the 3577A. Figure 5.30 shows the test set up.

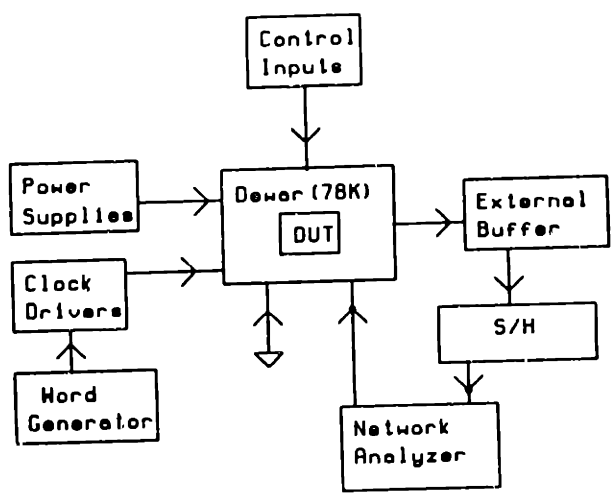


Fig. 5.36 Crosstalk test setup

Figure 5.37 shows a measurement of the magnitude of the signal channel itself.

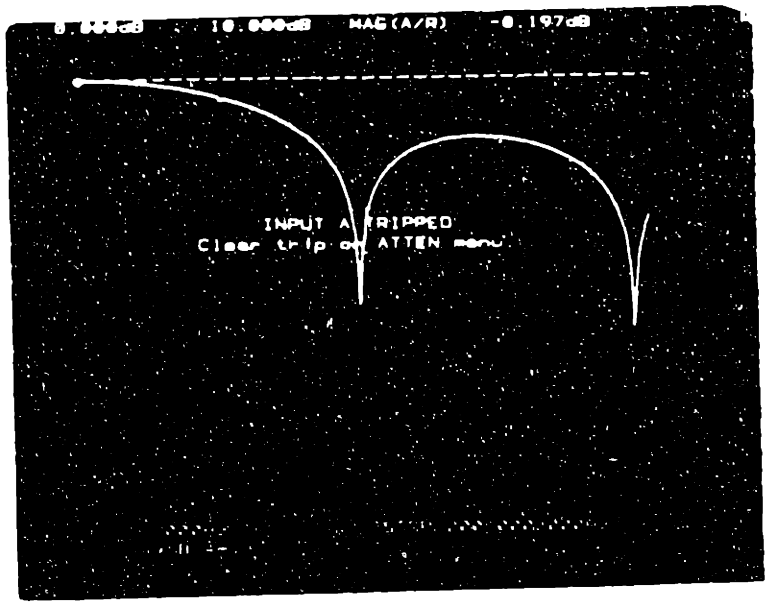


Fig. 5.37 Signal measurement, magnitude, Part 3 (78°K)

The vertical scale is 10dB/div and the horizontal scale is 20kHz/div. The plot shows the gain of 1 at low frequency as expected, and the sampling nature of the output buffer is evident in the sinx/x behavior. The sampling null at 100 kHz corresponds to the sampling frequency for one channel. Fig 5.38 show the phase plot for the same measurement.

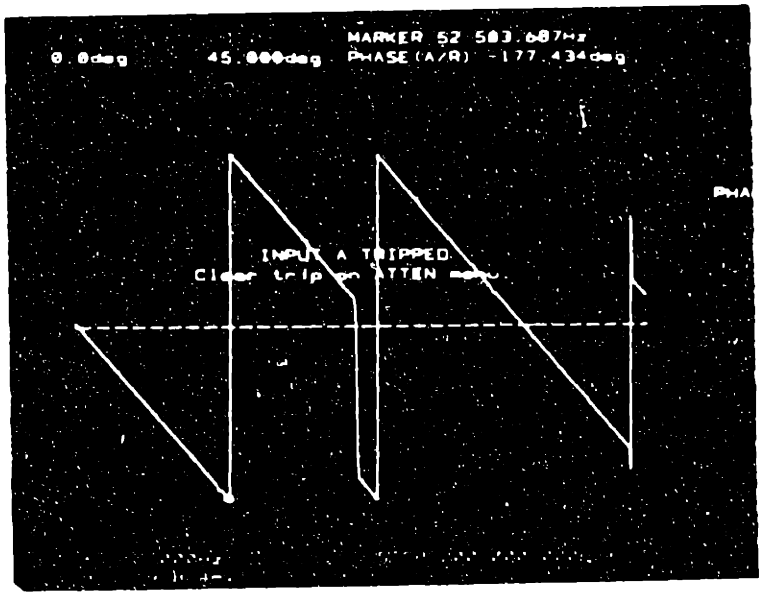


Fig 5.38 Signal measurement, phase, Part 3 (78°K)

The vertical scale is 45°/div and the horizontal scale is 20kHz. The phase at low frequency is 0 degrees. The output buffer is inverting so that the phase at low frequency should be 180°, but an additional inversion takes place in the external sample and hold. The phase is linear with frequency indicating a delay of approximately 10 μs. The output buffer itself delays the signal by 5 μs because of the sample and hold action of the input capacitors. The external sample and hold which is necessary to isolate a single output sample introduces and additional 5 μs. Use here has been made of the observation that sampling and holding a signal every T seconds introduces a delay of T/2. Figure 5.39 shows the crosstalk measurement magnitude.

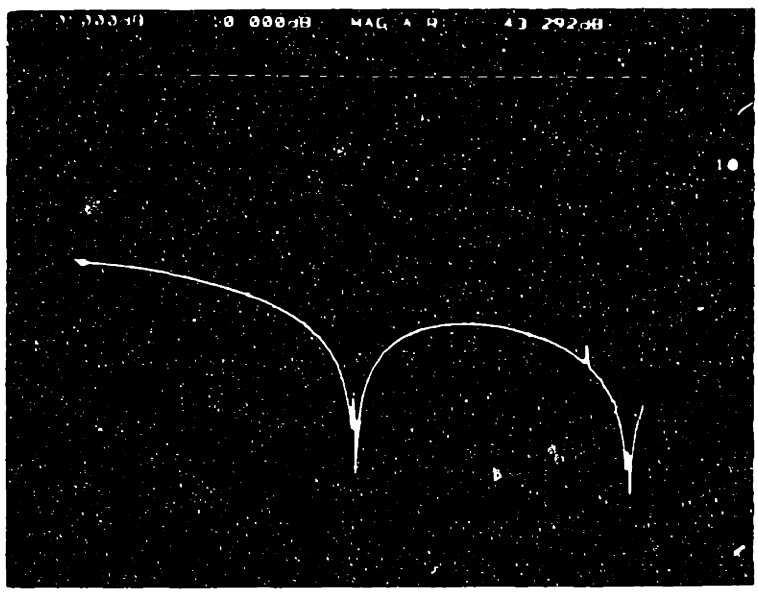


Fig. 5.39 Crosstalk measurement channel 1 → 2, magnitude, Part 3 (78°K)

The measurement is for channel 1 leaking into channel 2. The scales are identical as before. The crosstalk magnitude is -43 db. This is much larger than anticipated. Figure 5.40 shows the phase plot.

