A Quantitative Voltage Contrast Test and Measurement System.

by

B. Gilhooley, B.Sc.

A thesis submitted to the Faculty of Science
of the University of Edinburgh,
for the degree of Doctor of Philosophy

University of Edinburgh

1984

ABSTRACT

With the shrinking of the minimum feature size of an integrated circuit(i.c.) to 2-5μm and the adoption of high impedance, charge storage circuit technology, the traditional method of tungsten probe testing internal circuit nodes has been rendered impossible. This is a direct consequence of the relatively large probe tip diameter(10-20μm) and the associated probe capacitance(0.05-10pF).

The Electron Beam Probe offers a practical solution to this problem and has enabled qualitative voltage measurements to be made on small geometry devices by exploiting the phenomenon of Voltage Contrast and the Electron Beam Probe's inherent advantages of high spatial resolution(5nm) and high impedance.

A Quantitative Voltage Contrast Test and Measurement System is described and the conditions of device structure under which accurate measurements can be made investigated. The surface electric fields present on a planar i.c. modulate the mean energy of the secondary electrons, emitted when a high energy electron beam is incident on a device. Utilising a high pass energy filter to obtain a collected current proportional to the shift in the mean energy and a feedback loop to alter the filter cut-off energy, a voltage directly proportional to the potential of the probed track was obtained. Since the operating characteristic of the energy filter is monotonic and non-linear with one point of inflection, corresponding to the maxima of the secondary electron distribution, the inflection point was tracked to obtain an operating point insensitive to primary beam current variations. This was achieved by applying sinewave modulation to the filter cut-off energy and suppressing any second harmonic distortion

present in the output by adjusting the mean filter cut-off energy via the feedback loop.

An automatic test system was designed and constructed to provide all the necessary stimuli to an i.c. to execute a functional test. This consisted of a 40 channel exerciser with tri-state outputs, power supplies and a two phase clock. A 16 channel data-acquisition system was also constructed to allow data-logging of additional circuit parameters such as current consumption and digitisation of the measured voltage to enable further off-line signal processing such as averaging. All were programmable via an IEEE 488 instrumentation bus and controlled by a Superbrain microcomputer using the UCSD Pascal language.

ACKNOWLEDGEMENTS.

I would like to thank Dr. A.R. Dinnis for his help and guidance throughout the duration of the work and Dr. P.M. Grant for his useful discussions on the signal processing aspects of the project.

Mr.P.N. Nye's advice concerning the microcomputing and IEEE instrumentation was greatly appreciated, as was the help of Mr.J. Goodall with respect to the SEM.

Finally, I would like to thank the Science and Engineering Research Council for the financial support for this work.

DECLARATION OF ORIGINALITY.

I declare that this thesis has been composed by me and that the work it contains is my own.

(B.Gilhooley)

To my parents who made it possible,

and my wife Tricia,

without whose help and support

it could not have happened.

CONTENTS.

List Of Symbols and Abbreviations.

ADC : Analogue-to-digital converter.

A_o : Open loop gain.

DAC : Digital-to-analogue converter.

DIL : Dual-in-line package.

DSR : Diagnostic shift register.

DUT : Device under test.

E_{pe} : primary electron beam energy.

FET : Field effect transistor.

GPIB : General purpose interface bus.

LSB : Least significant bit.

LSSD : Level sensitive scan design.

MLW : Minimum line width.

MOS : Metal-Oxide-Semiconductor.

M:S : Mark-to-space.

MSB : Most significant bit.

MSI : Medium scale integration.

Q : Quality factor.

RAM : Random access memory.

S : Laplacian operator

SAR : Successive approximation register.

SE : Secondary electron.

SED : Secondary electron distribution.

SEM : Scanning Electron Microscope.

SNR : Signal-to-noise ratio.

SSI : Small scale integration.

VLSI : Very large scale integration.

V_p : Primary electron beam accelerating voltage.

V_{pp} : Peak-to-peak voltage.

V_r : Retardation voltage.

V_s : Surface voltage.

V_t : Threshold voltage.

CHAPTER 1.

INTRODUCTION.

The main emphasis of the research aimed at realising VLSI has been concentrated on the major problems of process development and design methods. However, the task of testing the resultant devices accurately has belatedly been recognised as a coequal problem and one which will be accentuated by increased device complexity. This thesis will present a VLSI test system based on electron beam techniques and the phenomenon of voltage contrast, as applied to digital NMOS devices. This introductory chapter will highlight the extent of the VLSI test problem and discuss the most widely accepted solution with its inherent drawbacks. An outline of the NMOS process and circuit technology will also be given, since these have a direct influence on the design of the electron beam system.

1.1) VLSI DEVICE TESTING.

In microelectronic device testing, the aim is to determine most efficiently if a device is functioning as specified. Secondary constraints are that, the number of tests required to make this decision is a minimum and that these tests are performed as quickly as possible. Torrero(1977) stated that the testing should be performed as soon as possible on the newly fabricated device, as the economic penalty for not detecting a faulty i.c. increased by a factor of ten with every subsequent stage in the manufacturing process.

A general test philosopy has evolved throughout the semiconductor industry to achieve the goals set out in the previous paragraph. For a mature product, the first stage of testing is a DC parametric test, performed at the wafer stage prior to dicing. That is the DC

bias levels within an i.c. are measured and used as a basis to decide if the device will function or not. After the good devices have been bonded and packaged, a Logic Integrity Test is excecuted at a sub-maximum clock frequency to ensure the device is functioning correctly. Finally, a maximum clock frequency AC parametric test is carried out to determine to what extent the device exceeds the lower limit specifications.

If a new product is being introduced into a mature process such as the 5μm, +5V NMOS, and the yield falls then the actual design of the circuit may be at fault and therefore DIAGNOSIS TESTING or FAILURE ANALYSIS has to be introduced. This consists of internally probing and fully evaluating the circuit until the design or processing error is discovered and understood. This is the major area of application of an Electron Beam Test System.

1.1.1) Types Of Testing.

Testing can be categorised into either Parametric or Functional. Parametric implies that an accurate measurement of one or several parameters is carried out, while a functional test involves excecuting a test which has a member of a predefined set as its outcome. i.e. pass/fail or 1/0. When a functional test is carried out on a digital semiconductor device a test vector is applied to the input and the resultant output recorded. This output is then compared to the expected output and a pass/fail flag is generated.

When a DC parametric test is carried out on a device, the output drive, input sense capabilities and power requirements are measured under differing conditions. When an AC test is carried out the device is in dynamic operation and parameters measured include set-up

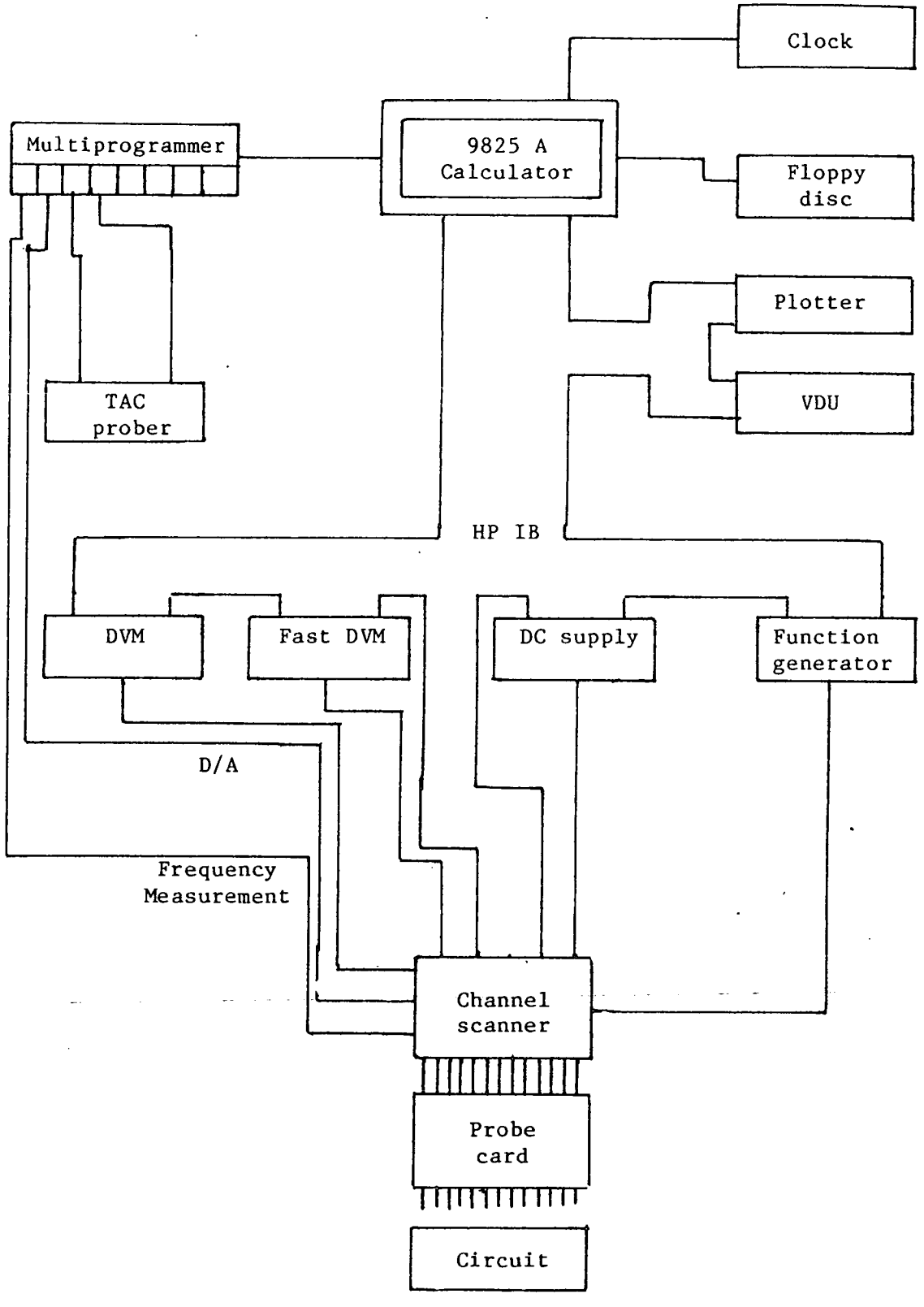


Figure 1.1.1. The Edinburgh Microfabrication Facility parameteric test system. (from Walton, 1982)

time(if the device is clocked), hold times, propagation delays, transition times and pulse widths.

1.1.2) Probe Testing.

Inherent in the need to test devices quickly and efficiently is the ability to access internal circuit nodes. Traditionally, this has been implemented by mechanically positioning a tungsten probe above the node to be accessed and bringing the two into physical contact.

A requirement of probe testing, is that each new circuit design must have its own custom probe card, consisting of up to 70 tungsten probes arranged in a circle around the i.c. The probe card is mounted on a jig, which has stepper motor controlled X, Y and Z direction movement. The jig is originally aligned with a circuit on a wafer and thereafter steps automatically across the wafer. The probes are connected through coaxial cables to a reed relay switching matrix to the appropriate programmable test equipment. The system would be controlled by a low-level high speed computer, using a test language e.g. Automatic Test Language And System(ATLAS). The instrumentation, as illustrated by the system shown in Figure 1.1.1, would typically include a 6.5 digit digital voltmeter, plus/minus power supplies, clock generator, word generator and rate counter.

A scanning electron microscope(SEM) study on a typical tungsten probe is shown in Figure 1.1.2(a)-(c). and demonstrates the normal tip diameter of between 12 and 20μm. The probe diameter varies between 0.25 and 0.5mm, while a probe pad, on the surface of the i.c. of about $360\mu m^2$ is necessary to ensure a good probability of contact. A positional accuracy of 2.4μm can be readily achieved. The system is normally single channel. i.e. only one measurement can be made at the

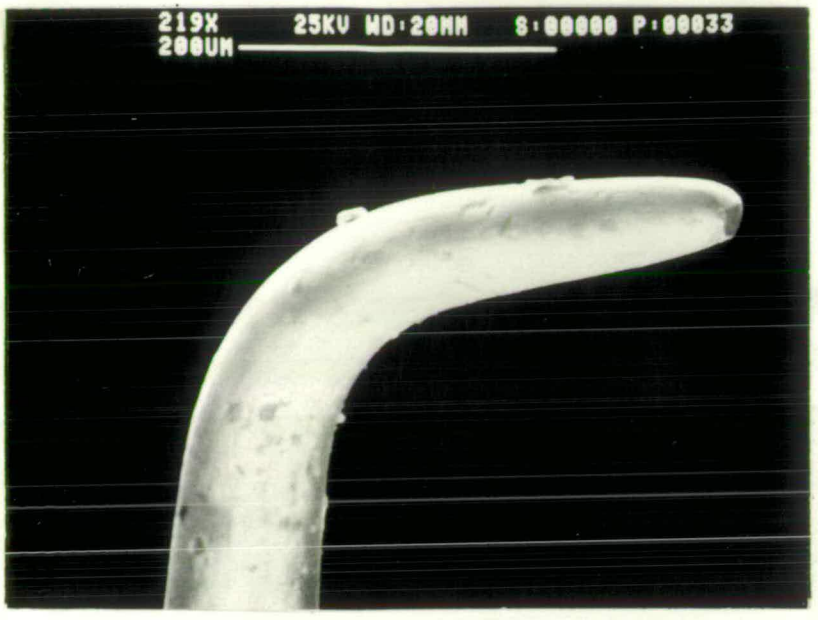
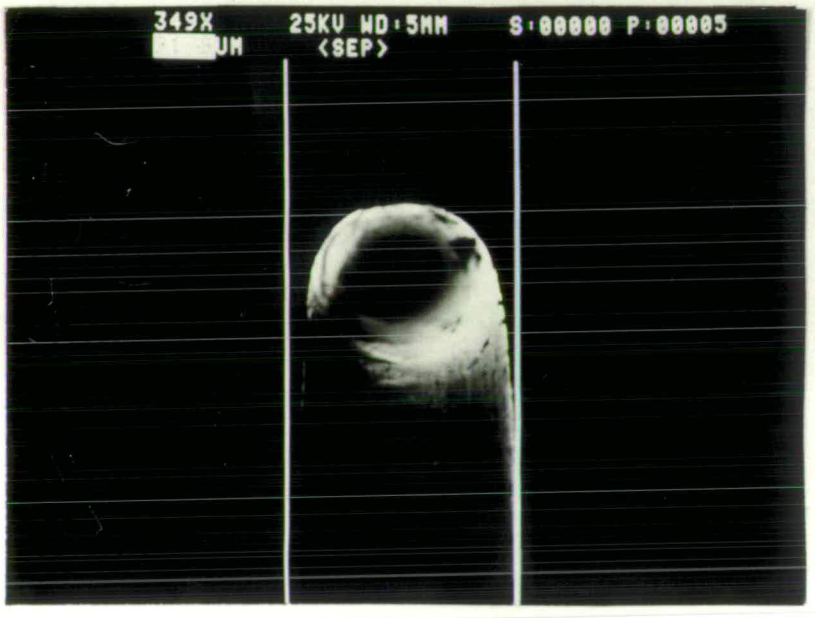


Figure 1.1.2. A study of a tungsten probe.

(a) A typical tungsten probe taken from a probe test station.



(b) Probe diameter measurement using the SEM. Measurement length is that bounded by the vertical lines on the micrograph : 20μm

(c) Probe tip diameter : 12μm

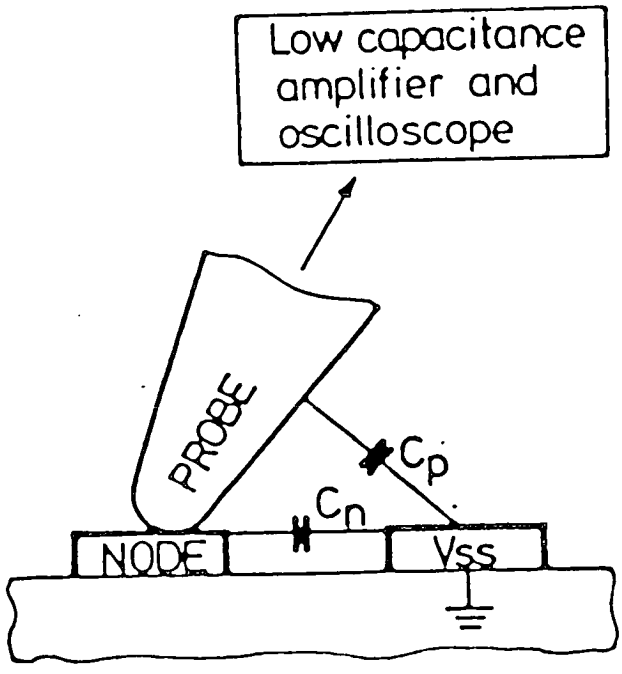
one time. Ardezzone(1974) in his review of probe parameters stated that the probe card is the limiting part of the test system, since its maximum bandwidth is 10MHz at $10V_{pp}$, giving a settling time of 35ns.

Wolfgang et al.(1979) investigated the loading effect of the mechanical probe and found that the capacitance associated with the probe modifies the measured risetime of the waveform, as shown in Figure 1.1.3.(a) and (b). The measured rise time(T_m) can be expressed in terms of the real rise time(T_r), the nodal capacitance(C_n) and the probe capacitance(C_p) as follows:

$$T_m = T_r(1 + \frac{C_p}{C_n}) \qquad (1.1.1)$$

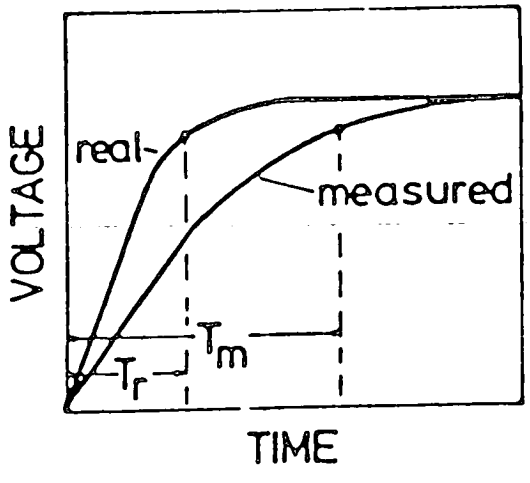
From this it can be seen that accurate measurements can only be made at nodes where $C_n \rangle C_p$, otherwise the waveform would not reach full amplitude and a malfunction would occur in the circuit. Under typical operating conditions of T_r = 10-20ns, $C_p \rangle$ 2pF and C_n = 0.05-10pF only nodes with $C_n \rangle$ 2pF can be probed accurately. The difference in real and measured rise time is illustrated by Figure 1.1.3.(b).

As an example of the time taken for functional and DC test at the wafer level, Torrero(1977) reported that taking wafer, chip registration, up/down and measurement times into consideration an average test time for functional tests is 1 node/second. An additional disadvantage of implementing functional tests at this stage is that the test can be destructive owing to either probe skid and/or punch through of the test node.



(a)

Capacitances associated with probe measurements.
C_n : nodal capacitance.
C_p : probe track capacitance.



(b)

Difference between real(T_r) and measured(T_m) risetimes.

Figure 1.1.3. Node measurement errors caused by probe capacitances. (from Wolfgang, 1979)

1.1.3) The VLSI Test Problem.

As stated by Muehldorf and Savkar(1981), there is no practical method of probing individual logic gates within an LSI chip. This is caused by the minimum device feature size being less than 5μm while the minimum probe tip size is 7-10μm. Further, with the use of high impedance and small geometry circuitry, the loading of the metal probe will become significant and make accurate measurements impossible. An alternative approach is to insert test pads into the important signal paths. However, as stated previously, the test pads have to be about $360\mu m^2$ in order to ensure accurate contact and this is a relatively large portion of a VLSIC in terms of latent processing power.

Exhaustive testing to confirm a circuit's truth table was used successfully for SSI and MSI devices as a means of accessing all the internal nodes of a circuit and testing by inference. However, for a 20 pin input combinational device, the number of total patterns or test vectors is 10^6(Muehldorf and Savkar, 1981). Further, if the circuit is sequential, the number of patterns (N_p) becomes

$$N_p = (2.S)^n \qquad (1.1.2)$$

where S is the number of stable states and n the number of inputs. An upper limit for S is

$$S_{max} = 2^r \qquad (1.1.3)$$

r being the number of feedback lines in a recursive machine. Clearly, this is not a practical alternative for VLSI circuits, since the test times would be excessive and the cost prohibitive. This is emphasised by the graph of computer time for test vector generation

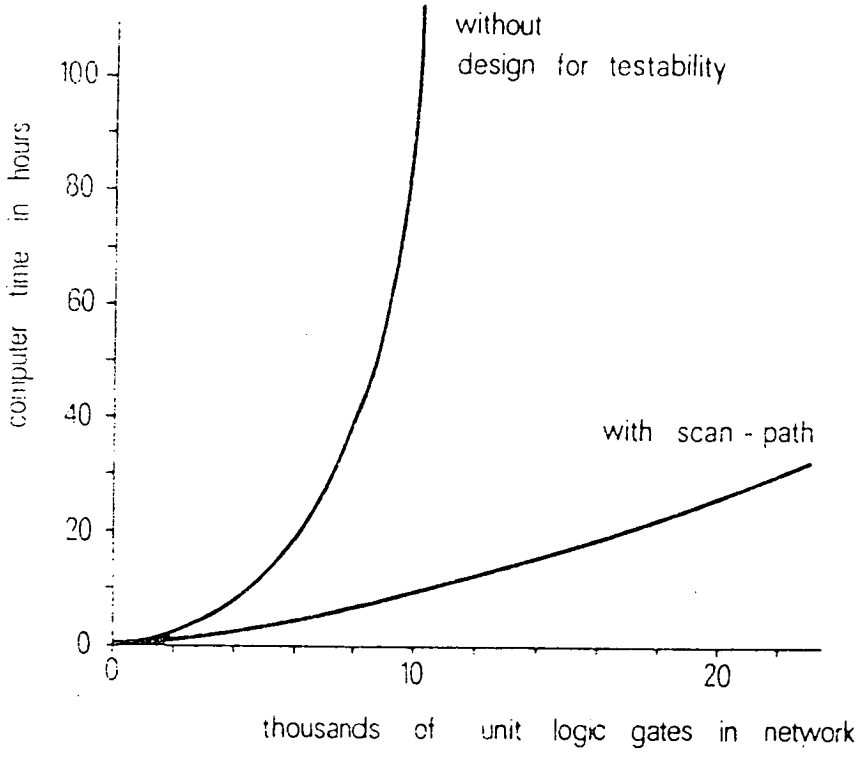


Figure 1.1.4. Growth of test generation cost, expressed in computer time versus the number of logic gates contained in the network (from Stewart, 1978).

against logic gates in Figure 1.1.4. Therefore, there are two major test problems: the inability to physically probe a VLSI device and obtain accurate results and the cost effective generation of test vectors.

1.1.4) Industry's Answer.

The semiconductor industry has attacked the VLSI test problem in two ways for functional test: Firstly, by developing algorithmic generation of minimised test vectors and secondly, by altering their design ethic, to conscienciously design for testability.

In combinational logic, a single input vector may exercise several internal nodes, which may be observable from the output pins. This gives rise to the idea of a minimal number of test vectors exercising a large proportion of the internal circuitry. That is the TEST PATTERN has a high FAULT COVERAGE and consequently the test time for a device can be reduced. A test pattern, with 100% fault coverage, can be generated for a combinational machine by using a formulated optimisation algorithm, such as the D-ALGORITHM (Roth, 1966), which in effect generates a set of minimised test vectors which exercise the internal nodes and propagate the switching action at these nodes to the outputs of the DUT.

Test Pattern Generation for VLSI machines can be either very expensive or even impossible, with the actual choice of vectors being largely empirical. However, large semiconductor manufacturers have approached the problem systematically using a DESIGN FOR TEST APPROACH. The general philosophy of the varying techniques is the same and was first introduced by International Business Machines (Eichelberger and Williams, 1977, 1978) as their LEVEL SENSITIVE SCAN

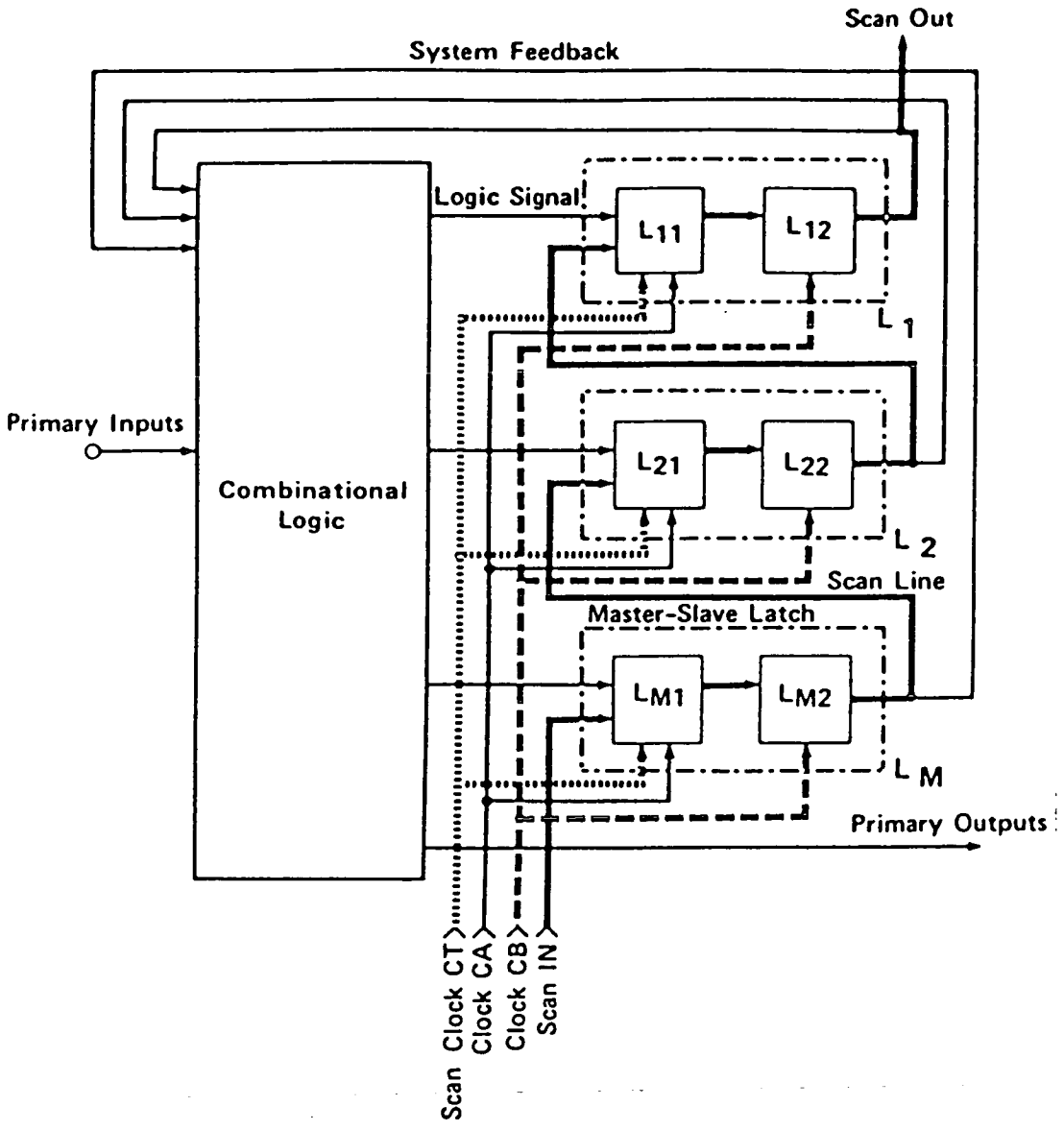


Figure 1.1.5. Structure of logic designed for testability using the Level Sensitive Scan technique (from Eichelberger and Williams, 1978).

DESIGN (LSSD). This approach requires that the machines be both clocked and synchronous (edge triggered), which eliminates any signal spikes upsetting the circuit function. Thus the logic is said to be LEVEL SENSITIVE. All the combinational logic must be partitioned into modules with inputs being derived from bistables and outputs being clocked into bistables as illustrated by the functional block diagram of Figure 1.1.5. That is all feedback paths must be broken to prevent an error being propagated back into the machine. All the bistables within a module must have the ability of linking with all others in a serial fashion to form a Diagnostic Shift Register(DSR), each of which can be accessed externally. The machine must have an automatic reset function which places the device into a known state on power up and when initialised by an external source.

The function of a LSSD can be explained as follows: The machine in normal operation is oblivious to the DSRs and operates as specified. However, when it is forced into the test mode, the circuit clock is disabled, the device reset and the DSRs linked. A test vector is inserted serially into the DSRs(SCAN IN) and the circuit clock enabled. After one clock period the circuit clock is again disabled and the resultant vector shifted serially out by the test clock(SCAN OUT). The device can now be returned to the circuit function mode and the resultant vector analysed. The power of the technique can be realised once it is appreciated that the well understood combinational test generation methods can now be applied to very complex machines and 100% test coverage can be achieved. Owing to the serial nature of the test architecture, additional interface pins are minimal and the additional circuitry is only between 5 and 20%. The main disadvantage of this method is its serial method, which negates

any maximum clock frequency tests being implemented.

1.1.5) Test Summary.

The approach to the test problem adopted has been to test by inference and relatively recently, to dictate design rules which ensure a device is 100% testable. The technique discussed has the disadvantage of necessitating additional circuitry and therefore reducing the latent processing power of an i.c. of a given area. If it was possible to access the nodes internally, the test pattern generation could be simplified since the faults would not have to be propagated to the output pins and thus the cost and time taken to execute a test would be reduced. In addition if nodal testing could be executed without the use of test pads, the circuit design would be less complicated and less silicon area would be used. To this end the research presented in this thesis replaces the mechanical probe with an electron beam probe and exploits the phenomenon of VOLTAGE CONTRAST to perform internal nodal testing.

1.2) NMOS Device Technology.

The N-Channel Metal Oxide Semiconductor(NMOS) process is based on monopolar Field Effect Transistors(FET) and was the process used to fabricate the first generation VLSI devices such as the 256k RAMs described by Fuji et al. (1983) and Fujishima et al. (1983). Although CMOS may supercede NMOS as the process technology used to implement VLSI, owing to the former's superior packing density and thermal dissipation properties, the test system was aimed at NMOS since this type of device was readily available and can be thought of as a subset of CMOS.

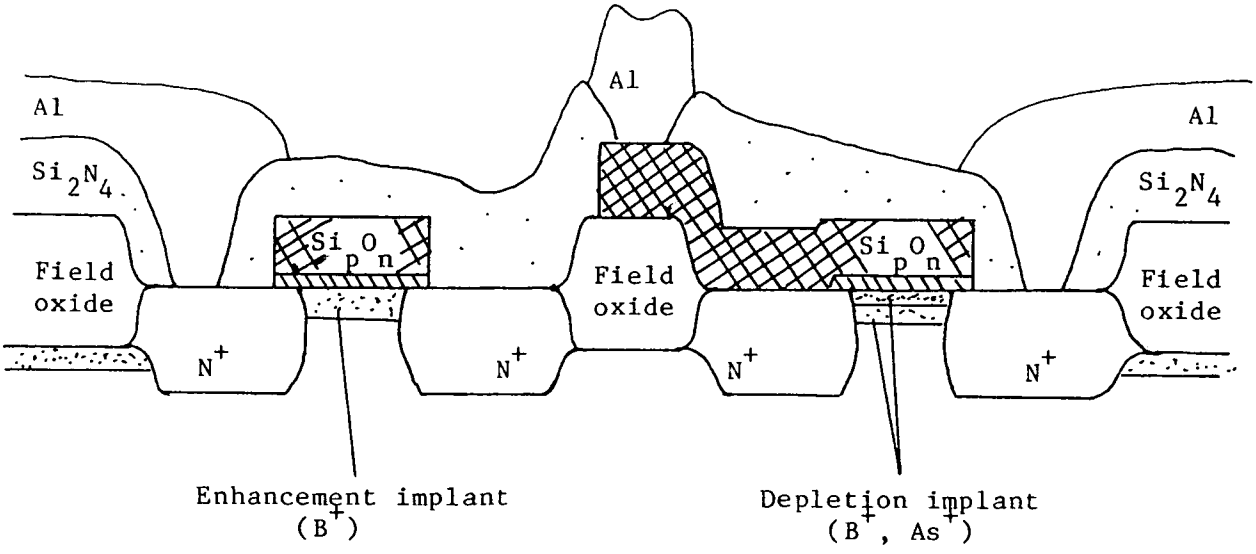


Figure 1.2.1.(a) Cross-sectional view of NMOS inverter.(from Emmerson, 1982)

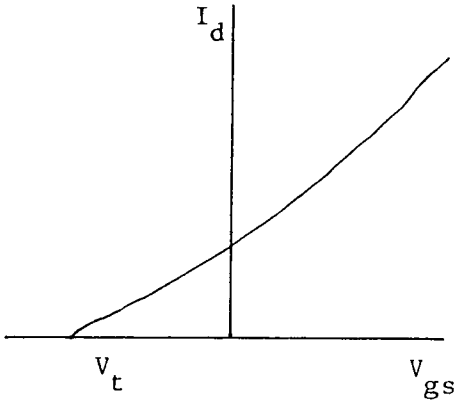


Figure 1.2.2(a) Depletion transistor.

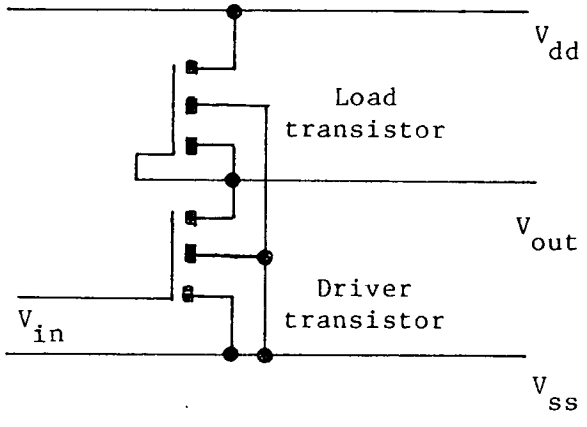


Figure 1.2.1.(b) Schematic of NMOS inverter.

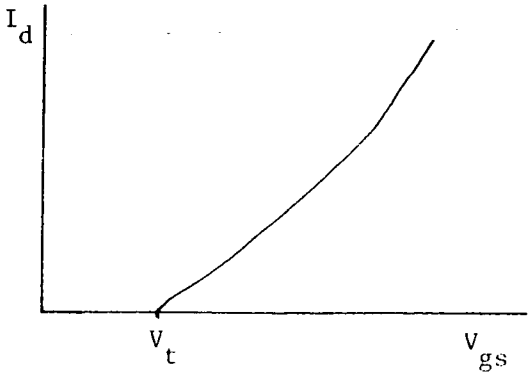


Figure 1.2.2.(b) Enhancement transistor.

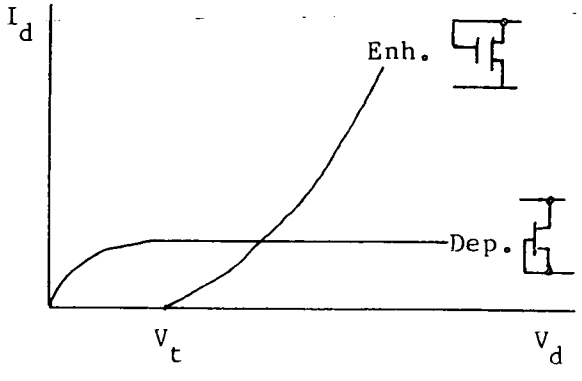


Figure 1.2.2.(c) Load characteristics of transistors.

Figure 1.2.2. NMOS inverter characteristics.

1.2.1) Inverter Device Structure and Characteristics.

The +5V, 5μm polysilicon gate process is at present a mature fabrication method, having superceded PMOS and metal gate NMOS. The +5V refers to the standard digital circuit power supply level and the 5μm refers to the smallest geometric dimension which is termed the Minimum Line Width (MLW), used in the process. i.e. The drain to source well spacing of the smallest FET. The basis of all digital NMOS circuitry is the inverter shown in cross-sectional view and circuit schematic in Figures 1.2.1(a) and (b) respectively. An N channel process is one in which the diffused wells are heavily doped with negatively charged ions such as Arsenic or Phosphorus. The conduction occurs at the silicon-polysilicon interface between the two channels. Direct metal gates were superceded by polysilicon gates owing to the latter's inherent self aligning properties, which resulted in smaller device size and reduced gate capacitance, which in turn, led to increased device switching speed.

The structure is said to be isoplanar since all the conduction mechanisms operate across the device horizontally and not vertically. Devices are isolated from each other via the relatively large field oxides, grown by oxidising the silicon substrate to form silicon dioxide. This layer penetrates deep into the silicon bulk, since the substrate sources the necessary silicon. In addition, the entire device is coated with a protective layer of insulator, such as silicon nitride or silicon dioxide.

Interconnections between conductive channels can be made in three different ways. Figure 1.2.1(a) shows a BURIED CONTACT made between the depletion transistor's gate and source. This is a

polysilicon interconnection to the conductive diffusion channel. Polysilicon is used because it has a low substrate capacitance and more reliable and efficient contacts can be made due to the absence of any etching. The power in the circuits is usually distributed via an aluminium evaporated layer, because the aluminium has low resistivity and therefore, a high power capacity. The aluminium is connected to the the transistor's diffused regions by a BUTTING CONTACT , which can be the source of step-coverage faults and electromigration, causing open circuits. Finally, there may be metal to polysilicon, as occurs when the signal and clock lines are connected to the package pins by way of the bonding pads. The significance of the three types of conductive material will become apparent in Chapter 2.

The inverter comprises an enhancement mode and a depletion mode device, as shown in Figure 1.2.1(b). An enhancement device requires a positive gate voltage with respect to the source to make it conduct from source to drain while the depletion mode can only be turned off by a negative gate voltage, as shown in Figure 1.2.2.(a) and (b). The gate-source voltage(V_{gs}) at which conduction commences is termed, the THRESHOLD voltage(V_t) and is a function of the gate oxide capacitance and bulk semiconductor charge density, among other parameters.

The depletion transistor is realised by using an Arsenic implant beneath the gate to increase conduction. Thus in the inverter circuit the depletion device acts as an active pull-up instead of a resistor, which would require a comparatively large amount of silicon area. Also owing to the fact that the depletion mode transistor takes a constant current, as shown in Figure 1.2.2.(c), power supply problems are suppressed, and faster switching speeds are obtained. Full output

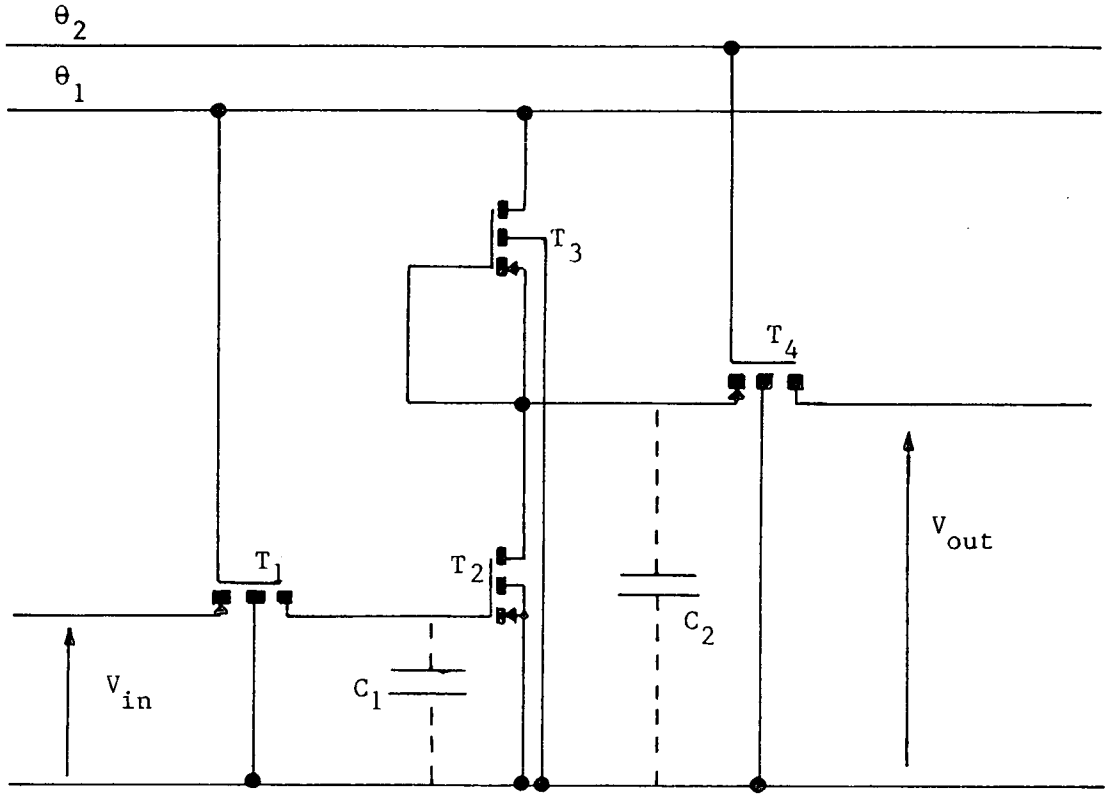


Figure 1.2.3. A Dynamic NMOS inverter using two phase Ratioless Logic.

voltage swing is acheived since the load transistor does not turn off as the load capacitance charges up, as is the case when an enhancement device is used as a load transistor.

1.2.2) Dynamic NMOS.

Digital logic circuitry is founded on the universal properties of the NAND or NOR gates and the ability to form a memory element such as a bistable from these circuits. Inherent in static logic circuitry is the setting of DC bias currents, giving rise to high power dissipation which consequently limit device size. Hence most high density circuits use RATIOLESS DYNAMIC LOGIC , which incorporates two small geometry, high impedance transistors, a clocked power supply and two transmission gates to realise the basic inverter structure. This configuration also has a higher noise margin and a faster operating speed.

The circuit, shown in Figure 1.2.3. does not rely on setting DC bias currents but rather charging and discharging capacitors. The capacitance is not realised by lumped components but utilises the inherent gate capacitance of the transistors (about 0.5pF). Initially, with θ_1 high and θ_2 low the capacitances C_1 and C_2 are charged to V_{in} level. When θ_2 goes high, θ_1 low C_2 is discharged through T_2 which is on. Conversely, if V_{in} is 0 , the gate is zero and C_2 is charged when θ_2 is high, through T_3.