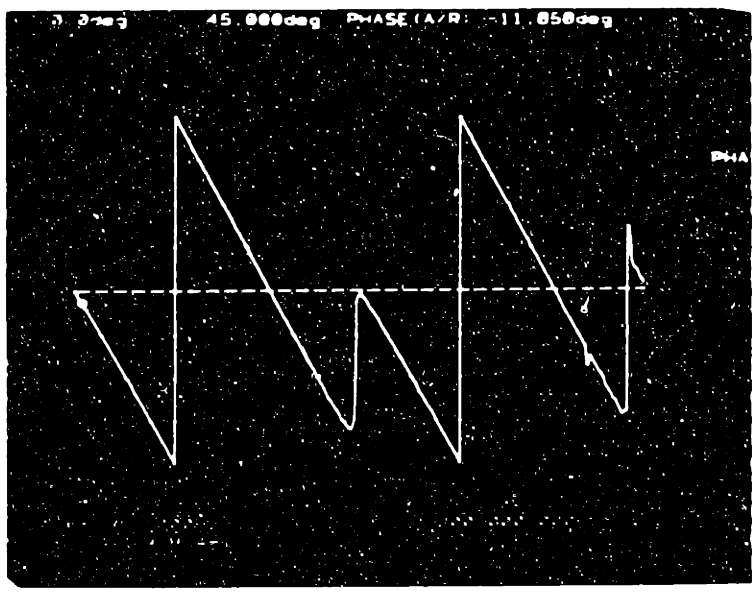


Fig. 5.40 Crosstalk measurement channel 1 → 2, phase, Part3 (78°K)

The delay in this case is approximately 15 μs. The crosstalk, when measured with respect to the other channel, produced similar results (-44 dB,15 μs delay).

In order to determine if the crosstalk was due to incomplete bus settling the bus reset switch was employed but no improvement was noted. Along the same line of inquiry, the OTA bias current was increased in order to improve settling but no improvement was noted. It was possible to lower the bias current until the crosstalk became even worse, presumably due to incomplete settling. Under these conditions the operation of the bus reset switch did improve the crosstalk measurement but only up to the 43 dB figure. Thus the bus reset switch was shown to be operative, but the source of crosstalk was shown to be due to another mechanism.

Four stray capacitances were identified as possible coupling mechanisms. These are shown in figure 5.41 as CS1, CS2, CS3, CS4.

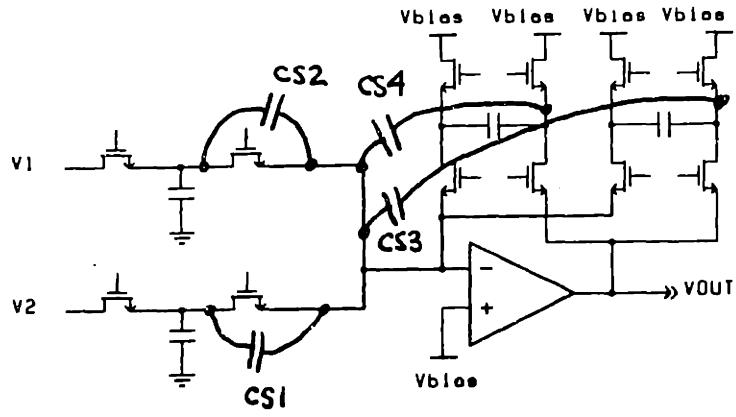


Fig. 5.41 Possible stray capacitances

It is possible to measure the strays CS1 and CS2 by stopping the clocks in a particular state and measuring the transmission of the channel which is not selected. This measurement indicates that CS1 = 8 fF and CS2 = .9 fF. There is no delay in the OTA for this mode of transmission so that the overall delay in the crosstalk mechanism for these strays would be 5 μs. The magnitude for the the crosstalk produced by these mechanisms would be -51 dB for channel 2 leaking into channel 1 and -70 dB for channel 1 leaking into channel 2. They are not sufficient to explain the crosstalk measurement.

It is not possible to measure CS3 and CS4 individually due to their location. CS3 and CS4 introduce crosstalk because each of them store the previous signal when they are removed from the feedback path. When they are reset they couple the previous signal into the integrator where it becomes part of the present signal. The delay characteristics of CS3 and CS4 are different than those of CS1 and CS2. CS1 and CS2 couple in a signal yet to be while CS3 and CS4 couple in signal already past. The delays in the OTA associated with the CS3 and CS4 crosstalk mechanism is therefore 10 μs, and their total delay in the crosstalk measurement

is 15 μs. This delay matches the delay observed in Fig. 5.34 and identifies CS3 and CS4 as the dominant source of crosstalk.

CS2 and CS4 both contribute for the crosstalk of channel 1 leaking into channel 2. CS2 is 1 fF, so CS4 must be approximately 20 fF to explain the -43 dB crosstalk measurement. CS1 and CS3 both contribute for the crosstalk of channel 2 leaking into channel 1. CS1 is 8 fF, so CS3 must be approximately 11 fF to explain the -44 dB crosstalk measurement. Since CS1 and CS3 are of comparable magnitude but with different delays, one would expect that as the frequency increases they would go out of phase with each other and produce and additional dip in the crosstalk magnitude plot for channel 2 leaking into channel 1. Figure 5.42 shows this plot.

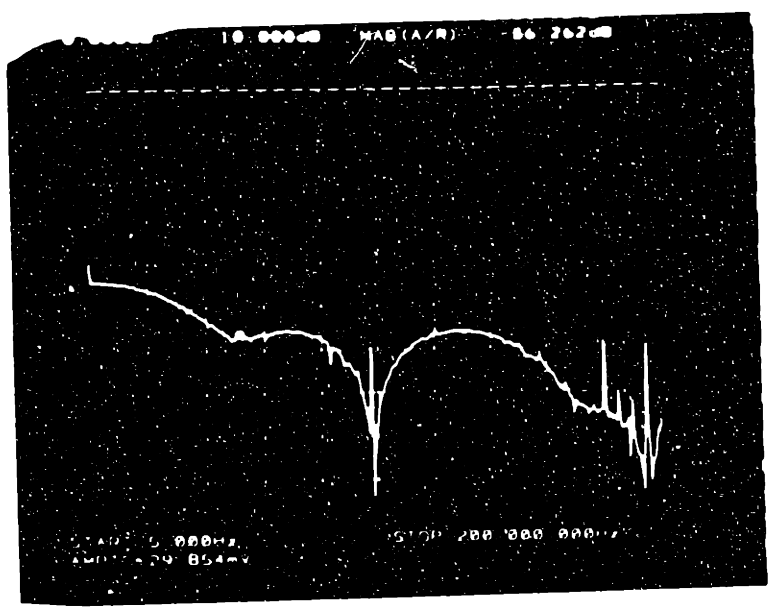


Fig. 5.42 Crosstalk measurement channel 2 → 1, magnitude, Part 3 (78°K)

The dip is present at 50 kHz exactly as expected. Inspection of the layout further confirms the presence of these strays. In this case the strays have caused a serious reduction in the achievable crosstalk with this buffer.

5.4.7 Supply Rejection

For completeness PSRR+ and PSRR- was measured using the 3577A Network Analyzer. The setup was the same as for the crosstalk except the source from the 3577A was placed in series with each of the two supplies. The rejection of the OTA itself was first measured by configuring it as a follower. Figure 5.43 shows the rejection from the positive supply while figure 5.44 from the negative.

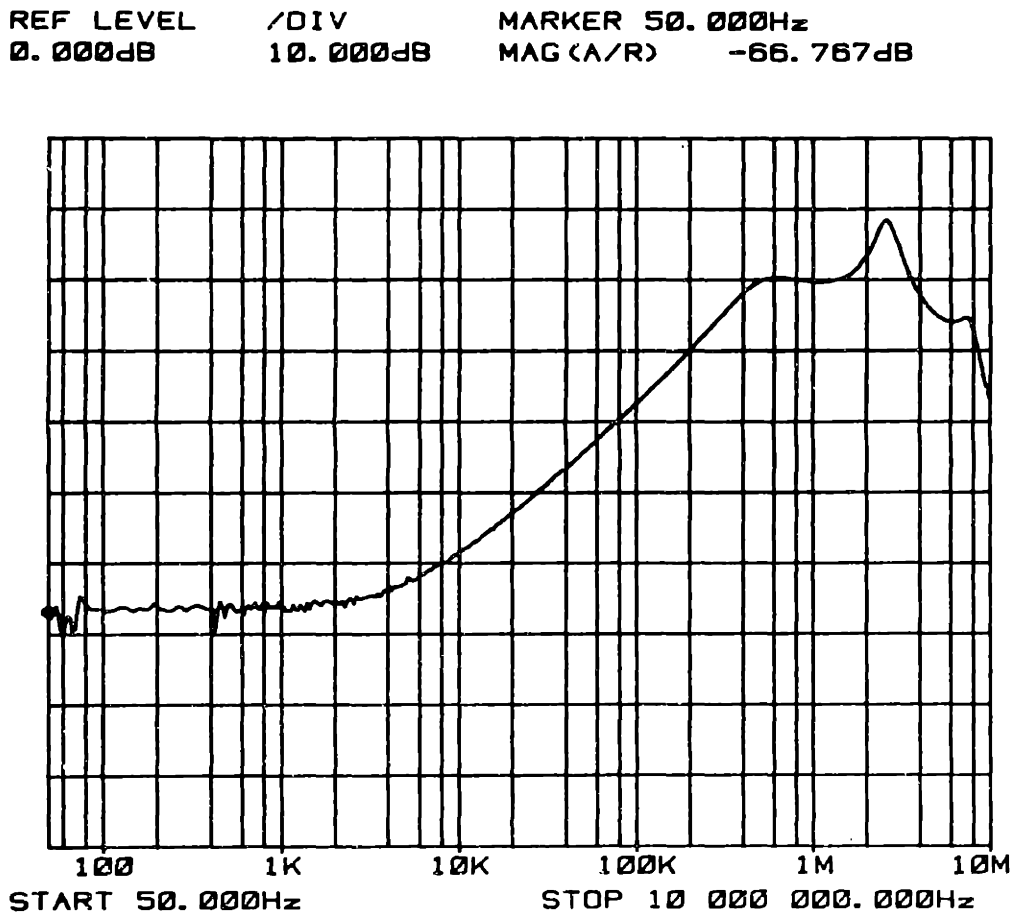


Fig. 5.43 OTA PSRR+ Part 3, (78°K)

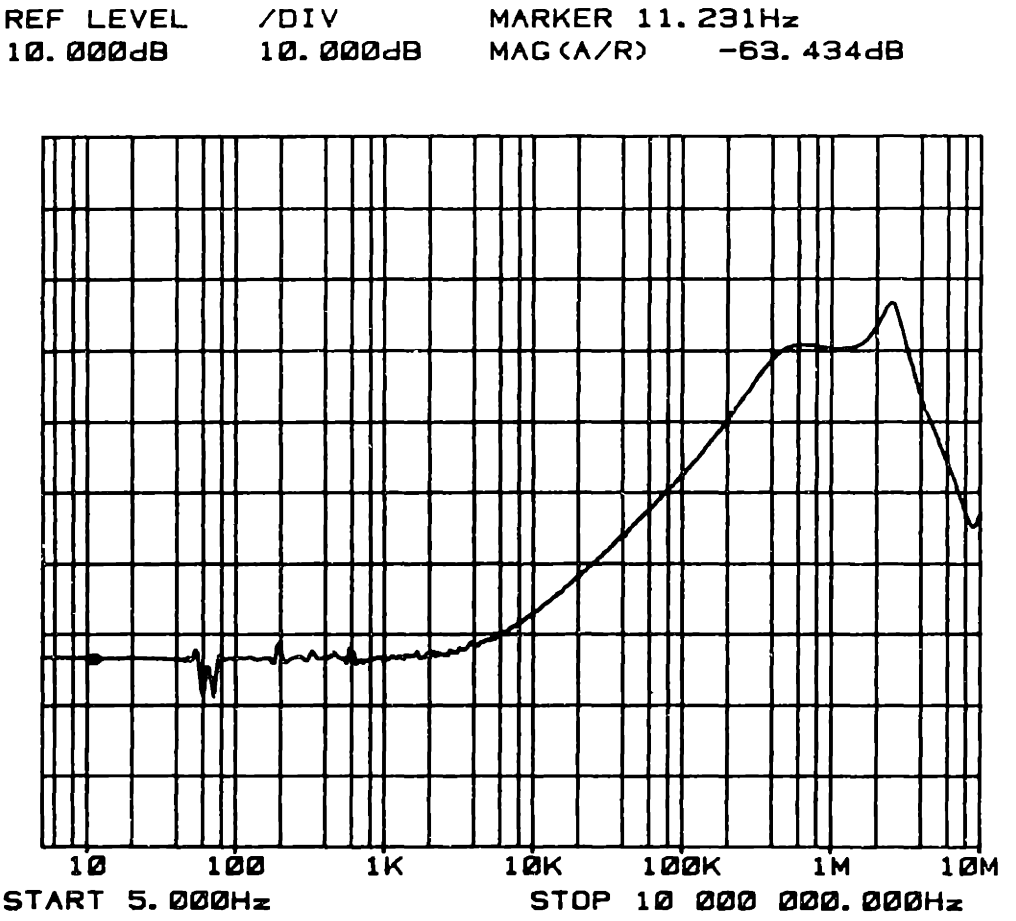


Fig. 5.44 OTA PSRR- Part 3, (78°K)

The measurements with the output buffer operating in its normal mode are shown as figure 5.45 and figure 5.46. As was the case with the crosstalk investigation, the intrinsic rejection of the amplifier is lost in the complete buffer. There were found to be overlaps of both supplies with the inverting node of the OTA consistent with the reductions observed above.