Since the gate capacitance is small, the charge will leak away via the insulation which supports the gate on the drive transistor and to a greater extent through the reverse-biased junction formed between the substrate and the drain of the transmission gate. Typi-

cally, the gate can retain the charge to give a good level for about 1ms, giving a minimum clock frequency for the process of 1kHz. The propagation delay associated with an NMOS inverter is about 20ns, while the transition time is 10ns. For dynamic NMOS the gate delay is determined by the system clock, which for practical systems is about 15MHz.

1.2.3) Process Summary.

The discussion of NMOS device processing has been necessarily brief, since only the aspects which directly affect electron beam test methods have been covered. These aspects can be summarised as:

a) 3 types of interconnect material.
b) Protective passivation insulator.
c) The signal is a stored charge at a high impedance($10^{12}\Omega$) low capacitance(0.5pF) node.
d) The device voltage threshold depends on the gate oxide. (an insulator)
e) The process has a minimum(1kHz) and a maximum(15MHz) clock frequency.

For a detailed discussion of semiconductor processing and FET circuit design techniques Allison(1975), Crawford(1967), and Glasser and Subak-Sharpe(1979) should be consulted.

1.3) Overview.

Concluding from the introductory chapter, an electron beam probe test system would have to satisfy the following specifications to provide a practical means of testing +5V, digital NMOS i.c.s:

1) Spatial resolution $< 1\mu m$.

2) Spatial positioning $< 0.5\mu m$.

3) Settling time $< 1ns$.

4) Dynamic range $> 10dB$.

5) Voltage range $= 0 \rightarrow +5V$.

6) Probe capacitance $< 0.5pF$.

7) Probe impedance $> 10^{12}\Omega$.

8) Fully automated.

Having introduced the problem which stimulated the research presented in this thesis in Chapter 1, Chapters 2 and 3 will describe to what extent the test system specifications have been acheived to date and why an electron beam test system is particularly suited to testing VLSI devices. Although a detailed account of the SEM was thought to be superfluous, the important characteristics are explained in Chapter 2 in conjunction with a brief review of qualitative voltage contrast. An in depth discussion of quantitative voltage contrast as developed to date is given in Chapter 3 and together with Chapter 2 can be considered the review section of the thesis.

The theory of the instrumentation system is set out in Chapter 4, in addition to a description of design philosophy, while Chapters 5 and 6 present the design and performance of the system itself. These three chapters constitute the original work carried out and can be summarised as the development of an original measurement theory and the design and construction of a measurement system based on this theory.

In conclusion Chapter 7 discusses to what extent the work presented has contributed to the solution of this problem and which

areas of research should be pursued to this end in the future.

The two publications resulting from this work are included at the end of the thesis.

CHAPTER 2.

QUALITATIVE VOLTAGE CONTRAST.

Voltage Contrast was first observed by Oatley and Everhart(1957) using the SEM and since then has evolved into a sophisticated tool with its main application being in the diagnostic testing of i.c.'s. This chapter will briefly review the instrumentation techniques, developed to enhance the usefulness of Qualitative Voltage Contrast, which are directly applicable to Quantitative Voltage Contrast and the pertinent details of the SEM.

2.1) The Scanning Electron Microscope.

Although the SEM is not an integral part of an Electron Beam Test System, it does provide the necessary finely focussed electron beam in a convenient and ready made form. Since there are numerous texts on the principles of operation of the scanning electron microscope (Oatley 1972, Wells 1974, Thornton 1968, Hearle et al. 1972), only essential aspects will discussed. A general block diagram of an SEM is given in Figure 2.1.1.

The electron beam formed by the SEM is the electron crossover present at the electron gun demagnified by the magnetic lens system and can be as small as 5nm. This is more than adequate for probing i.c. conductors. The current sourced by the SEM can vary between 10 and 100μA. The beam is scanned in two spatial dimensions by the magnetic fields produced by the scan coils and if the monitor has identical coils, the magnification is determined by the ratio of the currents in the coils. The final lens controls the focal length of the electron optical system and is about 5mm for optimum SNR, while the two preceding lenses control the resolution or spot diameter.

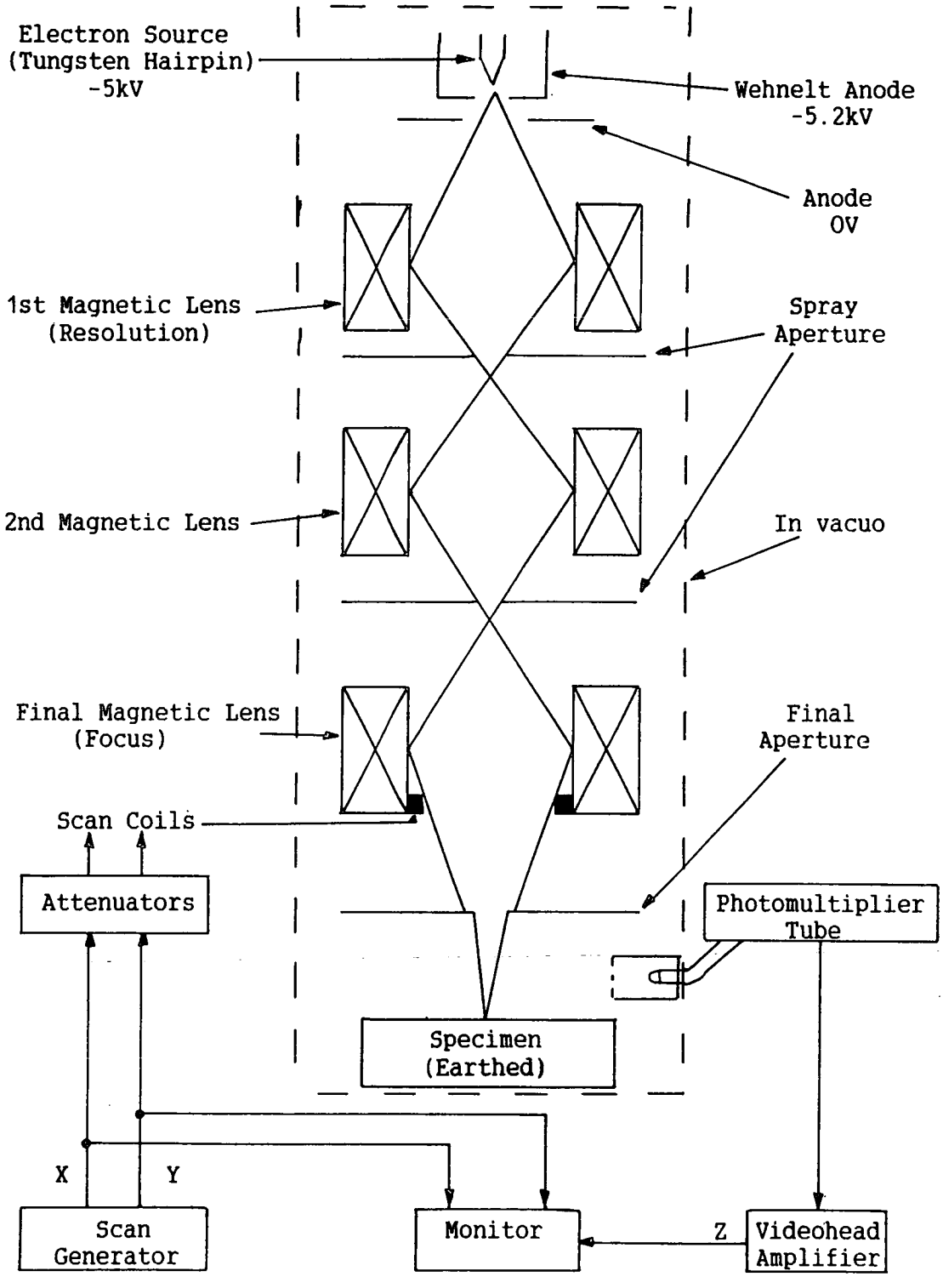


Figure 2.1.1. The Scanning Electron Microscope.

When a high energy beam of electrons interacts with a bulk specimen, among other interactions, electrons are emitted from the specimen. The intensity of these electrons is dependent upon the surface topography and composition and therefore, if these electrons are collected and used to modulate the intensity of a cathode ray tube, a two-dimensional image will be formed. The electrons are collected by a Faraday cage, which is held positive with respect to the specimen. When the electrons enter the cage, they are accelerated onto a photo-scintillator coated light pipe, which is held at 10kV for maximum conversion efficiency. The signal is transferred from the vacuum of the specimen chamber to a photomultipier, within which it is amplified. Finally, the current signal is converted into a voltage by the high gain transimpedance videohead amplifier. This signal is then used to modulate the luminance of the monitor.

Owing to its high resolution and depth of focus, the SEM in its own right has become an indispensable tool in the areas of quality assurance and inspection within the semiconductor industry. It is mainly used for qualitative feedback from topographic images for process control purposes(Nicolas, 1974), monitoring such parameters as step-coverage, electro-migration and pin-hole gate-oxide failures, for which optical microscopes do not have the necessary resolution. The three micrographs shown in Figures 2.1.2(a)-(c), illustrate the magnification , resolution and depth of focus obtained with the SEM for inspecting a process control i.c.

2.2) Secondary Electron Emission.

Although it is unnecessary to understand the interaction mechanism of the electron beam and the solid in great detail, certain

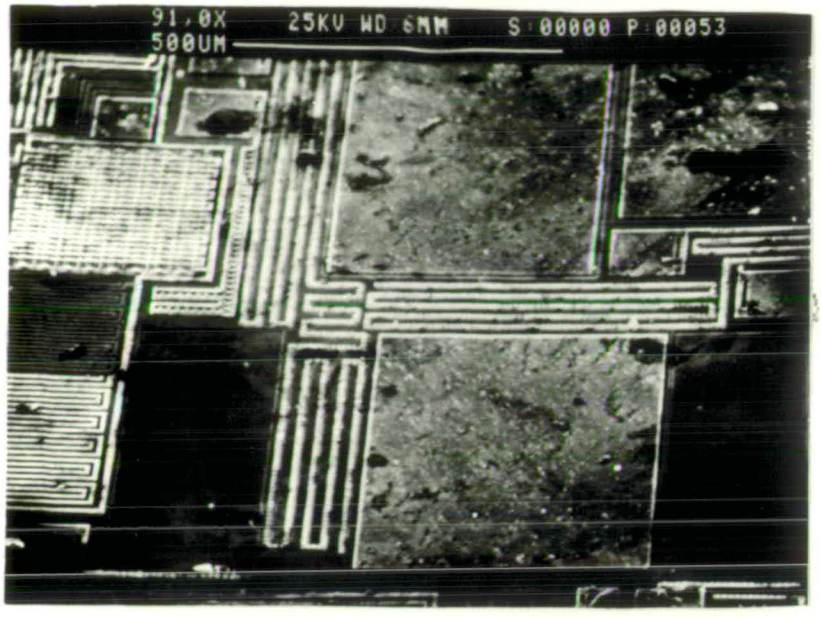
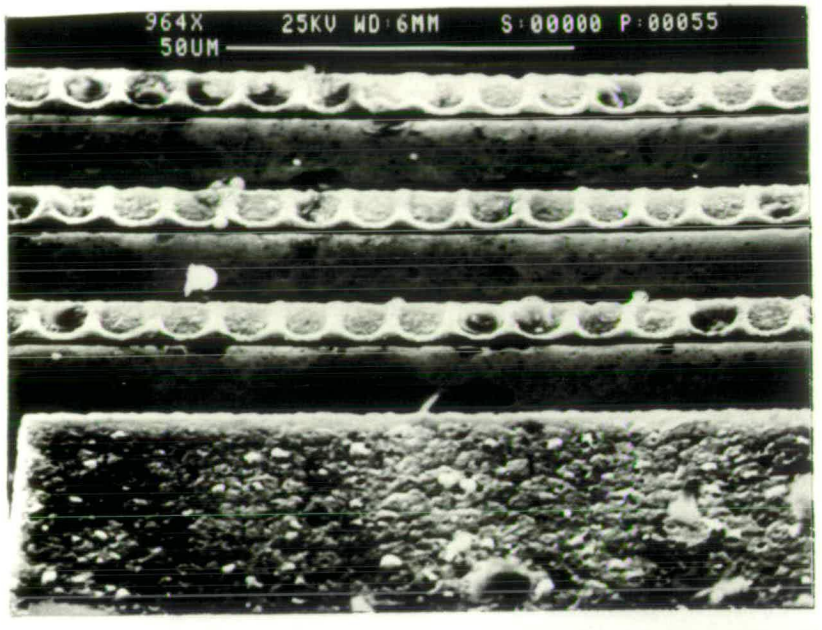
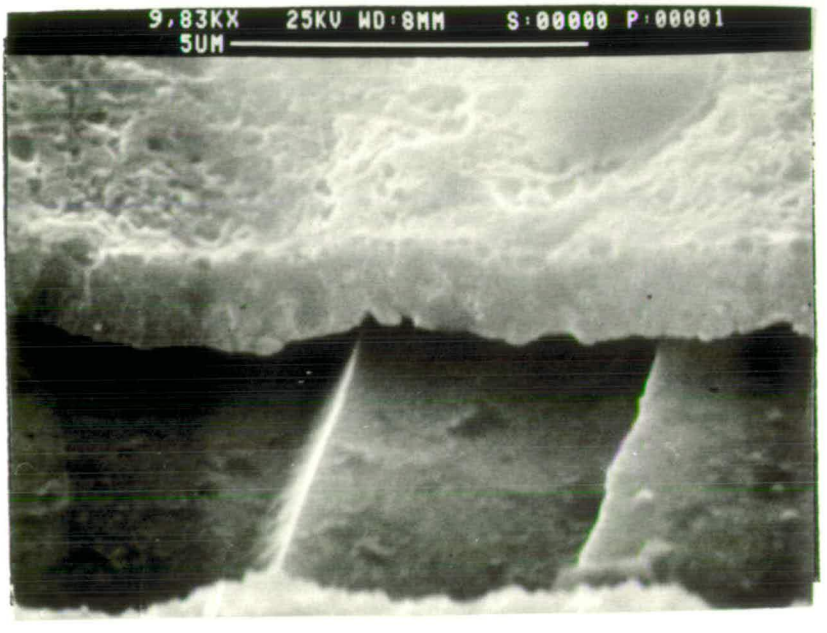


Figure 2.1.2. A study on a process monitoring i.c., inspecting contact chains (centre of (a) for step coverage continuity

(a) Magnification : 10^2



(b) Magnification : 10^3



(c) Magnification : 10^4

Note over-etching of SiO_2 giving undercutting of aluminium.

characteristics of the electrons emitted are important. For a more exhaustive discussion of secondary electron(SE) emission McKay(1948), Bruining(1954) or Gibbons(1966) should be consulted.

2.2.1) Energy Distribution.

The energy distribution of the emitted electrons is shown in Figure 2.2.1., illustrating a high energy peak, corresponding to the specimen's atomic composition and the low energy group, which is normally used to form the topographical image. Electrons with energies below 50eV are arbitrarily termed SEs and those above 50eV, backscattered electrons.

The maximum of the secondary electron distribution(SED) can be attributed to the manner in which the electrons escape from the surface of the specimen. The minimum energy required for an electron to escape the bulk is the work function of the material. Therefore, electrons with greater energy than the work function are more likely to escape and hence the increase in the number of electrons emitted as a function of energy. However, high energy electrons must be close to the surface to escape since their high energy will make them more likely to collide with other electrons or atoms, if generated in the bulk of the specimen. This would result in more low energy electrons being generated. Hence the reduction in the number of electrons collected as a function of energy.

The maximum corresponds to the crossover energy when the latter mechanism dominates the former. The amount of high energy electrons decreases since, the high energy electrons generated within the bulk of the material are more likely to undergo collisions producing low energy tertiary electrons, which cannot escape from the material.

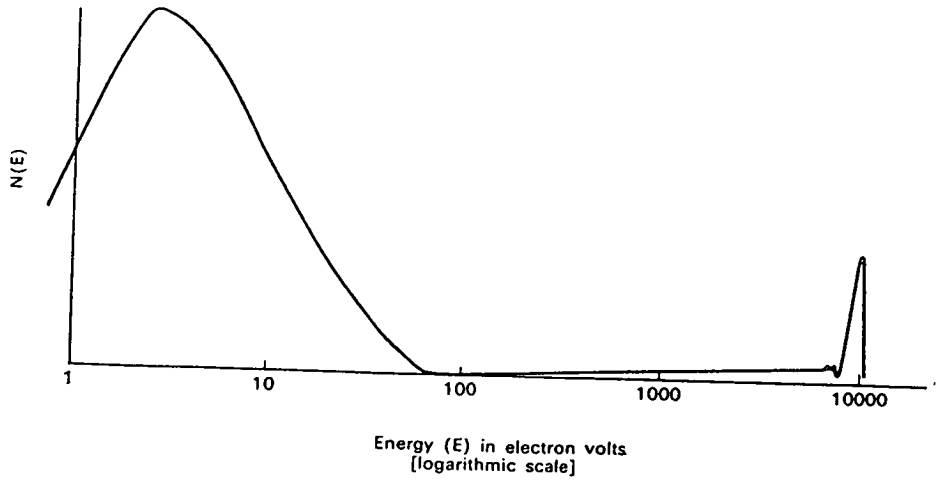


Figure 2.2.1. Energy distribution of secondary electrons induced by 10keV primaries (from Hearle et al., 1972)

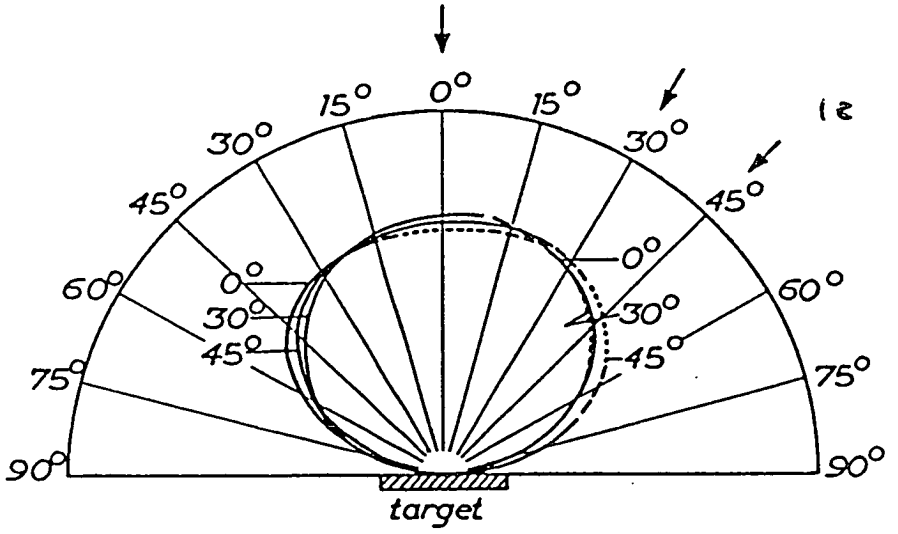


Figure 2.2.2. Angular distribution of seconday electrons (5-50eV) from polycrystalline nickel (V_p = 100V) (from Jonkers, 1912)

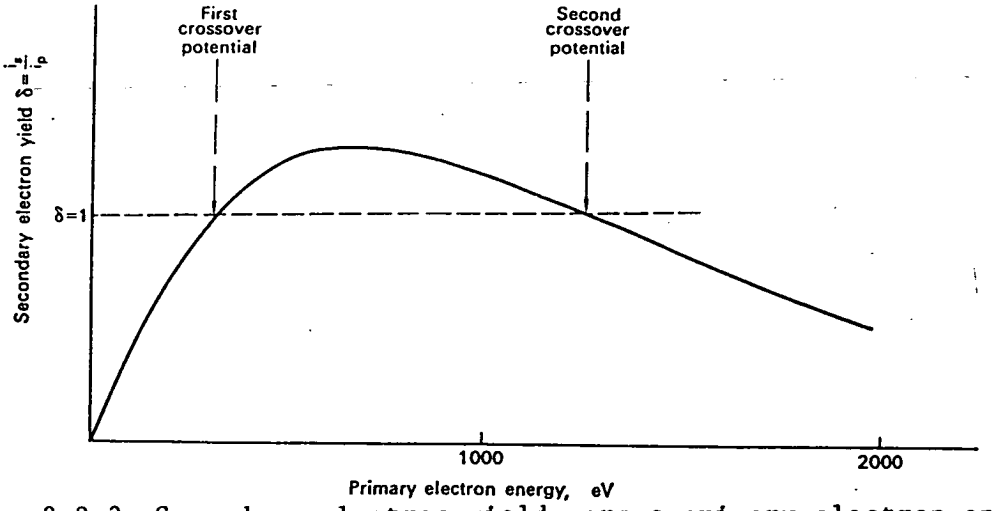


Figure 2.2.3. Secondary electron yield versus primary electron energy (from Hearle et al., 1972)

The SED has a similar shape for all materials with only the peak energy and intensity of electrons varying. However, a universal expression has not yet been formulated.

2.2.2) Angular Distribution.

As shown by Jonkers(1912) in Figure 2.2.2. the polar SE intensity distribution is generally cosinusoidal with a homogeneous polar energy distribution.

2.2.3) Yield.

The total number of electrons emitted(N_{tot}) can be obtained by integrating the SED(N(e)) weighted cosinusoidally over three dimensions:

$$N_{tot} = \int_{\omega=0}^{2\pi} \int_{\theta=0}^{\frac{\pi}{2}} \int_{e=0}^{50eV} N(e).de.Cos\theta.d\theta.d\omega \qquad (2.2.1)$$

The SE yield(δ) for a material is defined as

$$\delta = \frac{(\text{No. o f secondary electrons emitted})}{(\text{No. o f primary electrons})} \qquad (2.2.2)$$

The yield is dependent on the primary electron beam accelerating voltage and can be both greater and less than unity. The yield-energy curve, as shown in Figure 2.2.3, has the same general characteristic for all materials, but the energy values for unity yield vary. These values are important since they imply that the specimen will neither charge or discharge and therefore, the electron probe will not electrically load the specimen.

Feuerbaum(1979) plotted the current loading of an electron beam on an aluminium track as a function of primary beam energy(E_{pe})

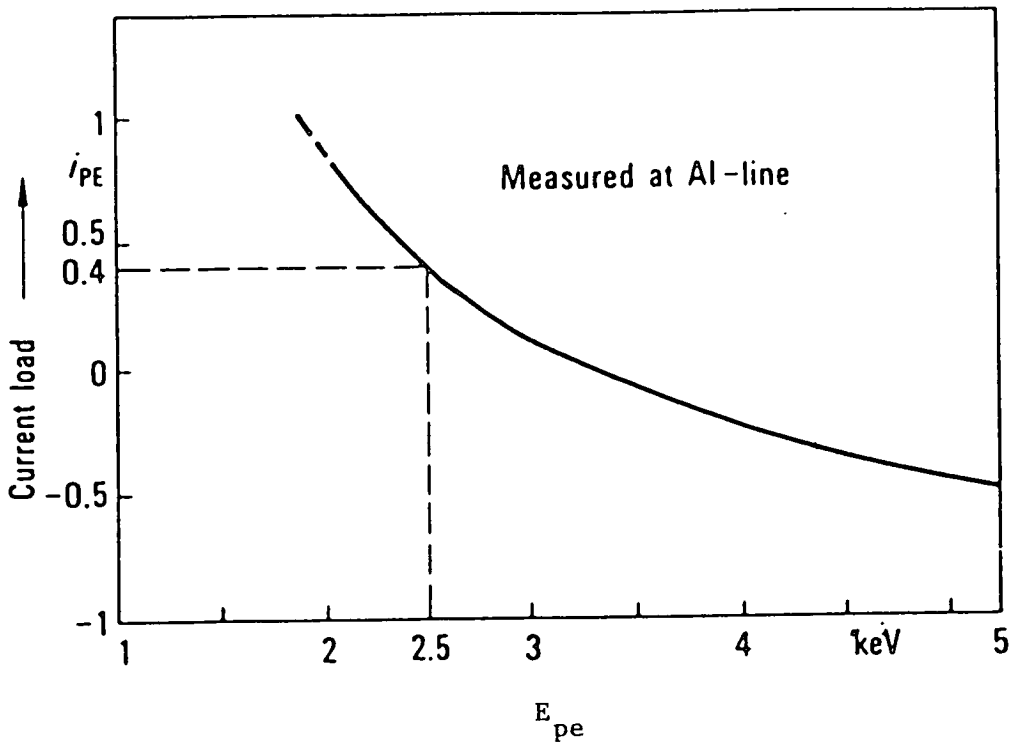


Figure 2.2.4. The current load of an electron beam probe as a function of the primary beam energy (from Feuerbaum 1979)

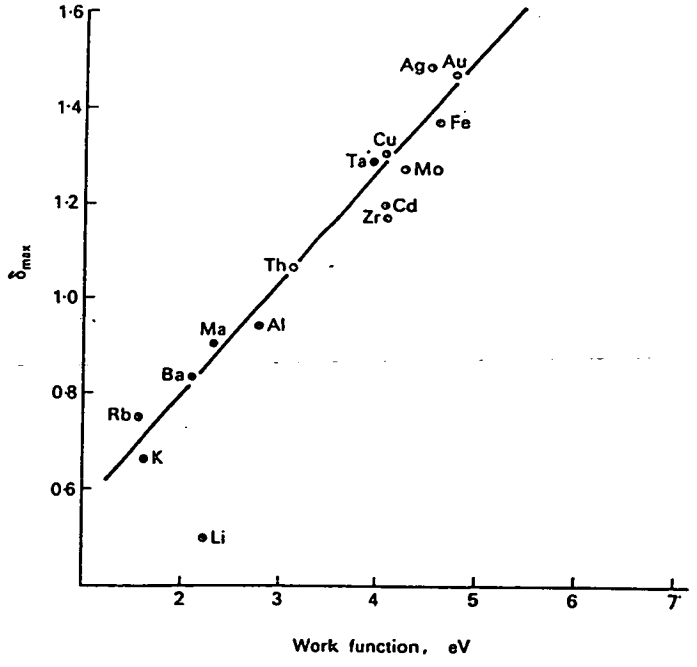


Figure 2.2.5. The variation of δ with work function of pure metals (from Hearle et al., 1972)

(Figure 2.2.4) and found that although the probe was nonloading at about 3keV, operation was preferable at 2.5keV to minimise the charging of surrounding field oxide.

As shown by Figure 2.2.5, the maximum Yield is dependent on the work function of the material and is therefore another source of contrast from which voltage contrast has to be isolated. Note especially the Yield for aluminium(0.9) and for gold(1.4).

2.2.4) Penetration Profile.

The penetration of the primary beam into the specimen can be an important parameter, as will be discussed in chapter 3. Everhart and Hoff(1971) investigated the penetration of the primary beam into a MOS capacitor as a function of primary beam energy as shown in Figure 2.2.6 and formulated the projected range(R_g) as

$$R_g = 4.E_{pe}^{1.75} \qquad (2.2.3)$$

The figure demonstrates that even with a primary electron beam energy of 6kV, some electrons will penetrate the gate oxide of a FET and modify the device characteristics. The units of projected range are mass thickness i.e. the product of material density and thickness($\mu g/cm^2$).

2.3) Voltage Contrast.

Qualitative voltage contrast manifests itself as the modulation of the brightness of the video monitor of the SEM, by an applied specimen voltage. This phenomenon is demonstrated by the two micrographs, Figures 2.3.1(a) &(b) displaying a gold bond wire earthed(a) and with a $5V_{pp}$ square voltage waveform applied(b). The frequency of

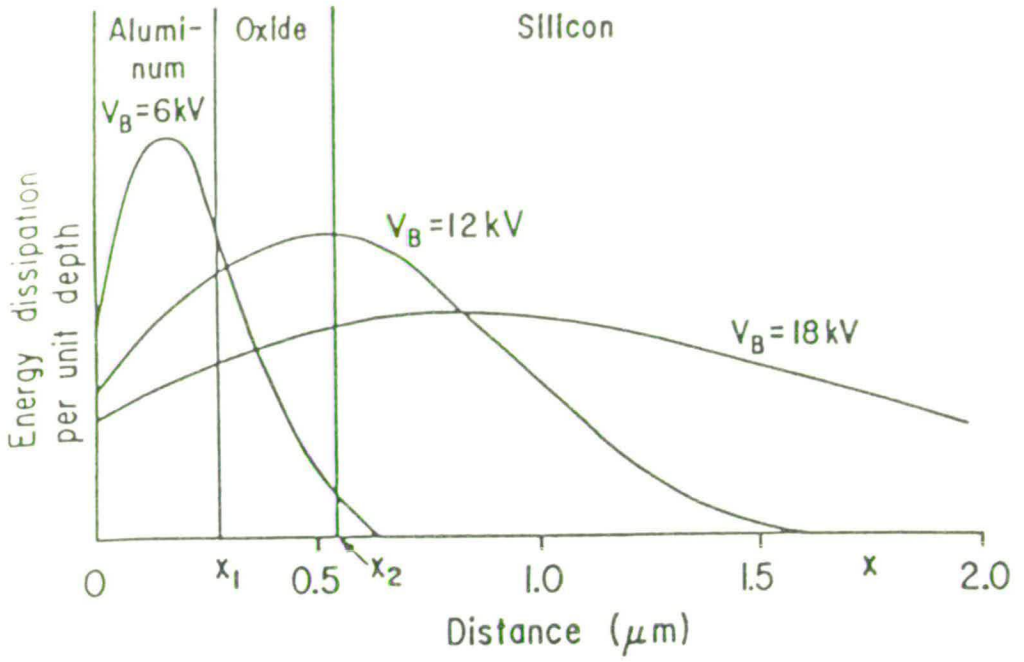
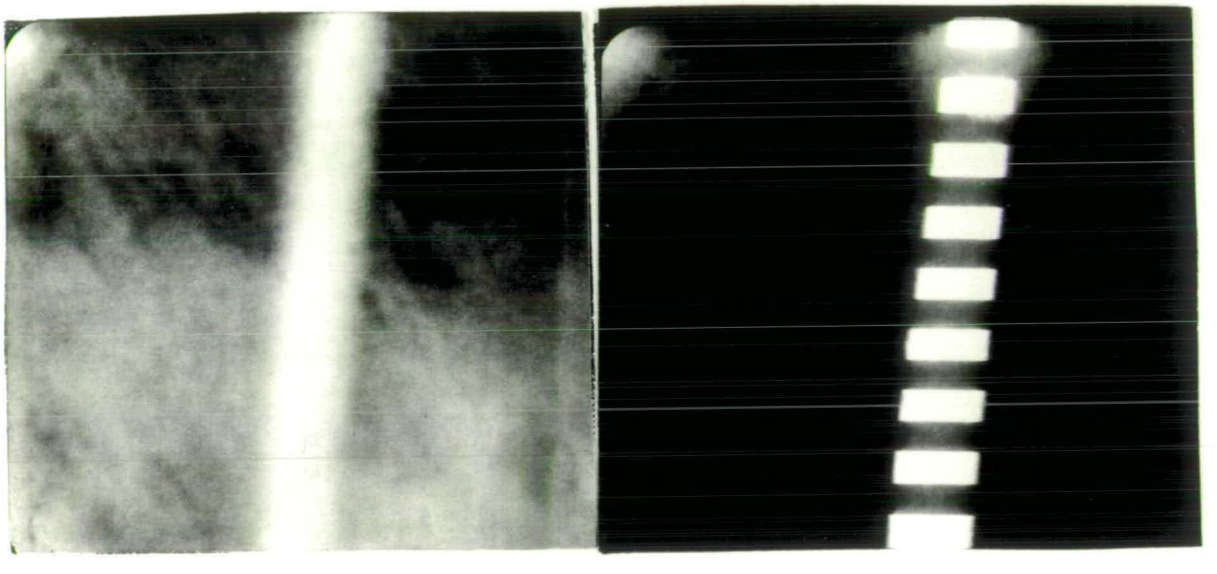


Figure 2.2.6. Penetration profile as a function of primary beam voltage for a MOS capacitor. (from Everhart and Hoff, 1971)



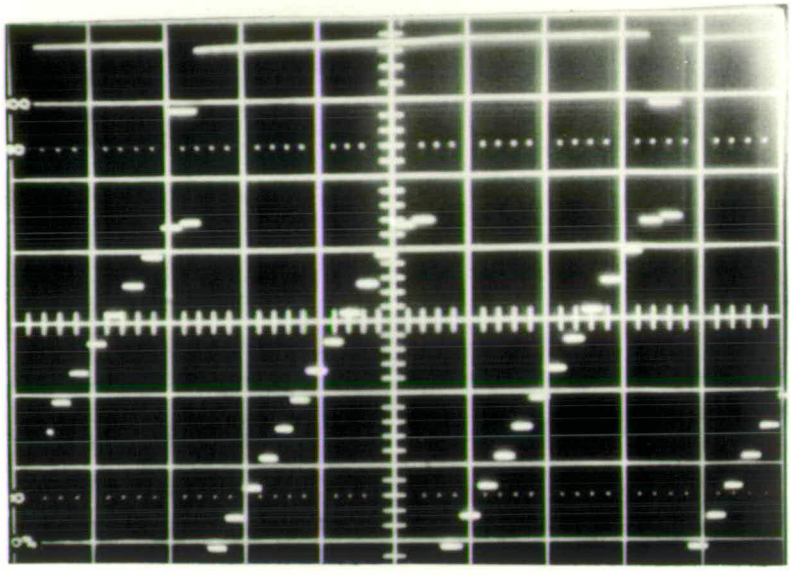
(a) (b)

Figure 2.3.1. Qualitative Voltage Contrast demonstrated by micrographs taken with a Cambridge Instruments' Stereoscan MKII SEM
(a) Gold bond wire held at 0V.
(b) Gold bond wire with a 5V amplitude square wave applied.

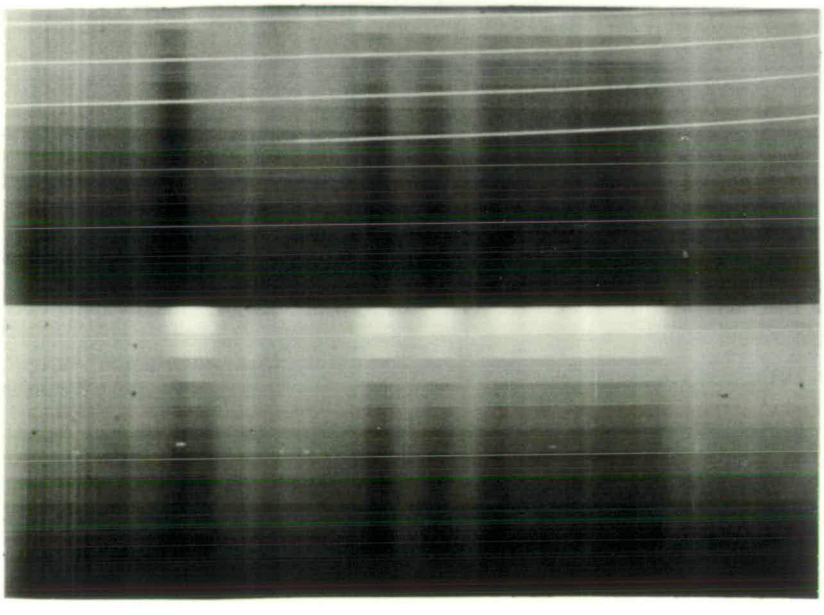
the applied waveform is greater than the field scan frequency but less than the linescan frequency, giving rise to a straition effect displayed in Figure 2.3.1(b). The variation of the SEM's luminance signal is attributed to a change in the number of collected electrons or detector current. This in turn can be explained by considering the interaction of the electrostatic field of the standard Everhart-Thornley(1960) detector, the dynamic electric fields present on the surface of a planar integrated circuit and the low energy SEs emitted by the bulk semiconductor.

When a negative voltage is applied to the specimen, the SEs are repelled from the surface of the specimen and diverted into the detection system by the relatively weak positive electric fields, giving a large collected current and therefore, a bright spot on the monitor. Alternatively, when a positive voltage is applied to the specimen, the electrons are attracted back to the specimen, resulting in a low collected current and a corresponding dark spot on the SEM's monitor.

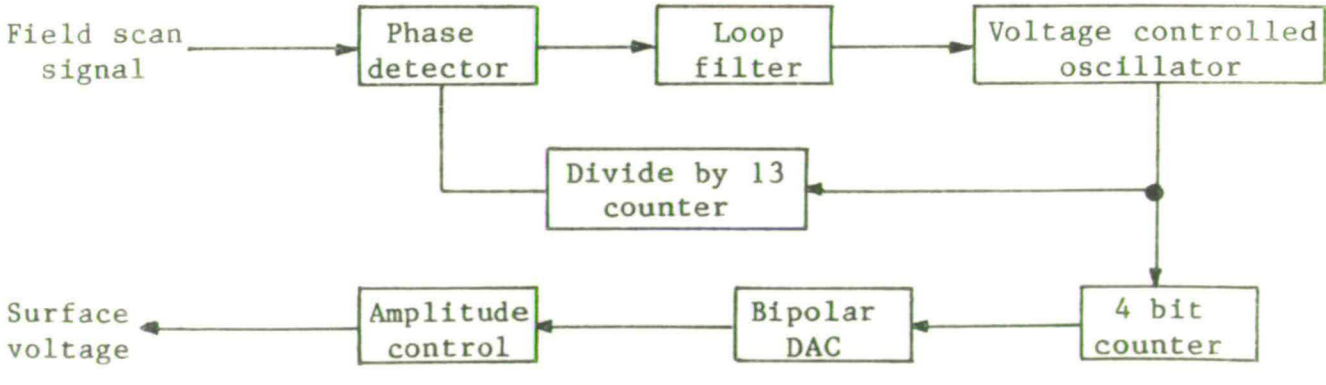
A corollary of this might be that the contrast mechanism must be bistable. i.e. the collected current is either zero or a maximum value. However, the collected current is not only a function of the polarity of the surface fields but also the relative strengths of the surface, detector and stray specimen chamber fields in conjunction with the energy and polar intensity distributions of the secondary electrons. This gives rise to a complex and nonlinear relationship between the collected current and the applied specimen voltage. However, it is this basic phenomenon which is exploited to make quantitative surface voltage measurements.



Upper trace : Scan generator's field synchronisation waveform
Lower trace : Applied surface voltage. Both traces 5V/cm, 20ms/cm



SEM monitor's output, illustrating a nonlinear response. Note the applied voltage is synchronised to twice the field scan frequency.



Schematic of the instrumentation required to obtain above traces.

Figure 2.3.2. Nonlinear Voltage Contrast.

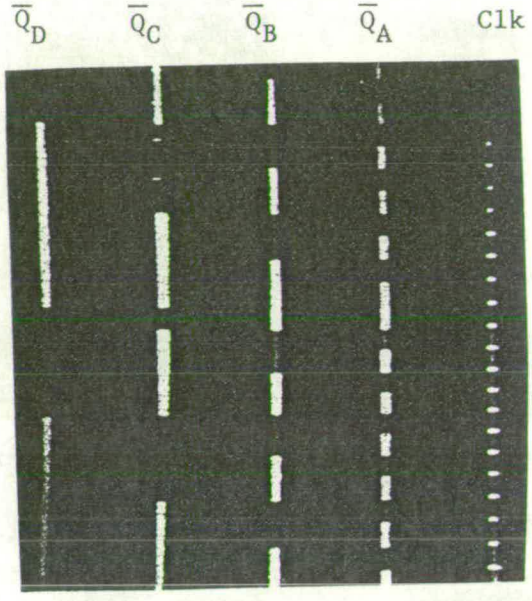
The nonlinear relationship is demonstrated visually in Figure 2.3.2, which shows the SEM monitor's response when a voltage staircase waveform is applied to a specimen in synchronism with the monitor's field frequency. The brightness of the micrograph should increase in 12 well defined steps if the relationship was linear. However, it is difficult to differentiate between levels at either end of the scale even allowing for the phosphor's inherent nonlinearity.

2.4) Instrumentation Techniques.

The fact that the SEM can produce a two dimensional surface voltage intensity modulated map of an i.c. is in itself a very useful tool to the device tester, since it allows the location of open and short circuits which cannot be resolved by either an optical or electron microscope. However, various instrumentation techniques have been developed to enhance this basic mode of operation and will now be discussed.

2.4.1) Voltage Coding.

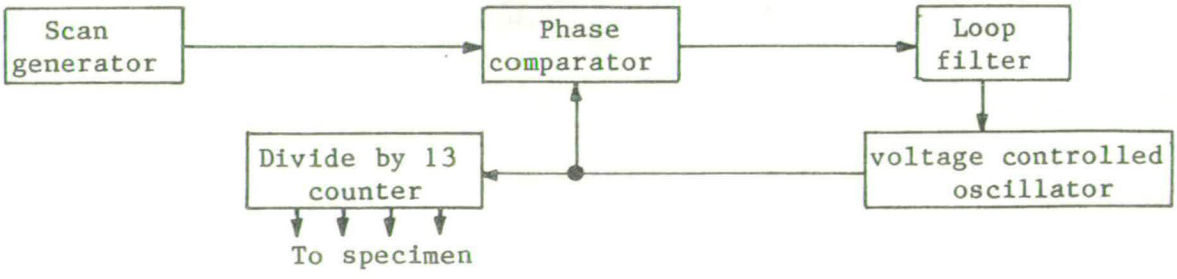
Qualitative Voltage Contrast can be enhanced by synchronising the DUT's clock frequency with the SEM's field scan frequency, after Lukianoff(1975), producing a stationary image with black and white striations which correspond to the clock and linescan beat frequency. An example of this technique is shown in Figure 2.4.1, in which a two dimensional display of the truth-table of a divide by 13 counter is shown. The counter was a four bit synchronous programmable type with feedback to the parallel load input. The clock, derived from the field frequency of the scan generator by a phase locked loop, is shown on the extreme right and the most significant bit(MSB) to the



Visual output obtained from the SEM monitor

Decimal Equivalent	Q_D	Q_C	Q_B	Q_A	Clock Cycle
5	0	1	0	1	1
6	0	1	1	1	2
7	0	1	1	1	3
8	1	0	0	0	4
9	1	0	0	1	5
10	1	0	1	0	6
11	1	0	1	1	7
12	1	1	0	0	8
13	1	1	0	1	9
14	1	1	1	0	10
15	1	1	1	1	11
3	0	0	1	1	12
4	0	1	0	0	13
5	0	1	0	1	14

Logic truth table derived from the above display.
i.e. a divide by 13 synchronous up counter.



Schematic of instrumentation required to obtain above display

Figure 2.4.1. Voltage Coding.

least significat bit(LSB) from left to right. This type of image could be exploited in tracing internal dynamic signals and identifying latched transistors in a dynamic signal path.

This technique is termed VOLTAGE CODING , since the time varying voltages are coded into the spatial domain of the image. The main disadvantage of this type of display is that it is only intelligible to the human eye if the clock frequency is well below the linescan frequency, which is about 15kHz for a television rate display. Thus this technique can only be applied for a small frequency window to dynamic circuits.

2.4.2) Stroboscopy.

Plows and Nixon(1968, 1969) introduced STROBOSCOPY to the SEM, to observe the voltage intensity maps of high frequency i.c.s. As illustrated by the block diagram in Figure 2.4.3., this technique is implemented by pulsing the primary electron beam in synchronism with the signal applied to the specimen and thus constructing a two dimensional voltage map of the specimens surface at a particular time instant. By varying the relative phase of the primary beam pulse and the specimen signal, voltage maps of the entire specimen at different time instants can be viewed. This is a useful technique since a voltage map of the entire circuit can be clocked through and evaluated. The clock frequency of the DUT must be much greater than the linescan frequency of the SEM to ensure that enough samples are taken on every linescan to form an image.

The primary beam pulse must have a small mark-to-space ratio(M:S)($<$0.1) in order to approximate an impulse function in comparison with the specimen signal. This approximate impulse function

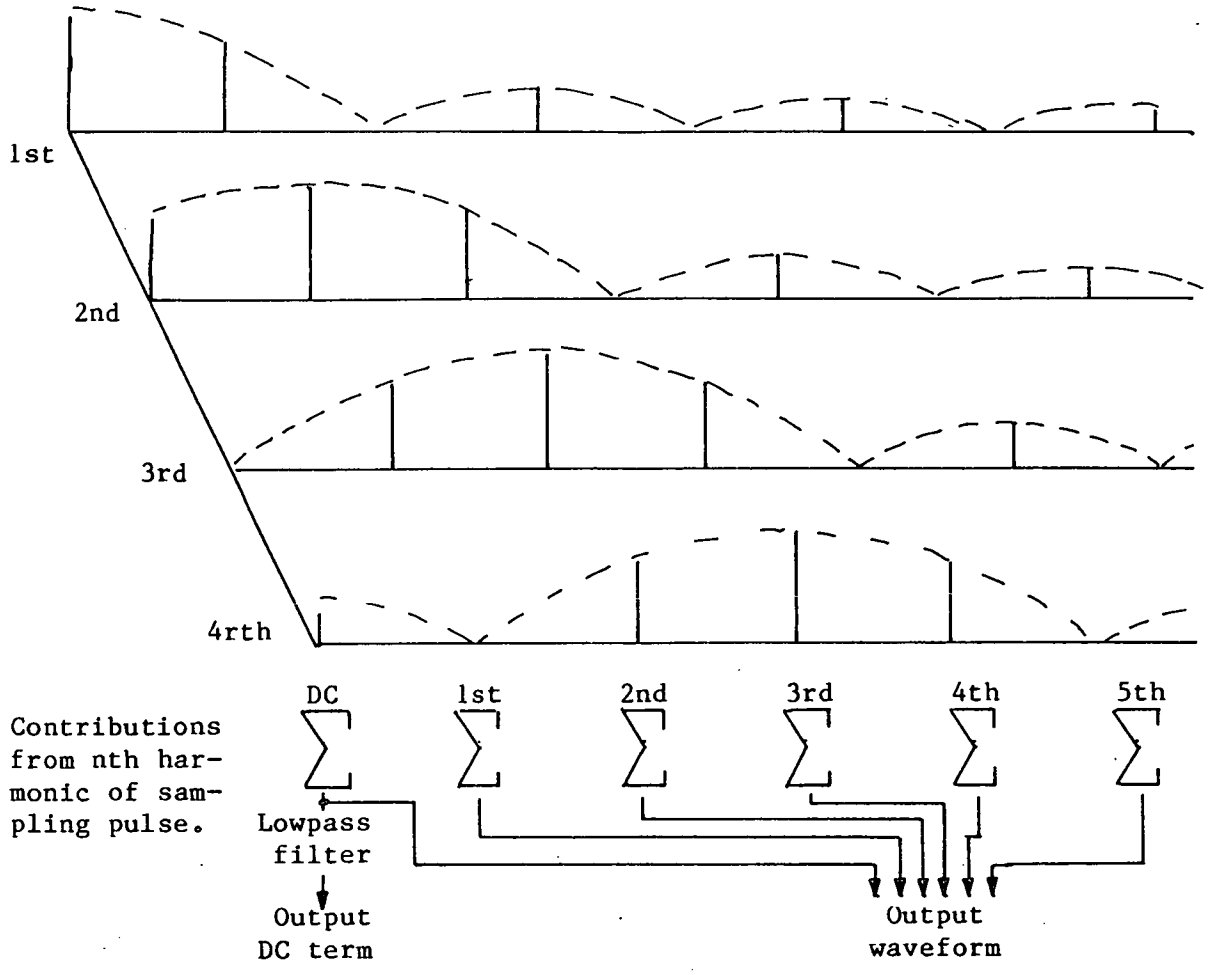


Figure 2.4.2. Frequency components of the output signal's power spectrum. (not to scale)

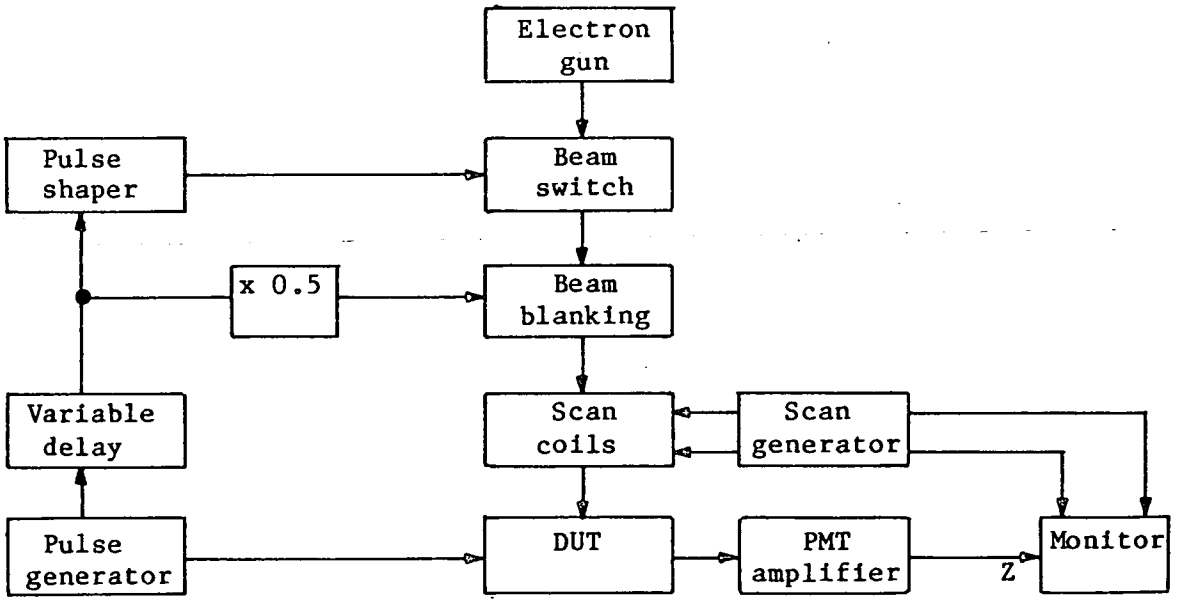


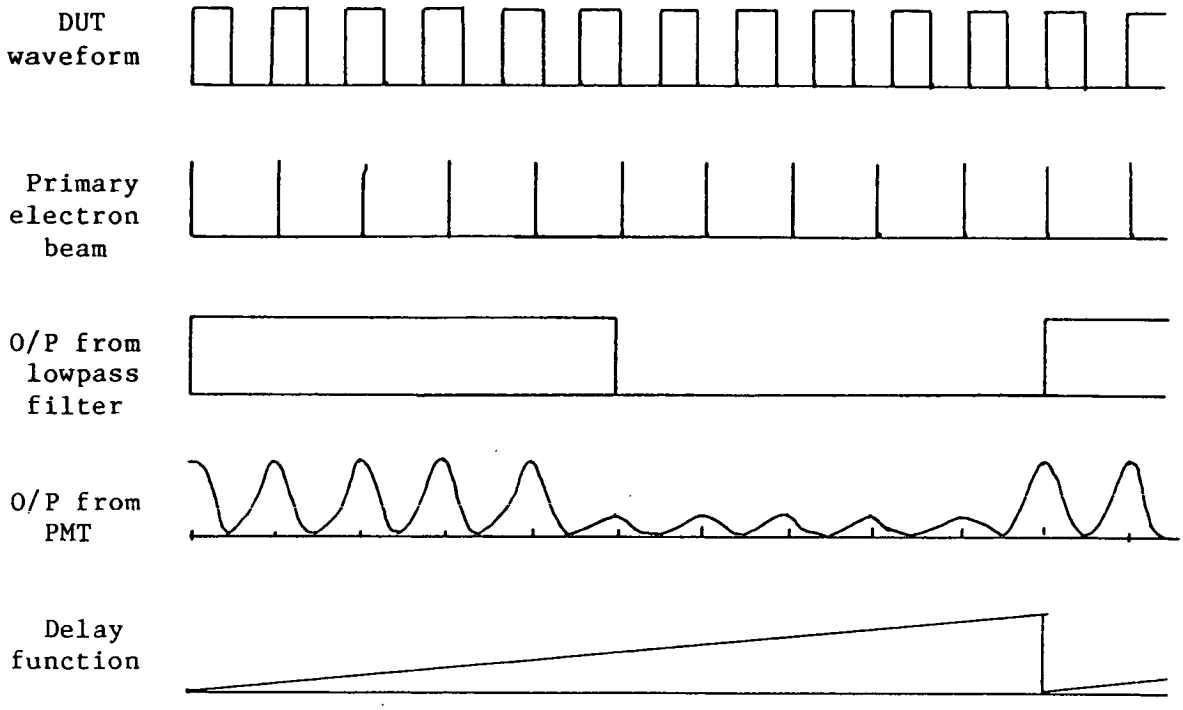
Figure 2.4.3. Instrumentation required to implement Stroboscopy.

is multiplied with the sum of the topographical contrast and the voltage contrast, resulting in gross aliasing, since the sampling frequency is the same as the specimen signal fundamental as shown in Figure 2.4.2. The power at each harmonic and subharmonic is a function of the relative phase of the primary beam pulse and the specimen signal. Since the topographical information is not time dependent, it appears at harmonics of the line frequency and therefore if the resultant signal is lowpass filtered(by the videohead amplifier) the output image is a voltage map at one particular instant superimposed on the normal topographical image.

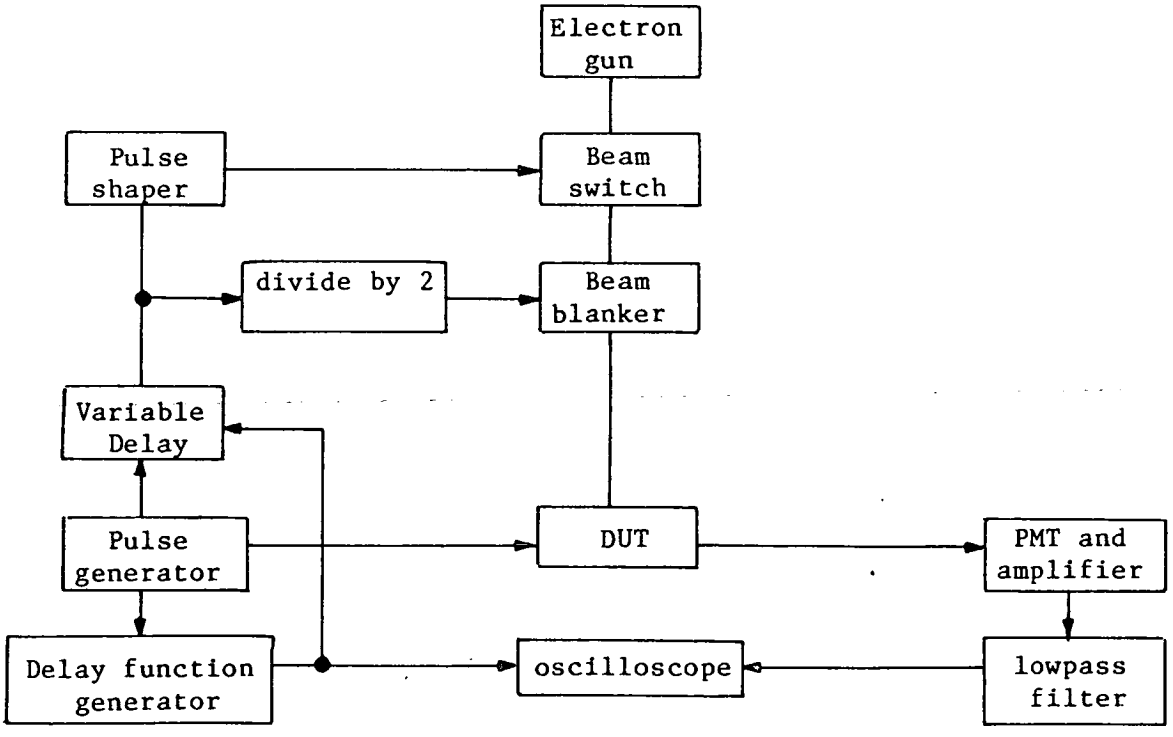
The most popular method for realising the beam switch is to use parallel plates before the first magnetic lens of the SEM and deflect the beam across the final spray aperture prior to the final magnetic lens. Menzel and Kubalek(1979) have reported pulse widths of 10ps at frequencies of 10GHz, although Gopinath and Hill(1973, 1977) obtained 5ps pulses at a frequency of 9GHz using an inter-digital deflection structure. The parallel plate method is particularly suitable since it can be utilised over a wide range of frequencies unlike re-entrant cavity beam pulsers(Hosokawa et al., 1978). For a more in depth discussion of Stroboscopy, Gopinath and Hill(1974) should be consulted.

2.4.3) Sampling.

An important disadvantage of both the previous types of displays is that the resultant images incorporate a large amount of redundant information and do not lend themselves to fast computer data-extraction and logging techniques. In addition, no information concerning the shape of the waveforms present on the device is available. i.e. risetimes. This problem can be resolved by using the



Theoretical waveforms for an electron beam sampling system.



Schematic of the instrumentation required to realise a Sampling SEM.

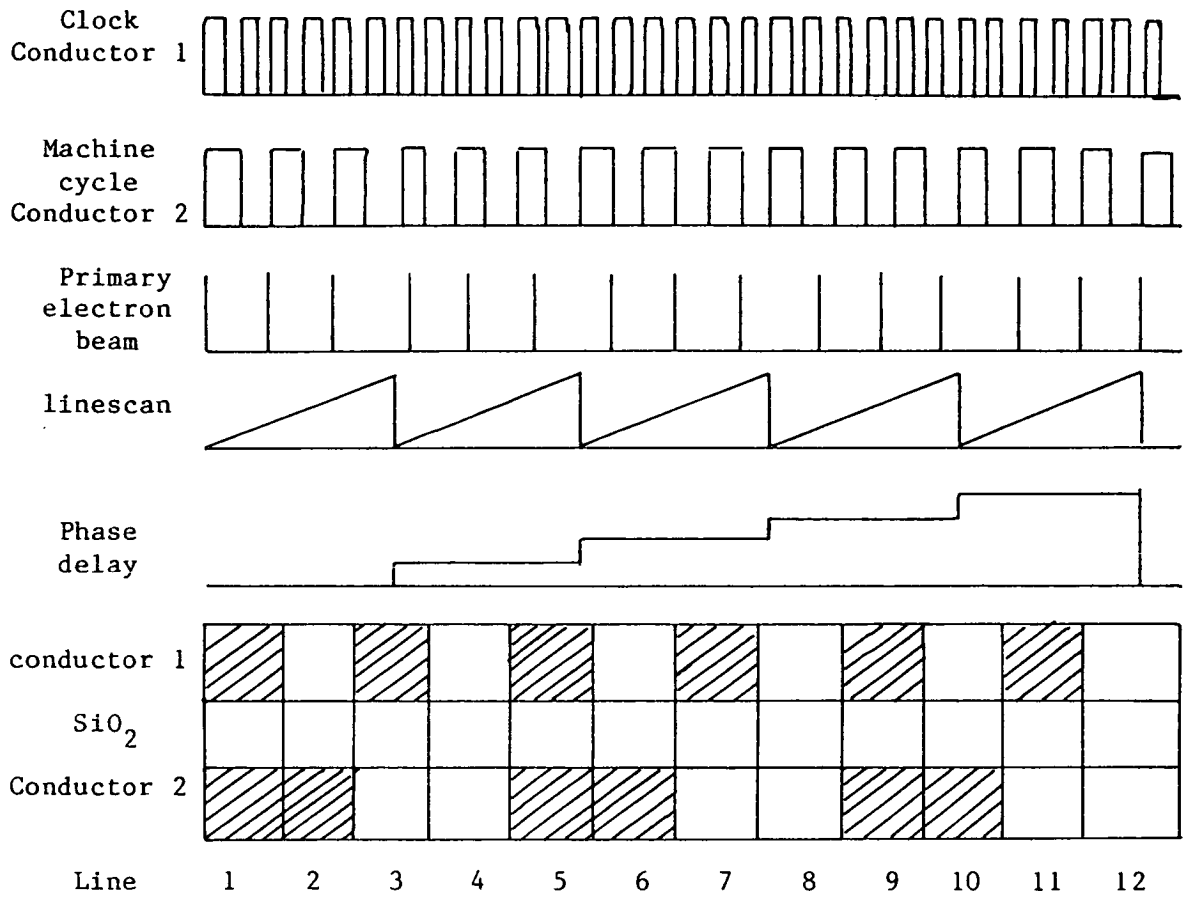
Figure 2.4.4. The Sampling SEM.

WAVEFORM RECONSTRUCTION technique (Thomas et al.,1976) which is similar to Stroboscopy except that the primary beam is maintained stationary on a single node of a circuit with the result that the topographical contrast collapses to a DC level. When a periodic time function is convolved with an impulse function, the result is a time expanded function, if the convolution time is greater than the period of the time function. This principle is demonstrated in Figure 2.4.4, where the convolution time is several cycles of the observed waveform. The convolution time is determined by the frequency of change of the relative phase of the sampling pulse and the specimen waveform, which is linearly varied through one period of the specimen waveform. Hence, the system bandwidth is determined by the frequency of the phase change and not the clock frequency of the device under test, enabling very fast waveforms to be reconstructed.

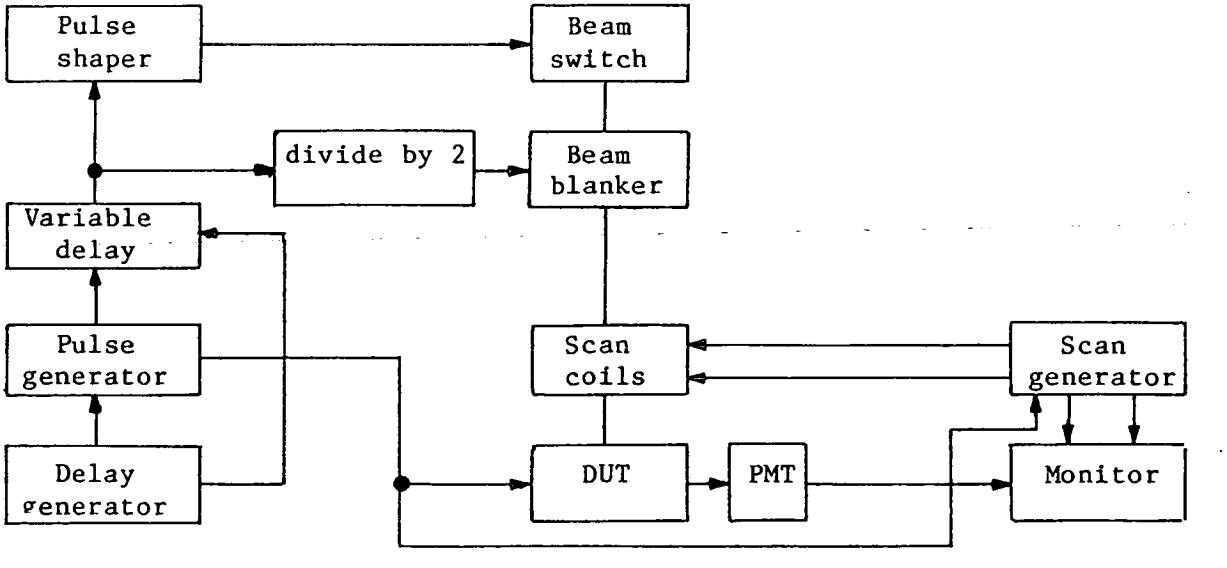
A consequence of small primary beam pulse M:Ss is that the output signal power is reduced and therefore the output signal-to-noise ratio(SNR) is degraded. To counteract this problem the number of samples taken at each phase increment can be increased, allowing the system bandwidth to be reduced and hence, improving the system SNR. However, the time taken to reconstruct an entire waveform can take up to 15 minutes, as reported by Lische et al.(1983)

2.4.4) Logic State Mapping.

If the phase relationship between the specimen signal and the sampling is varied at an integer sub-multiple of the linescan frequency over a set number of signal periods and the number of periods is synchronised to the field frequency, then a two dimensional truth-table can be generated on a video monitor. Crichton(1980)



Theoretical waveforms and display for 2 parallel conductors intersperced with an insulator. The electron beam is scanned at right angles to the conductors



Schematic of the instrumentation required to realise Logic State Mapping.

Figure 2.4.5. Logic State Mapping.

demonstrated that this method was particularly suited to devices which incorporated bus structures such as microprocessors. Figure 2.4.5. shows a block diagram of the necessary instrumentation and the theoretical waveforms for two adjacent conductors. LOGIC STATE MAPPING is in essence a sampled data version of voltage coding which extends the bandwidth of the technique.

2.5) Voltage Contrast Summary.

It has been shown in this chapter that an electron beam test system is feasible and offers the following advantages over the mechanical probe: greater spatial resolution(5nm), an electrically non-loading probe and a high temporal resolution(10ps). These facts qualify it as the only method available to internally probe VLSI devices in situ(i.e. without test pads).

The disadvantages are that high temporal resolution can only be acheived for synchronous devices and for a greater nodal measurement time. Hence asynchronous waveforms such as microprocessor internal handshakes cannot be resolved. However, the critical drawback of all the electron beam methods described so far is that they are all qualitative and hence, no accurate amplitude measurements can be made.

CHAPTER 3.

QUANTITATIVE VOLTAGE CONTRAST.

In order to make quantitative measurements with the Electron Beam Probe, it is first necessary to derive or approximate a linear relationship between the applied surface potential and the output voltage of the videohead amplifier. The contrast mechanisms must be fully understood before a linear system model can be formulated and an instrumentation system constructed. Hence this chapter will review the mechanisms underlying voltage contrast and examine the techniques currently utilised to obtain a linear relationship.

3.1) Voltage Contrast Mechanisms.

As stated in the previous chapter, the collected current is a complex function of all the electric and magnetic fields present in the specimen chamber. The three underlying contrast mechanisms will be discussed.

3.1.1) Trajectory Voltage Contrast.

When the low energy electrons leave the surface of a specimen, their trajectories are determined by the local electromagnetic field, their initial energy and their angle of exit from the specimen. The electromagnetic field includes contributions from final lens magnetic leakage, the Faraday Cage attraction voltage and the specimen surface voltage. The shape of the specimen chamber itself, usually earthed, affects the local electromagnetic field. Hence not all electron trajectories which are initiated on the specimen surface terminate within the electron collection system. However, if the potential difference between the specimen and the Faraday cup is made large

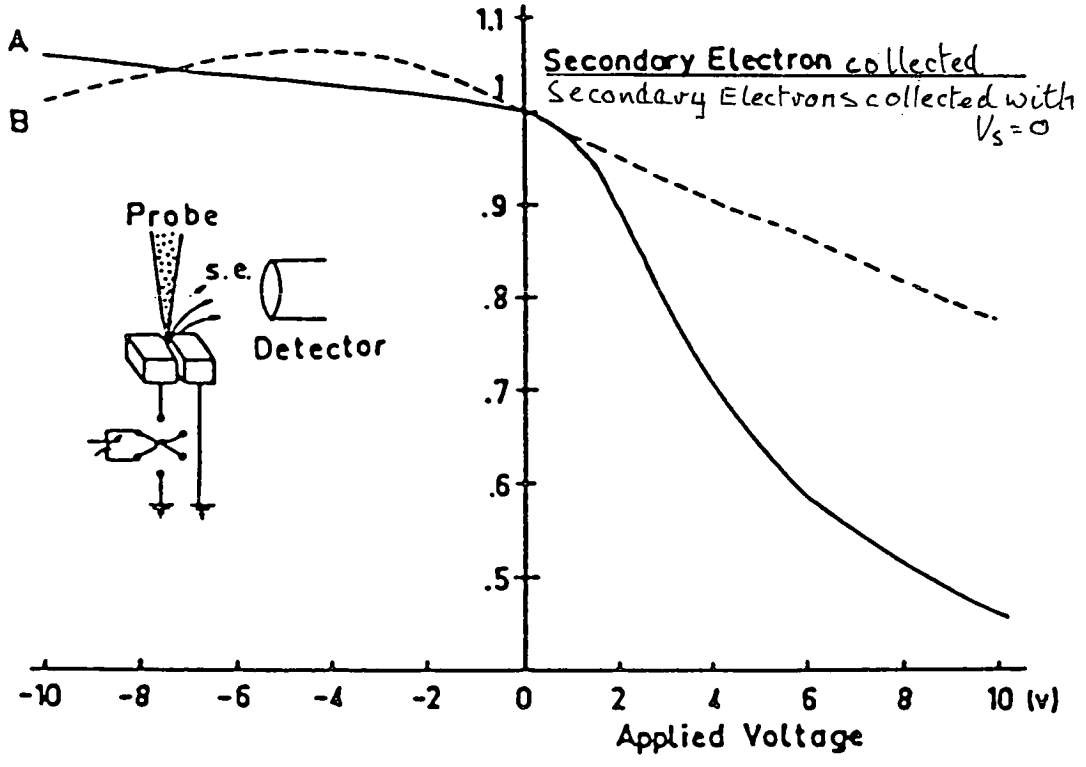


Figure 3.1.1. The intensity variation of secondary electrons as a function of specimen voltage. (from Kimoto, 1968)

enough, the electric field generated will dominate the electromagnetic field within the chamber and a high proportion of the electrons generated will be collected.

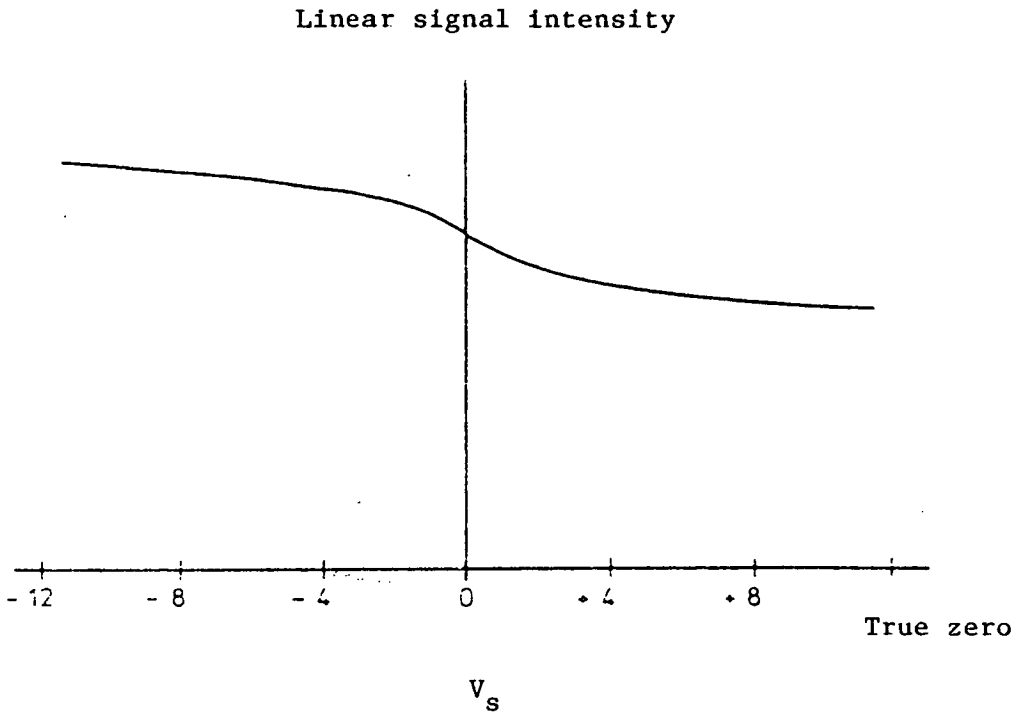
Since the applied surface voltage modulates the local electromagnetic field, the collected current will be a function of the applied surface voltage. As shown in Figure 3.1.1, Kimoto(1968) confirmed that the collected current was a monotonic function of the applied negative surface voltage and that the number of collected electrons increased as the magnitude of the surface potential increased.

Everhart(1969) exploited the trajectory contrast by moving the detector closer to the specimen and demonstrating that a voltage resolution of at least 50mV was attainable by placing a small slit across the detector. However, by introducing the slit the SNR was reduced since fewer electrons were collected.

3.1.2.) Yield Voltage Contrast.

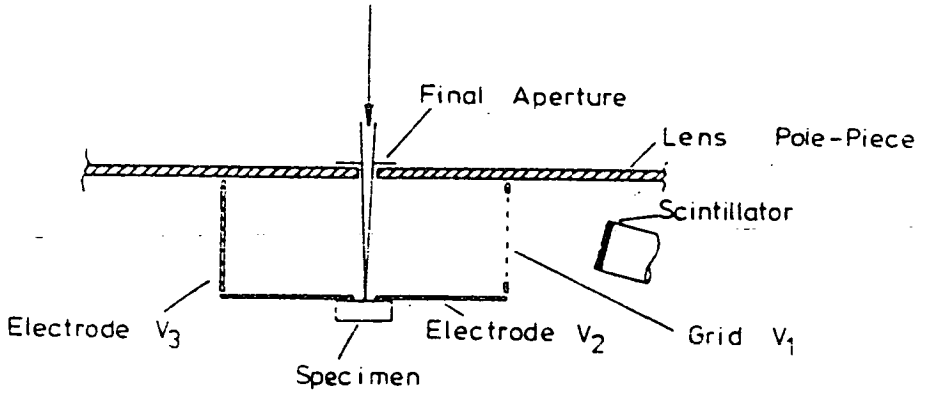
If a positive voltage is applied to the specimen, the surface field will attract electrons back to the surface, and subsequently only the electrons with sufficient energy and momentum will escape the effect of the surface field, and be collected. This appears as if the secondary electron yield has been altered by the applied surface voltage and adjusting the external electric fields, such as that of the Faraday cage, has little effect. Again Kimoto(1968) Figure 3.1.1, showed that the collected current decreased rapidly as the magnitude of an applied positive surface voltage increased.

If the yield and trajectory contrast mechanisms were mutually



(a)

Transfer function of detector, obtained with stationary primary beam.



(b)

Structure of the detector arrangement. $-300 \leq V_{1-3} \leq +300$. Cylindrical about electron beam axis.

Figure 3.1.2. Voltage Contrast detector used by Banbury and Nixon, 1969.

exclusive, a linear monotonic relationship should be derived when a bipolar surface voltage was applied. However, Kimoto(1968) recorded a very nonlinear monotonic relationship and offered the explanation that both mechanisms were present, but only one was dominant in either of the two regimes. That is when a positive or negative voltage was applied to the specimen, Yield or Trajectory contrast dominated respectively. Joy and Titchmarsh(1970) attempted to isolate the dominant of the two contrast mechanisms by collecting the specimen current as well as the emitted current and found that the yield mechanism was dominant.

In order to isolate the yield contrast, Banbury(1969, 1970) attempted to suppress the trajectory contrast by applying electrostatic fields around the specimen, thereby ensuring a controlled electric field around the specimen by eliminating the effects of any stray fields. As shown by Figure 3.1.2., a more linear relationship(cf. Figure 3.1.1.) was obtained between the applied surface voltage and the collected current, with only 8% nonlinearity. However, the dynamic range of the signal was very small with a maximum signal amplitude of 80mV.

3.1.3) Energy Voltage Contrast.

If secondary electrons are accelerated through a potential difference, their energy will increase accordingly. A phenomenon exploited by the Everhart-Thornley(1960) detector system. Therefore, the position of the SED on the energy axis will be modulated by the surface electric field, which is proportional to the applied specimen voltage.

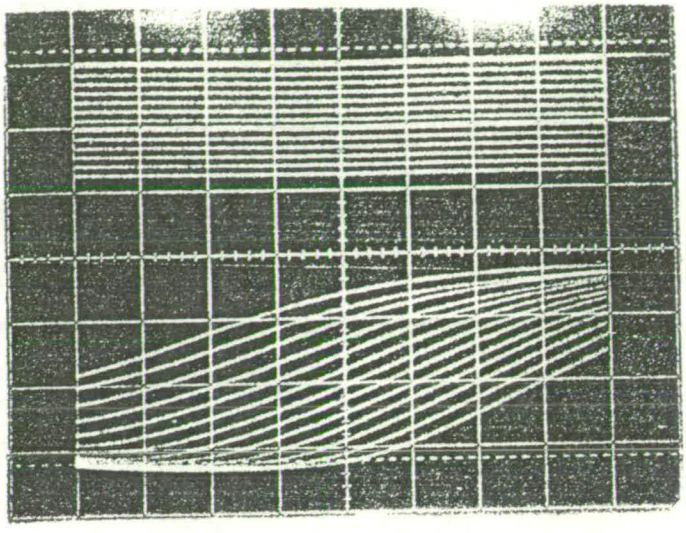


Figure 3.1.3. Energy Filter characteristic curves.
Upper trace: Applied surface voltage 1V/mm.
Lower trace: Energy curves 2V/mm.
Retardation waveform period: 0.5ms.

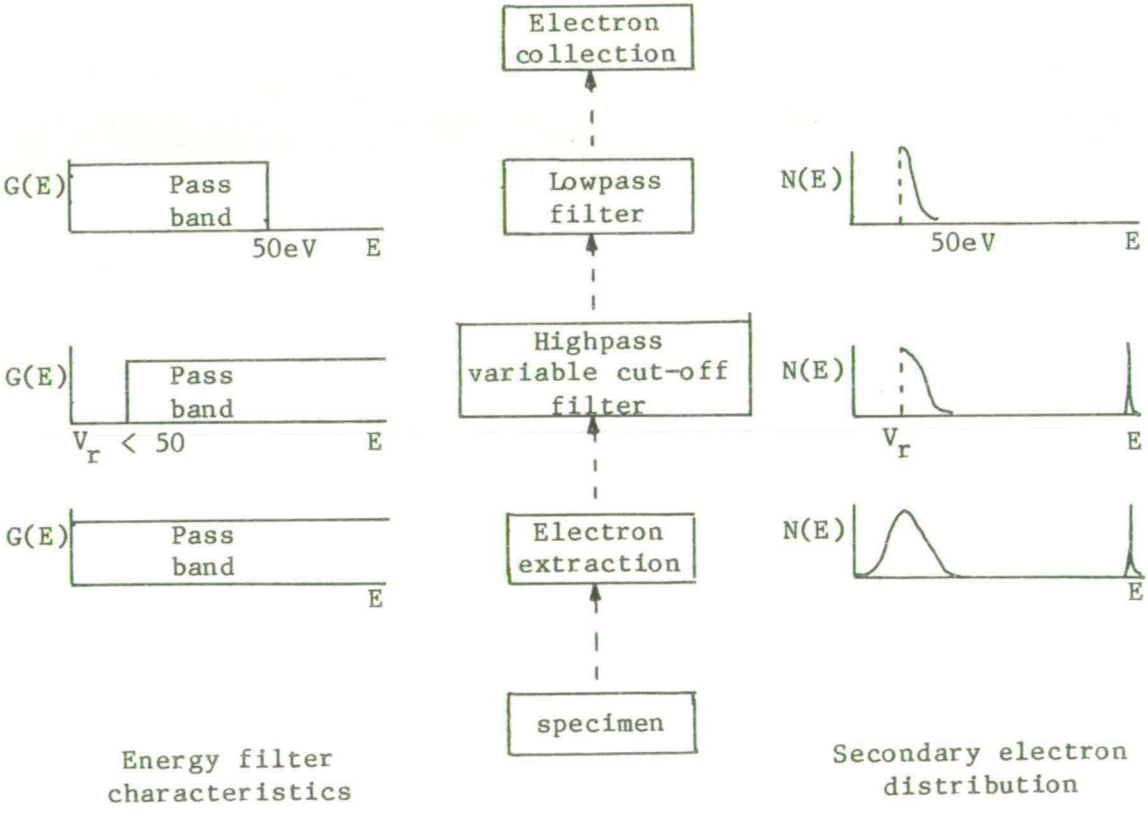


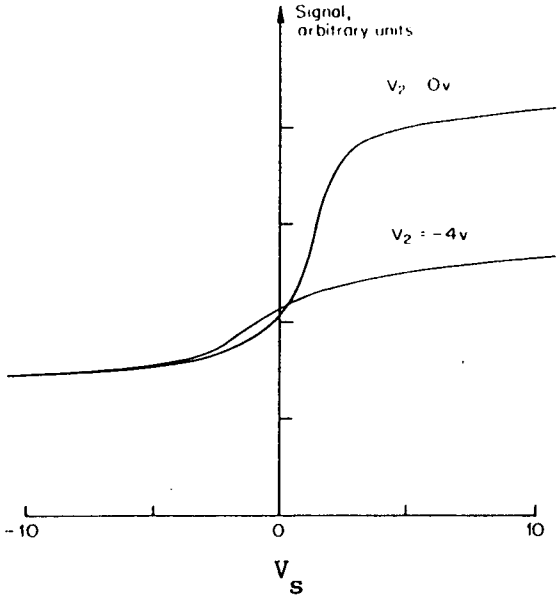
Figure 3.2.1. Block diagram of ideal energy filter.

An example of energy voltage contrast is shown in Figure 3.1.3, in which the integral of the SED is plotted for differing surface voltages. The upper trace is the applied surface voltage, going positive to negative from top to bottom. The corresponding translation of the cumulative SED("S"-curve) is shown in the lower trace, with a left to right translation for negative to positive voltages.

In order to exploit this inherently linear type of contrast, all or at least a fixed fraction of the emitted secondary electrons have to be collected, to suppress the Yield Contrast and since they must undergo uniform energy modulation, a controlled electrostatic field has to be provided to suppress the Trajectory Contrast. Further, modulation in the energy domain has to be mapped onto a time variation of current or voltage. All three objectives are attained by utilising a device termed an "Energy Analyser" or more correctly a Velocity Analyser.

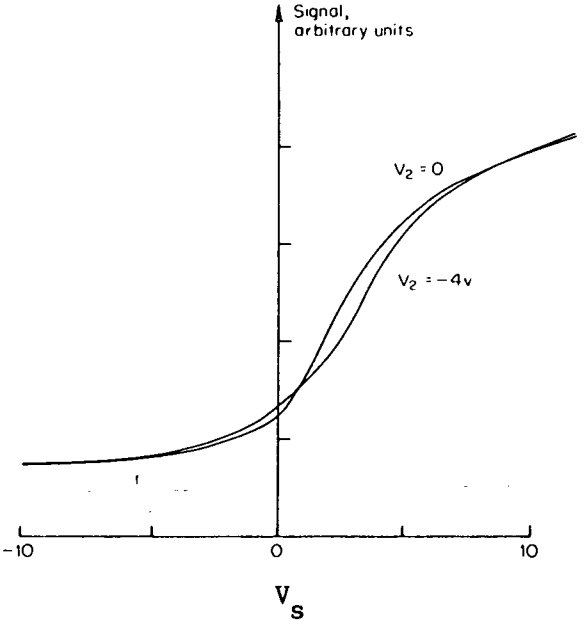
3.2) Energy Filter.

The energy analyser, as shown in Figure 3.2.1, can be thought of systematically as comprising three parts: a primary lens to suppress local field effects and extract the electrons, a lowpass energy filter to attenuate all high energy backscattered electrons and a time variant energy filter to implement the translation from the energy domain to the time domain. Since the practical design of the energy filter itself is a research area(Khursheed(1983), Menzel, 1983) the discussion will be limited to an analyser with ideal characteristics.



(a)

V_e: +49V; Field strength at specimen: 0.5kV/mm.



(b)

V_e: +0.5kV; Field strength at specimen: 10kV/mm.

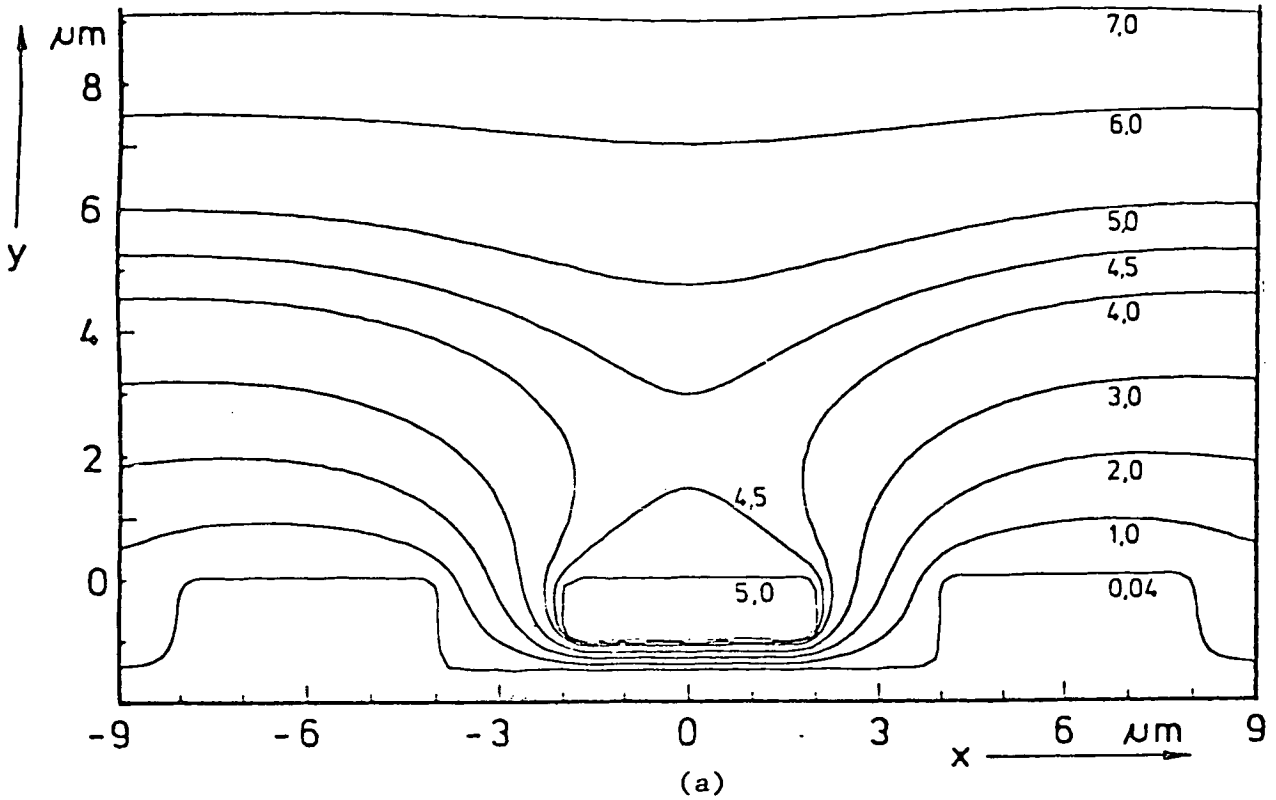
Figure 3.2.2. The compensation acheivable when an extraction field is used with respect to the effects of local fields. V_2 is the potential of a parallel neighbouring track with respect to the probed track. (from Plows, 1969).