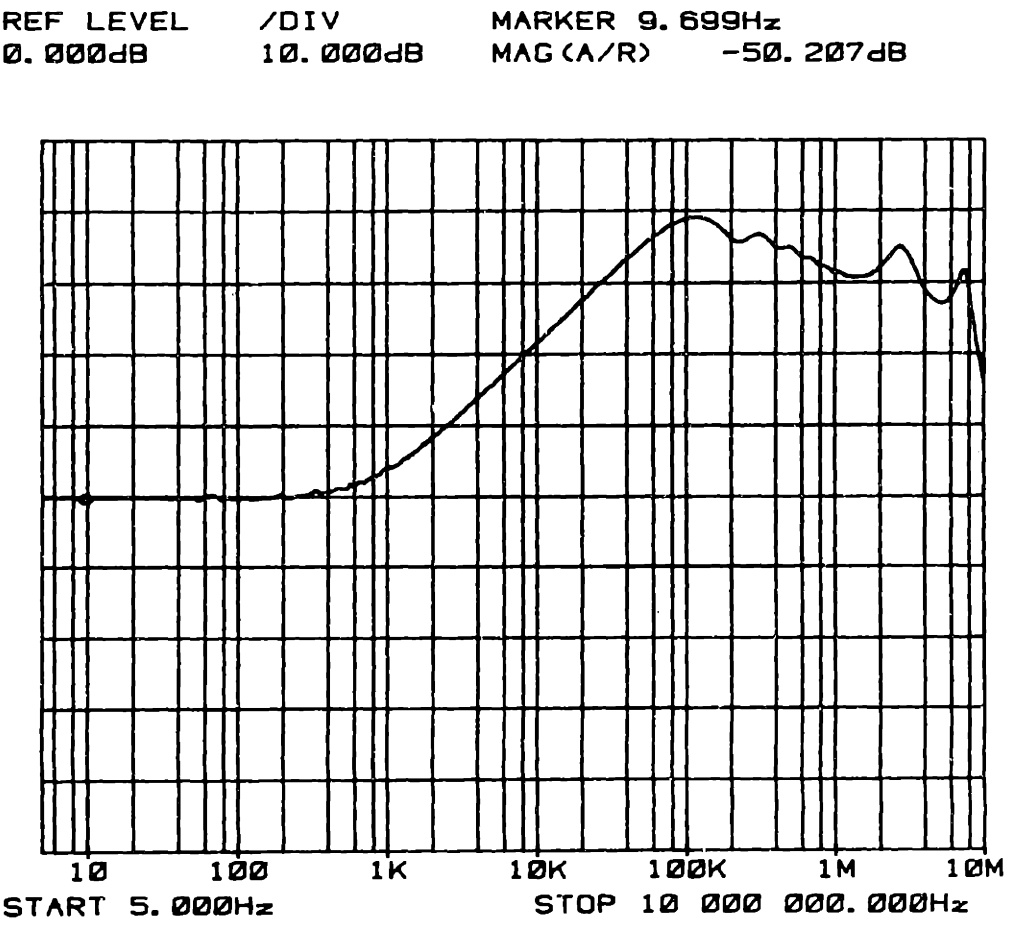


Fig. 5.45 Output buffer PSRR+ Part 3, (78°K)

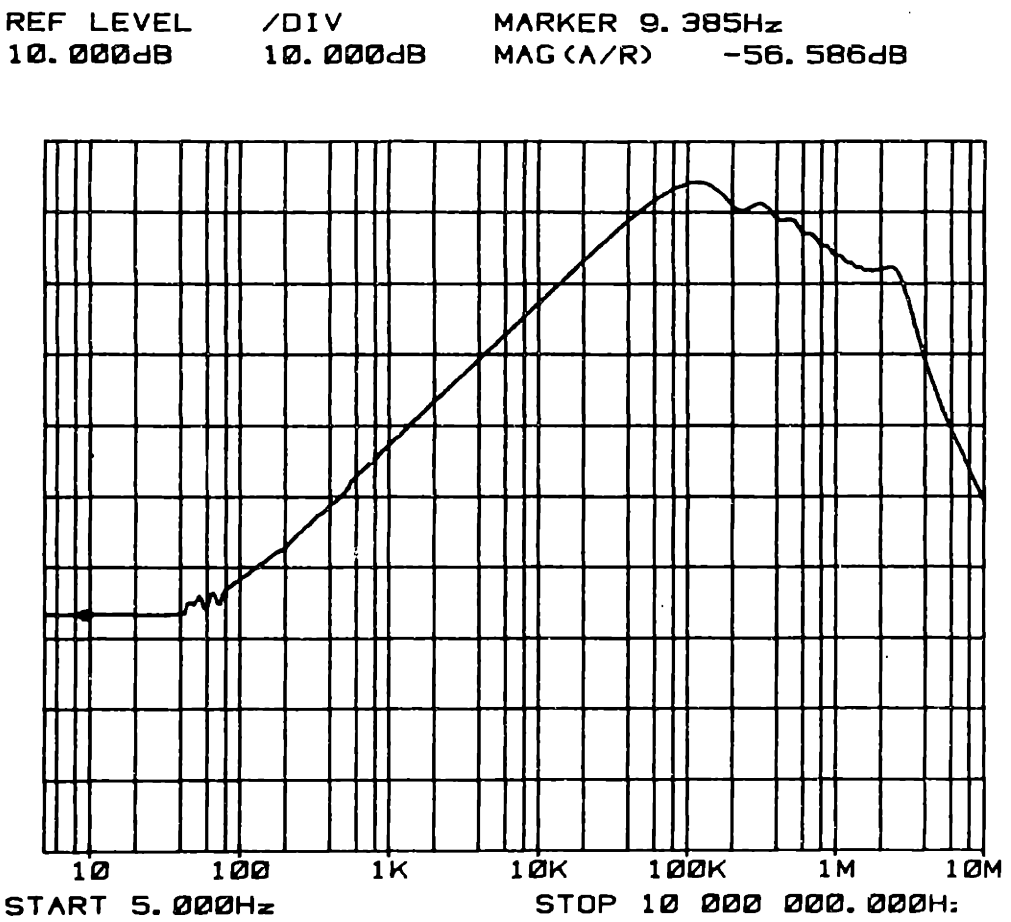


Fig. 5.46 Output buffer PSRR- Part 3, (78°K)

6. SUMMARY AND CONCLUSIONS

6.1 Result Summary

A summary of measured and predicted OTA characteristics is presented as Table 6.1. Measured values are averages of all units tested, with the exception of the output current limits. Those units affected by mismatch limited output current are not included. Predicted performance is based on a nominal bias current of 1 μA for the input stage current sources. Due the width offset factor, WD, in the fabrication process the bias current for the measured data was probably low by 25%. The test chip did not allow for direct measurement of bias currents and the low current condition was not suspected until testing was complete. Low bias current is probably responsible for the low measurements of transconductance and unity gain frequency.

	Measured	Predicted
Transconductance	500 μS	590 μS
DC gain	91 dB	91 dB
Unity gain frequency	900 kHz	940 kHz
Positive current limit	720 μA	> 140 μA
Negative current limit	450 μA	> 210 μA
Input referred thermal noise	34 nV/√Hz	40 nV/√Hz
Input referred 1/f noise (@1hz)	1.2 μV/√Hz	1.5 μV/√Hz
PSRR+ (DC)	67 dB	-
PSRR- (DC)	63 dB	-

Table 6.1 Measured and predicted OTA characteristics

The predicted steady state maximum output currents are lower limits, yet it is rather unusual that the measured values exceed the limits by as much as they do. Also, the measured positive current limit is significantly larger than the negative, while the predicted results would indicate that it should be the other way around. This remains an open issue.

The noise predictions are really estimates because measured noise data was not available from the N channel transistors on the device test chip. N channel characterization was done with a 150x5 transistor from the circuits test chip. N channel high frequency noise was assumed to be a factor of three higher than theoretical thermal noise, as was observed in the case of the P channel devices. The degree of agreement with the noise measurements is quite reasonable.

	Measured	Predicted	Required
Output Impedance	3.8 kΩ	3.9 kΩ	
Positive slew rate	5.4 V/μs	> 1.4 V/μs	
Negative slew rate	6.1 V/μs	> 2.1 V/μs	
Settling time, positive (3V step, .03%)	3.4 μs	3.2 μs (linear)	< 4.0 μs
Settling time, negative (3V step, .03%)	3.2 μs	3.2 μs (linear)	< 4.0 μs
Output noise, RMS	170 μV	109 μV	< 500 μV
Integral nonlinearity	.03%	< .1%	< .1%
Output Range	5.5 V	5.4 V	> 3.0 V
Crosstalk	-43dB	-99 dB	
PSRR+ (DC)	50 dB	-	
PSRR- (DC)	56 dB	-	

Table 6.2 Measured, predicted, and required characteristics of Output Buffer

Table 6.2 shows the measured, predicted, and required characteristics of the output buffer. Measured results are averages of all units tested, with the exception of those parts affected by mismatch limited output current. Here again, the measured slew rates are much larger than the lower limits established in Section 4.4.2. The measured noise is significantly larger than the predicted value. This discrepancy is due to a sampled noise contribution from the switch elements 3.5 times larger than theoretical. Measured crosstalk was much greater than should be achievable with this buffer configuration. This degradation was identified as being caused by two distinct stray capacitances introduced in the layout.

6.2 Suggested Improvements

The analysis and measurements of this output buffer indicate that it can be improved in a number of areas. As it stands, a significant (40%) fraction of the units have a serious output current limit caused by device mismatch. This current limit is severe enough to prevent the affected units from meeting the settling time specification. There are other areas where the design is less than optimum, yet the performance is acceptable.

The most likely cause for the mismatch in the M4An, M4Bn current mirror is threshold mismatch caused by small gate area. This threshold mismatch is magnified by low (Vgs-VT) on these transistors. Increasing the length of all three M4 transistors is called for. This has three beneficial effects. First, the (Vgs - VT) of the M4 current mirror is increased. Second, the gate area of the transistors is increased, so the threshold match is improved. Thirdly, but unrelated to output current, the input referred noise contribution from M4An is reduced. This noise reduction is significant for the OTA because M4A dominates the input referred 1/f noise of the entire amplifier and makes a significant contribution to the input

referred thermal noise. The noise reduction is insignificant for the output buffer, however, whose dominant noise contribution is the sampled contribution from the switch elements. Increasing the length of M4 also has the effect of lowering the frequency of the nondominant pole in the N channel input stage and could be expected to slow the buffer. In this case, however, the zero value time constant contribution from node 4 is not dominant. Furthermore, the nondominant pole of the OTA is determined by the P channel input stage, so a slight reduction in the nondominant pole from the N channel input stage will have negligible impact. Increasing the length of all M4 transistors in the N channel input stage to 5 or 6 μm should have the desired affect of eliminating the mismatch induced output current limit while having negligible effect on the speed of the buffer.

In fact, the speed of the buffer can be improved by a second change. Reducing the length of the M2 current mirror in the P channel input stage will increase the frequency of the nondominant pole of the output buffer. As it stands, the time constant from node 2 in the P channel input stage is dominant while the noise contributions from M2 are negligible. Reducing the length from 30 μm to something in the vicinity of 20 μm is therefore called for.

The N channel cascode in the P channel input stage is not in its own well but should be. Placing it in its own well with the source tied to the well will decrease the threshold voltage and increase the positive output current limit as shown in Section 4.3.4. As it stands the positive output current limit is quite good, but this is one of the areas where the measured performance does not agree with predictions.

Crosstalk performance can be drastically improved by minor rerouting in the layout to avoid the OTA inverting input.

6.3 Suggestions for Further Work

The output buffer presented here meets all of the design goals which were established at its inception. If additional requirements or better performance became necessary the present design can be modified in a number of ways.

In order to achieve a lower output impedance a new design is under way which utilizes a two stage amplifier. This new design is projected to achieve a 300Ω output impedance with an increased power budget of 750 μW.

Another possible modification would be a completely differential OTA. The major advantages of such a design would be lower noise, better PSRR, better CMRR, and higher output range. In the present design, a differential output is provided by having a reference output. The reference output buffer is a duplicate of the signal output buffer but carries a dc signal. With a completely differential OTA, each output of the pair would move with the signal. In the present design there is an increase of $\sqrt{2}$ in noise due to the reference output. In a completely differential design, noise from both outputs is referred back to a single differential input stage. The output range would be doubled because both positive and negative output would be possible. Completely differential amplifiers also have better high frequency CMRR and PSRR. A completely differential OTA is not called for with the present requirements because the noise and output range are already much better than required. CMRR and PSRR requirements are not well established at this point. Another point is that a completely differential amplifier would require a truly differential input. The present focal plane architecture does not provide a differential signal and would have to be modified to accomplish that.

The noise performance of the present design is limited by the sampled noise of the switch elements. This noise was 3.5 times the theoretical noise limit. Investigation into the nature of this noise source would be appropriate. Because

this effect has been observed before[34] it may be of general concern for low temperature CMOS design.

A1. MINOR LOOP STABILITY

The local feedback loop that regulates input stage bias current provides positive feedback. The loop gain can be shown to be less than one at all frequencies if the bias voltage for the cascode transistor is constant [21]. A constant voltage bias is the simplest and has the lowest power consumption. For the case of constant voltage biasing, a small signal model of the input circuit is shown in figure A.1. Small signal transistor models with transconductance, output conductance and gate to source capacitance have been used. Output conductances for the current sources have been included as g_{oA} and g_{oB}. Diode connected transistors have been reduced to resistors. The back gate transconductance is also modeled if g_{m1A}, g_{m1B}, and g_{m3} are understood to represent the sum of the gate to source transconductance g_m and the bulk to source transconductance g_{mb}. Parallel capacitances have been grouped together. Notation is taken to agree with [21].