3.2.1) Local Field Suppression.

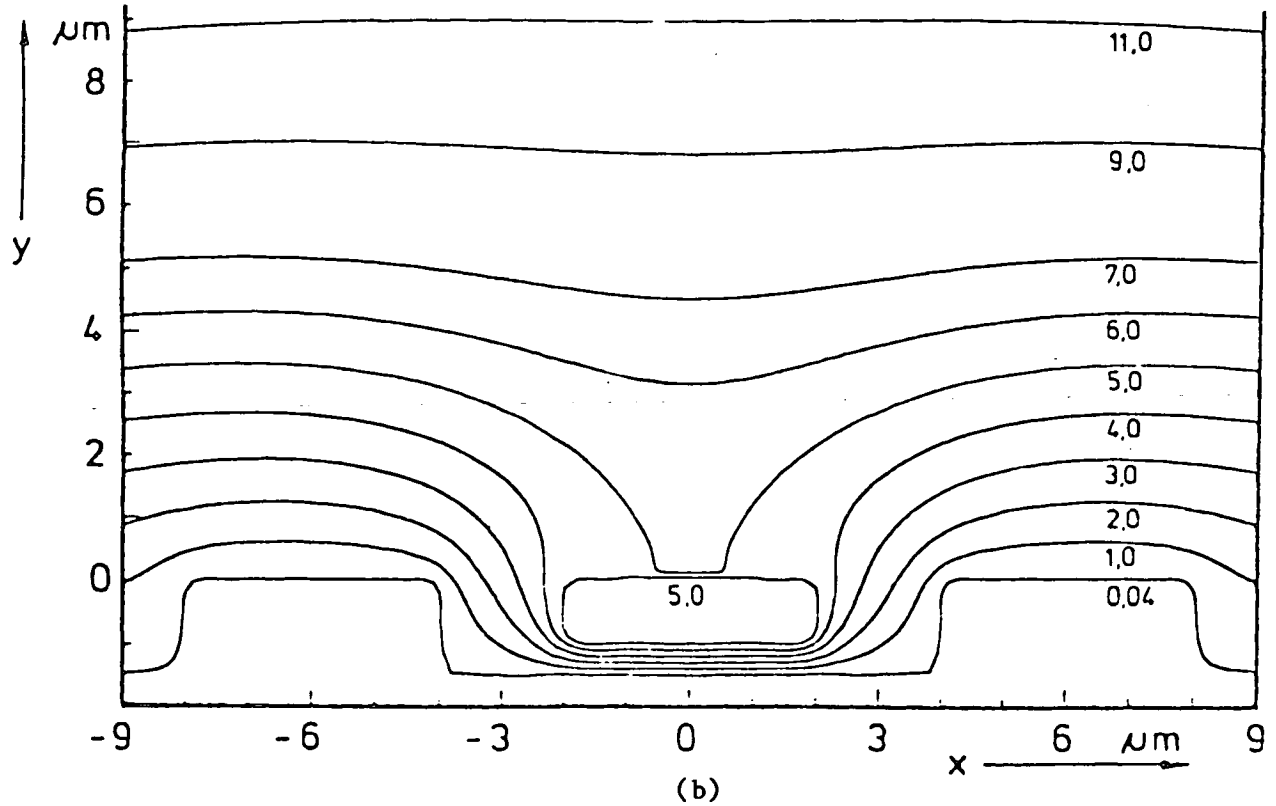
Local fields are characterised by the fact that they divert generated electrons away from the electron collection system and hence introduce errors into a voltage contrast measurement. One type of local field is caused by voltages on a neigbhouring conductor, inducing a large transverse velocity component in the electrons, deflecting them horizontally parallel to the specimen surface. Another type, is that caused by yield contrast discussed previously.

Two methods have been adopted to suppress the effects of the transverse surface fields. Wells(1968), Plows(1969) and Hannah(1974) all employed cylindrically symmetrical immersion lenses within which the specimen was placed. The high field strengths induced above the specimen compressed the surface fields onto the specimen and minimised the effects of the transverse fields, since few electrons passed through them. In addition, since these lenses were constructed using positive potentials, the SEs were also attracted into the internal fields present in the analyser. This type of suppression gave good results on the electron beam axis. However, when the target point moved off axis, the field strength was reduced and aberrations in the lens compromised performance. Plows(1969) demonstrated the effect local fields had on "S"-curves as depicted by Figure 3.2.2(a) and the compensation acheivable by introducing an extraction field of 10kV/mm at the specimen surface, as illustrated in Figure 3.2.2(b).

To give large area coverage, Haas(1966) introduced the concept of utilising a mesh screen above the specimen and applying a large positive voltage of 800V to generate fields of 4kV/mm a few mm from the specimen surface. Although giving a larger working area, the



(a)
Equipotential lines at field strength of 0.6kV/mm.



(b)
Equipotential lines at a field strength of 1kV/mm.

Figure 3.2.3. The effect of an extraction field on surface fields. (from Menzel, 1983)

actual specimen viewing area is reduced by this method and the electron beam has to be aligned with the grid. In addition the actual M:S of the grid is important in maintaining a linear extraction field.

As shown in Figure 3.2.3, Menzel(1983) simulated the effect, increasing the extraction voltage has on the surface fields and found that a field strength of 1kV/mm was necessary to eliminate the measurement errors introduced by surface fields at the measurement point. However, increasing the extraction voltage reduces the spatial resolution of the measurement system, especially at low primary beam energies since the electron beam undergoes defocusing, caused by the extraction field. In addition, Fujioka et al.(1981) reported that increasing the extraction field strength indefiniteley did not result in proportional improvements in local field suppression.

The extraction electrode in effect carries out the first integral of the SE current equation, attenuating the two velocity components in the plane of the specimen and collecting all the emitted electrons independent of their direction of exit.

3.2.2) Primary Filter.

The primary filter is necessary to attenuate high energy backscattered electrons, emitted from the specimen and prevent them from entering the main energy filter, where they would impinge on the analyser walls, generate more secondaries and corrupt the signal. This effect can be minimised by coating the analyser with a material with a low SE coefficient such as carbon black. Although these backscattered electrons are modulated by the surface fields, the percentage modulation is small, in the order of 1% or less for electrons with

energies in the 1kV regime. Therefore SEs are preferred as the information source, since the modulation levels can reach 100%. In addition, the backscattered electron distribution is heavily dependent on specimen composition. Thus a lowpass energy filter with a cut-off around 50eV is necessary.

3.2.3) Secondary Filter.

The secondary energy filter performs the function of energy discriminator and passes only those electrons in a predetermined energy range. Therefore, the output current of the energy filter is the integral of the product of the energy filter function(H(E)) and the SED.

$$I_{out} = I_s \, T \int_0^{50eV} \frac{N(E-V_s).H(E-\psi)de}{N(E)} \qquad (3.2.1)$$

Where ψ is the energy cut-off of the filter and is a function of the retardation voltage(V_r), V_s is the surface voltage and T is the transmission factor of the analyser. Equation 3.2.1. can be rewritten in terms of primary beam current(I_p):

$$I_{out} = T\delta I_p K_a \qquad (3.2.2)$$

Where K_a equals

$$K_a = \int_0^{50eV} \frac{N(E-V_s).H(E-\psi)dE}{N(E)} \qquad 0=\langle K_a=\rangle 1 \qquad (3.2.3)$$

If the filter cut-off energy is linearly modulated by a time dependent function, the output current is the convolution of the filter function and the SED. If ψ is independent of time, the analyser becomes a discriminator with its quiescent output level determined by the value of K_a when V_s is zero.

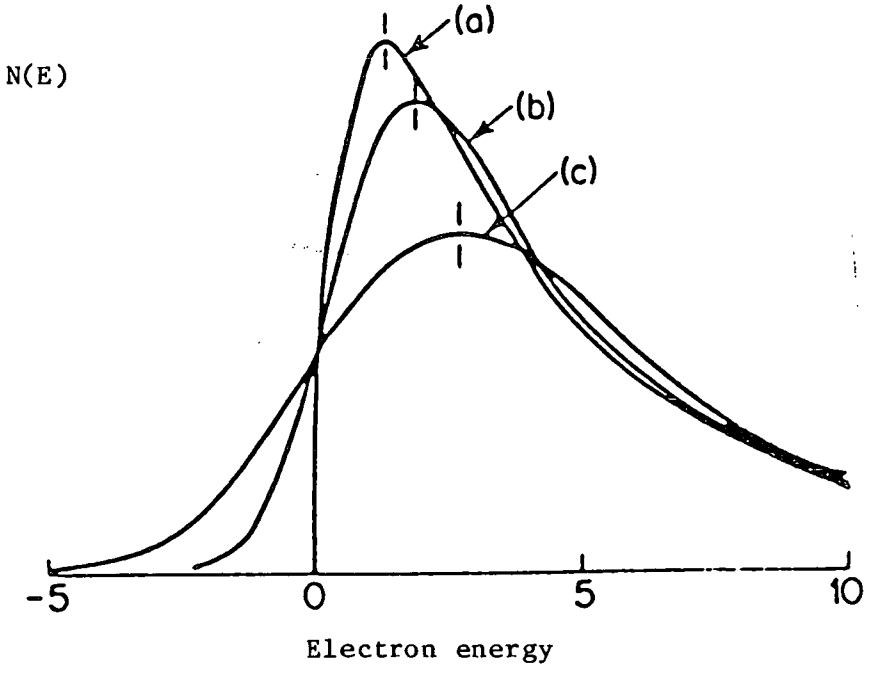


Figure 3.2.4. The effect of bandpass energy analyser resolution on the apparent secondary electron distribution. (from Wells, 1974)
(a) Typical curve. (from Kollath, 1947).
(b) 2.5V analyser resolution
(c) 5V analyser resolution.

The filter function has either a highpass or bandpass characteristic, since the lowpass characteristic is equivalent to the highpass. If a bandpass filter function is desired it has to approximate an impulse function in order to reduce the smoothing effect it has on the SED. For a practical system the energy width of the filter function should be a factor of 100 less than the half-height width of the SED. The smoothing effect was demonstrated empirically by Wells(1974) as shown in Figure 3.2.4. The output of the filter is a reproduction of the SED as a function of time. The narrow energy width required for a bandpass characteristic produces a very low SNR and therefore the signal has to be averaged to give accurate results.

By comparison, a highpass energy function produces an output signal with a high SNR and has a unit step filter function. The discriminator characteristic is the cumlulative SED and although inherently nonlinear it is monotonic, unlike the bandpass filter.

If the relevant filter characteristic is chosen with a cut-off energy of less than 50eV and in order that it bisects the SED, obtained with the specimen earthed, a two quadrant transfer function will be obtained with a dynamic range equal to the standard deviation of the distribution. Outwith these limits the surface voltage modulation will either be zero or undetectable. However, since the transfer characteristic of the bandpass filter does not give a 1:1 relationship between the applied surface voltage and the recorded output voltage, it requires more sophisticated instrumentation to achieve a practical measurement system. Although, the cumulative distribution function is grossly non-linear, it can provide quantitative information. Both characteristics suffer from the disadvantages of being nonlinear and having a relatively low dynamic range.

3.3) Open Loop Systems.

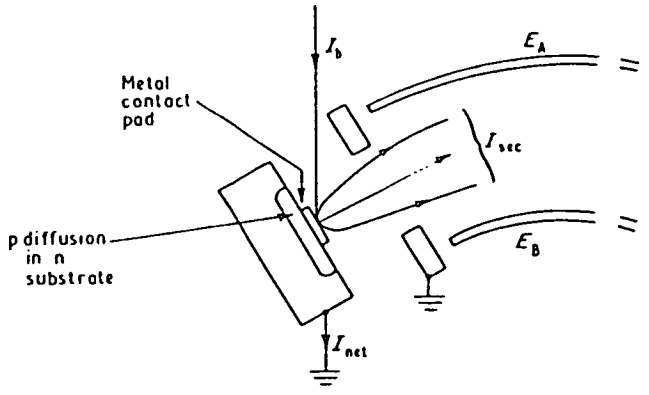
To overcome the problem of dynamic range inherent in energy discriminators, measurement systems were developed which relied on the actual position of the SED on the energy abscissa.

3.3.1) Scanned Systems.

Wells and Bremer(1968 and 1969) instigated energy analysis as a means of measuring surface voltages. They employed a 84^{o} cylindrical plate analyser with a bandpass characteristic, as shown in Figure 3.3.1(a). By applying appropriate bias potentials to the analyser plates only electrons within a specific energy range were transmitted to the Everhart-Thornley detector. Further by adjusting the applied plate potentials with respect to the specimen voltage, the filter characteristic was made tunable. Therefore, the required coordinate transform could be attained by varying the voltage applied to the bias plates and observing the corresponding variation in output voltage from the videohead amplifier.

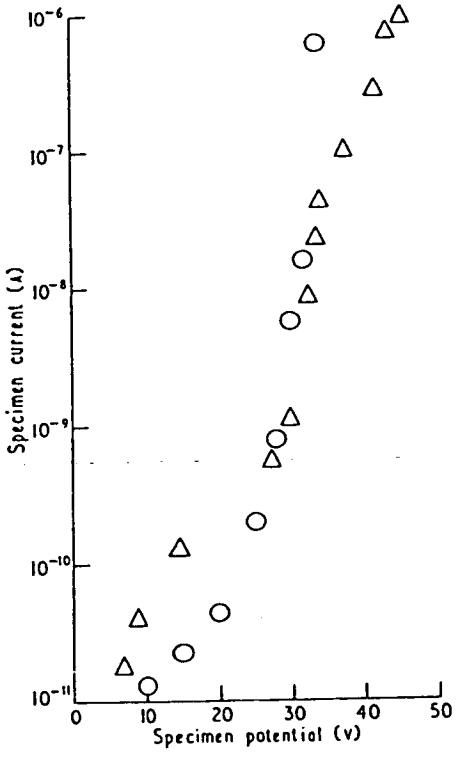
Wells and Bremer demonstrated their system by measuring the reverse characteristic of a p-into-n device, by both electrical and electron beam methods as shown in Figure 3.3.1(b). A 20dB reduction in SNR of the video output signal was noted by Wells and attributed to the fact that only a small fraction of the secondary electron current was actually collected. This was caused by the increase in focal distance of the final magnetic lens brought about by the insertion of the analyser which reduced the beam current for a specified electron beam diameter.

Hannah(1974) employed the converse philosophy of ramping the



(a)

Input arrangement to cylindrical plate, bandpass energy analyser when measuring reverse p-n junction characteristics (device is shown enlarged).(Wells and Bremer, 1969)



(b)

Reverse characteristic of p-into-n device: circles as measured electrically; triangles as measured by contactless electron beam.

Figure 3.3.1. Bandpass energy analyser used by Wells and Bremer, 1968

specimen reference voltage with respect to the static voltages applied to the plates. This ensured that all the collected electrons traversed identical electrostatic fields independent of their own energy and thus produced a constant energy resolution filter. The passband resolution could be altered by varying the potentials applied to the bias plates. The transverse electrostatic fields were suppressed in Hannah's case by a cylindrical immersion and focusing lens system utilised immediately above the specimen.

In order to make quantitative measurements of the surface voltage, peak detection was employed in both cases. Wells plotted the entire secondary electron distribution for each specimen potential using a pen recorder and manually measured the displacement between two corresponding peaks. This enabled quantitative relative measurements to be made and required the system to be calibrated with an accurate voltage source.

Hannah's system, shown in Figure 3.3.2, was fully automated employing a digitally controlled ramp waveform generator. The video-head output was bandlimited and differentiated before a zero-crossing detector was used to determine the position of the peak. The comparator, used for zero crossing detection, automatically inhibited the DAC used to generate the ramp. The measurement was taken as a fraction of the scanning period with respect to zero applied voltage. This was made possible by recording the output of the 10 bit, binary coded decimal counter, used to generate the ramp voltage, when the zero-crossing was detected and subsequently subtracting the two numbers for consecutive measurements, giving a numerical result corresponding to the difference of the two applied voltages. A measurement accuracy of 0.4% was reported, that is 0.1V in a ±12.5V

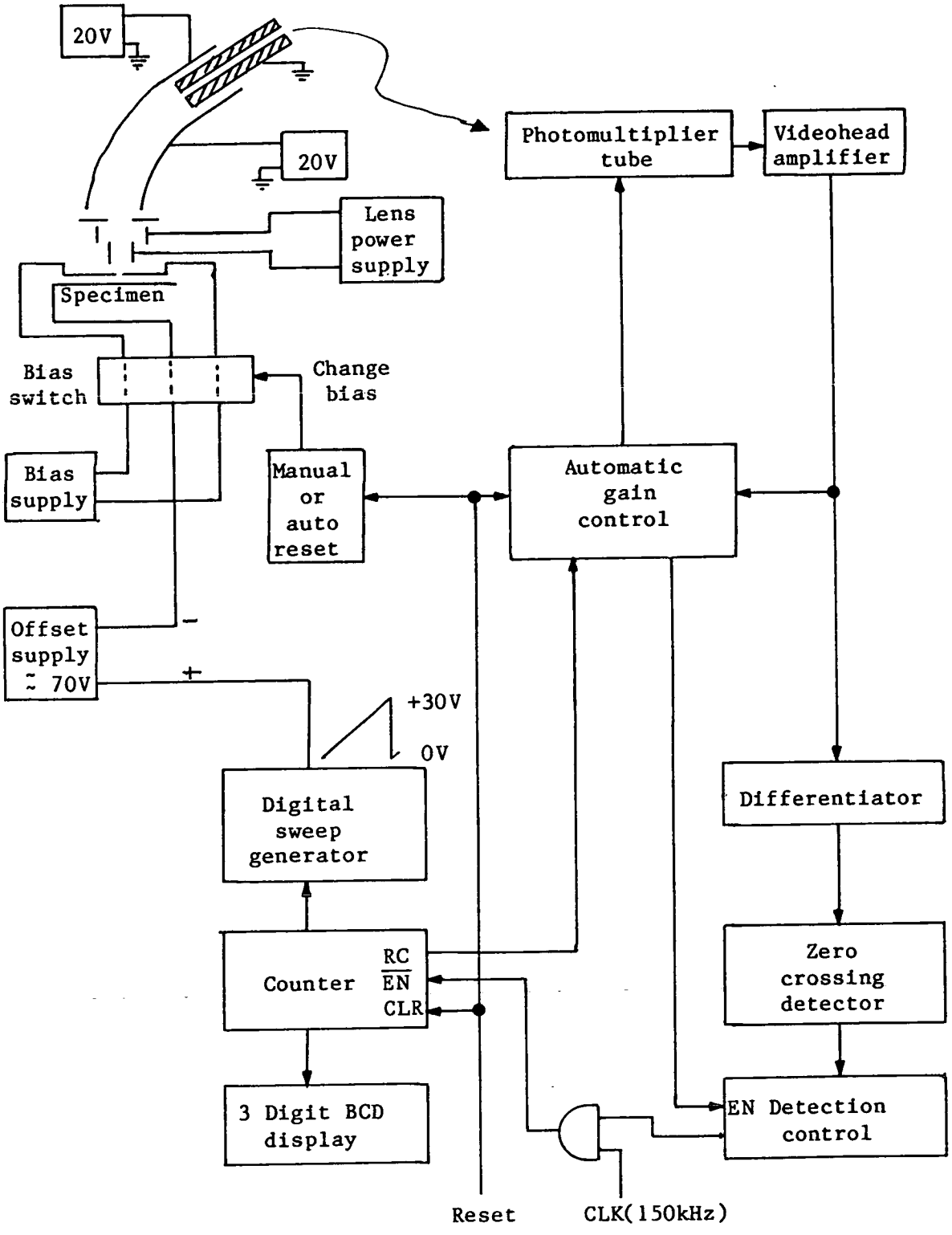


Figure 3.3.2. An automatic voltage measurement system(from Hannah, 1974, simplified)

range, with a measurement time of 8ms.

Hannah also recognised the problem of differing secondary electron yields for different materials and automatically adjusted the gain of the photomultiplier by using an iterative feedback method. This consisted of generating the SED and detecting the peak height. If the amplitude exceeded a preset value then the photomultiplier gain would be automatically reduced by an increment. This procedure would continue until the amplitude of the SED fell within preset limits. The main sources of noise recorded by Hannah were 50Hz mains hum and higher harmonics coupled into the instrumentation system either via the primary electron beam source or the low level circuitry and the primary beam shot noise.

Retardation field energy analysis was introduced to surface voltage measurement by Driver(1969). The measurement system, as shown in Figure 3.3.3. consisted of two concentric hemispheres; the inner hemisphere(M_i) being a highly transparent grid and the outer(M_o) fashioned from solid copper. The specimen was placed at the centre of the hemispheres. By adjusting the relative voltages of the hemispheres with respect to the specimen and collecting the currents from both of them, a bandpass filter could be obtained.

In this type of filter the inner grid was negative with respect to the outer shield and the specimen voltage while the outer shield was positive with respect to the inner grid and the specimen. Thus the outer shield collected electrons with energies greater than the outer shield potential and the inner grid electrons with energy greater than the inner grid potential. By subtracting these two signals a bandpass characteristic is arrived at and if both bias poten-

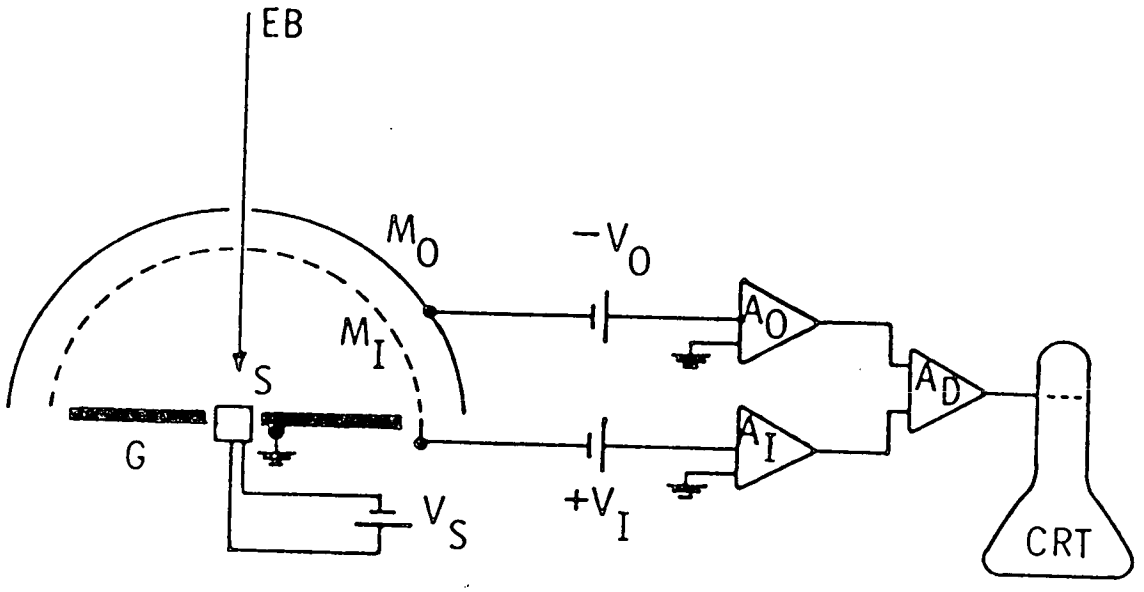


Figure 3.3.3. Bandpass energy analyser used by Driver(1969).

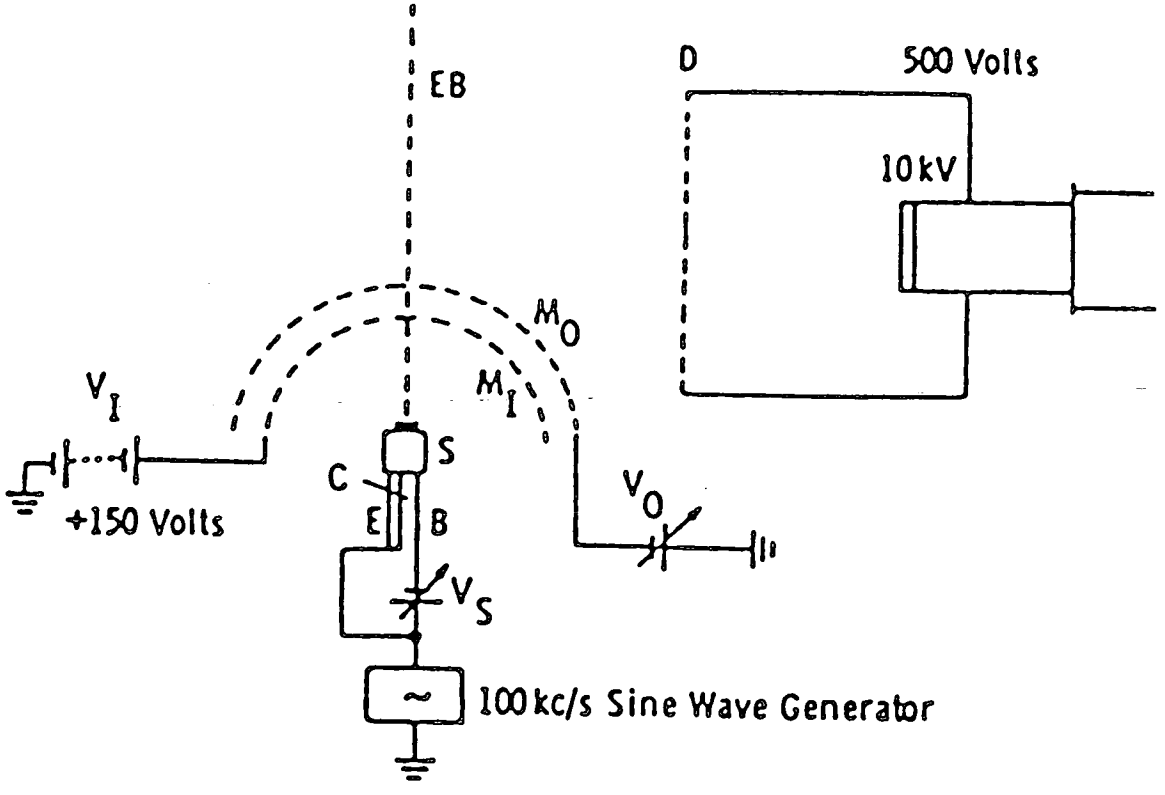


Figure 3.3.4. Highpass, retardation field energy analyser introduced by Driver(1969). M_o is the retardation and M_i the extraction grid.

tials are ramped with respect to the specimen the secondary electron distribution is reproduced at the output of the subtracting operational amplifier. This type of arrangement gives easy control of the analyser and does not suffer from the nonhomogeneous fields inherent in the deflection type analysers.

Another analyser configuration introduced by Driver, shown in Figure 3.3.4 was two concentric grid hemispheres with the signal collection carried out by an Everhart-Thornley detection system. In this analyser the inner grid(M_i) was biased positively with respect to the specimen and the outer grid(M_o) was biased with a continuously variable negative voltage source. The inner grid performed the function of yield and trajectory contrast suppression, while the outer grid allowed only electrons with a greater potential than that on the grid to pass through. Further with only a weak attraction field(+500V) on the collector cage, the high energy backscattered electrons would not be collected and thus the object of the primary high pass filter was realised. If the potential of M_o was scanned the characteristic "S"-curve would be generated. An advantage the latter analyser has over all the others is that it was poosible to scan a large area of the specimen when not making measurements.

All of the systems discussed were inherently open loop and through the requirement of a scanning function, more suited to measuring DC voltages, rather than AC voltages. The required bandwidth for the system was mainly dependent on the frequency of the fundamental of the ramp waveform and the power spectrum of the SED, since variation in the specimen potential manifested itself as a delay in the time domain of the SED. Thus the delay corresponded to a modulation of the SED phase spectrum or more correctly the

derivative of the phase spectrum with respect to frequency. Therefore the system bandwidth was independent of the bandwidth of the specimen signal, although a linear low distortion phase response was necessary. With bandwidth reduction the main source of process gain, the effect of increasing the bandwidth to include all the signal components, to improve the accuracy of peak position detection, inherently degraded the output SNR.

3.3.2) Computer Software Calibration Techniques.

Touw et al.(1977) introduced the concept of using a software inverse mapping technique to linearise a voltage contrast measurement system, giving an output signal directly proportional to the surface voltage. In this technique the highpass retardation energy filter used, was biased with a static potential to the middle of the required dynamic range, giving an output current that was proportional to the shift in the secondary electron distribution. The calibration was executed by using three different probe pads; one as the calibration pad, one as a reference pad and the other as the measurement pad.

A digitally generated scan was used to position the primary electron beam on the required pad and a lanthanum hexaboride electron gun used to improve the primary beam current density and therefore the resolution of the system. A 0.1Hz ramp waveform was applied to the calibration pad and the beam switched between the calibration and reference point, which was held at 0V, at a frequency of 200Hz. The output signal was detected using a lock-in amplifier and filtered with a lowpass filter which had a 3dB frequency of 10Hz. 512 samples of the output signal were taken during the scan period and stored in

a memory partition.

Next, with the calibration point earthed, the measurement sequence was repeated and the resultant 512 samples subtracted from the original samples, giving a discrete "S"-curve, independent of topographical contrast, stored in memory. With the measurement point firstly earthed and then with an applied voltage, the measurement sequence was again repeated, resulting in 512 samples of the applied voltage independent of topographical contrast, stored in memory adjacent to the discrete "S"-curve.

To derive the measured voltage, the 512 samples of the applied voltage were averaged and a search of the "S"-curve memory implemented to find the "S"-curve value nearest the average value of the applied voltage. Since the values of the "S"-curve were stored consecutively in the memory, the address of the "S"-curve value nearest the applied voltage value was directly proportional to the applied voltage.

The measurement technique could be simplified by using one point for both the calibration and reference point and using a pulse amplitude modulated ramp waveform. The constraints on the calibration point were that, it was of the same metalisation layer as the measurement point and an external voltage could be applied to it, while the reference point had to be constrained to have a constant potential, although this potential need not necessarily be known.

A voltage resolution of 100mV through 2.5μm of passivation and a measurement range of ±13V was reported.

Fujioka et al(1978, 1981) used a computer controlled SEM and

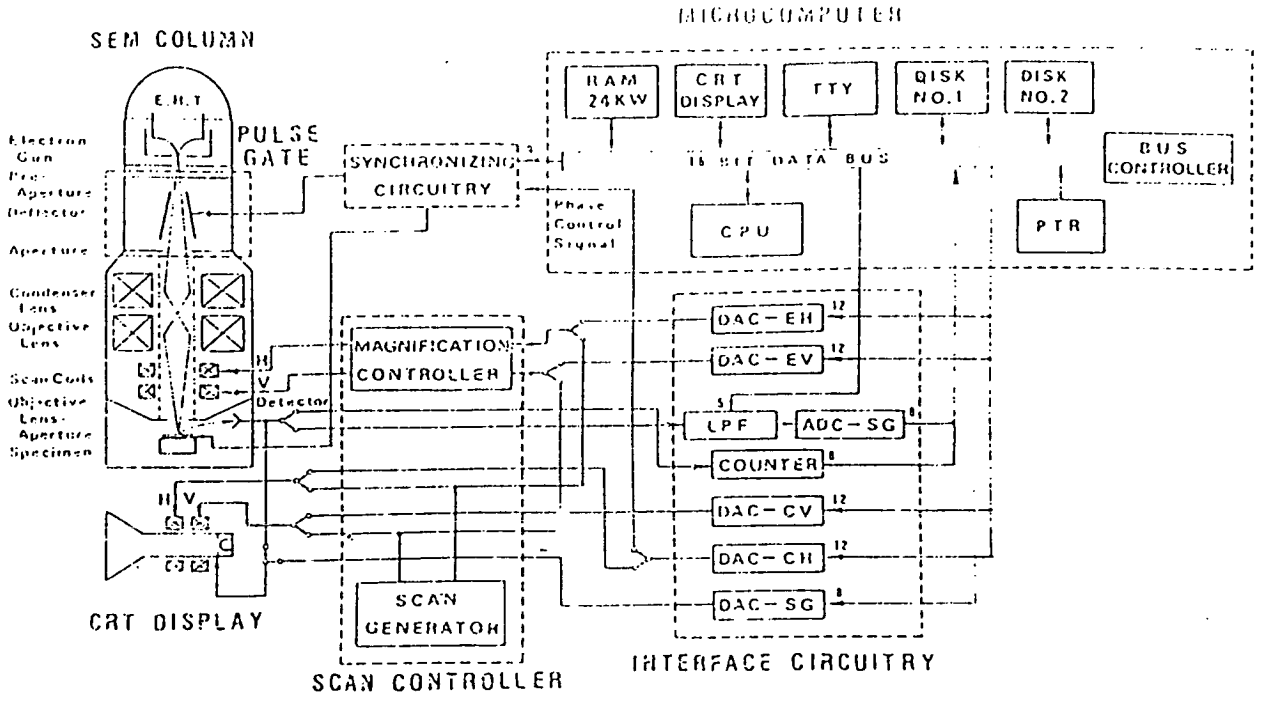
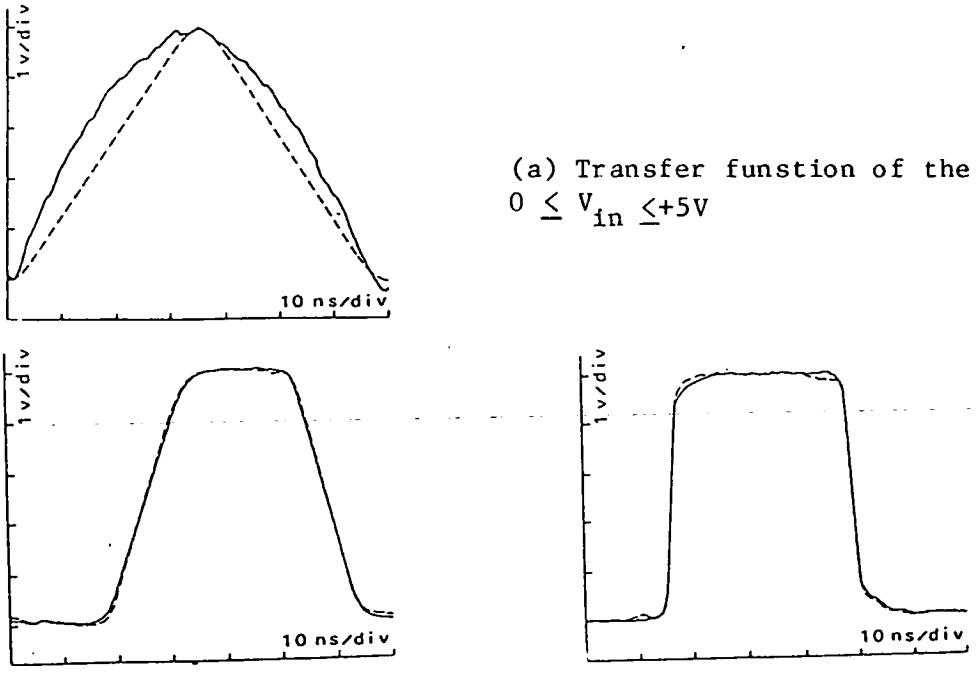


Figure 3.3.5. Block diagram of computer controlled SEM, used for software calibration.(Fujioka, 1981).



(a) Transfer funstion of the system.
$0 \leq V_{in} \leq +5V$

(b)
squarewave
12ns risetime

(c)
squarewave
3ns risetime

Figure 3.3.6. Voltage contrast characteristics of the system prior to calibration. By oscilloscope: broken line, by Electron Beam Probe: unbroken line. (Fujioka, 1978)

measurement system to implement software linearisation of the collected current-surface voltage relationship. The computer controlled SEM and measurement system is shown in Figure 3.3.5. No energy analyser is included since the system was recalibrated for every measurement point and a discrete compensation function derived. The system utilised an electron beam sampling system to resolve high frequency waveforms.

The raw video signal was firstly lowpass filtered and then digitised using a 12-bit ADC converter. The resultant video samples were stored in the computer's 8kbytes of RAM. To carry out the necessary linearisation of the surface voltage-collected current relationship, a $5V_{pp}$ triangular waveform was applied to the node to be tested. The peak voltage of the calibration waveform defined the dynamic range of the measurement system. The sampling system was stepped through 600 phase delays and the sampled data stored. This could be repeated up to 100 times and mean-squared averaging implemented to improve the signal to noise ratio and therefore, the calibration accuracy. The difference between the applied waveform and the recorded waveform was computed and stored in the computer. Figure 3.3.6 shows the triangular calibration waveform(a) and waveforms after inverse filtering(b) and (c), when a $5V_{pp}$ squarewave of varying risetimes is measured with the system.

The minimum detectable voltage for this system was 50mV and the data acquisition time per node(T_n) was formulated as

$$T_n = (T_c + T_s.N.M) \qquad (3.3.1)$$

Where M is the number of sampling positions, N the number of averages, T_c the calibration time (0.72s) and T_s the sampling time,

which is 0.0012s for M=600, N=1 and T_c=0.72s. Hence the total measurement time must take into account the calibration time and at least one sample. An additional constraint was caused by the size of the computer's memory. This was that the number of sampling positions(P) which could be stored was determined by the number of phase delays(M) and can be formulated as

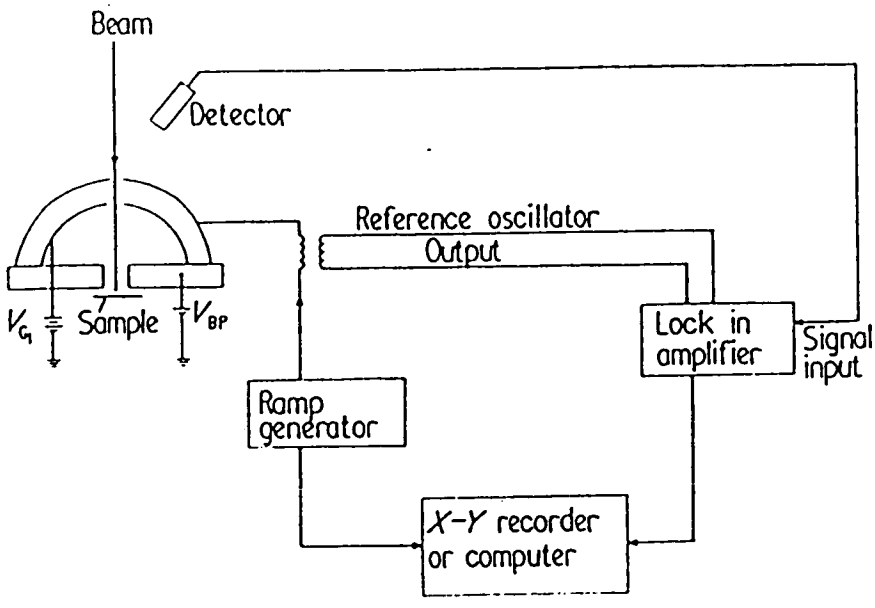
$$(M+2)P < 8192 \qquad (3.3.2)$$

The main disadvantages of such systems is the need to calibrate the system for each individual measurement point, resulting in long measurement times, and the effect of time variant nonlinearities introduced by time dependent waveforms on adjacent nodes. However, Fujioka measuring four different points on a quadruple multiplexer i.c., all at the same potential, deduced that the error in calibration due to the effects of local fields (up to 25kV/mm) was negligible, although the error was not quantified.

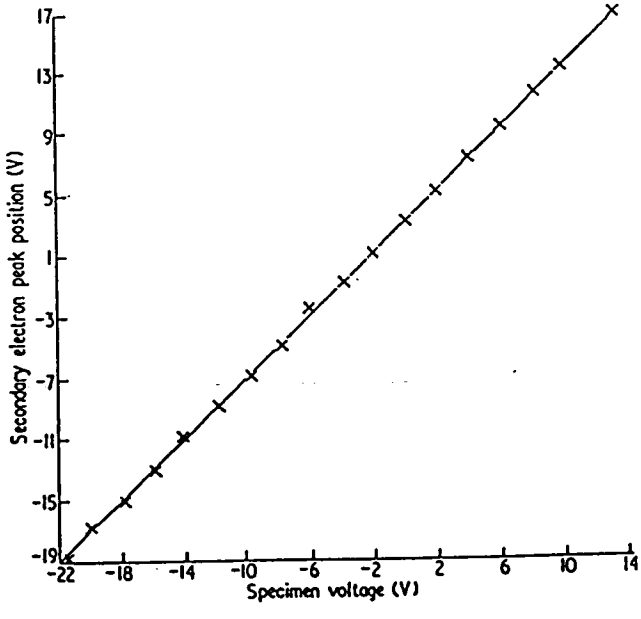
3.3.3) Modulation Techniques.

Although modulation techniques were extensively used as a method of isolating voltage contrast from topographical contrast, they were not used in the measurement process itself until Hardy et al.(1975).

This system utilised a hemispherical retardation energy analyser, with both a retardation and extraction grid. The system is shown in Figure 3.3.7(a) and demonstrates the method of coupling a modulation waveform onto the scanning, ramp waveform via an isolating transformer. The output signal was detected by a lock-in amplifier, after having been filtered in a bandpass filter with a Q of 10, centred at 1kHz. By applying small signal theory to the analyser it can



(a)
Electrical circuit for voltage contrast detector.