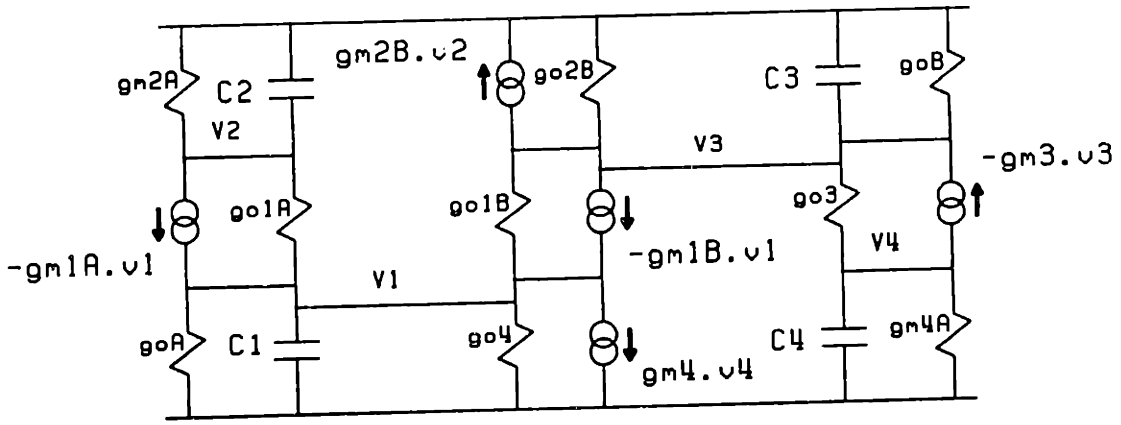


Fig. A.1 Small signal model of local feedback loop

Callewaert and Sansen solve for the loop gain by breaking the feedback loop at the M4A, M4B current mirror. A different new technique is used here. The new technique solves for the loop gain without breaking the feedback loop. The loop gain of a circuit can be defined as

$$LG = 1 - Det(G)/\Pi(Gii) \quad . \qquad (A.1)$$

is derived. Det(G) is the determinant of the nodal conductance matrix, and $\Pi(Gii)$ is the product of its diagonal elements. This equation results when the nodal conductance equations are used to form a linear block diagram for the circuit. In forming the block diagram, each nodal KCL equation is divided by its diagonal coefficient so that the node voltage in question is represented as a linear combination of the remaining circuit node voltages. Equation (A.1) is then the loop gain of the entire diagram[40]

It can be seen from equation (A.1) that the term of the determinant which arises from the product of its diagonal elements will cancel the leading 1 on the right hand side of the equation. The remaining terms are all the nondiagonal contributors to the determinant divided by the product of the diagonal elements. The loop gain will, therefore, have poles at the zeros of the diagonal terms in the nodal conductance matrix. This technique for calculating loop gain is best suited for machine computation. It is used here to demonstrate its use, and to compare results with reference [21]. The nodal conductance matrix for the small signal circuit of figure A.1 is shown in Table A.1 .

$g_{o1A}+g_{o1B}+g_{oA}+g_{o4}+g_{m1A}+g_{m1B}+C_1 \cdot S$	$-g_{o1A}$	$-g_{o1B}$	g_{m4B}
$-g_{o1A}-g_{m1A}$	$g_{m2A}+g_{o2A}+g_{o1A}+C_2 \cdot S$	0	0
$-g_{o1B}-g_{m1B}$	g_{m2B}	$g_{o2}+g_{o1B}+g_{o3}+g_{oB}+g_{m3}+C_3 \cdot S$	$-g_{o3}$
0	0	$-g_{m3}$	$g_{o3}+g_{m4A}+C_4 \cdot S$

Table A.1 Nodal conductance matrix

Each diagonal term determines a pole in the loop gain. Assuming that all output conductance terms are negligible by comparison with transconductance terms, and using the equality of tranconductances of matched devices, the poles of the loop gain can be given as

$$p_1 = -2 \cdot g_{m1} / C_1 \qquad (A.2)$$

$$p_2 = -g_{m2} / C_2 \qquad (A.3)$$

$$p_3 = -g_{m3} / C_3 \qquad (A.4)$$

$$p_4 = - g_{m4} / C_4 . \qquad (A.5)$$

All frequencies are expressed in radians per second. These expressions agree with the corresponding expressions in [21].

The calculation of the zeros in the loop gain is more involved. Of the 23 terms from the determinant that determine the zeros of the loop gain, 7 are nonzero in this case. The polynomial is second order in s so that two zero's exist in the loop gain. The zeros are located by the equation

$$a \cdot S^2 + b \cdot S + c = 0$$

$$a = -C_4 \cdot G_{13} \cdot C_2 \cdot G_{31} - C_4 \cdot G_{12} \cdot G_{21} \cdot C_3 - G_{43} \cdot C_1 \cdot C_2 \cdot G_{34}$$

$$b = G_{43} \cdot G_{14} \cdot C_2 \cdot G_{31} + C_4 \cdot G_{13} \cdot G_{21} \cdot G_{32} - G_{43} \cdot G_{11} \cdot C_2 \cdot G_{34} - G_{43} \cdot C_1 \cdot G_{22} \cdot G_{34}$$

$$\begin{aligned} c = & - G_{43} \cdot G_{14} \cdot G_{21} \cdot G_{32} + G_{43} \cdot G_{12} \cdot G_{21} \cdot G_{34} - G_{43} \cdot G_{11} \cdot G_{22} \cdot G_{34} \\ & + G_{43} \cdot G_{14} \cdot G_{22} \cdot G_{31} + G_{44} \cdot G_{13} \cdot G_{21} \cdot G_{32} - G_{44} \cdot G_{13} \cdot G_{22} \cdot G_{31} \\ & - G_{44} \cdot G_{12} \cdot G_{21} \cdot G_{33} \end{aligned} \qquad (A.6)$$

The above expressions can be simplified considerably by using the fact that the transconductance terms are much larger than the output conductance terms. The bias current is the same for all transistors in the small signal model, so the differences in transconductances are attributable to device type (p or n) mobilities and device geometry (W/L). Because the dependance on mobility and geometry is as a square root, device transconductances are unlikely to differ by as much as an order of magnitude. The difference between a device transconductance and its output conductance on the other hand is likely to be

more than 2 orders of magnitudes. It is valid, therefore, to order terms in powers of gm. All three terms in the expression for a are the same order in transconductance.

$$a = - C_4 \cdot G_{13} \cdot C_2 \cdot G_{31} - C_4 \cdot G_{12} \cdot G_{21} \cdot C_3 - G_{43} \cdot C_1 \cdot C_2 \cdot G_{34}$$

$$= - C_4 \cdot C_2 \cdot g_{o1} \cdot g_{m1} - C_4 \cdot C_3 \cdot g_{o1} \cdot g_{m1} - C_1 \cdot C_2 \cdot g_{m3} \cdot g_{o3} \qquad (A.7)$$