(b)
Voltage transfer function of the measurement system.

Figure 3.3.7. The highpass energy filter and instrumentation system used by Hardy et al.(1975)

be shown that the amplitude of the fundamental in the output signal is directly proportional to the first derivative of the operating characteristic. This principle was not new to energy analysis, having been used by Taylor (1974) as applied to Auger spectroscopy. A more extensive discussion of the relating theory will be given in the next chapter, since it forms the basis of the work presented.

Hardy et al. made quantitative measurements by noting the shift in the position of the peak of the SED as the surface voltage was measured. They claimed a greater dynamic range and sensitivity than from Banbury's detector(Figure 3.1.2) and demonstrated the inherent independence of this measurement technique to either variations in collected or beam current. In addition, the compatibility of this technique with a sampling system was shown. The linearity and dynamic range of this system is demonstrated by the voltage transfer function plotted in Figure 3.3.7(b). Note however, the measurement system is inherently manual.

3.4) Closed Loop Systems.

Flemming and Ward(1970) concluded that if measurement of an externally applied voltage was to be executed, as opposed to work functions (see Haas(1966)), the measurement process was comparative and did not require plotting the SED exhaustively. To avoid this, they proposed a feedback system to linearise the energy filter characteristic which was constrained to be a highpass type, since it alone was a monotonic function. Subsequently, highpass energy filters were adopted without exception. The advantages of this type of system were that the output SNR was improved, since more current was being collected and the bandwidth was increased since the scan-

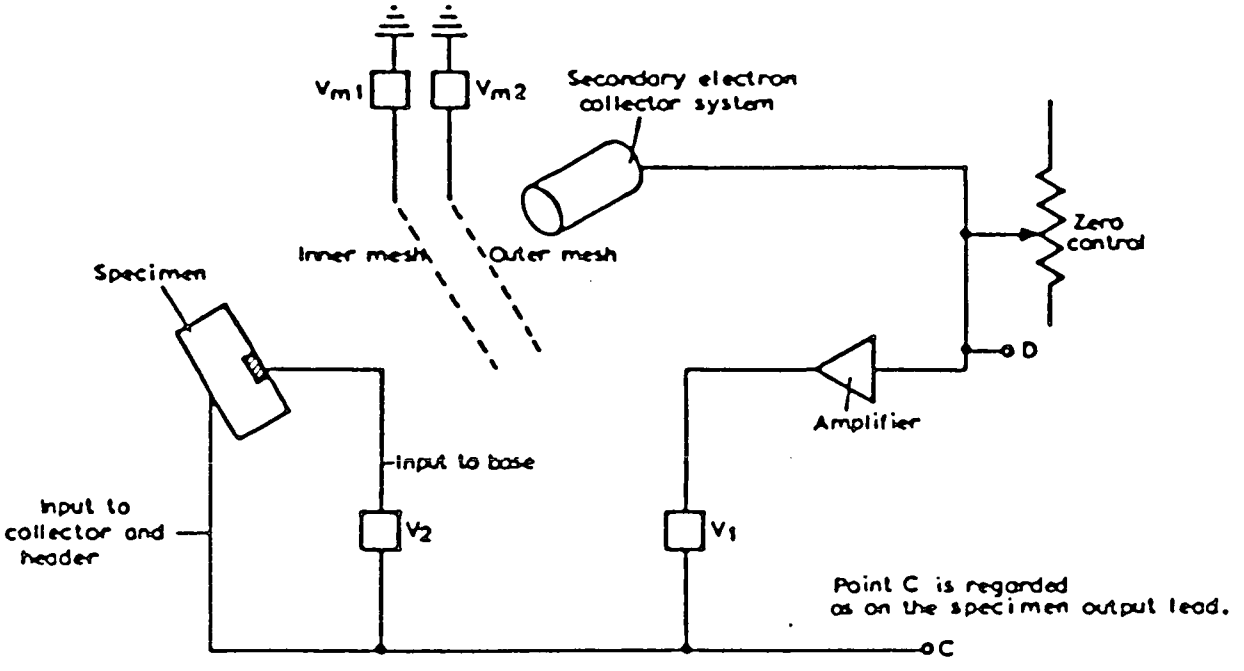


Figure 3.4.1. Floating ground system of feedback introduced by Flemming and Ward(1970). $V_{m1} >> V_{m2}$.

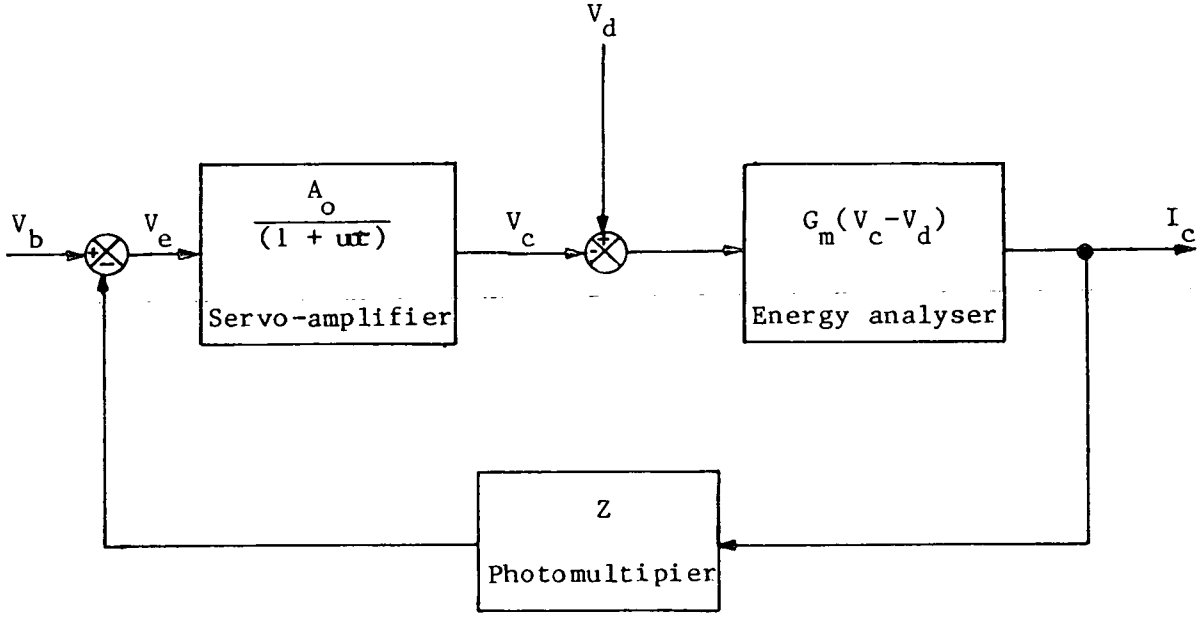


Figure 3.4.2. System block diagram of feedback loop.

ning function became redundant.

3.4.1) Linearisation.

The linearisation principle adopted by Flemming and Ward forms the basis of all present day measurement systems and therefore, it will be discussed in detail. A schematic of their system is shown in Figure 3.4.1.

Using an extraction mesh held at 1kV to suppress trajectory contrast and a retardation grid held at 0V, Flemming and Ward exploited a retardation type energy analyser to make quantitative surface voltage measurements with a voltage resolution of 0.25V. By employing a control loop which sensed the collected current and applied a correction voltage to a floating specimen to maintain the collected current constant, the correction voltage became directly proportional to the surface voltage. This in effect altered the position of the SED on the energy abscissa to ensure that the area of overlap between the filter characteristic and the SED was maintained fixed. Consequentially, a constant current was passed.

In applying feedback theory to the system, the shift in the energy distribution can be modelled as a disturbance input(V_d) summed with a correction input(V_c), as shown in the block diagram of Figure 3.4.2. Assuming a linear analyser characteristic with a transconductance, G_m and a single pole instrumentation system, the collected current(I_c) can be formulated as:

$$I_c = \frac{G_m A_o V_b}{(\omega\tau + 1 + A_o Z G_m)} + \frac{G_m V_d}{(\omega\tau + 1 + A_o Z G_m)} \qquad (3.4.1)$$

Where A_o is the loop gain of the system, Z the transimpedance of

the photomultiplier tube and the 3dB frequency(F_{3dB}) can be expressed as

$$F_{3dB} = \frac{1 + A_o Z G_m}{2\pi\tau} \tag{3.4.2}$$

If the forward loop gain fulfills the condition

$$A_o \gg \frac{1 + \omega\tau}{Z G_m} \tag{3.4.3}$$

equation 3.4.1. reduces to

$$I_c = \frac{V_b}{Z} + \frac{V_d}{Z A_o} \tag{3.4.4}$$

and simplifies to

$$I_c = I_q + I_e \tag{3.4.5}$$

The collected current is composed of a quiescent current(I_q), which is proportional to the reference voltage and an error current(I_e) which is inversely proportional to the forward loop gain of the system. Hence the system is a classical regulator, maintaining the collected current fixed with proportional control and dominant pole compensation.

The correction voltage(V_c), which is applied to the retardation grid can be expressed in similar terms:

$$V_c = \frac{A_o V_b}{1 + A_o Z G_m} + \frac{A_o Z G_m V_d}{1 + A_o Z G_m} \tag{3.4.6}$$

which for the condition of equation 3.4.3 reduces to

$$V_c = \frac{V_b}{Z G_m} + V_d \tag{3.4.7}$$

The correction voltage consists of a DC bias voltage term, proportional to the reference voltage and a term proportional to the disturbing surface voltage. If the DC bias voltage term is ignored, the following linear relationship is deduced.

$$V_c = V_d \tag{3.4.8}$$

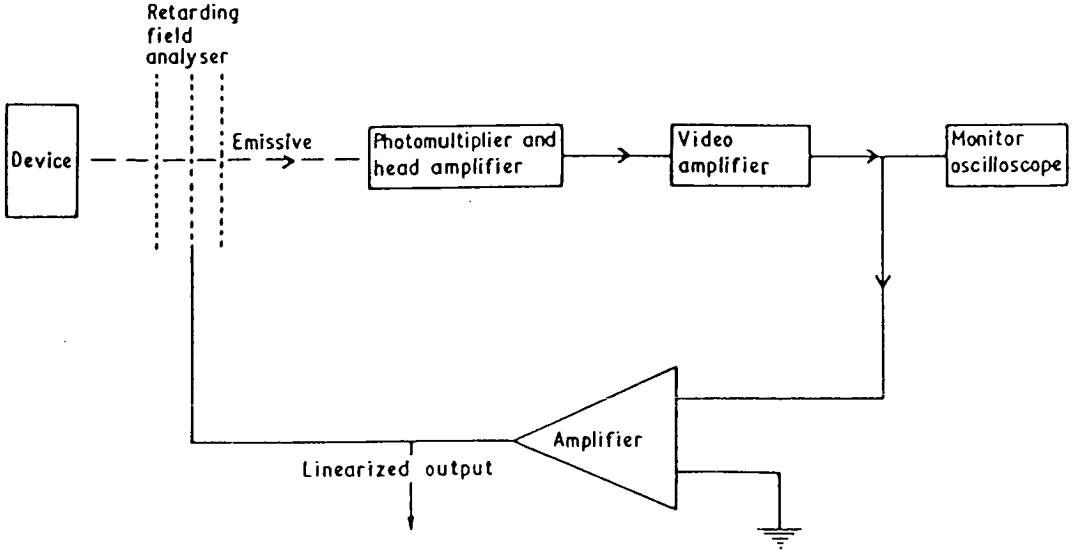
Thus the correction voltage is directly proportional to the surface voltage and is in phase with it. This can be explained intuitively as follows: For a given correction voltage a positive applied surface voltage decreases the magnitude of the collected current. To restore the collected current to its nominal value, the correction voltage must be made more positive or more correctly less negative.

The error voltage(V_e) is the difference between the applied surface voltage and the correction voltage and can be expressed as follows:

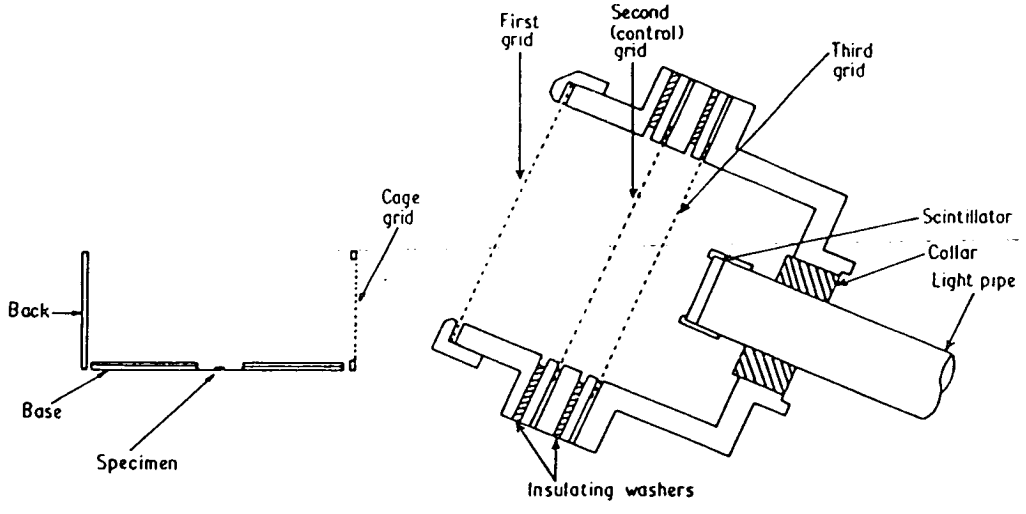
$$V_e = V_d - V_c = \frac{V_b}{ZG_m} + \frac{V_d(\omega\tau + 1)}{A_o ZG_m} \tag{3.4.9}$$

Therefore, the error voltage comprises a DC term proportional to the reference voltage and a term inversely proportional to the forward loop gain and directly proportional to the 3dB frequency of the system.

The nonlinearity of the analyser characteristic is attenuated by a factor proportional to the forward loop gain. This is an inherent property of negative feedback, exploited in audio amplifiers (Black, 1934, Roberge, 1974). This can be interpreted intuitively in that the feedback maintains a small enough error current to ensure that the "S"-curve is linear over the perturbation range.



(a)
Feedback loop.



(b)
Energy analyser.

Figure 3.4.3. Floating retardation grid configuration adopted by Gopinath and Sanger(1971).

3.4.2) Floating Grid Systems.

Although floating earth type systems were utilised by Flemming and Ward(1970), Hannah(1974) and Balk(1976), these were superseded by the floating grid type configuration, introduced by Gopinath and Sanger(1970, 1971). The main reason for the adoption of the floating grid was that adjusting the specimen reference voltage was not practically feasible for LSICs since customers require the circuits to be tested at the specified working condition voltages. In addition the floating ground technique has the inherent problem of coupling essentially DC waveforms of digital circuits to a continuously varying floating device. This would require the entire test system to be floated and modulated at the same frequency as the observed signal.

The system utilised by Gopinath and Sanger(1971), shown in Figure 3.4.3 incorporated a Banbury type deflection cage and a parallel grid retardation structure, with large deflection voltages(±100V) which suppressed the yield contrast mechanism exploited by Banbury.

Fentem and Gopinath(1974) utilised a hemispherical type energy analyser similar to Driver's, but with the inner grid used as an extraction electrode and a highpass characteristic. The analyser set-up, shown in Figure 3.4.4., originally incorporated a solid external hemisphere as the electron collector, with an aperture at the top for the primary electron beam. As a consequence of this construction, the specimen could not be viewed prior to probing and the detection process was inherently noisy becuase SEs were generated by the backscattered electrons and the high impedance electrometers. Dominant pole or bandwidth reduction compensation was employed in the feedback system which had to remain stable even when no electron

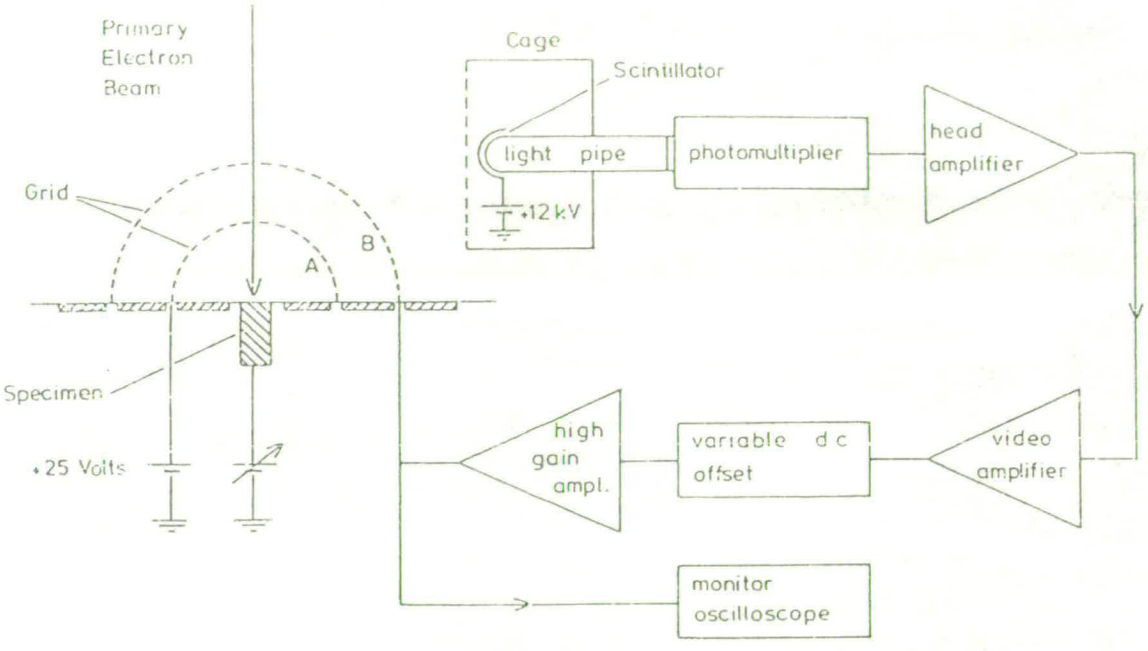
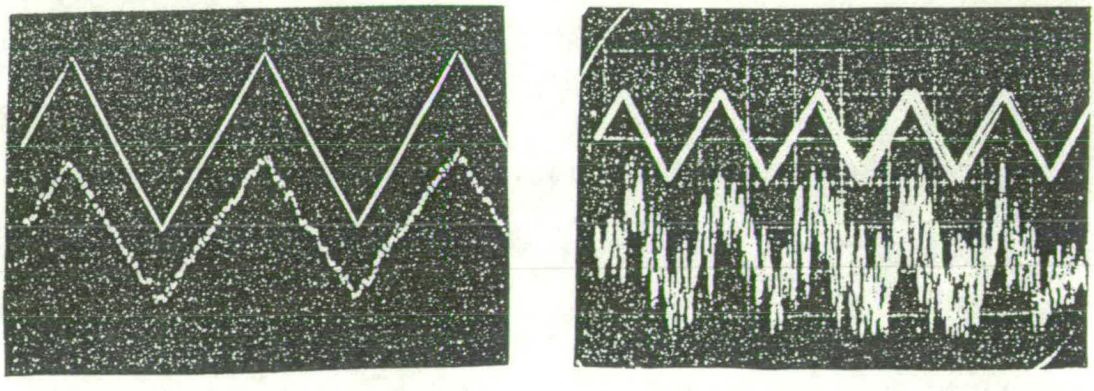


Figure 3.4.4. Schematic diagram of the measurement system used by Fentem and Gopinath(1974).



(a) (b)

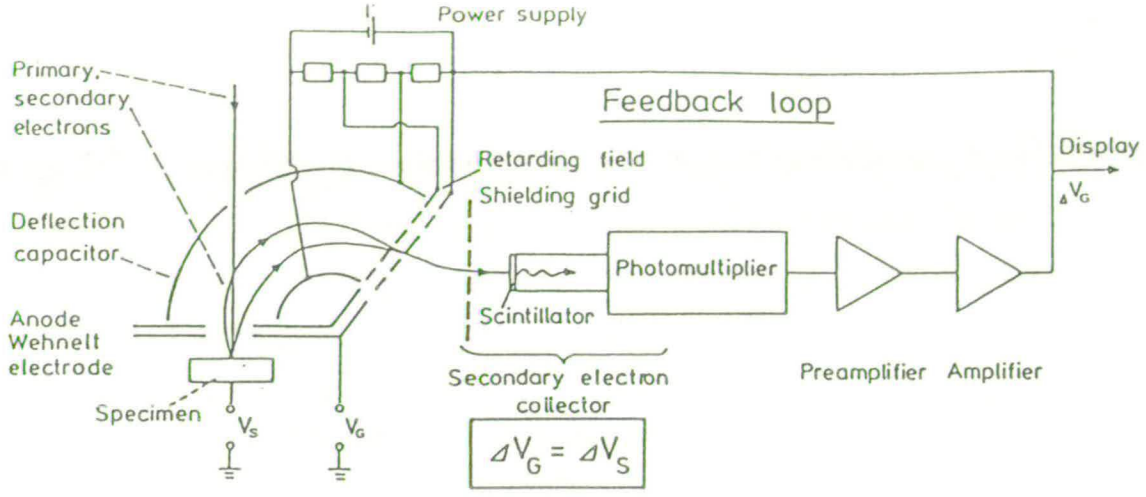
Figure 3.4.5. Comparison of the applied surface voltage(Upper Trace) and that obtained from the feedback loop(Lower trace), for the following triangular waveforms: (a) $\pm$10V, (50/mm); (b) $\pm$0.5V, (50V/mm). (from Fentem and Gopinath, 1974).

current was collected. A limited bandwidth of 1Hz was attained with this type of arrangement because of the capacitive coupling between the concentric grids which led to positive feedback.

This analyser was rejected and an alternative constructed, based on Driver's two grid arrangement with collection by the normal SEM collection system. Some linearity in a 20V bipolar range was sacrificed for an increase in bandwidth of 3 orders of magnitude and a voltage resolution of 80mV. This arrangement allowed the specimen to be viewed while measurements were not being taken. The waveforms presented in Figure 3.4.5 demonstrate the performance of the analyser and instrumentation system.

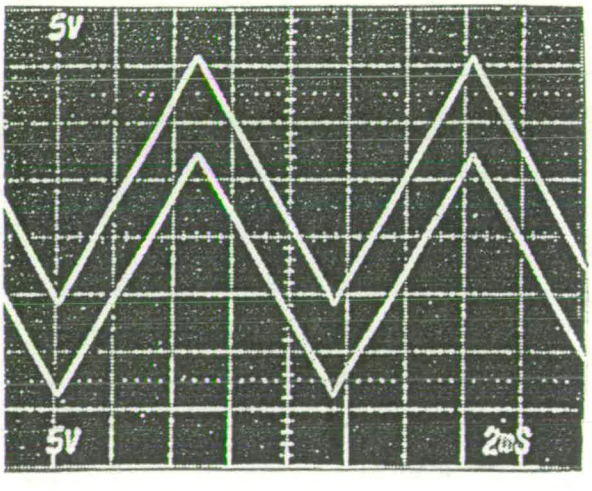
Using the sectoral energy analyser with parallel grid retardation, shown in Figure 3.4.6(a) Feuerbaum et al.(1978) achieved a 20V dynamic range with a 100Hz triangular waveform and a voltage resolution of 10-100mV. The system not only fed the correction voltage to the retardation grid, but also applied a fixed fraction of it to all the other electrodes. With an improved system using a cylindrical parallel plate analyser and negative feedback, Feuerbaum(1979) reported the best system performance to date in terms of bandwidth and voltage resolution, with an estimated 3dB bandwidth of 120kHz and a maximum voltage resolution of 1mV. The system is shown in Figure 3.4.7(a) while the bandwidth estimating waveforms are displayed in Figure 3.4.7(b).

A disadvantage shared by all of the systems discussed is that a sometimes lengthy and intricate, manual set-up procedure preceded a measurement. This included adjusting the photomultiplier black level and gain and the analyser bias voltages to ensure a high collected



(a)

Schematic representation of energy analyser and signal processing unit.



(b)

Triangular signal applied to an aluminium platelet. The upper curve was measured with the oscilloscope and the lower obtained using the electron beam probe.

Figure 3.4.6. Voltage Contrast measurement system adopted by Feuerbaum et al.(1978).

current. Thereafter these parameters had to remain constant throughout the experiment.

3.5) Measurement Errors.

Fentem and Gopinath(1974) stated three assumptions made to qualify the operation of their analyser system. These were:

a) The applied voltage does not alter the shape of the SED.

b) That the origin of the distribution moves linearly with applied voltage with respect to ground.

c) That the applied potential does not result in any field enhanced emission of SEs.

To this a fourth should be added:

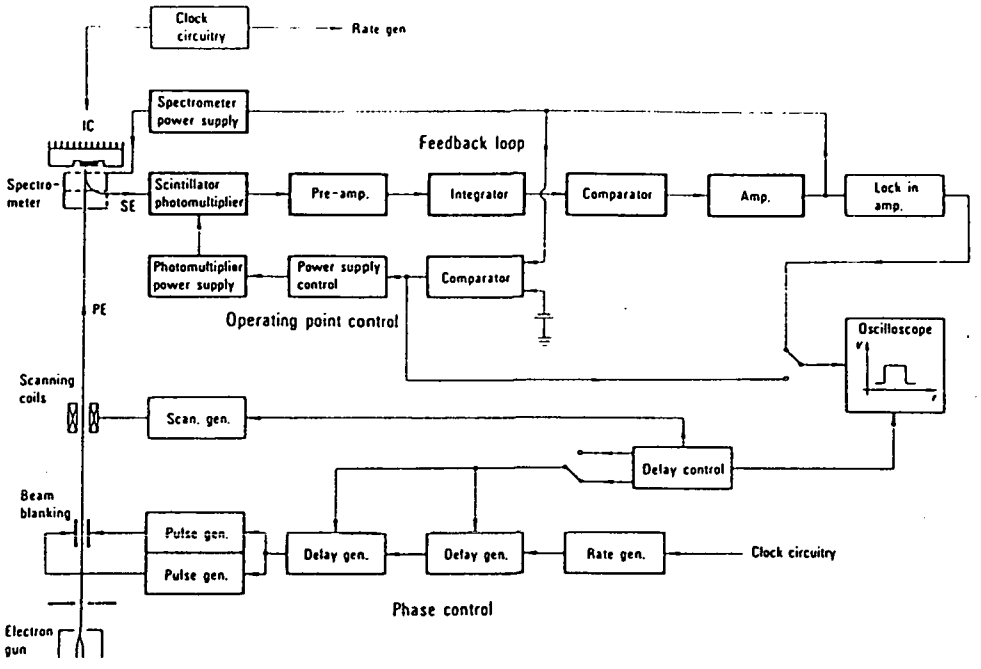
d) That all collected electrons should undergo acceleration and decceleration through homogenous electromagnetic fields, from leaving the specimen to entering the Faraday cup collection system.

These tenets are implicit in the principles of operation of all the measurement systems discussed and if held, measurement errors can be reduced to two types: a) current coupled and (b) energy coupled.

3.5.1) Current Coupled Errors.

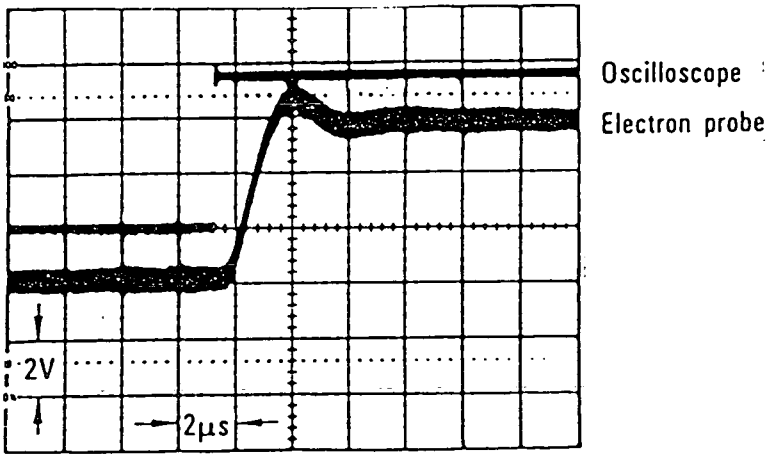
When the emitted SE current is varied the effect on the SED is shown in Figure 3.5.1, curves (a) to (b). There is a proportional increase in the area under the curve which results in a deviation of the "S"-curve from its original shape as displayed in Figure 5.3.2, curves (a) to (b).

For a given quiescent current the net effect is a variation in



(a)

Schematic of instrumentation system adopted for measuring waveforms using the sampling mode.



(b)

Measurement of rise time (1.2μs) of the feedback loop from the smearing of a measured squarewave signal.

Figure 3.4.7. Voltage Contrast measurement system used by Feuerbaum(1979).

the transconductance of the energy filter characteristic. From equation 3.4.7. if the transconductance varies, the correction voltage must alter accordingly to maintain a fixed quiescent current. Hence an error voltage(ΔE_i) is introduced into the measurement process, as illustrated in Figure 3.5.2, curves (a) to (b). Theoretically, the significance of this error could be minimsed by operating at low quiescent currents. However, this would result in a degradation of the SNR. Peak detection systems are inherently immune to this type of error since the electron energy at which the peak of the SED occurs is not a function of the SE current.

3.5.2) Current Error Sources.

Any parameter which causes the collected SE current and hence, the transconductance of the energy analyser to vary will induce a current coupled error. Since from chapter 2

$$I_{sec} = \delta I_p \qquad (3.5.1)$$

the transconductance is a function of the SE yield of the material under test and the primary beam current. i.e.

$$I_c = G_m(\delta, I_p)(V_c - V_d) \qquad (3.5.2)$$

Topographical contrast will also influence the transconductance, but can be minimised by selecting a topographically oblique test point.

3.5.2(a) Yield.

If measurements are made on each of the types of interconnect discussed in chapter 1 and compared, errors introduced by the variation of SE yield will be recorded. Feuerbaum(1979) reported a system,

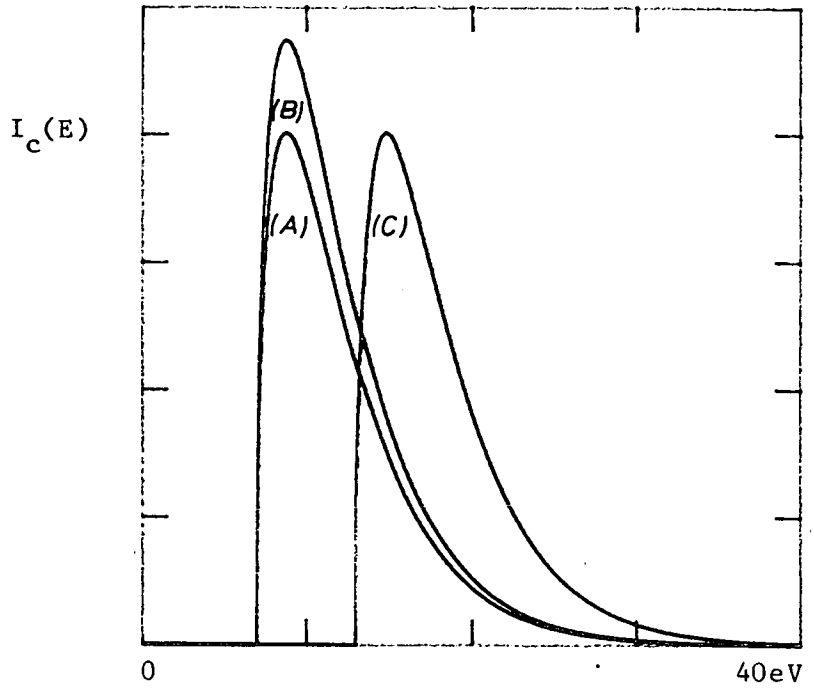


Figure 3.5.1. Variations in the SED caused by deviations in emitted SE current, curves (a) to (b), and peak energy, curves (a) to (c).

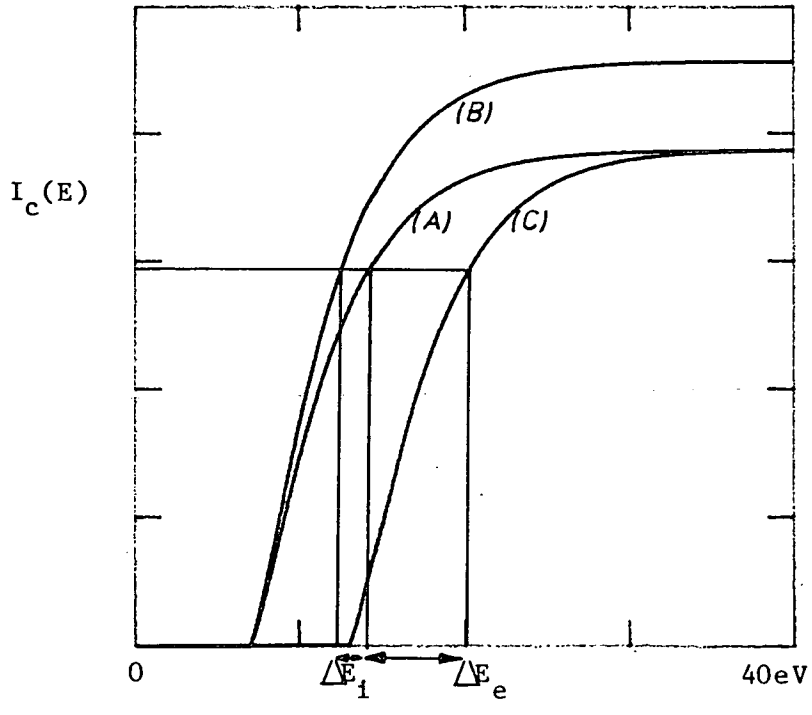


Figure 3.5.2. The effect changes in the SED has on the energy analyser operating characteristic, and the resultant errors.

ΔE_i : current coupled error, curves (a) to (b).

ΔE_e : energy coupled error, curves (a) to (c).

shown as part of the block diagram in Figure 3.4.7. which could minimise these errors. A second feedback loop was incorporated around the photomultiplier, its power supply and the energy analyser. For each measurement point the DUT was earthed and a DC voltage(+6V) applied to the retardation electrode. The secondary feedback loop was then closed and the output from the videohead amplifier compared to an arbitrary reference voltage. The gain of the photomultiplier tube was servoed via its supply voltage until the videohead amplifier output voltage equaled the reference voltage. The relationship between the videohead output voltage(V_o) and the reference voltage(V_{ref}) can be expressed

$$V_{out} = \frac{KG_m V_c V_{ref} A_o}{(1 + KG_m V_c A_o)} \qquad (3.5.3)$$

Where A_o is the open loop gain of the error amplifier and both the correction voltage and photomultiplier gain proportionalty constant(K) are constant. For high loop gain, this reduces to

$$V_{out} = V_{ref} \qquad (3.5.4)$$

independent of any collected current variation. This forms an automatic gain control system which compensates for any change in energy analyser transconductance between measurement points by adjusting the transimpedance of the photomultiplier tube. Hence for each measurement point, the quiescent current is varied but the operating point on the "S"-curve remains the consistent.

The advantage of this system is that it automates the set-up procedure and avoids saturation of the photomultiplier tube when changing measurement points. Also it inherently calibrates each measurement point and DC measurements can be made by subtracting the

reference voltage from the correction voltage.

The major disadvantage of this error correction procedure is that it implicity assumes that the transconductance of the analyser does not vary throughout the measurement period. Hence it is a low bandwidth system and is only effective if the measurement time is much less than the rate of change of the transconductance. As will be shown this is not normally the case.

3.5.2(b) Contamination.

Contamination arises in the SEM owing to the condensation or molecular deposition of hydrocarbon films, which are subsequently polymerised by the high energy primary electrons. The hydrocarbons originate from the oil used in the diffusion pumps, employed to reduce the vacuum to a workable level(10^{-5}Torr), the "O"-ring vacuum seals and the lubricant oil of the rotary roughing pump. Further sources of contamination are the outgassing of materials used within the specimen chamber and the hydrocarbons introduced by the operator's hands.

Seiler and Stark(1965) reported that contamination attenuated the secondary electron yield, while Conrue and Laberge(1975) stated that the formation rate of polymerised oil by volume was approximately linear with time for a specified beam current. The fact that these films could be conductive was indicated by Yakowitz et al.(1972). Contamination manifests itself in the measurement process, as a steadily increasing error voltage or if a waveform is being observed, as a gradual shift in the baseline. No effect on the voltage swing of an AC waveform will be observed, since the shift in the SED is independent of collected current.

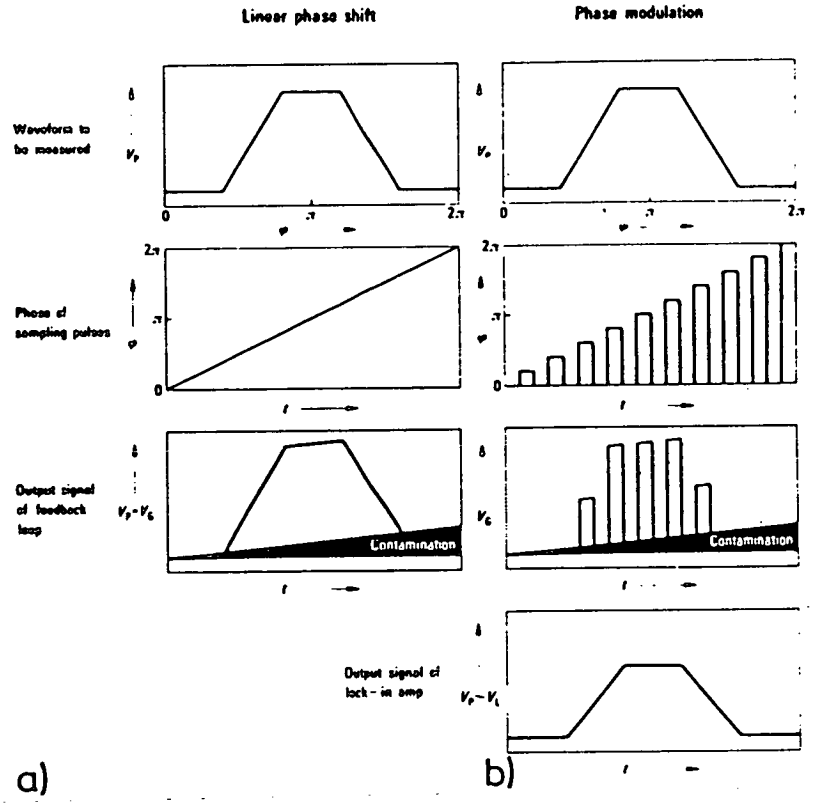


Figure 3.5.3. Comparison of theoretical waveforms for linear phase delay and phase modulation, utilised by Feuerbaum to eliminate baseline drift. (from Feuerbaum, 1979).

The effect of contamination induced errors can be arrested by a combination of two methods: a) good SEM vacuum practice and b) signal processing. The rate of deposition of the hydrocarbon layer can be reduced by taking several preventive measures: using oil free pumps such as turbomolecular or cryogenic, utilising materials that do not outgas at low pressure such as polytetrafluoroethylene insulated cables and ceramic packaged i.c.s, backfilling with dry nitrogen when bringing the specimen chamber up to atmospheric pressure to prevent moist and contaminated room air being sucked into the chamber and finally, ensuring that the operator wears surgical gloves while working in the chamber. If the contamination has to be reduced further, nitrogen traps can be used in the diffusion pump line and a cold finger placed around the specimen. All these methods work by preventing the contamination reaching the specimen.

The contamination problem is especially acute in high accuracy sampling systems which as stated in chapter 2, can take up to 15 minutes to record a waveform. Feuerbaum(1979) presented a solution to both contamination and long term instrument drift. Figure 3.5.3(a) shows schematically the effect contamination has on a measurement when the phase delay is varied linearly over one period of the observed waveform. The baseline drift is evident. In Figure 3.5.3(b) the waveform corresponding to discrete variation of phase delay is shown without the baseline drift. This is achieved by taking a zero phase delay calibration sample after each phase delay sample. Subsequently, the output from the feedback system is in a sampled data form and since the peak-to-peak variation of the waveform is reproduced, independent of contamination, the low frequency drift component can be eliminated by amplifying this signal with a lock in

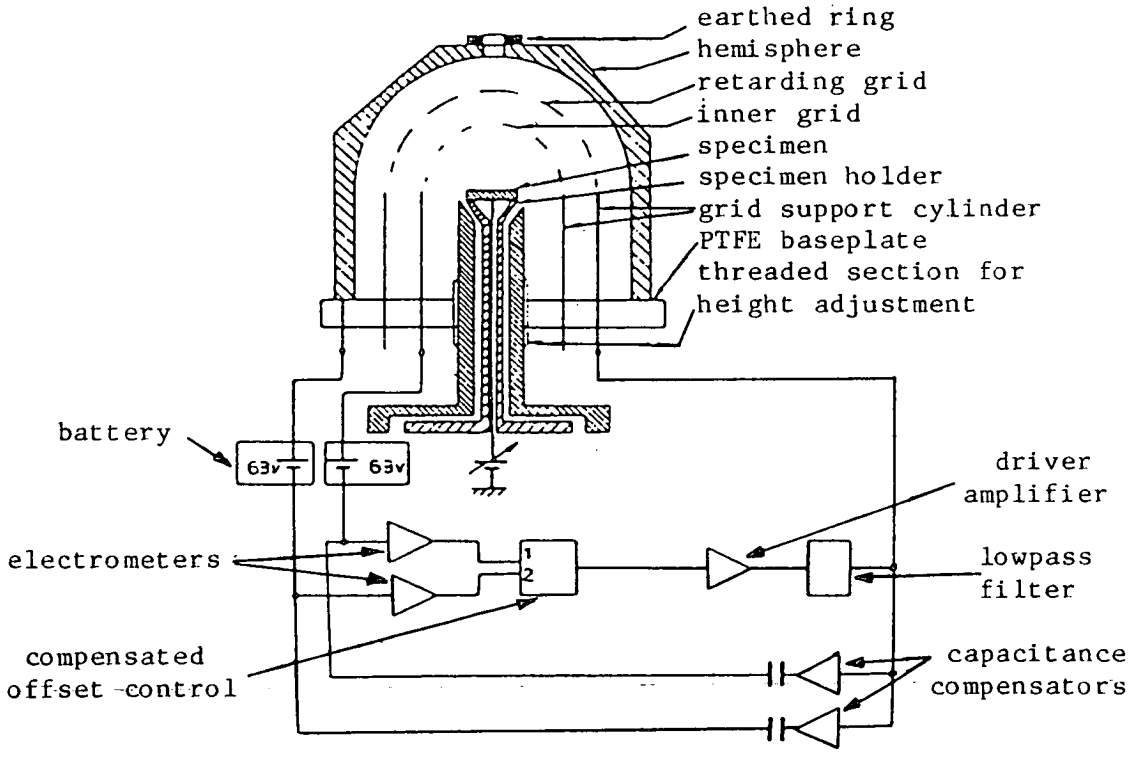


Figure 3.5.4. Schematic diagram of the modified Voltage Contrast measurement system adopted by Tee and Gopinath(1976).

amplifier. The disadvantage of this system is that it effectively halves the bandwidth of the system.