One term in the expression for b is of higher order than the others.

$$b = G_{43} \cdot G_{14} \cdot C_2 \cdot G_{31}$$

$$= C_2 \cdot g_{m3} \cdot g_{m4} \cdot g_{m1} \qquad (A.8)$$

The expression for c has two terms that are of higher order than the others.

$$c = - G_{43} \cdot G_{14} \cdot G_{21} \cdot G_{32} + G_{43} \cdot G_{14} \cdot G_{22} \cdot G_{31}$$

$$= - g_{m3} \cdot g_{m4} \left(g_{m1A} \cdot g_{m2B} - g_{m2A} \cdot g_{m1B} \right)$$

$$= - g_{m3} \cdot g_{m4} \left(- g_{m2} \cdot \Delta g_{m1} + g_{m1} \cdot \Delta g_{m2} \right) \qquad (A.9)$$

With matched devices, the first order expression for c is equal to zero. In equation (A.9) the mismatch quantities are defined as

$$g_{m1} = \left(g_{m1A} + g_{m1B} \right) / 2 \qquad (A.10)$$

$$\Delta g_{m1} = g_{m1B} - g_{m1A} \cdot \qquad (A.11)$$

Completely analogous expressions apply for g_{m2} and Δg_{m2}. A term second order in the mismatch quantities has been neglected in (A.10). A question arises as to whether the neglected terms in equation (A.11) are larger than the term that remains. The answer to this question depends on the degree of match in devices M1A, M1B and M2A, M2B. The neglected terms are of order $g_m \cdot g_m \cdot g_m \cdot g_o$. If included, equation (A.9) can be replaced by

$$c = - g_{m3} \cdot g_{m4} \left(- g_{m2} \cdot \Delta g_{m1} + g_{m1} \cdot \Delta g_{m2} \right) - 2 \cdot g_{m1} \cdot g_{m2} \cdot g_{m3} \cdot g_{o3} + g_{m1} \cdot g_{m3} \cdot g_{m4} \cdot g_{o2} \cdot \qquad (A.12)$$

The mismatch terms are likely to be dominant but all terms are retained for completeness.

The coefficients in the quadratic equation for the zeros of the loop gain have been simplified. The equation could be solved by using the quadratic equation, but the roots turn out to be well separated so that it is easier to approximate their location as

$$z_1 = - c / b$$

$$= \frac{- g_{m2} \cdot \Delta g_{m1} + g_{m1} \cdot \Delta g_{m2} - g_{m1} \cdot g_{o2} + 2 \cdot g_{m1} \cdot g_{m2} \cdot g_{o3} / g_{m4}}{g_{m1} \cdot C_2} \qquad (A.13)$$

$$z_2 = - b / a$$

$$= \frac{C_2 \cdot g_{m1} \cdot g_{m3} \cdot g_{m4}}{C_4 \cdot C_2 \cdot g_{o1} \cdot g_{m1} + C_4 \cdot C_3 \cdot g_{o1} \cdot g_{m1} + C_1 \cdot C_2 \cdot g_{m3} \cdot g_{o3}} \qquad (A.14)$$

The second pole is seen to be at a much larger frequency than the first. If the mismatch terms are dominant in the expression for z_1 then the ratio z_2/z_1 is of order

$$z_2 / z_1 \approx (g_m/g_o)\,(g_m/\Delta g_m)\,. \qquad (A.15)$$

The expression for z_1 agrees with the expression given in [21] for the limit in which the mismatch terms dominate. The higher order terms differ because the approach here is more exact than that given in [21] . Only one zero is given in [21] but two zeros result in this more precise treatment.

The dc value of the loop gain A_{dc} is easily arrived at by setting S=0.

$$A_{dc} = \frac{-g_{m2} \cdot \Delta g_{m1} + g_{m1} \cdot \Delta g_{m2} - g_{m1} \cdot g_{o2} + 2 \cdot g_{m1} \cdot g_{m2} \cdot g_{o3}/g_{m4}}{2 \cdot g_{m1} \cdot g_{m2}} \qquad (A.16)$$

Once again this agrees with the result in reference [21] in the limit $g_o \ll \Delta g_m$, but the first order terms in go disagree. The dc value of the loop gain is much less than one under quiescent conditions, and its sign may be positive or negative depending on the relative weight of the terms involved. The magnitudes of the poles and zeros are in the relation

$$z_1 \approx \Delta g_m / C$$

$$p_i \approx g_m / C$$

$$z_2 \approx (g_m/g_o)\ g_m/C$$

$$z_1 \ll p_1 , p_2 , p_3 , p_4 \ll z_2 \qquad \text{(A.17)}$$

The largest value that the loop gain will achieve will be in the region between the first and second pole. A plot of the magnitude of the loop gain is given in figure 2.3 in which p_2 is assumed to be the dominant pole. The maximum magnitude of the loop gain is then

$$A_{max} = A_{dc} \cdot p_2/z_1 \ .$$

$$= C_2/(2 \cdot g_{m2}) \cdot p_2 = 1/2 \ . \qquad \text{(A.18)}$$

If a pole other than p_2 is dominant the loop gain is found to be even less. The loop gain is always less than one, and the feedback is always stable.

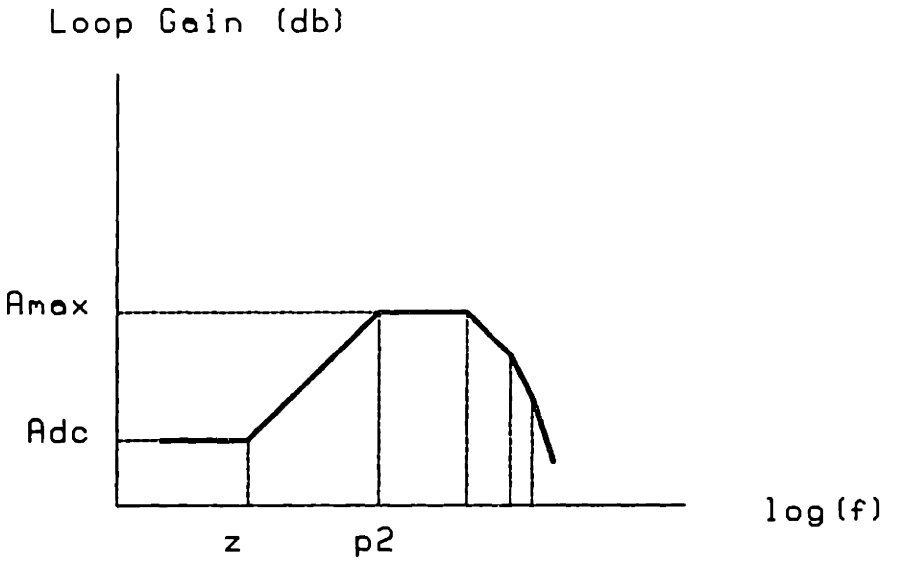


Fig. A.2 Loop Gain

The technique presented here is rather tedious to accomplish by hand. It is done here for demonstration purposes only. It is well adapted for numerical simulation where the nodal matrix for the closed loop configuration is what is at hand. To break a loop in a SPICE simulation in order to calculate loop gain is less attractive than implementing the algorithm described here for two reasons. First and foremost, it is only approximate, because when a loop is broken it is impossible to create the proper loading conditions as well as the proper bias conditions. Secondly, it is extra work just to prepare a circuit in the open loop configuration.

REFERENCES

1. M.T.Chahine "AIRS : The Atmospheric Infrared Sounder" May, 1991.

2. W.S. Boyle and G.E. Smith, "Charge coupled semiconductor devices," *Bell Syst. Tech. J.* (Briefs), Apr. 1970, p.587

3. J.R.Janesick et al., "New advancements in charge-coupled device technology-sub-electron noise ...", *SPIE* vol. 1242 Charge-Coupled Devices and Solid State Optical Sensors (1990) p.223-237.

4. W.F.Kosonocky and J.E.Carnes, "Charge-coupled digital circuits," *IEEE J. Solid-State Circuits*, vol. SC-6, Oct. 1971, p.314-322.

5. W.F.Kosonocky and J.E.Carnes,."Two phase charge-coupled devices with overlapping polysilicon and aluminum gates," *RCA Review*, vol. 34, Mar. 1973, p.164-202.

6. M.H.White et al., "Characterization of surface channel CCD image arrays at low light levels", *IEEE J. Solid-State Circuits*, vol. SC-9, Feb. 1974, p.1-13.

7. H.Kim, "Effects of transistor geometry on CCD output sensitivity", *SPIE* vol. 1242 Charge-Coupled Devices and Solid State Optical Sensors (1990) p.195-203.

8. D.D.Wen, "Design and operation of a floating gate amplifier",*IEEE J. Solid-State Circuits*, vol. SC-9, Dec.1974, p.410-414.

9. R.J.Brewer, "The low light level potential of a CCD imaging array", *IEEE Trans. Electron Devices*, Vol. ED-27, Feb. 1980, p.401-405.

10. R.J.Brewer, "A low noise CCD output amplifier",in *Dig. Tech. Papers, Int. Electron Device Meet* (Washington, DC, Dec. 1978) p.610-612.

11. Y.Matsunaga, et al., "A high sensitivity output amplifier for CCD image sensor", *IEDM tech. dig.*, 1987, p.116-119.

12. Y.Matsunaga et al., "Ultra high sensitivity on-chip amplifier for VLSI CCD image sensor", *1990 Symposium on VLSI Circuits*, p.59-60.

13. S.Oosawa et al., "An electron counting detector for CCD image sensors", Extended abstracts of the 1988 conference on solid state devices and materials,Tokyo, 1988, pp355-358.

14. C.E.Chandler et al., "Sub-electron noise charge coupled devices",*SPIE* vol. 1242 Charge-Coupled Devices and Solid State Optical Sensors (1990) p.238-251.

15. W.C.Black, "A high performance low power CMOS channel filter", *IEEE J. Solid-State Circuits*, vol SC-15, no.6, Dec. 1980, p.929-938.

16. M.A.Copeland and J.M.Rabaey, "Dynamic amplifier for MOS technology",*Electronics letters*, vol.15, no.10, May 79, p.301-302.

17. B.J.Hosticka, "Dynamic CMOS amplifiers", *IEEE J. Solid-State Circuits*, vol SC-15, no.5, Oct 1980, p.887-894.

18. M.G. Degrauwe, "Adaptive biasing CMOS amplifiers", *IEEE J. Solid-State Circuits*,vol SC-17, no.3, Jun 1982, p.522-528.

19. R.Castello et al.,"A high-performance micropower switched-capacitor filter," *IEEE J. Solid-State Circuits*, vol. SC-20, no.6, pp.1122-1132, Dec. 1985.

20. S.L.Wong and C.A.Salama, "An efficient CMOS buffer for driving large capacitive loads", *IEEE J. Solid-State Circuits*, Vol SC-21, no.3, Jun 1986.

21. L.G. Callewaert and W.M.Sansen, "Class AB CMOS amplifiers with high efficiency", *IEEE J. Solid-State Circuits*, Vol 25, no.3, Jun 1990 p.684-691.

22. K.E. Brehmer and J.B. Wieser, "Large swing CMOS power amplifier", *IEEE J. Solid-State Circuits*, vol SC-18, no.6, Dec 1983, p.624-629.

23. B.K.Ahuja et al.,"A programmable CMOS Dual channel interface processor for telecommunications applications", *IEEE J. Solid-State Circuits*,vol. SC-19, no.6 Dec. 1984, p.892-899.

24. Gray,P.R., and R.G. Meyer. Analysis and Design of Analog Integrated Circuits. p.714, John Wiley and Sons, New York, 1984

25. ibid. p.452

26. M.J.M. Pelgrom et al.,"Matching properties of MOS transistors", *IEEE J. Solid-State Circuits*, Vol 24, no.5, Oct 1989 p.1433-1440

27. R.K.Lakshmikumar et al.,"Characterization and modeling of mismatch in MOS transistors for precision analog design", *IEEE J. Solid-State Circuits*, Vol. SC-21,no.6,Dec 1986 p1057- 1066

28. Gray,P.R., and R.G. Meyer. op. cit., p.647

29. The concept of switching two capacitors in and out of the feedback path was first suggested to me by Neil Butler.

30. B.Furrer et al.,"Noise analysis of sampled data circuits",*Arch. Elek. Ubertragung*, Vol. 35, no.11, pp.426-430,1981

31. J.H.Fischer,"Noise sources and calculation techniques for switched capacitor filters",IEEE *J. Solid-State Circuits*,Vol. SC-17,pp.742-752, Aug. 1982

32. R.M.Fox et al., "MOSFET behavior and circuit considerations for analog applications at 77 K",*IEEE Trans. on Elec. Dev.*,Vol ED-34,pp114-123,Jan. 1987

33. S.M.Sze. Physics of Semiconductor Devices. John Wiley and Sons,New York,1981

34. Y.P.Tsividis. Operation and Modeling of the MOS Transistor. McGraw-Hill,New York,1987

35. J.J.Tzou et al.,"Some CMOS Device Constraints at Low Temperature",*IEEE Electron Device Letters*, Vol. EDL-6, No. 1, Jan. 1985

36. K.Takagi et al., "Excess high frequency noise and flicker noise in MOSFETS", *Solid-State Electronics*, Vol. 22, pp. 289-292, Jul. 1978

37. J.Stobie et al.,"High linearity measurement technique for infrared FPA applications", *SPIE* , Aug. 1991

38. J.K.Roberge, Operational Amplifiers : Theory and Practice. p.441,John Wiley and Sons, New York,1975

39. P.J.Cobler, "Low temperature noise analysis of a switched capacitor preamplifier", MIT Masters Thesis, Jan. 1990

40. J.K.Roberge, op. cit. p.44