3.5.2(c) Primary Beam Current.

The primary beam current can be considered to be composed of a DC component and a noise component. Owing to the statistical nature of the electron generation the primary beam current can have a SNR as low as 4dB for a tungsten filament. For a high enough photomultiplier gain the SE current noise can be assumed to be comprised of Shott noise emanating from the primary beam current and therefore to be Gaussian and white. However, low frequency instrument drift and 60Hz mains interference can constitute secondary noise components. The noise manifests itself in the measurement process as a high frequency statistical variation of the analyser transconductance, which limits the measurement accuracy and the acceptable bandwidth of the system. The primary beam current and the quiescent current should be made as large as possible to maximise SNR. An alternative method of improving the system's accuracy is to limit the system bandwidth, evoking the familiar trade-off between settling-time and accuracy.

Lanthanum hexaboride filaments are gradually displacing the Tungsten hairpin(Cambridge Instrument's S200) since they offer one to two orders of magnitude increase in current and brightness as described by Wells(1974).

Tee and Gopinath(1976,1977) incorporated an instrumentation scheme in their system, shown in Figure 3.5.4., to suppress beam current variation effects. Using a two grid hemispherical analyser with an outer solid collector hemisphere, both the highpass current on the outer hemisphere and the "retarded" current on the inner

extraction grid were collected and amplified in electrometers. The output voltages, corresponding to the collected currents were then passed into a computational circuit to produce an output of the form:

$$V_{out} = V_1 - K(V_1 + V_2) \qquad (3.5.5)$$

Where K is a proportionality factor which determines the operating point and is controlled by the compensation unit, V_1 voltage and V_2 the lowpass voltage. Hence the usual time invariant reference voltage has been replaced by a FIXED FRACTION of the total collected current, independent of primary beam current variations. In this way the operating point on the "S"-curve is maintained constant, although the quiescent current is continuously altering. V_{out} will only vary when a time dependent surface voltage is applied.

In order to avoid the instabilities which limited the bandwidth of Fentem and Gopinath's system(1974), noise cancelling techniques were employed. A fraction of the retardation voltage was inverted and added to the electrometer inputs giving a net interference signal equal to zero. The beam current can only affect operation of the system if it is large enough to saturate the electrometers or small enough to reduce the loop gain to produce an unacceptable loop gain accuracy. In addition to beam current variations, changes in collected current due to surface topography or composition are also suppressed by a factor of between 4 and 5, while beam current variations are suppressed by up to a factor of 20. Utilising a loop gain of 40-45dB a bandwidth of 50Hz was achieved with a voltage resolution of 2mV and a linearity of 1%.

3.5.3) Energy Coupled Errors.

Energy coupled errors arise when the position of the SE distribution on the energy abscissa corresponding to zero applied surface voltage varies. This is demonstrated in Figure 3.5.1, curves (a) to (c). This generates an error component, ΔE_e in the correction voltage to which both constant current and peak detection type systems are susceptible, as shown in Figure 3.5.2, curves (a) and (c).

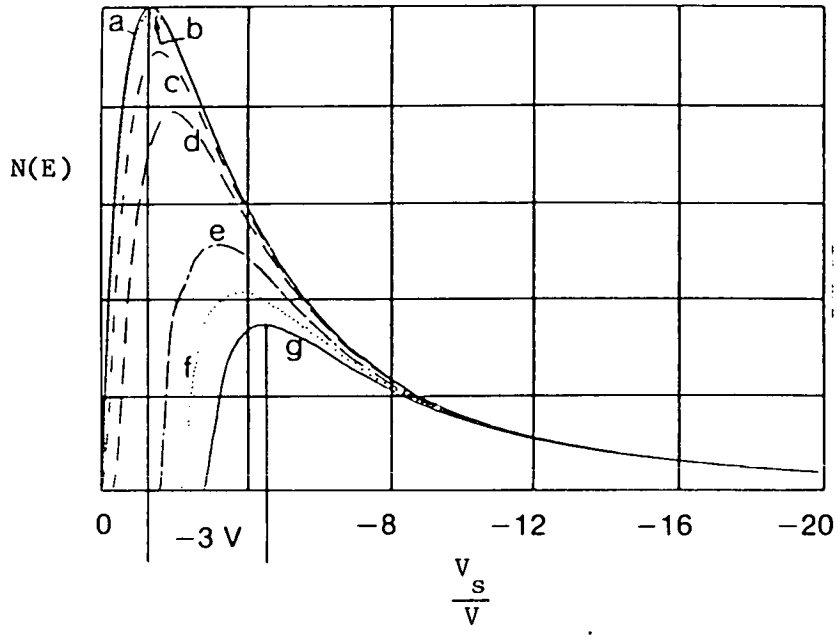
The major source of this type of error occurs when voltage measurements made on different materials, such as polysilicon and aluminium, are compared. This error can be eliminated by either of two methods: a) calibrate each measurement point, increasing the measurement time or b) limit all measurements to be made on the same type of material. i.e. the metallisation layer.

3.5.4) Secondary Electron Distribution Distortion.

The conditions set-out at the beginning of the section are generally only encountered in specimens which are conductive, of a large area and featureless, such as aluminium or copper stubs. However, a test system must be able to make measurements on real i.c.s, without loss of accuracy to be practically useful. In particular, the closely spaced signal carrying conductors present on i.c.s can seriously compromise point (d) of the conditions noted in section 3.5.1. and result in measurement errors which can be time-variant and only repeatable under certain conditions.

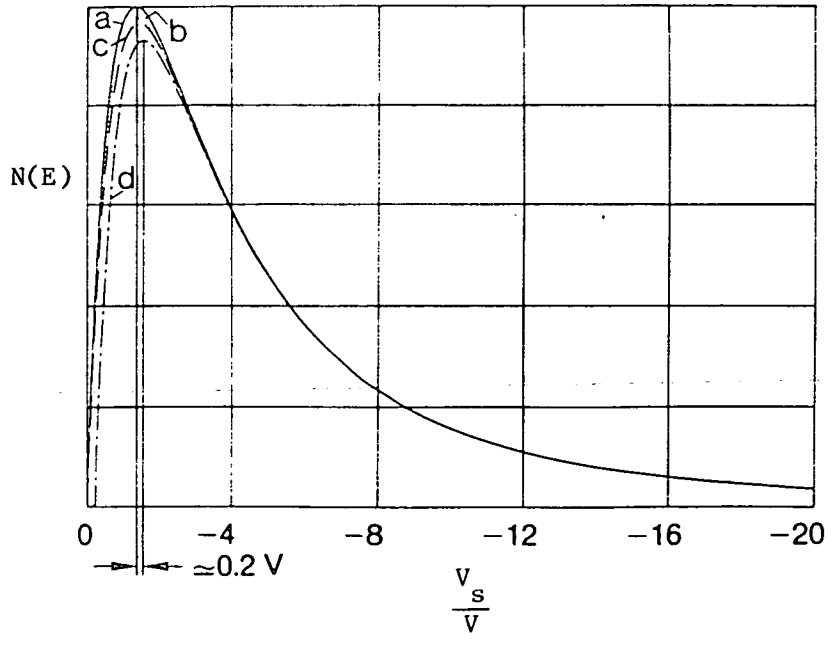
3.5.4(a) Local Fields.

Local fields were discussed in section 3.2.1. qualitatively and with a view to suppressing their effects. The magnitude of the



(a)

Extraction field = 0.4kV/mm. (a) V_s = 0; (b) V_s = +2V; (c) V_s = +3V; (d) V_s = +4V; (e) V_s = +6V; (f) V_s = +7V; (g) V_s = +8V.



(b)

Extracting field = 1kV/mm. (a) V_s = 0V; (b) V_s = +6V; (c) V_s = +7V; (d) V_s = +8V.

Figure 3.5.5. The influence of the local retarding field on the secondary electron distribution. (from Menzel and Kubalek, 1983).

measurement errors caused by them will be discussed here. Menzel and Kubalek(1983) simulated the effect the local electric field of the specimen, shown in Figure 3.2.3, has on the SED. With the two external conductors earthed, a positive voltage was applied to the central electrode, which was then probed and the resultant SED recorded. This was repeated for a positive voltage of varying magnitude and an extraction field of 0.4kV/mm. The results are presented in Figure 3.5.5(a) with the distributions normalised in the energy domain by subtracting the surface voltage from the energy displacement. These curves show that fewer and fewer low energy electrons are collected as the magnitude of the surface voltage is increased and the net result is gross distortion of the SED. By increasing the extraction field to 1kV/mm, the effect of the local field is minimised but not eradicated, as shown in Figure 3.5.5(b). Both constant current and peak detection systems would suffer from these effects and the errors deduced from Figure 3.5.5(b) are 0.05V and 0.2V respectively.

This demonstrates that a constant current type system has a lower sensitivity to local field effects than a peak detection system. However, in the measurement range 0 to +5V the difference is minimal. Feuerbaum(1979) stated that the operating point should be chosen on the high energy side of the SED to minimise these effects.

3.5.4(b) Transverse Fields.

Figure 3.5.6 shows the effect that a time-varying potential applied to an adjacent electrode has on an electron beam measurement. No clock feed-through was apparent when the test point was mechanically probed and when the extraction field was increased to 1kV/mm a residual 10% error was observed.

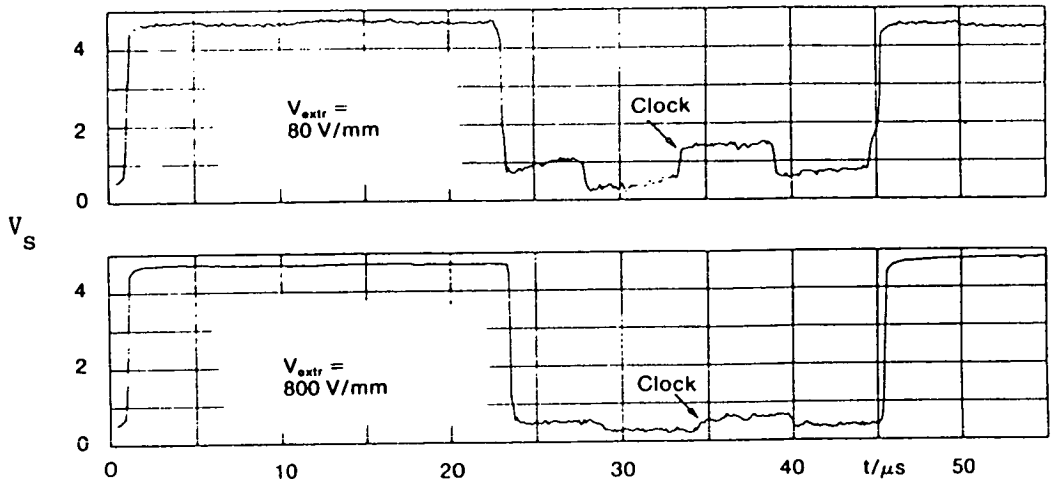
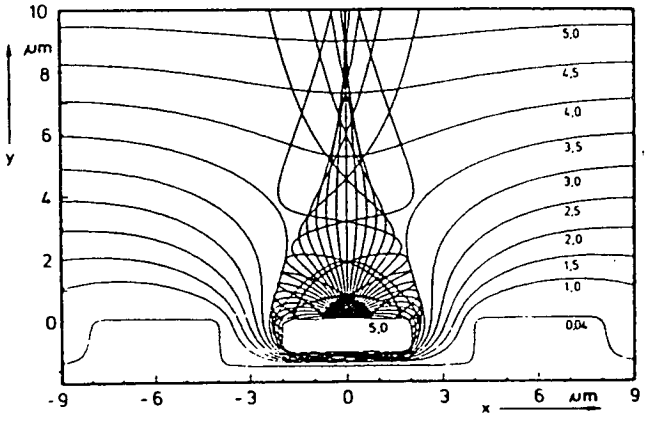
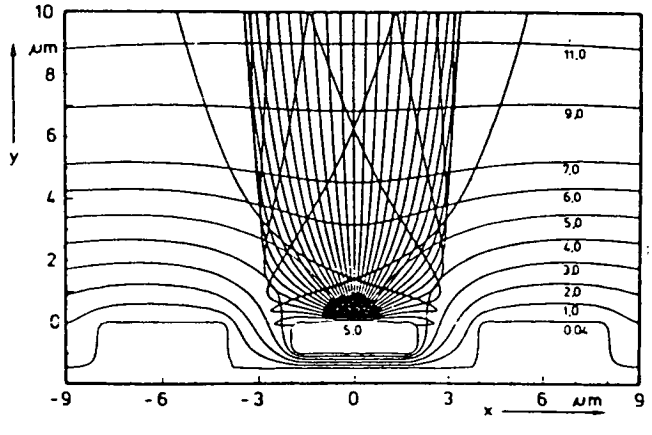


Figure 3.5.6. Feedthrough of adjacent electrode voltage into electron beam measurement.



(a)

Focussing in the x-y plane caused by transverse fields. Extraction field = 0.4kV/mm.



(b)

Defocussing in the plane of the electrode caused by the extraction field. Extraction field = 1kV/mm.

Figure 3.5.7. The effect of transverse fields on a 1.5eV electron. (from Menzel and Kublek, 1983)

Again Menzel and Kubalek(1983) simulated the effect with the test specimen of Figure 3.2.3. and again the external electrodes were held at OV, while the internal electrode was held at +5V and probed. It was found that the emitted SEs were focussed as shown in Figure 3.5.7(a) by the transversal fields while increasing the extraction voltage resulted in defocussing along the length of the electrode, as shown in Figure 3.5.7(b).

The effect of focussing was to increase the number of detected electrons of a certain energy, while the defocussing reduced the overall number of electrons collected. It was also found that the focussing effect dominated and the net result was an increase in the number of detected electrons. The distortion found in the corresponding "S"-curve is displayed in Figure 3.5.8. The ideal "S"-curve is that represented by (a), while (b) is the result of analyser transmission losses. The "S"-curve obtained with an extraction field of 0.4kV/mm is constant at low energies owing to local field effects but at higher energies more electrons are collected than in curve (b) as a result of the focussing effect. With an extraction field of 1kV/mm the effects of local fields and focussing are attenuated, but still give rise to measurement errors of 10 to 15%, as shown in Figure 3.5.9.

3.6) Practical Constraints.

Electron beam testing can be rendered impossible and terminate with the destruction of the DUT if certain precautions and limitations are ignored. High energy electrons can damage MOS transistors enough to induce functional failure. Field oxide and passivation can charge up generating an unpredictable electric field around the test

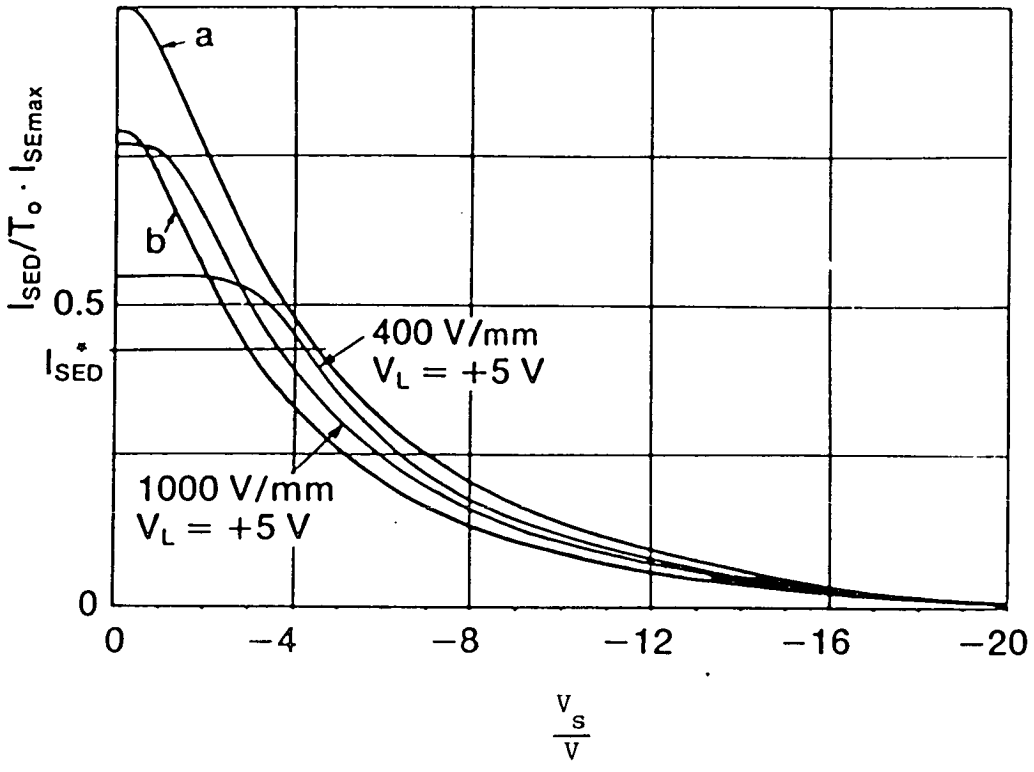


Figure 3.5.8. Changes of the integral SE spectrum at different extraction field strengths and at +5V line voltage. (from Menzel and Kubalek, 1983)

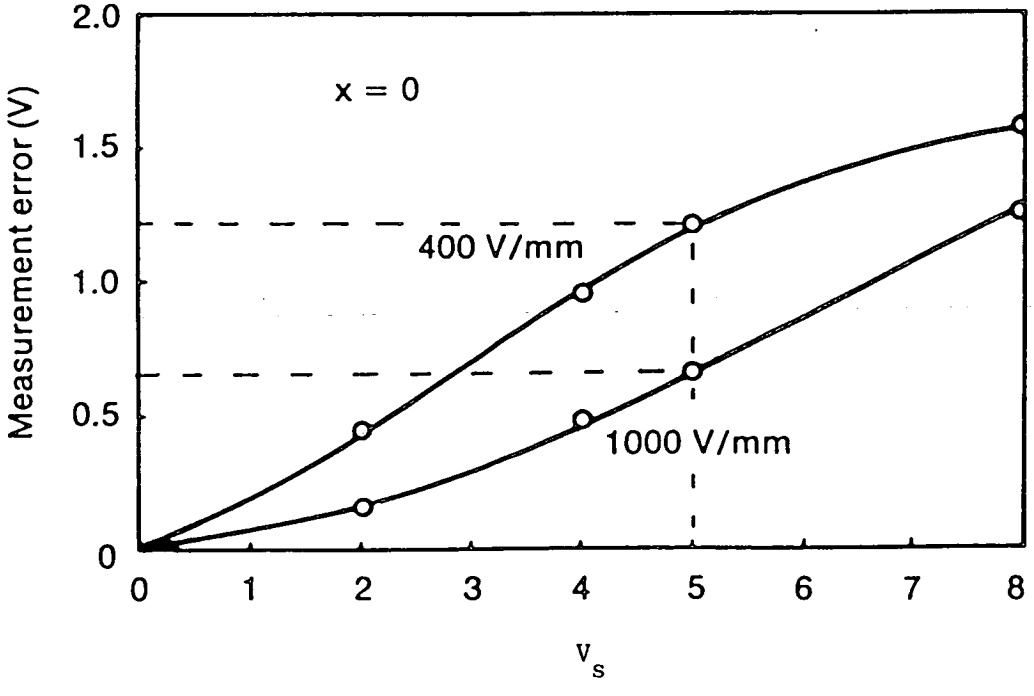


Figure 3.5.9. Measurement error due to the focussing of SEs in the x-y plane and defocussing in the plane of the electrode above positively biased i.c. lines. (from Menzel and Kublek, 1983)

point compromising any quantitative measurements made. These topics are discussed below.

3.6.1) Electron Beam Damage.

Although Green et al.(1965) had previously performed studies of electron beam damage to bipolar transistors, Speth and Fang(1965) were the first to report the effects of electron beams on metal-gate MOS field effect transistors. Their results, shown in Figure 3.6.1, show that for a primary electron beam accelerating voltage of 5kV and with the gate positively biased with respect to the substrate, the threshold voltage of the device shifted linearly with applied gate bias under irradiation. For a given gate bias, the threshold voltage would shift, reaching a steady state value and remain stable unless the gate bias was increased.

MacDonald and Everhart(1968) offered the explanation that for a given accelerating energy the primary electron beam penetrated the gate oxide, where it generated electron hole-pairs. The electrons, owing to their greater mobility were swept out of the gate under the influence of the gate bias, while the holes became trapped at the oxide-silicon interface, giving rise to an electric field. In addition, it was reported that the electric field strength is linearly related to the electron dosage or fluence(charge per unit area) until a saturation value is reached. Also the time taken for the saturation value to be reached is dependent on primary beam current.

As stated by Keery et al.(1976) other transistor parameters affected include drain-junction avalanche breakdown voltage, transconductance and drain-junction leakage current. Utilising the Evehart-Hoff(1971) equation to determine the minimum primary beam

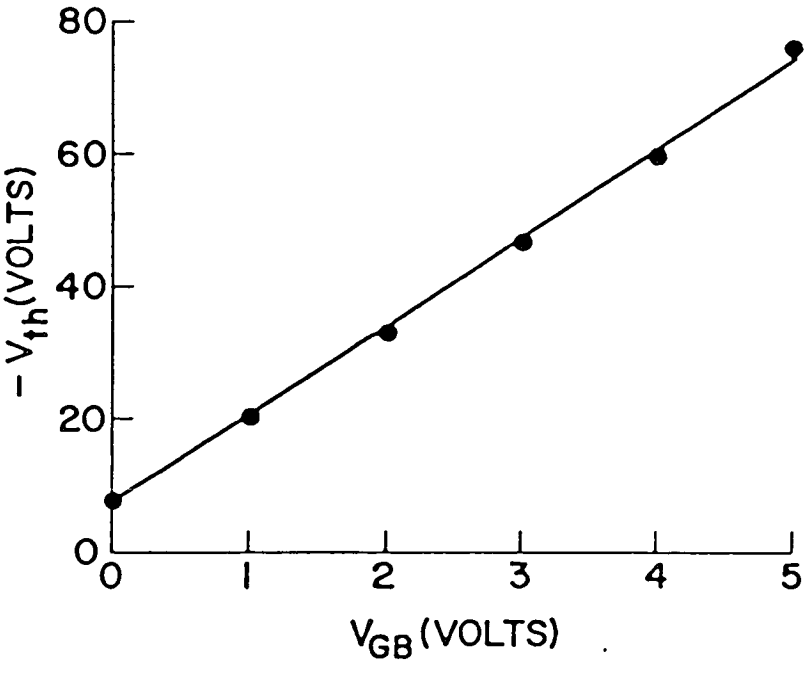


Figure 3.6.1. Steady-state threshold voltage as a function of the gate bias during electron irradiation. (from Speth and Fang, 1965)

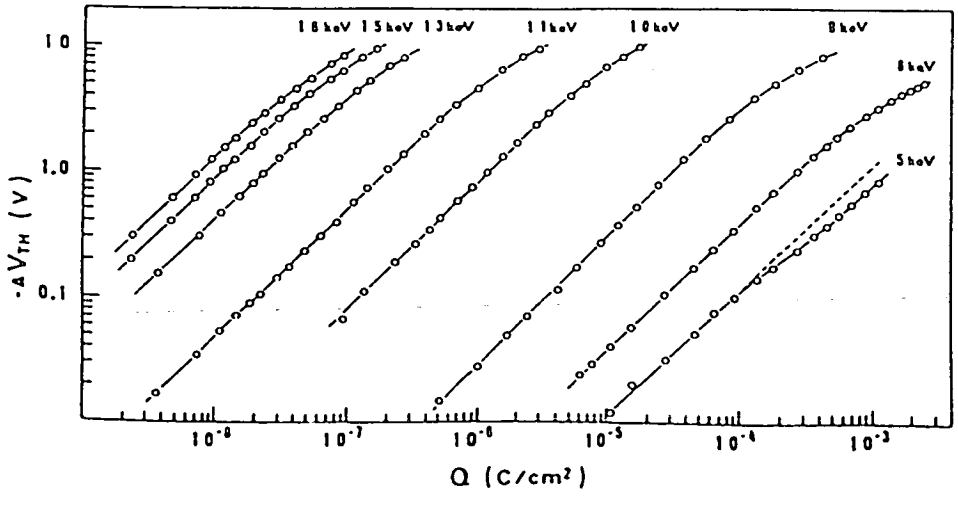
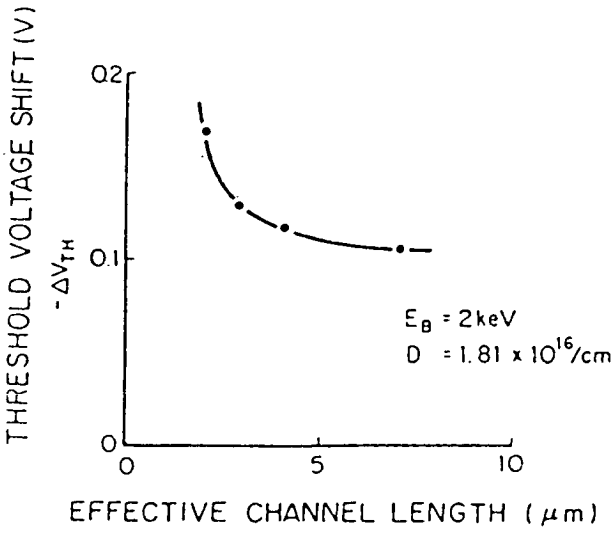


Figure 3.6.2. Threshold voltage shift as a function of electron beam fluence with the nominal electron energy as a parameter. The calibrated electron energy is 1.07 times as high as the nominal one.

accelerating voltage(12.3kV), which would not result in beam penetration of the gate oxide of a polysilicon gate device, Nakami et al.(1981) showed that even at accelerating voltages of 5kV, electron beam damage was apparent. The results are displayed in Figure 3.6.2. This was attributed to range straggling of the primary beam and secondary radiation of the primary beam.

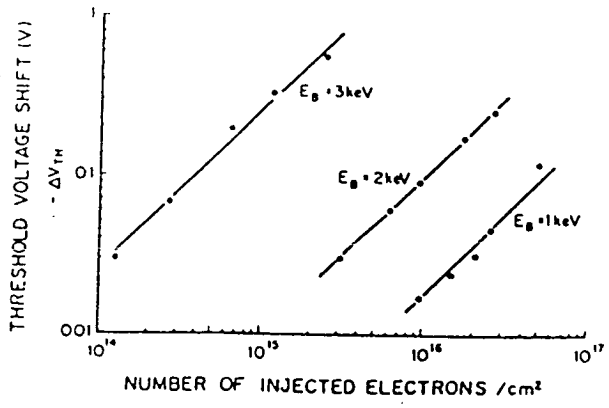
Again using a polysilicon gate device Miyoshi et al.(1982) irradiated long(4.1μm) and short(2μm) channel devices with a primary accelerating beam energy of 2kV and recorded the change in several parameters, as shown in Figure 3.6.3. A relationship was deduced between channel length and shift in threshold voltage, as shown in Figure 3.6.3(a) and it was concluded that devices with long channels were not as greatly affected as those with a short channels. Threshold shifts with primary beam accelerating voltage of 1kV were also reported, as shown in Figure 3.6.3(b) and attributed to secondary radiation of x-rays generating electron-hole pairs in the gate oxide. Although the threshold shift is proportional to the number of injected electrons, low primary beam accelerating voltages result in less threshold voltage shift for the same electron dosage. A pulsed primary beam was found to result in less damage than a continuous beam owing to the holes discharging during the beam off time, as shown in Figure 3.6.3(d).

Although, electron beam damage cannot be avoided it can be minimised by using low electron beam accelerating voltages, low beam current and the shortest possible measurement time on the gate of a device. The devices can be returned to their original condition by annealing in air at between 200 and 300^{o}C for a time dependent on the device of between minutes and hours.



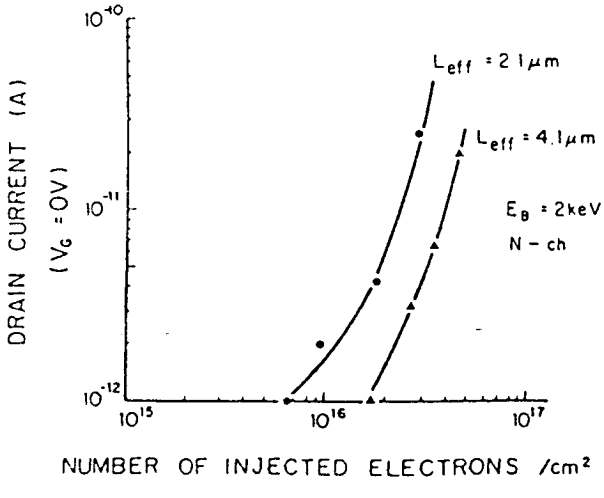
(a)

Relationship of the shift in threshold voltage of n-channel polysilicon gate MOSFET's and effective channel length.



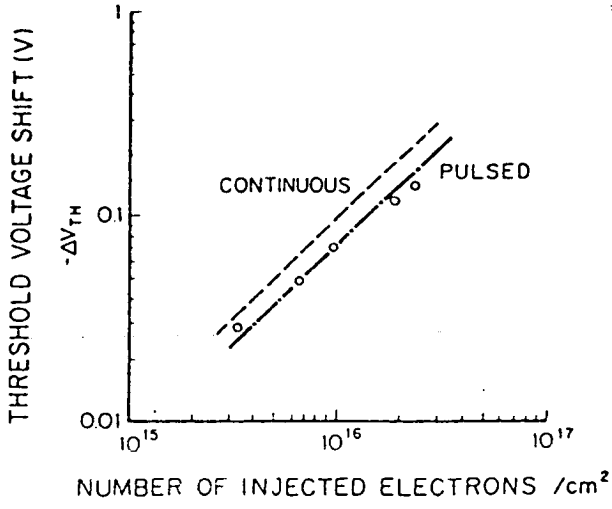
(b)

Shifts in threshold voltage of n-channel polysilicon gate MOSFET's as a function of the number of injected electrons with three different primary electron energies.



(c)

Shifts in drain leakage current of n-channel polysilicon gate MOSFET's as a function of the number of injected electrons.



(d)

Shift in threshold voltage on an n-channel polysilicon gate MOSFET as a function of the number of injected electrons with the pulsed irradiation.

Figure 3.6.3. The effect of electron bombardment on MOSFET parameters(Miyoshi et al.,1982)

3.6.2) Specimen Charging.

Normally, a finished device will be covered with a protective layer of insulator such as silicon nitride or silicon dioxide, about 1μm in depth. Taylor(1978) showed that voltage contrast could be observed through this protective insulator layer, if the primary beam accelerating energy(10keV) was high enough to enable the electron beam to penetrate the passivation. The high energy beam establishes a conductive channel through the insulator by means of Electron Beam Induced Conductivity (EBIC, Aris et al.1976) and therefore, the potential of the surface of the passivation layer assumes that of the underlying track. Further, it was reported that voltage contrast was observable if the primary beam did not penetrate the passivation layer, but that it was masked by the increasingly large negative potential of the surface as charge accumulated.

Alternatively, Kotorman(1980) was able to witness voltage contrast through the protective layer by reducing the primary beam energy to ensure that the electron yield was unity and therefore, negating any charging problems. In this way a capacitor structure is formed between the conductive track, the electron beam and the insulating protective layer and thus dynamic voltage contrast can capacitively couple through and be observed on the surface of the device.

Dinnis et al.(1982) reported another technique which involved coating the passivation with a thin layer of conductive material such as gold. The layer could not be so thick as to shield the electric fields nor should it make contact with the bond pads. Voltage contrast was clearly observable, although the accuracy of quantitative measurements was not confirmed.

3.7) Summary.

It has been shown in this chapter that quantitative measurements, exploiting the phenomenon of voltage contrast can be made and that the accuracy and settling-time of these measurements can satisfy the criteria for a test system presented in chapter 1. The bandwidth of such systems can be as large as 300kHz, while the resolution can be as low as 1mV. However, the accuracy of the measurements can be compromised by local field effects present on the surface of i.c.s and by primary beam current variations, giving measurement errors up to 15%. Hence, the best system performance can only be achieved with large area, topographically oblique specimens.

A voltage contrast instrumentation system must be based on the following four points to enable fast, accurate measurements to be made.

(a) OPERATING POINT : must be independent of the SE current. i.e. the normalised SED.

(b) A FEEDBACK loop should be incorporated to maximise bandwidth and dynamic range.

(c) AUTOMATIC GAIN CONTROL should be used to avoid saturation of the photomultiplier tube and to maximise SNR.

(d) CALIBRATION of each measurement point should be possible.

A hybrid of Feuerbaum(1979) and Tee and Gopinath(1976) systems would suffice these conditions. However, the latter system has a limited bandwidth and does not lend itself to photomultiplier type detection. Therefore, the alternative of a feedback loop based on peak detection was pursued in this work.

CHAPTER 4.

A VOLTAGE CONTRAST TEST AND MEASUREMENT SYSTEM.

To implement voltage contrast testing of +5V NMOS devices the three following constituent subsystems are necessary : a) a voltage measurement system with the property that the output signal has a 1:1 correspondence with the applied surface voltage, b) a utility subsystem capable of providing all the necessary stimuli to permit dynamic testing of the DUT and c) a data-logging subsystem to read and record relevant measurements including the output from the measurement system and any system parameters such as current consumption and supply voltage.

A major design constraint was that any test system constructed had to be compatible with the existing instrumentation as described by Dinnis et al., (1982). A complete system block diagram is shown in Figure 4.1 with the sections pertaining to this work in dashed outline. The original system consisted of a digital scan generator, a primary beam sampling system and a specimen stage position controller. All of these were programmable via an IEEE 488 instrumentation bus, under the control of a "Superbrain" microcomputer. Hence the IEEE instrumentation bus was a de facto instrumentation interconnect standard. Since the microcomputer supported the UCSD version IV Pascal programming language, this was used throughout for computer programming.

Synchronisation was necessary between the DUT clock frequency and both the sampling and scan generator subsystems to enable the implementation of the Voltage Coding and Logic State Mapping test modes, as discussed in chapter 2. Further, synchronisation with the

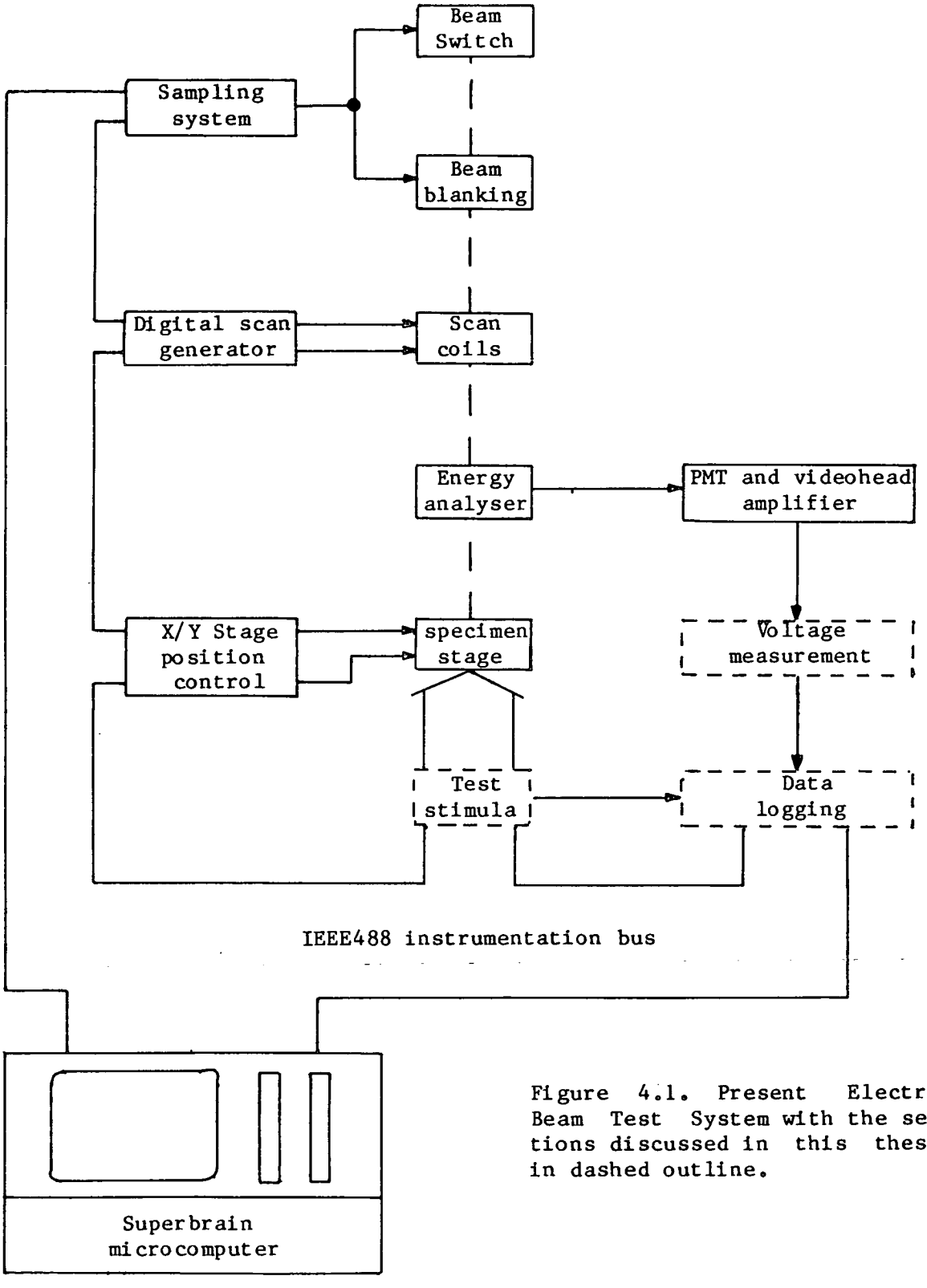


Figure 4.1. Present Electron Beam Test System with the sections discussed in this thesis in dashed outline.

data-logging subsystem should reduce the effect of additive noise in the system.

This chapter will present the design principles and preliminary experiments which led to the adopted system architecture.

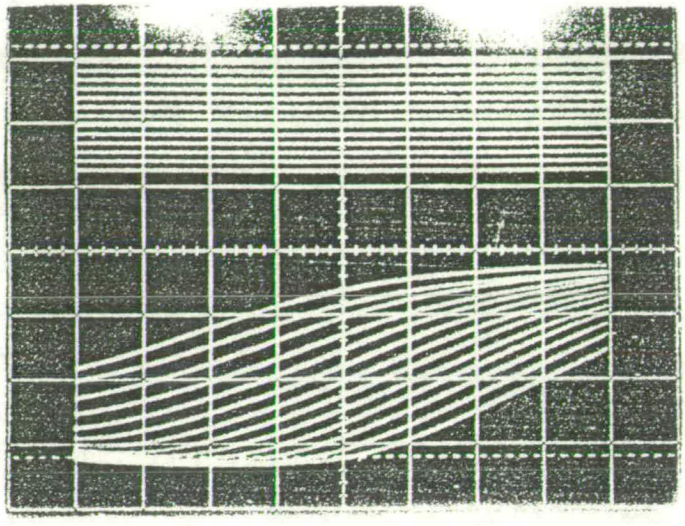
4.1) A Voltage Measurement System.

As a consequence of the arguments presented in the preceding chapter, it was decided to construct a measurement system incorporating the advantages of large dynamic range and bandwidth accrued from a feedback system, the beam current insensitivity of a peak detection system, and the good SNR of a highpass retardation type energy filter. Therefore, the voltage measurement system would consist of a high pass energy filter and a peak tracking servo-mechanism.

As discussed in chapter 3, the accuracy of the overall measurement system could be compromised, irrespective of the performance of the instrumentation, if the energy filter could not reproduce the integrated SED with a high degree of fidelity over the required applied surface voltage range and for a variety of i.c. surface structures. Therefore a series of experiments was implemented to evaluate the performance of the energy analyser intended to be used with the instrumentation system.

4.1.1) Energy Filter Evaluation.

The introduction of transverse electric fields between neighbouring conductors at different potentials can distort the "S"-curve reproduced by a high pass energy filter by varying the amount of electrons collected as a function of energy. In order to demonstrate any distortion present in the "S"-curves of the energy analyser



-ve V_r +ve V_r

Figure 4.1.1. Characteristic "S"-curves.
Upper trace: applied surface voltage 1 V per mm.
Lower trace: corresponding "S"-curves 0.1ms per mm.

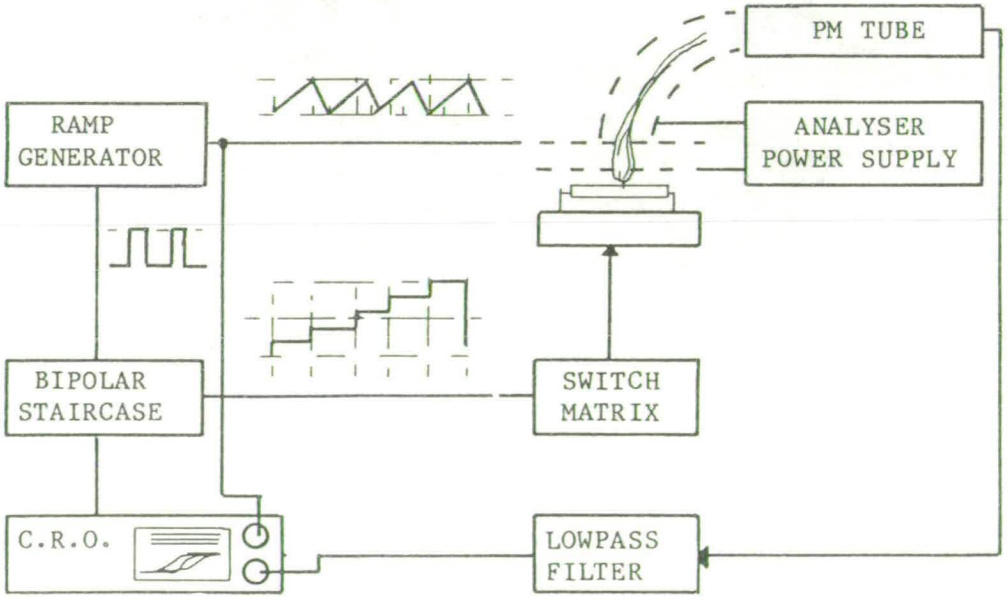


Figure 4.1.2. Instrumentation required to generate characteristic "S"-curves.

caused by local field effects, a series of linearly displaced "S"-curves, covering the dynamic range of the measurement system was generated for each measurement point. A typical display is shown in Figure 4.1.1, in which the upper trace is the applied specimen voltage, going vertically negative to positive and maps onto the horizontal axis of the lower trace which displays the displaced "S"-curves, going left to right.

A schematic of the instrumentation set-up utilised to obtain the analyser characteristics is presented in Figure 4.1.2. The waveform generator provided a $40V_{pp}$ linear ramp waveform, which was applied to the retardation grid of the analyser and a flyback pulse, which was used to trigger and blank a dual channel oscilloscope. In addition, the flyback pulse was used to synchronise the applied specimen voltage, which was a bipolar staircase waveform, in such a way that the period of the retarding waveform equalled the period of one amplitude increment. Both the number of steps and the step magnitude could be altered manually. The output from the videohead amplifier was lowpass filtered and applied to the dual channel oscilloscope, while the applied surface voltage was connected to the second channel. The time base was adjusted until all the "S"-curves appeared superimposed as in Figure 4.1.1.

The retarding field, highpass energy analyser as described by Ranasinghe and Khursheed(1983) was utilised throughout the series of experiments in conjunction with a Cambridge Instruments' S100 SEM with a stationary primary beam and an accelerating voltage of 1kV. The SEM's working distance was 27mm.

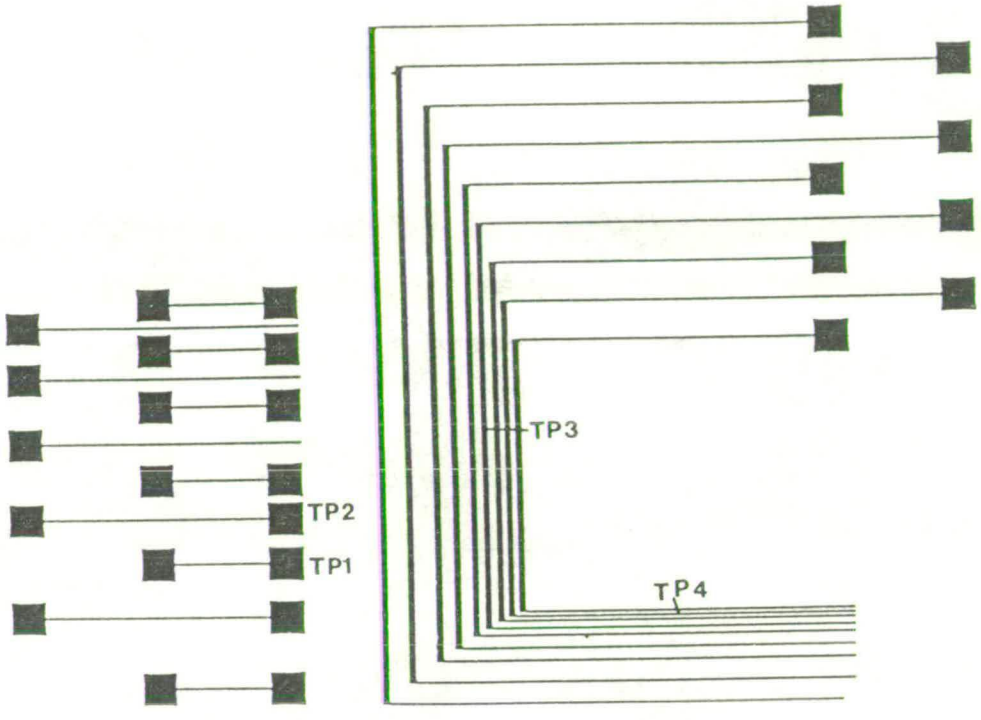


Figure 4.1.3. Floor Plan of Test Structure.

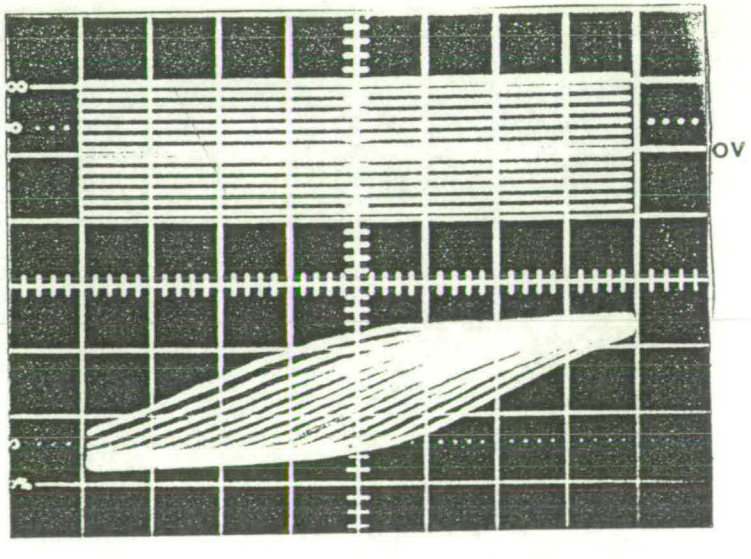


Figure 4.1.4. Gold test pad. Extraction voltage +100V.
Upper trace: Surface voltage applied to gold test pad.
1V per mm.
Lower trace: "S"-curve series generated.
50 μms per mm.

4.1.2) Test Structure.

A specially fabricated test structure, designed by Khursheed (1983) and shown in Figure 4.1.3. was utilised to allow the investigation of field effects along parallel aluminium electrodes with varying track and inter-track dimensions. The inter-track spacing ranges from 3μm to 50μm, while the track width varies from 3μm to 25μm. In addition, there are two arrays of 100μm by 100μm square test pads with interpad spacings varying from 10μm to 100μm. In the upper array of pads a 10μm wide electrode bisects the intergap space. The test structure has silicon dioxide deposited beneath the aluminium tracks to reconstruct typical device conditions. Each test integrated circuit contains two identical test structures; one passivated with silicon dioxide, the other unpassivated.

4.1.3) Test Pads.

Initially, a gold bond pad 0.4mm by 1mm was probed to ensure that linearly displaced "S"-curves could be obtained with the analyser without any apparent local field effects in order that, any subsequent distortion in the "S"-curves could be attributed to local fields and not the analyser. The series of "S"-curves obtained is displayed in Figure 4.1.4 for a surface voltage range of -10V to +10V and demonstrates that no apparent distortion is present.

An aluminium test pad 100μm by 100μm (TP1) was probed and a series of "S"-curves produced. A positive potential of 30V applied to the adjacent pad produced a 50% reduction in the amplitude of the "S"-curve corresponding to a surface voltage of +5V. However, with an applied potential of +20V no observable distortion was present. The inter-pad spacing was 25μm. Both these experiments were carried out

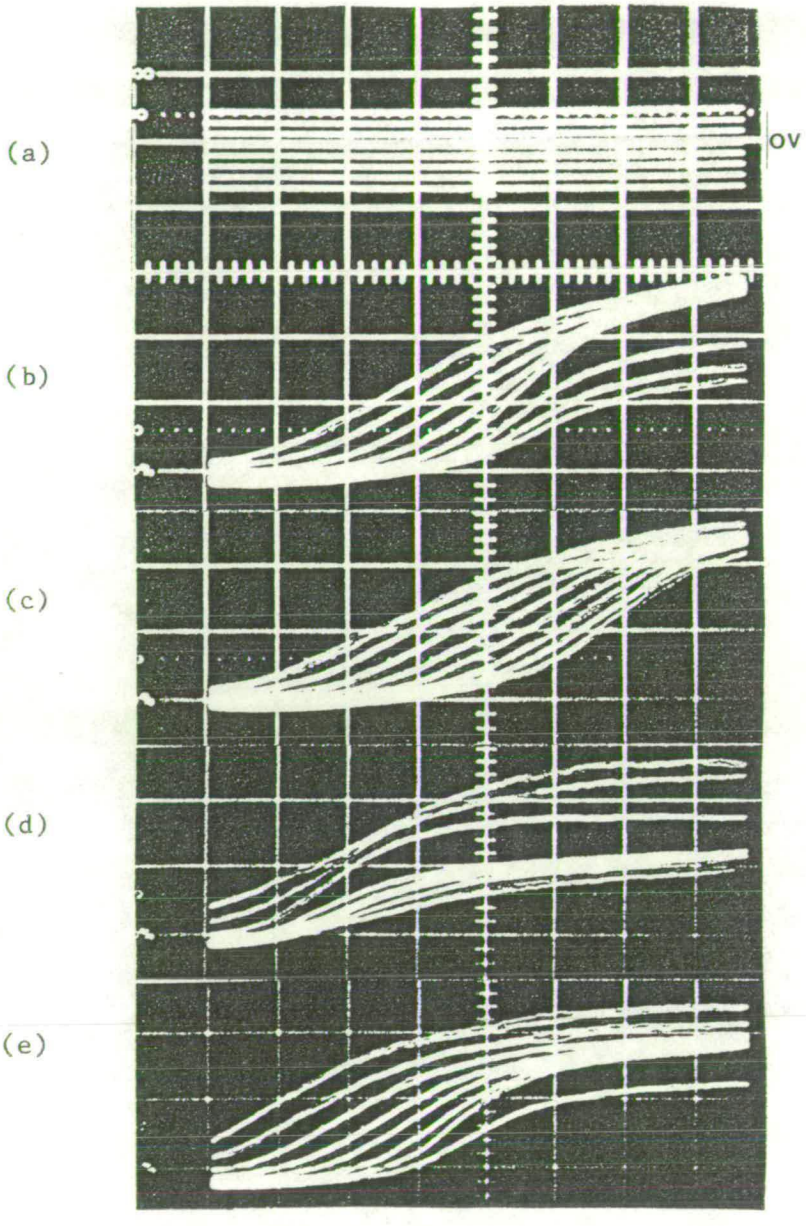


Figure 4.1.5. Pad edge probe. Extraction voltage: +100V
(a) Applied surface voltage 1V per mm
(b) Adjacent pad voltage 0V
(c) Adjacent pad voltage +5V
(d) Adjacent pad voltage -6V
(e) Adjacent pad voltage 0V, middle of pad.

with the primary beam positioned in the centre of the test pad. This demonstrates that normal test pads inserted into the signal path of a device could be used as electron beam test points without loss of accuracy owing to local field effects.

To highlight the effect of local fields on "S"-curves, an aluminium pad (TP2) was probed close to the edge bordering an adjacent pad with an interpad spacing of 10μm. A DC voltage was applied to the adjacent pad and a bipolar staircase waveform to the test pad after which the resultant "S"-curves recorded. This procedure was repeated for several DC voltages in the range -6 to +5V.

The series of "S"-curves shown in Figure 4.1.5(b) was obtained with 0V applied to the adjacent pad and demonstrates that "S"-curves corresponding to positive surface voltages are affected predominantly, reducing in amplitude and giving rise to an apparent shift from their original positions (cf. Figure 4.1.4). The "S"-curves corresponding to negative surface voltages also reduce in amplitude, but the 0V curve remains unaffected. These results are as expected, since the magnitude of the surface fields is dependent on the potential difference between the two pads. Hence distortions are present for positive and negative potential differences, but absent when the pads are maintained at the same potential. As shown in Figures 4.1.5(b) to (d), the degree of distortion increases with the magnitude of the potential difference of the test pads and is greater for positive potential differences.

Figure 4.1.5(c) demonstrates the effect produced when +5V is applied to the adjacent pad. The "S"-curves corresponding to surface voltages around +5V are seen to increase in amplitude and become more

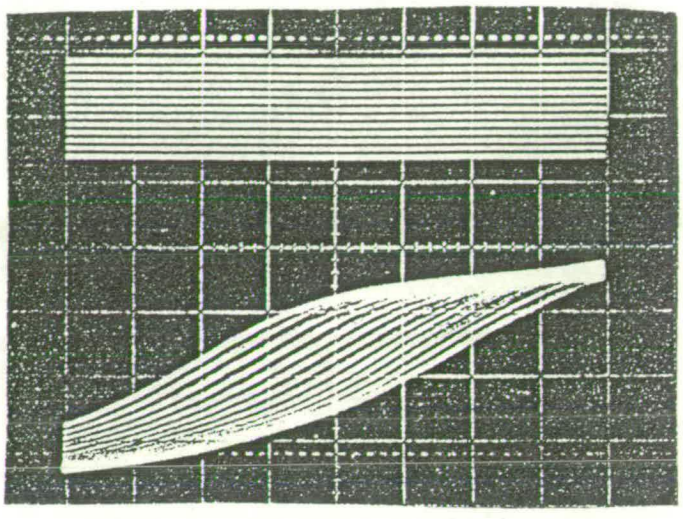
regular in shape, while the remaining "S"-curves reduce in amplitude. This is as expected since the relative voltage between the two pads for a surface voltage around +5V is very small and hence the effect of the surface fields is greatly reduced for these conditions. Figure 4.1.5(d) shows "S"-curves obtained for -6V applied to the adjacent pad and again, as expected the "S"-curves for surface voltages around -6V become more regular, while the other "S"-curves greatly decrease in amplitude.

Figure 4.1.5(e) demonstrates that as the electron beam is moved away from the pad edge towards the centre the effect of the local fields on the "S"-curve decreases. Although the effect of surface fields can be reduced by increasing the extraction voltage by a factor of four, it could not be eliminated.

4.1.4. Parallel Tracks: 5μm Wide, 3μm Spacing.

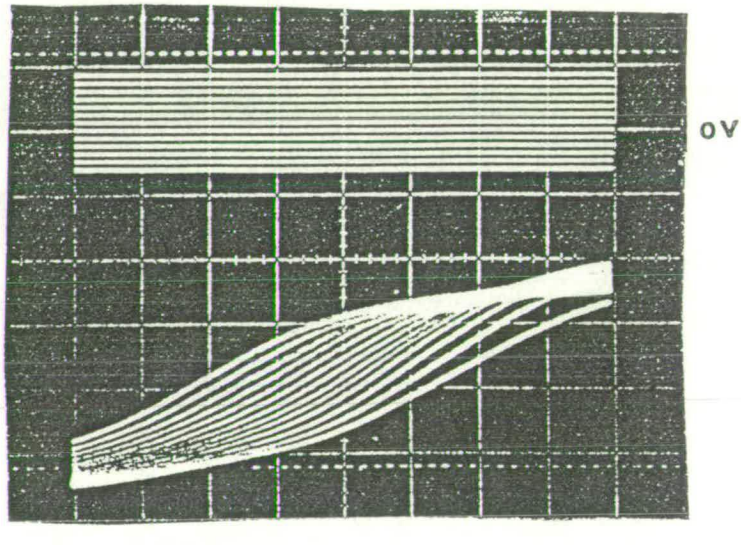
Initially 10μm wide electrodes with intertrack spacing of 25μm were probed at test point three (TP3). However, the "S"-curves obtained were grossly distorted and could not be the basis of a measurement system. This was thought to be caused by the inter-electrode oxide charging and producing complex surface fields.

Shown in Figures 4.1.6(a) and 4.1.6(b) are the results of an experiment in which three parallel conductors, 5μm wide and with an inter-track spacing of 3μm were utilised (TP4). A bipolar staircase waveform was applied to all three tracks and the central track probed, as shown in Figure 4.1.6(a). A linear set of undistorted "S"-curves was generated. When the outer two tracks were grounded the distortion apparent in the second display (Figure 4.1.6(b)) was observed. At positive surface voltages of about +10V and -10V a



(a)

Surface voltage applied to all electrodes.
Upper trace: Applied surface voltage.
Lower trace: "S"-curve group.
Extraction voltage: +200V.



(b)

Outer electrodes earthed.
Upper trace: Applied surface voltage 1V per mm.
Lower trace: "S"-curve group.
Extraction voltage: +200V.

Figure 4.1.6. Parallel electrodes.

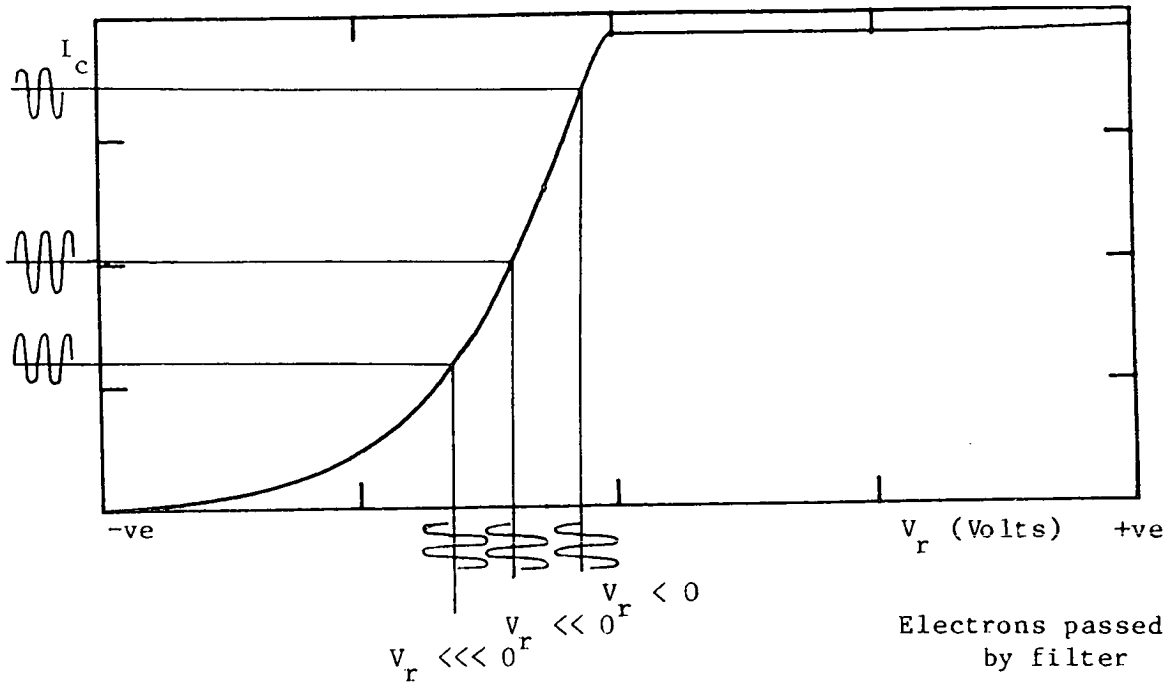
nonlinear translation of the "S"-curve occurred. In addition, the amplitude of the SE signal was reduced owing to the presence of large transverse surface fields which became greater as the potential difference between the central track and its neighbouring track was increased. However, the surface voltage/energy displacement relationship remains linear in the region of interest. i.e. from 0V to +5V and thus it is possible under these conditions to carry out quantitative measurements.

4.1.5. Conclusions.

It has been shown empirically that the effect of surface fields is most dominant when the probed conductor is positive with respect to close adjacent conductors and that when probing MOS structures the electron beam should be positioned in the middle of the widest track, in the area of least oxide. Also it has been demonstrated that, with this analyser, 5μm conductor tracks can be probed, but only under special conditions.

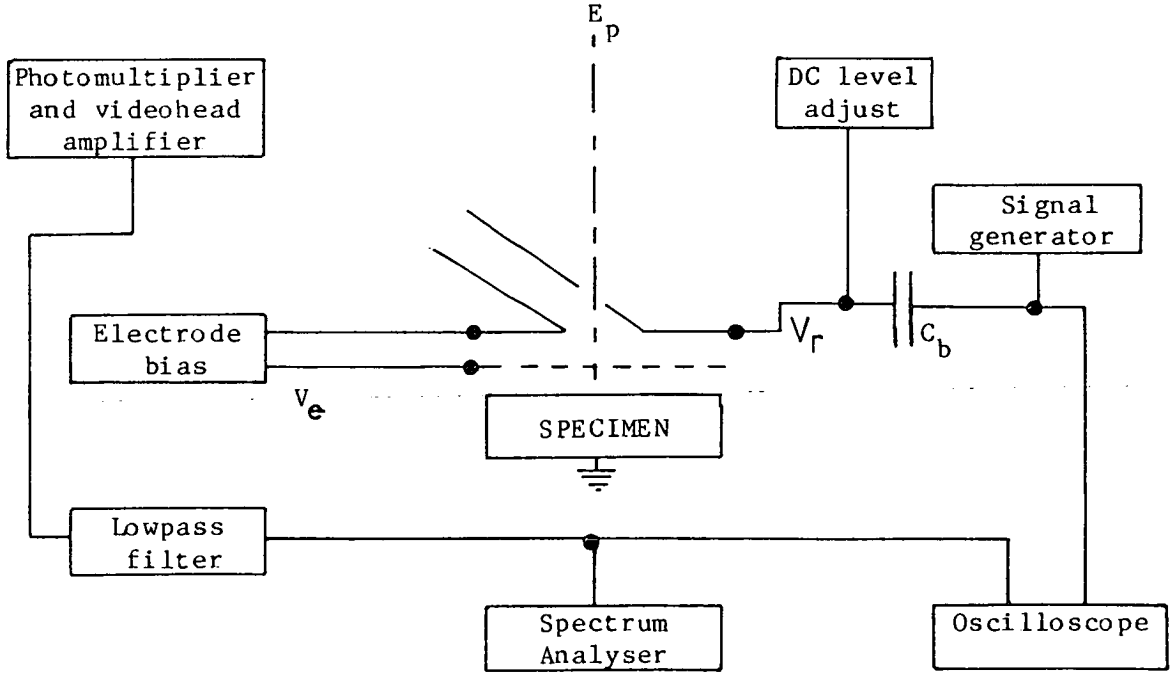
4.2) Instrumentation.

Although a peak tracking servo-mechanism was desirable, the SED could not be used directly as an operating characteristic, since it was not monotonic and was a many valued function. The point of inflection on the "S"-curve could be used as the operating point after Gopinath(1978) and would accrue the benefits of optimising the "S"-curve transconductance and avoiding the noise problem associated with differentiation. However, this type of system would remain in practice a constant current type instrument and would be prone to collected current variation errors.



(b)

Collected current as a function of the cumulative secondary electron distribution for three filter cut-off energies determined by the DC level of the retardation potential.



(a)

Instrumentation required for preliminary experiments.

Figure 4.2.1. Preliminary experiments.

Using the dispersion type analyser described by Dinnis(1982), a series of preliminary experiments were executed utilising a Cambridge Instruments' Stereoscan MkII SEM with an accelerating voltage of 1.5kV and a stationary beam. An earthed copper stub was used as a specimen.

4.2.1) Preliminary Experiments.

With the experimental set-up shown in Figure 4.2.1(a), a waveform was superimposed on a DC bias potential and applied to the retardation electrode of the energy analyser. The output of the videohead amplifier was lowpass filtered and applied to both an oscilloscope and spectrum analyser, enabling the frequency power spectrum and the time dependent output voltage to be observed simultaneously. The input waveform was also applied to the oscilloscope.

Using a sine waveform, three DC bias levels were chosen, corresponding to the cut-off energies indicated in Figure 4.2.1(b). Figures 4.2.2. to 4.2.4. illustrate the dispays obtained when a $1.2V_{pp}$, 0.5kHz was applied to the retardation electrode. In the (a) Figures the upper trace is the output waveform while the lower trace is the input waveform. In Figure 4.2.3 no distortion is apparent either in the power spectra or the output time waveform. This was achieved by reducing the amplitude of the modulating waveform to 1V. Under these conditions, the collected current is a linear function of the retardation voltage. However, if the DC bias voltage is altered above or below this value, second harmonic distortion is introduced as shown by Figures 4.2.2. and 4.2.4.

It can also be deduced from the traces 4.2.2 and 4.2.4(a) that the second harmonic distortion is inverted at either side of the

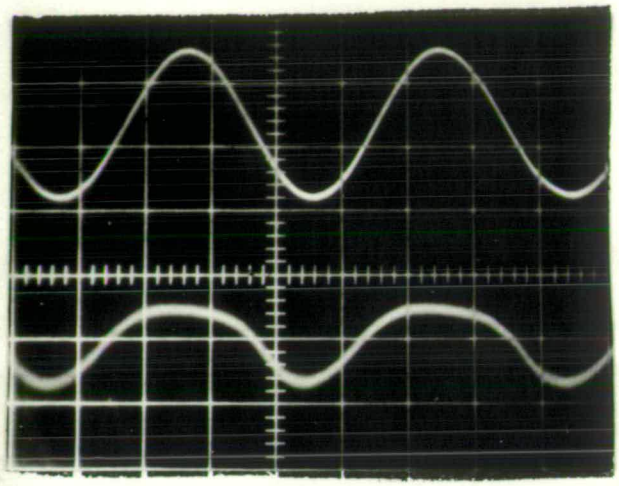
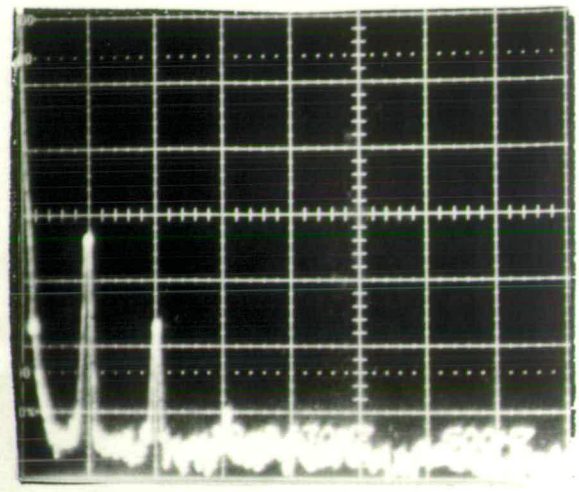


Figure 4.2.2(a) V_r <<< 0



(b) Power spectrum

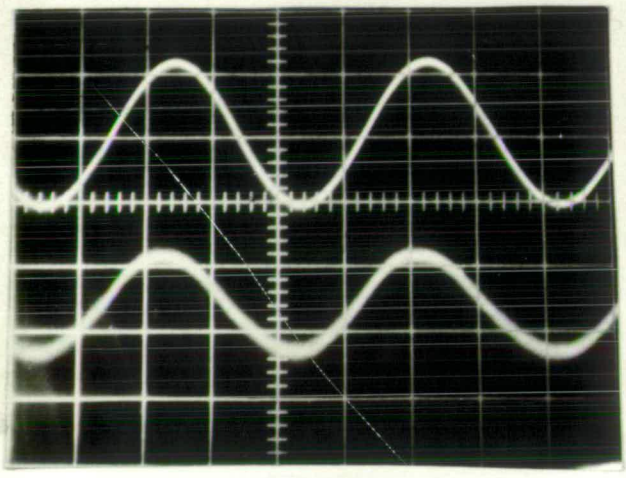
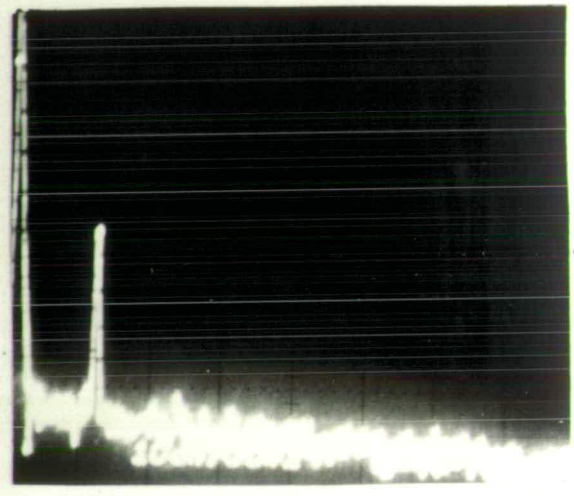


Figure 4.2.3(a) V_r << 0



(b) Power spectrum

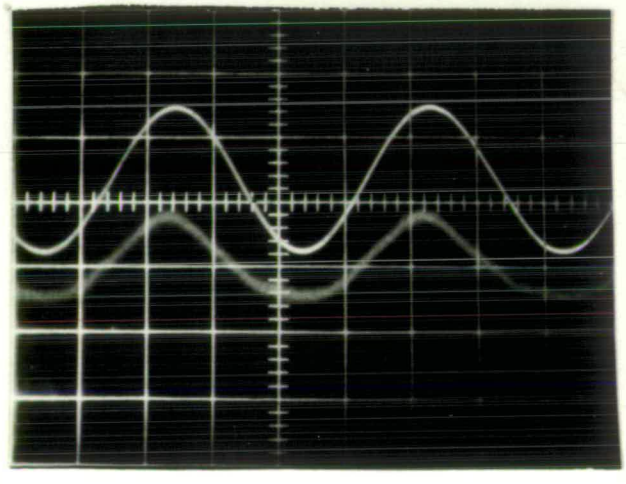
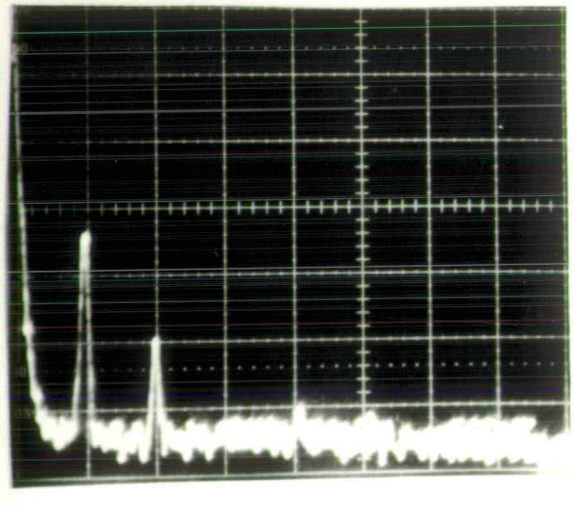


Figure 4.2.4(c) V_r < 0



(b) Power spectrum

Upper traces: input waveform 50mV per mm. Lower traces: output waveform 50mV per mm. Both 50μs per mm.

Power spectra 50Hz per mm, 30Hz resolution, 1 dB per mm.

Input and output waveforms and corresponding power spectra for three retardation voltages.

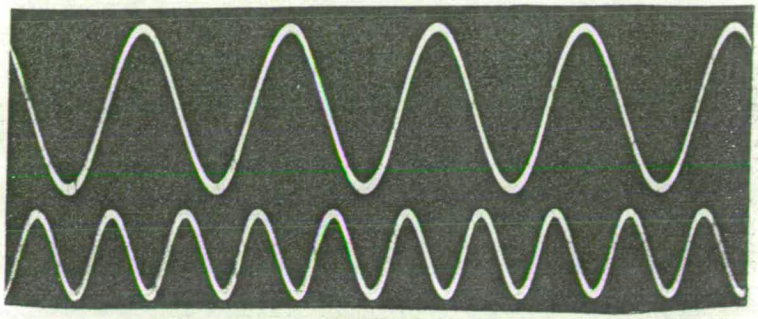
linear DC bias, since the "flattening" of the sinewave occurs at the peak and then the trough. Also the power of the second harmonic distortion is approximately similar in both cases.

To investigate these results further, a double channel nonlinear detection system was constructed as shown in Figure 4.2.5. Each channel consisted of two cascade multiple feedback high Q (20) active filters, followed by a full wave operational amplifier precision rectifier and lowpass filter. The output voltage of each channel was proportional to the the amplitude of the respective harmonic divided by π . The same experimental set-up was used and the resonant frequency of the bandpass filters tuned to the fundamental and second harmonic.

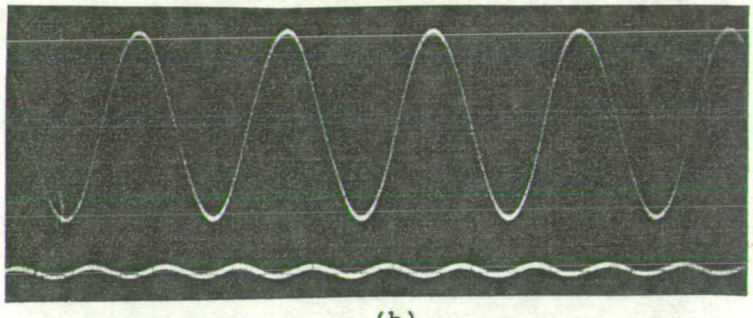
By adjusting the DC bias potential, Figures 4.2.6(a)-(c) were obtained corresponding to Figures 4.2.2(a)-(c) to 4.2.4(a)-(c). The traces were obtained by probing the detector prior to the detection stages. From the traces of the fundamental it can be seen that a slight increase and then decrease in amplitude is apparent as the bias voltage is varied through the linear bias potential. The second harmonic begins approximately 270^{o} out of phase with the fundamental and ends 90^{o} out of phase while its amplitude simultaneously goes through a minimum of zero.

4.2.2) Discussion of Results.

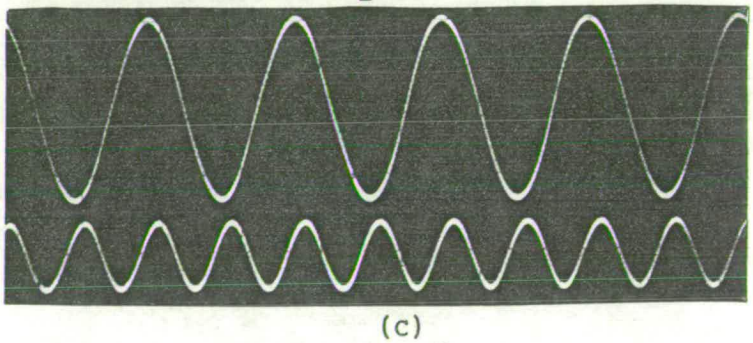
Extrapolating from the previous section's results, a servomechanism could be constructed to measure the displacement of the SED by applying a modulating sinewave to the retardation electrode and detecting the amount of second harmonic distortion. Thereafter a control voltage could be applied to the bias voltage applied to the



(a)
$V_r \lll 0$



(b)
$V_r \ll 0$



(c)
$V_r < 0$

Figure 4.2.6. The frequency components of the output waveforms obtained from three different retardation voltages with a superimposed sinusoid.

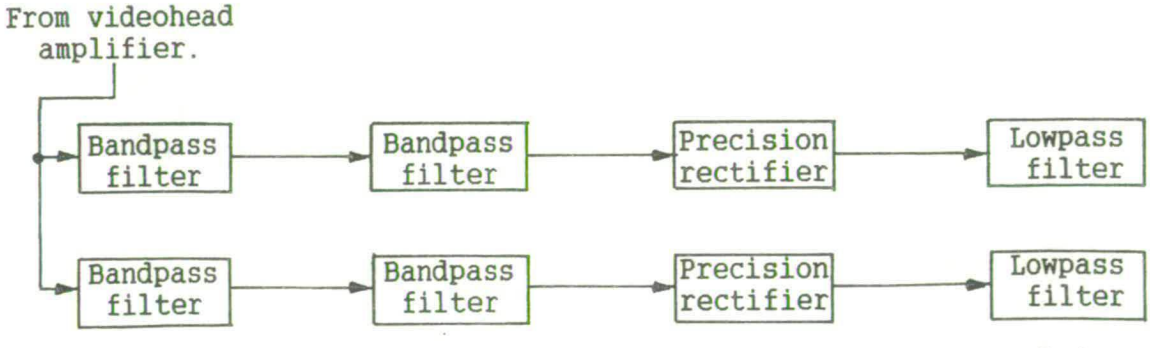


Figure 4.2.5. Double channel nonlinear detection system used to obtain the above traces.

retardation electrode to maintain the amount of second harmonic distortion at zero. Hence the most linear region of the SED could be tracked without the need to detect a constant current.

The results can be explained in terms of small signal theory and confirm, to a certain extent, the theory presented by Taylor(1969) for Auger spectroscopy, within which a Gaussian SED was assumed.

The collected current(I_c) is a function of the retardation voltage(V_r), which consists of a DC bias(V_b) and a modulating sinewave with amplitude V_m.

$$V_r = V_b + V_m \mathrm{Sin}\ \omega t \qquad (4.2.1)$$

The "S"-curve operating characteristic can be approximated by a Taylor expansion provided all the required terms have continuous and not zero derivatives. Therefore, the total collected current is given by

$$I_c = \sum_{n=0}^{n=\infty} \frac{I^n(V_b)(V_m \mathrm{Sin}\ \omega t)^m}{n!} \qquad (4.2.2)$$

Expanding and collecting terms for the fundamental gives

$$F_1 = \left[I'(V_b)V_m + \frac{I'''(V_b)V_m^3}{8} + \frac{I^v(V_b)V_m^5}{192} \right] \mathrm{Sin}\omega t \qquad (4.2.3)$$

and for the second harmonic term

$$F_2 = \left[I''(V_b)V_m^2 - \frac{I^{iv}(V_b)V_m^4}{48} + \frac{I^{vi}V_bV_m^6}{1536} \right] \mathrm{Cos}2\omega t \qquad (4.2.4)$$

If the amplitude of the modulating sinusoid is made small enough to ensure that the higher order terms tend to zero, equations 4.2.3.

and 4.2.4 reduce to

$$F_1 \simeq I_i(V_b)\mathrm{Sin}\omega t \qquad (4.2.5)$$

for the fundamental and similarly for the second harmonic

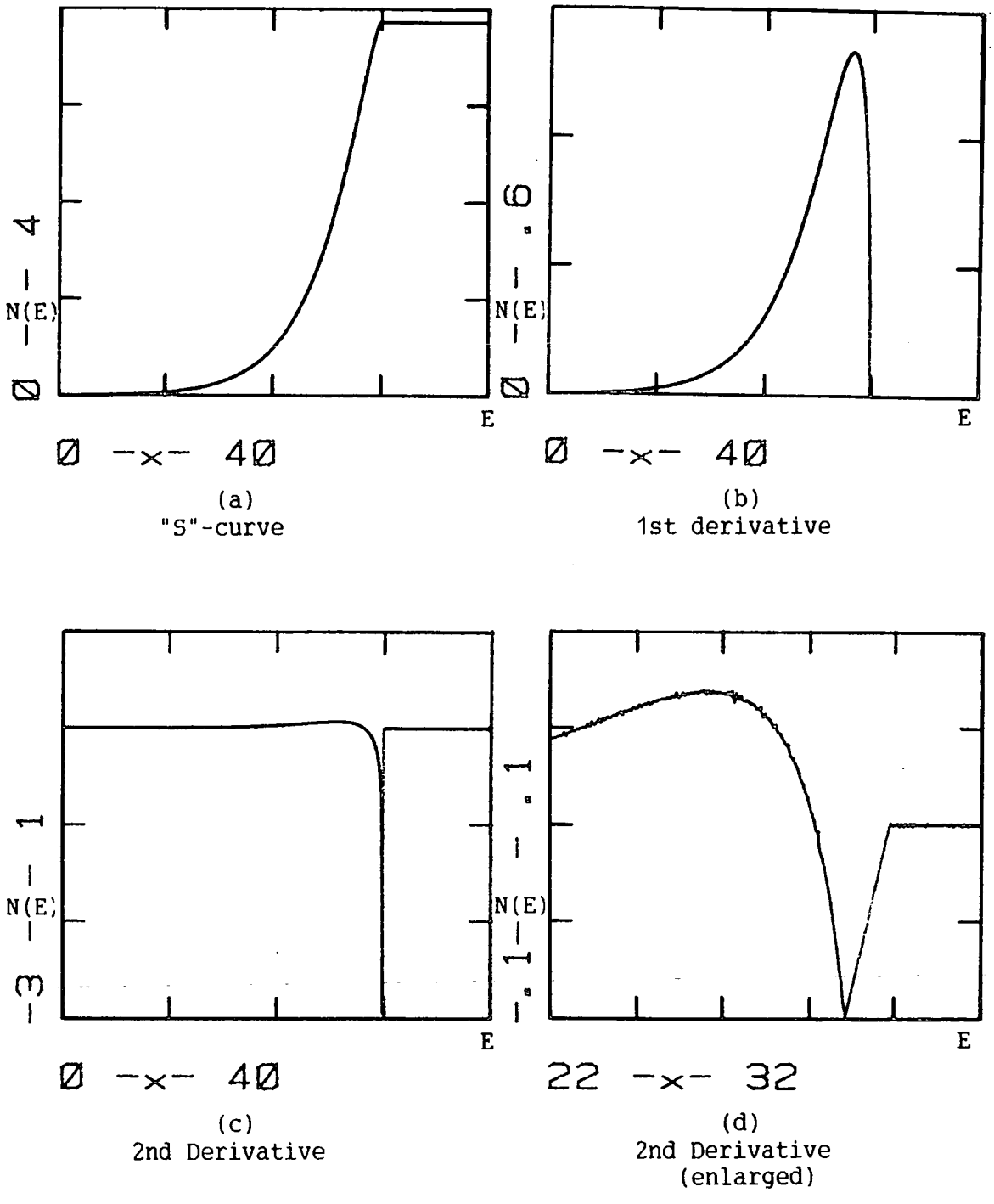
$$F_2 \simeq \frac{I^{ii}(V_b)V^2\mathrm{Cos}2\omega t}{4} \qquad (4.2.6)$$

This demonstrates that if V_m is made sufficiently small, the amplitudes of the fundamental and the second harmonic will be directly proportional to the first and second derivatives, respectively. Therefore, by detecting and demodulating these frequency components the first and second derivatives of the "S"-curve can be obtained without electronic differentiation.

From calculus the peak of the SED is coincident on the energy abscissa with the point of inflection of the "S"-curve and the zero-crossing of the second derivative. Therefore, by tracking the most linear part of the high pass energy filter characteristic, the peak of the SED can be tracked.

4.2.3) Operating Characteristic.

If a servo-mechanism was to be based on the suppression of second harmonic distortion, a deeper understanding of the proposed operating characteristic was necessary. To this end a suite of numerical analysis programs was written to implement convolution, integration, differentiation and filtering and thereby the shape of the second derivative of the "S"-curve was determined. The SED was approximated, after Everhart and Lin(1978), with a Rayleigh distribution of the form:



(a)
"S"-curve

(b)
1st derivative

(c)
2nd Derivative

(d)
2nd Derivative
(enlarged)

Figure 4.2.7. The "S"-curve and its derivatives.
V_r positive left to right for all curves.

$$N(E) = \sqrt{E}\, e^{-0.286E} \qquad (4.2.5)$$

The resultant curves, displayed in Figure 4.2.7, were generated with the assumption that the analyser cut-off energy was ramped from a negative to positive value. The SED was convolved with the highpass energy filter characteristic to give the "S"-curve shown in Figure 4.2.7(a).

Differentiating the "S"-curve resulted in the function shown in Figure 4.2.7(b), which can be seen to be the SED after it has been reflected in the vertical axis and shifted by an amount equal to the maximum filter cut-off energy. For a scanned system, the slewing direction of the retardation waveform should be consistent i.e. always negative to positive, since the polarity of the derivatives as a function of time is reversed in accordance with the slewing direction.

The second derivative, shown in Figure 4.2.7.(c) demonstrates that three different electron energies concur with the operating point criterion; that the first derivative should equal zero. Hence to avoid the servo-system locking onto either of the two false operating points, the feedback loop must only be closed when the loop error term corresponds to an electron energy value which lies between the points of inflection of the SED.

The nonlinearity and asymmetry of the second derivative about the zero-crossing is illustrated by the enlarged zero-crossing region of Figure 4.2.7(d). The asymmetrical transconductance could lead to instability and therefore, the loop error term should be limited to an approximately linear region about the zero-crossing.

The transconductance of the linear region of the operating characteristic is an order of magnitude less than that of the "S"-curve, as demonstrated by Figures 4.2.7(a) and (c). This in turn will lead to a degradation of the SNR of the output signal of the video-head amplifier. By maximising the amplitude of the modulating waveform, the SNR of the detected signal can also be improved.

Equation 4.2.4. demonstrates that by increasing the modulation amplitude, the accuracy with which the second derivative can be reproduced is compromised. Hanisch et al.(1975) reported that if the first derivative of a Gaussian distribution was to be obtained with 1% accuracy, the modulation amplitude should be a maximum of 5% of the half-width between the points of inflection of the distribution. However the SED is asymmetrical and therefore, the modulation amplitude is determined by the displacement between the peak of the SED and the nearest point of inflection.

Since the conditions for the operating point were that it was both monotonic and four quadrant, the accuracy with which the second derivative is reproduced is of secondary importance and therefore, the modulation amplitude can be increased as long as these conditions are fulfilled. Hanisch et al.(1975) further showed that increasing the modulation amplitude instigated an increase in the half-width of the derivative. When applied to the SED, this results in an enlarged linear region with a reduced transconductance. In addition, equation 4.2.4. demonstrates that no displacement in the operating point is precipitated by increasing the modulation amplitude, since all even ordered derivatives have a zero-crossing co-incident with the point of inflection of the "S"-curve.

In the light of the arguments presented the modulation amplitude can be increased up to twice the half-width between the points of inflection of the SED. This corresponds to 10% of the full-width of the "S"-curve or ≃ 1-2V, in this case.

Summarising:

a) The feedback loop should only be closed when the error term corresponds to the point of inflection of the "S"-curve ±5% of the full-width of the "S"-curve.
b) The amplitude of modulation should be less than or equal to 10% of the full-width of the "S"-curve.
c) When implemented, scanning should be consistently in one direction. i.e. negative to positive.

4.3) Utility Subsystem.

The purpose of the Utility subsystem was to provide all the necessary stimuli to enable the DUT to function properly. A block diagram is presented in Figure 4.3.1. and illustrates the component parts of the utility subsystem which comprises an exerciser or pattern generator, a two phase clock, two power supplies and a specimen chamber interface. All of the constituent utilities were under computer control in order to minimise the manual set-up of a test sequence and to enable the archival storage of test parameters and results for specific i.c.s. This would also allow subsequent tests to be executed automatically.

Since economics dictated that the system be constructed as inexpensively as possible and a general purpose tester was thought superfluous to requirements it was decided that a dedicated exerciser

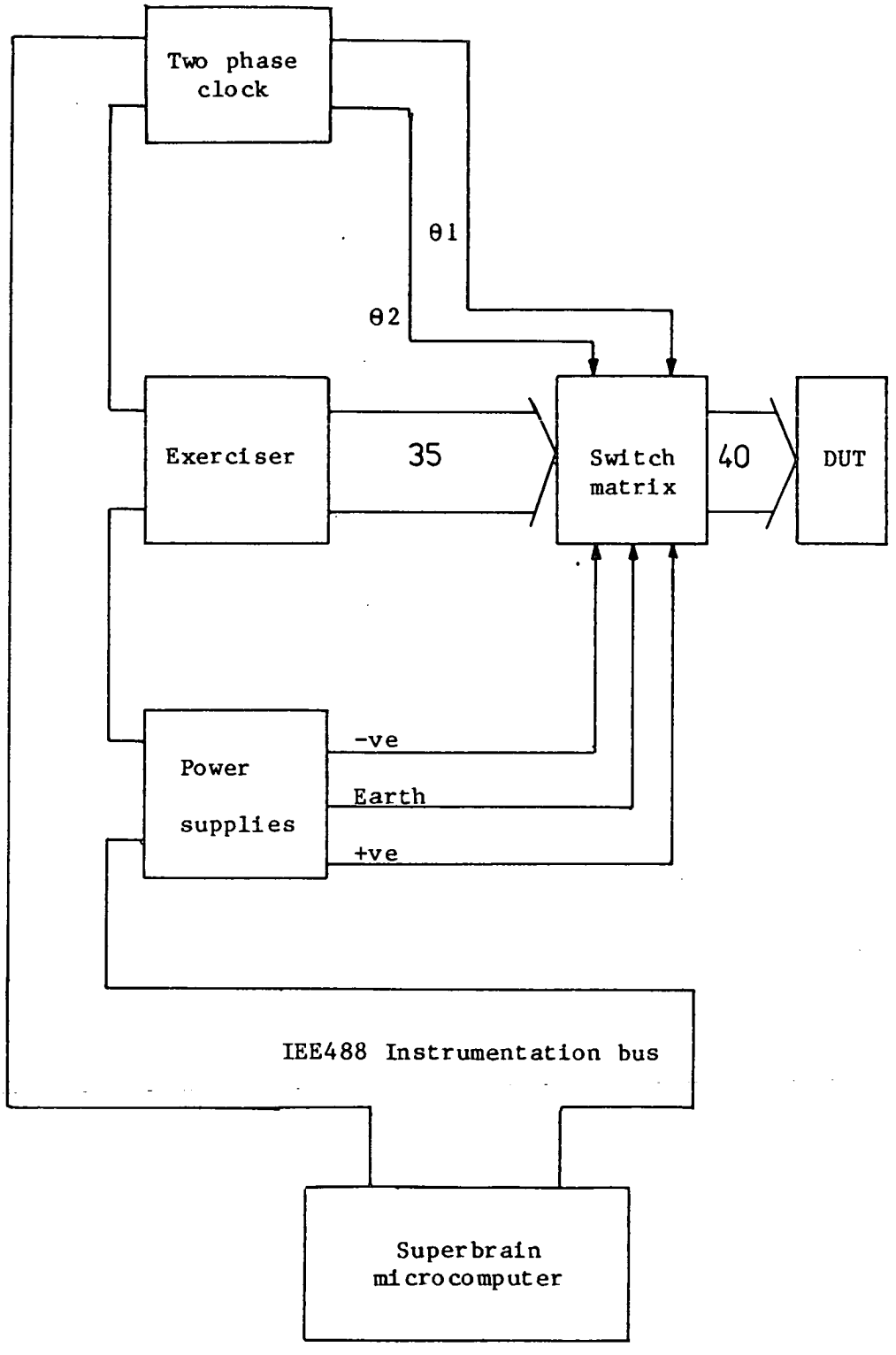


Figure 4.3.1. Block diagram of Utility Subsystem.

which would perform functional or Logic Integrity Testing only was more germane. The design philosophy was that the test vectors would not be generated by the exerciser itself, but either by independent software or possibly from the circuit simulation programs used in the design process. Subsequently, the vectors would be down loaded from the Superbrain into the exerciser memory and from there applied to the DUT. Owing to the restricted memory size, devices requiring a large amount of test vectors would have to be tested using overlapping burst test techniques; that is after the vectors contained in the exerciser memory were exhausted, the device clock would be halted and the Superbrain would perform a memory swap with the exerciser, after which the device clock would be restarted and the test sequence continued. Since the exerciser was a purely functional tester, it complemented the Electron Beam Probe, the main function of which was internal parametric testing. Constraints placed on the design of the exerciser were that the output stages had to be compatible with +5V NMOS logic levels and if possible, they should be bidirectional to facilitate the testing of bus architectured devices. In addition, the exerciser had to be synchronised with the clock generator.

The clocks were designed with non-overlapping level transitions and buffered outputs to drive clock loaded devices. Their frequency could be programmed from less than 100Hz to over 20kHz. The two power supplies were programmable from 0 to 10V magnitude in 100mV increments and were used to provide the 5V±10% tolerance power supply demanded by +5V logic and the -2.5V substrate bias necessary for NMOS devices. Since the power supplies were programmable, this would allow power supply sensitivity analysis for different device parameters to be investigated.

The specimen chamber interface performed the task of assigning specific output lines, such as clocks, to specific device pins. In addition, it had to be capable of shorting all unused lines to earth. Since a switching matrix consisting of 1000 bidirectional reed relays would be necessary, under computer control, it was decided to implement the programming of lines onto pins with a patch board. Forty single pole double throw toggle switches were used to enable all the specimen chamber input lines to be grounded. Lines were assigned to pins using banana connectors. The lines were fed into the specimen chamber through a vacuum interforce connect and were connected to a zero-insertion connector, mounted on top of the specimen stage. All wiring inside the specimen chamber used polytetrafluoroethylene covered cable for low out-gassing.

4.4) Data-Acquisition Subsystem.

The primary function of the data-acquisition subsystem was to digitise the output of the voltage measurement system to enable further processing and archival storage of the measured waveform. However, the secondary task of monitoring device parameters such as power supply voltage and current ripple and the DC current and voltage loading of the device was identified. In addition if the first and second derivatives of the operating characteristic were to be generated, the data-acquisition system would provide a means of collecting this data.

Hence it was decided that a multichannel system was necessary with an excess number of channels to enable future expansion and with the capability to quantise both AC and DC waveforms. A block diagram of the system is presented in Figure 4.4.1. and illustrates the com-

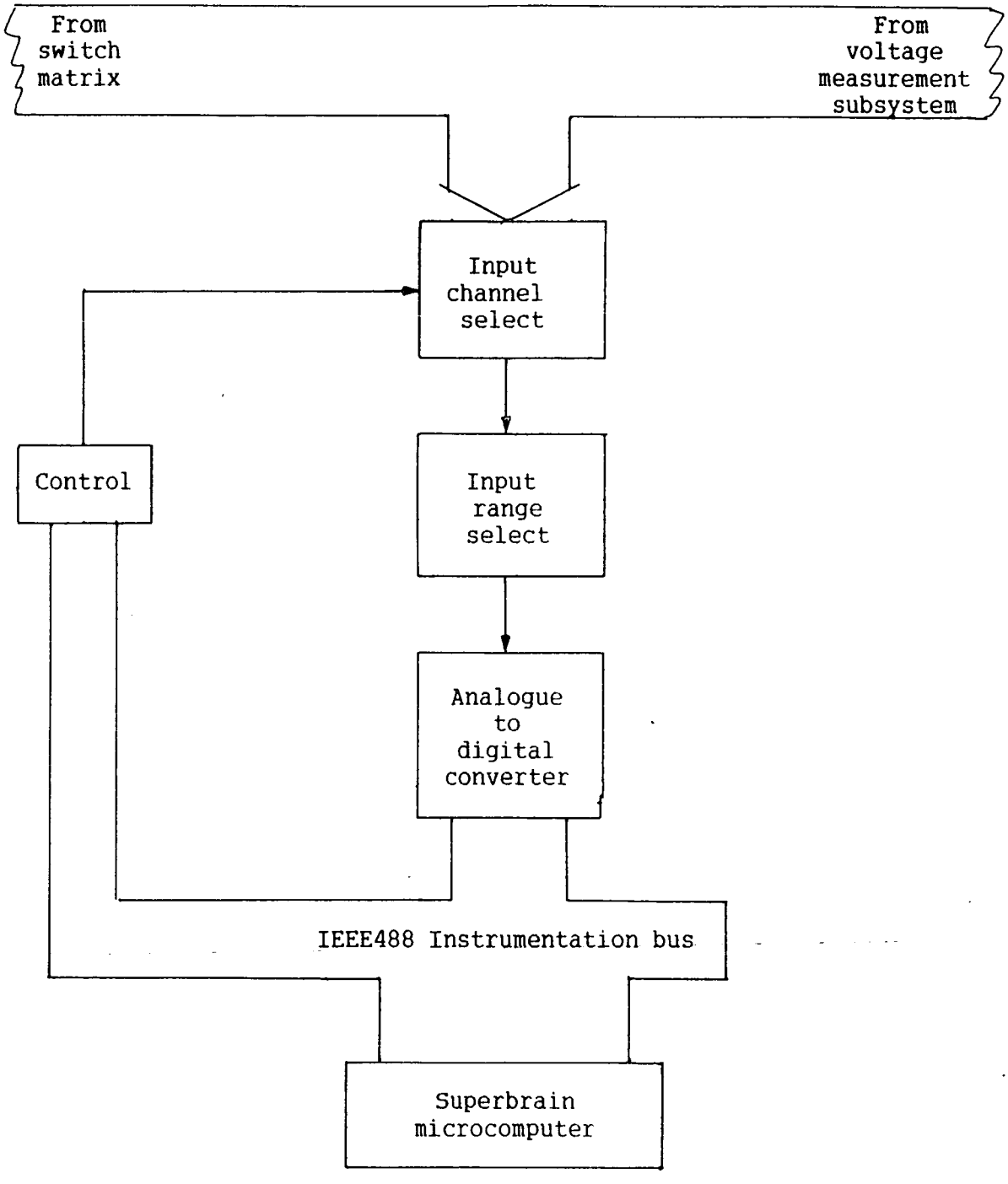


Figure 4.4.1. Block diagram of Data-Acquisition Subsystem.

ponent parts. To maximise the accuracy of the quantisation process a variable input voltage range was incorporated as well as polarity range selection. Since all the waveforms to be quantised were expected to be less than 5kHz, a one pole lowpass filter in conjunction with a high sampling rate was used to limit any aliasing effects.

Since it was envisaged that during the execution of a test sequence, the Superbrain would only be active if an exerciser memory swap was necessary, it was decided that the intermittent period could be utilised to collect data via the data-acquisition system. Therefore, the analogue-to-digital converter(ADC) had to be compatible with the IEEE488 instrumentation bus and was constrained to be eight bits wide. However, with a +5V maximum voltage, this gave an accuracy of 10mV.

CHAPTER 5.

AN AC VOLTAGE CONTRAST MEASUREMENT SYSTEM.

This chapter describes the voltage measurement system designed and constructed, in accordance with the principles outlined in the previous chapter. In addition, the performance of the system and the design aids utilised to realise it, will be discussed.

The servo-mechanism can be divided into two parts: a) the feedback system and b) the modulation system. The feedback system is a regulator, maintaining the amount of second harmonic distortion equal to zero and as a consequence, the retardation electrode voltage(V_r) the same as the surface voltage(V_s).

$$V_s = V_r \tag{5.1.1}$$

Hence the output signal of the modulation system(V_{mod}) must be directly proportional to the amount of second harmonic distortion as a function of both the the surface voltage(V_s) and the retardation voltage(V_r) about the operating point. Therefore, the equation

$$V_{mod} = K(V_s - V_r) \tag{5.1.2}$$

must be true for small deviations about the system operating point, where K is the gradient of the second derivative at the operating point.

5.1) Feedback System.

Since, as previously stated, the servo-system employs unity gain negative feedback, the forward loop gain should be made as large as possible to maximise the servo accuracy and be combined with some

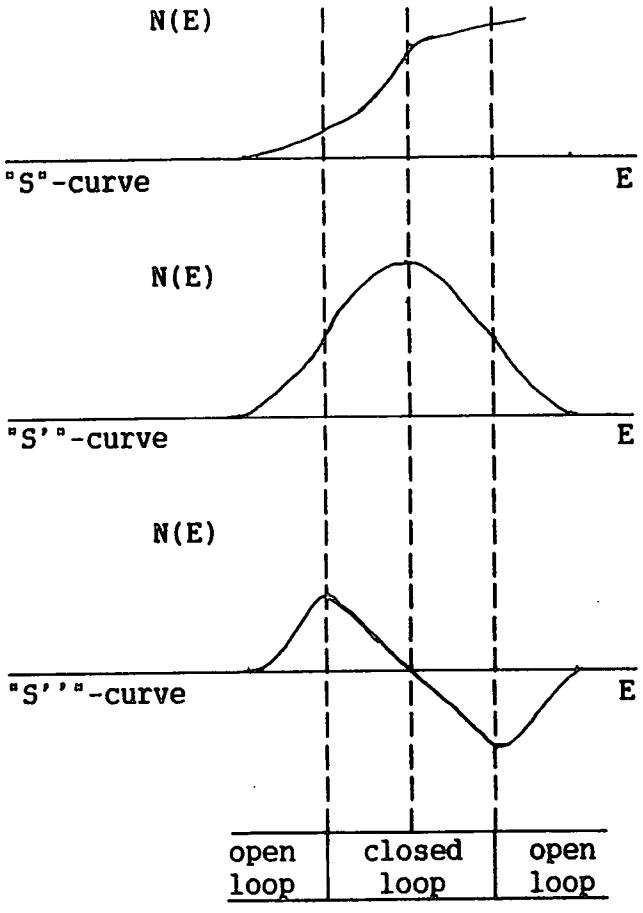


Figure 5.1.1. Voltage measurement system operating characteristics.

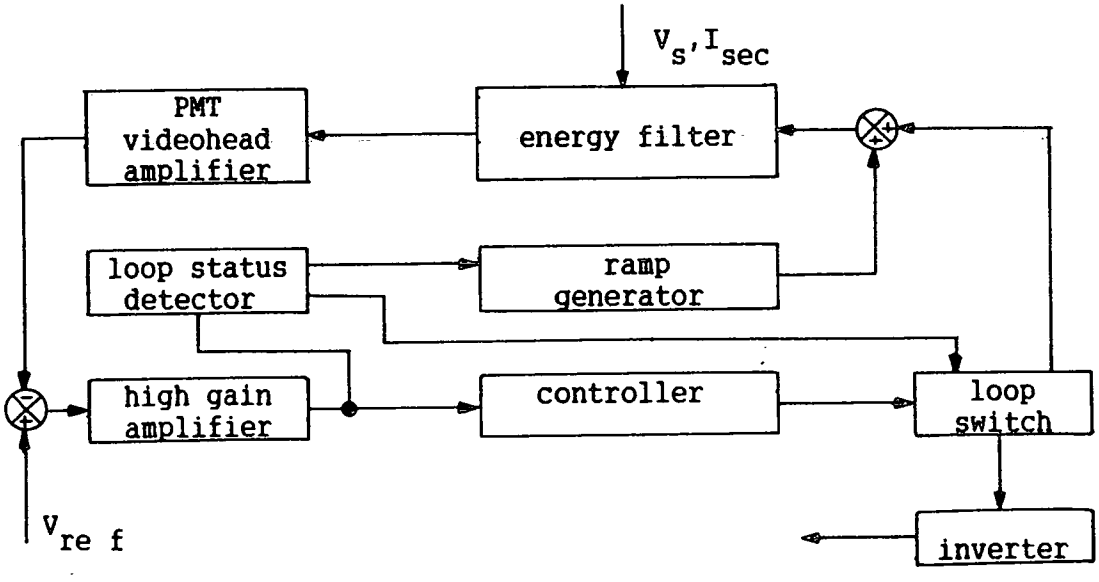


Figure 5.1.2. Block diagram of voltage measurement system.

form of compensation other than dominant pole to maximise the system bandwidth.

As a consequence of the operating characteristic discussed in the previous chapter, the instrumentation system had to overcome two inherent problems: initially determining the position of the operating point on the energy abscissa and preventing the feedback loop locking onto either of the two false operating points. As illustrated in Figure 5.1.1, the operating characteristic can be divided into a region within which the feedback system is inherently locked and two other regions within which lock will not be achieved. Unless the analyser cut-off energy lies within the lock region of the operating characteristic on start-up, lock will not be achieved and the feedback loop will saturate. If lock is accomplished, the surface voltage will be tracked by the retardation voltage subject to one criterion: that the loop error voltage does not exceed the "inherent" lock range. This may occur at high frequencies as the loop gain decreases and the error voltage increases.

These adversities were circumvented by designing the hybrid scanning and feedback system shown in Figure 5.1.2. Initially, the system is operated in an open loop mode and a linear ramp voltage applied to the retardation electrode of the energy filter. When the lock region of the operating characteristic is encountered the ramp voltage is held constant and the loop automatically closed. The retardation voltage is then servoed until it equals the surface voltage. Thereafter the retardation voltage is identical to the surface voltage.

If the loop error increases to such an extent that it impinges

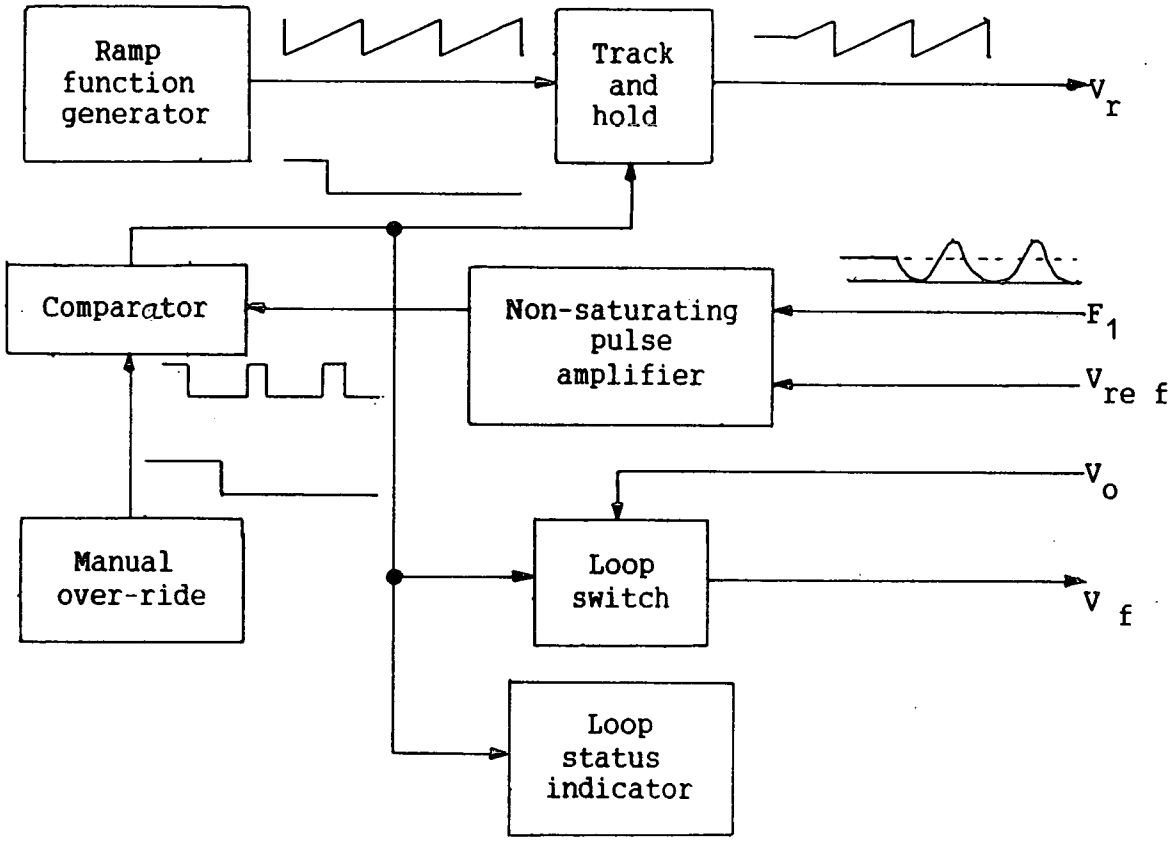


Figure 5.1.3. Block diagram of loop status control.

on the out of lock region of the operating characteristic, the feedback loop is automatically opened and returned to the seek mode.

5.1.1) Operating Characteristic Detection and Capture.

The lock region of the operating characteristic, defined by the points of inflection of the SED, was detected by monitoring the second derivative of the "S"-curve as the retardation voltage was swept. The value of the second derivative was compared to a DC bias level, which was adjusted to coincide with the points of inflection of the SED. The output of this comparison controlled the status of the control loop as shown in Figure 5.1.3. Since the collected current corresponding to a point of inflection of the SED is a function of the total emitted SE current, the DC bias level should be varied with the SE current to prevent false triggering of the control loop. However, since this was not feasible it was assumed that the "S"-curve was normalised by varying the gain of the photomultiplier tube before a measurement was taken.

During the seek mode of operation, the retardation voltage was required to linearly scan over a voltage range corresponding to the maximum voltage deviation of the surface voltage. This was analagous to linearly varying the cut-off energy of the energy filter. The linear scanning function was realised by applying a ramp waveform to the retardation electrode until a point of inflection of the second derivative was detected. Thereafter, the ramp voltage was sampled, using a sample-and-hold i.c. and the loop closed. Since the bandwidth of the sample-and-hold was 1MHz and the hold capacitor was selected to minimise acquisition time, the main source of error was the Hold step DC-feed through from the logic switching. This was reduced by

halving the switching waveform's amplitude. Since the retardation bias was scanned in the one direction constantly, the hold error was constant and not significant, as it did not exceed the operating characteristic's capture requirements.

The loop switch had to maintain low feed-through for large voltage swings when in the open loop mode and pass them undistorted when in the closed loop mode. The problem of switching a high voltage waveform was circumvented by designing a fraction of the loop gain forward of the loop switch and thus reducing the signal levels the switch was required to block. Manual override of the loop-switch was provided by routing the lock detection signal through a selector circuit. A loop mode indicator was also provided.

The detection circuitry consisted of a high gain non-saturating pulse amplifier and a comparator with a switching speed of 10ns. The output of the comparator was TTL level compatible and was used to close the loop switch and trigger the sample-and-hold. This minimised the delay in detecting the lock region of the operating characteristic.

The conditions for capture depended on the rise-time of the feedback system, the slew rate of the ramp voltage, the delay of the detection circuits, and the logic feed-through error of the sample-and-hold. This is because the chosen level of the second derivative has to be detected, inverted and applied to the retardation electrode, before the ramp waveform has slewed out of the lock region. This can be formulated as:

$$\left[\frac{S_r}{T_r + T_d} + E_1\right] \ll F_2 \qquad (5.1.3)$$

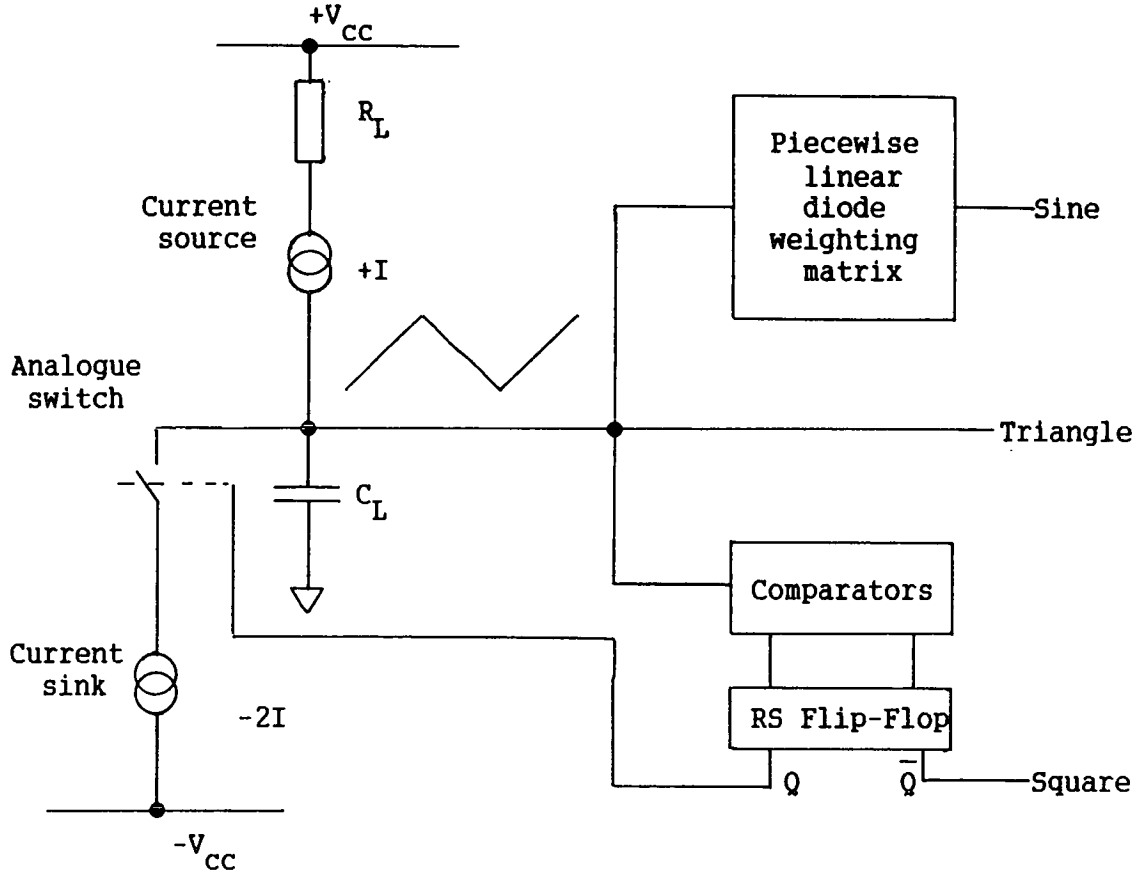


Figure 5.1.5. Schematic of Intersil 8083 function generator i.c.

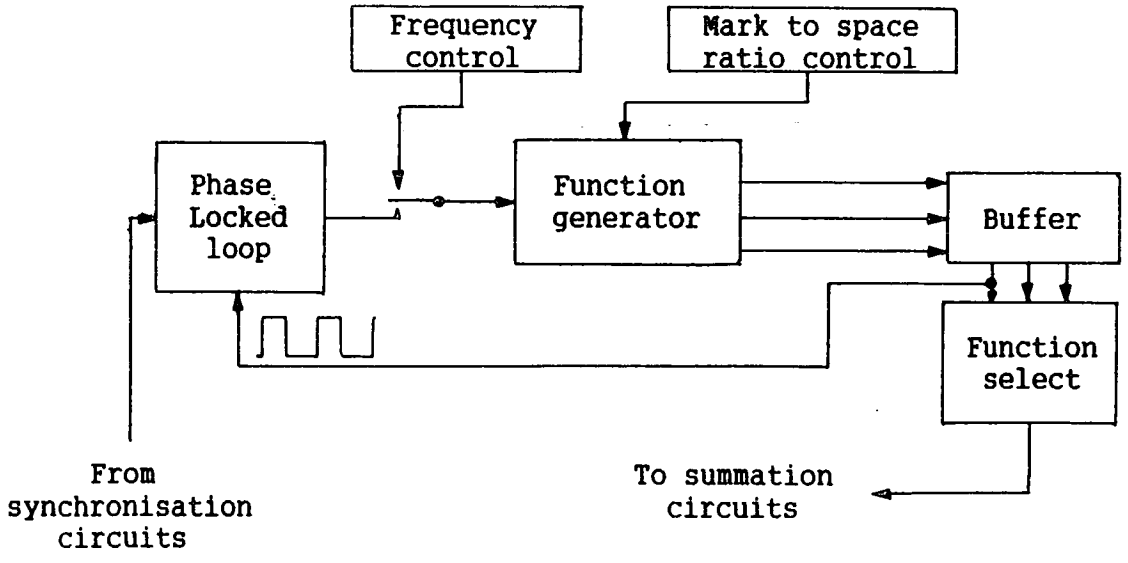


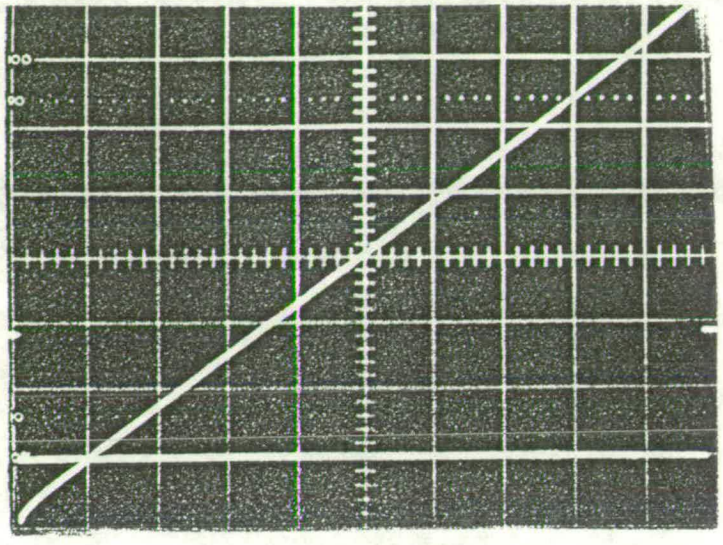
Figure 5.1.4. Block diagram of low frequency function generator.

Where E_1 is the logic feed-through error, T_d the detection delay error, S_r is the slew rate of the ramp,

T_r the system risetime and F_2 the width of the second derivative characteristic in volts. For a baseband system the problem is exacerbated since the slew rate of the baseband signal has to be taken into account, because it moves the operating characteristic with respect to the ramp voltage. For the worst case, in which the surface voltage is decreasing, S_r now becomes the sum of both the surface voltage slew rate and the ramp voltage slew rate. Therefore the system bandwidth has to be made as large as possible for baseband signals.

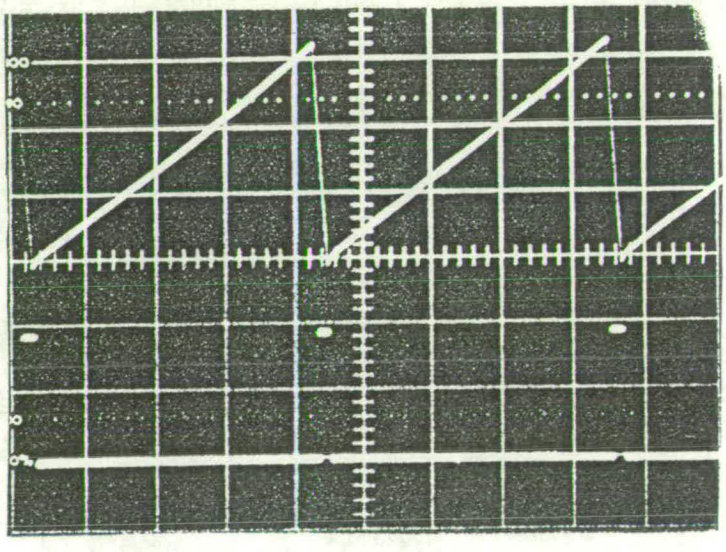
5.1.2) Bias Scanning.

A block diagram of the low frequency function generator utilised to synthesise the ramp waveform is shown in Figure 5.1.4. Central to the circuit is the Intersil 8038 monolithic waveform generator, incorporated because of its good thermal stability (50ppm/^{o}C) and linearity(0.1%) at low frequency. The waveform generation is based on charging and discharging a capacitor with two matched, externally programmable constant current sources, as shown in Figure 5.1.5. By varying the ratio of the two currents an asymmetrical waveform can be produced. Hence the ramp waveform was obtained. When the voltage on the capacitor reaches a specific value (about 1V), the second current source is switched in and discharges the capacitor with a current twice that of the original. The switching comparator provides a square wave output and a sinewave output is provided by applying the triangler waveform to a diode shaping circuit. In addition, the frequency can be controlled by a voltage applied to a transistor in the current source biasing network.



(a)

Ramp waveform generated by low frequency function generator.
Upper trace: ramp 0·5V per mm, 0.5ms per mm.
Lower trace: square 0·2V per mm, 0.5ms per mm



(b)

Ramp waveform generated by low frequency function generator.
Upper trace: ramp 0·5V per mm, 1.25ms per mm.
Lower trace: square 0·2V per mm, 1.25ms per mm.

Figure 5.1.6. Output waveforms obtained from low frequency function generator, illustrating the linearity of the scan waveform.

The linearity of the ramp generator and the output stages is demonstrated by the waveform traces presented in Figure 5.1.6, showing a $40V_{pp}$ ramp signal at a frequency of 30Hz.

All three waveform outputs of the oscillator were buffered and attenuated to $1V_{pp}$. A manual frequency control was incorporated by applying a voltage via a potentiometer to the constant current biasing circuitry. In addition, the square waveform was converted to a TTL level signal by a comparator for use as a trigger output.

Since the bias generator was to provide a low frequency scan function in the open loop mode, a facility for synchronising its frequency to that of the mains was incorporated. If the scan is an even multiple of the mains frequency, the mains interference can be reduced by averaging over an integral multiple of the mains period. This was achieved by exploiting the voltage controlled oscillator property of the 8038 integrated circuit and the exclusive "OR" type phase detector of the LM565 linear i.c.

Since the mains specification is 50Hz ±0.5Hz, a pull-in bandwidth of 1 to 2Hz was necessary. A phase lock loop was constructed in which the oscillator frequency locked to twice that of the mains frequency. Lead-lag compensation was used in order to obtain a second order loop response with a zero at the axis giving an integral response and thus a zero first order error. Thus the oscillator waveform was 90^{o} out of phase with the mains frequency. Since the derived mains synchronised waveform was rectangular, acute filtering of the fundamental was necessary since this component was generated from the first harmonic in the mains waveform.

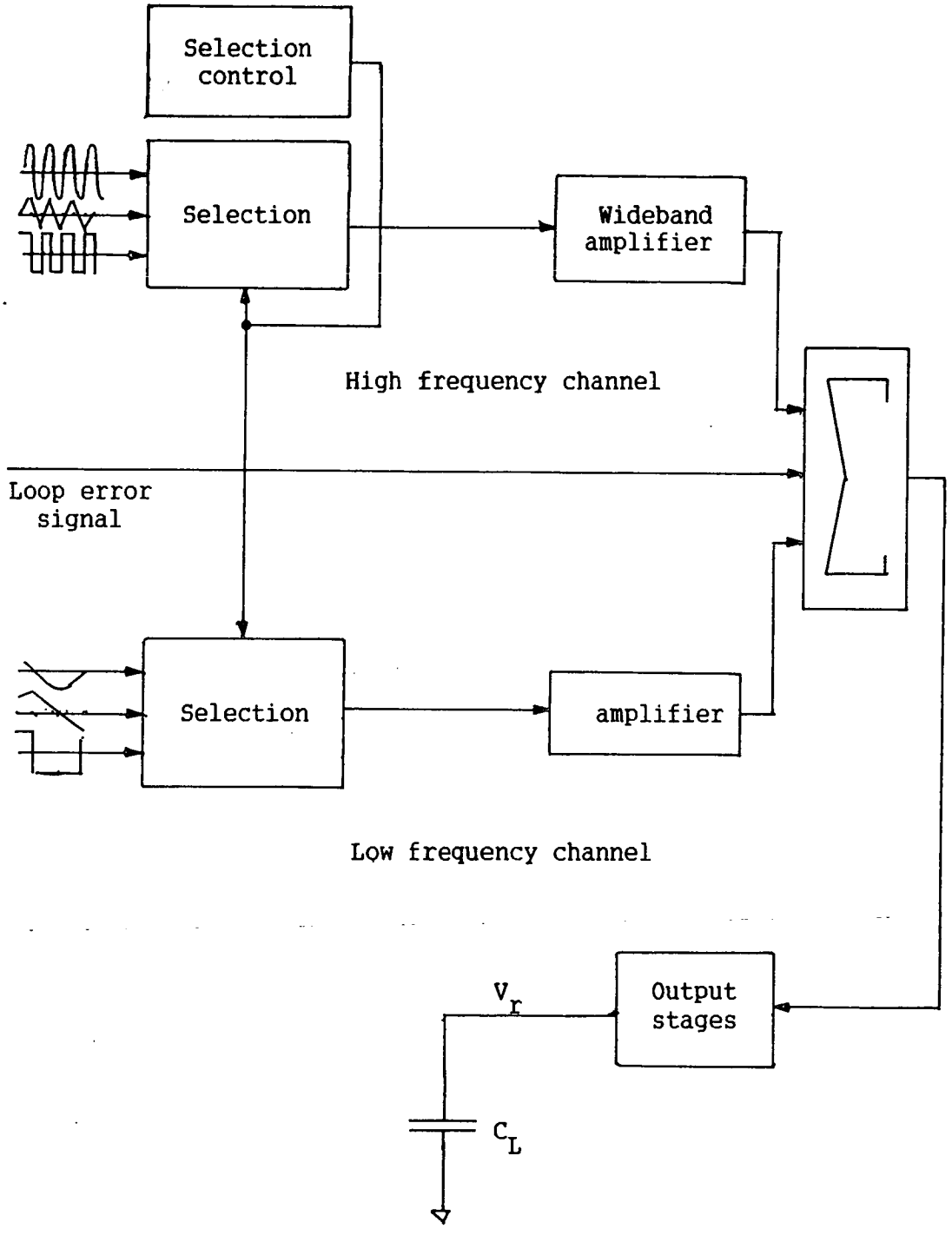


Figure 5.1.7. Block diagram of Selection and Summation circuitry.

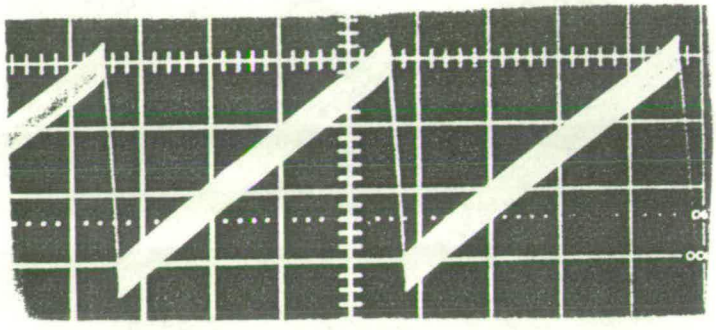
A circuit designed by Lefferts(1982) was used to provide the mains waveform and was founded on triggering a thyristor at the mains zero crossing by a delayed mains signal. This would happen only once per mains period due to the back biasing of the thyristor. The triggering of the thyristor was used to switch a transistor with an optocoupler as part of its load. The output of the opto- coupler was then detected by a schmitt trigger nand gate.

In order to investigate the characteristics of a retarding field energy analyser, as reported in chapter 4, a staircase function was necessary to increment the surface voltage. The trigger output of the low frequency function generator was used to clock a synchronous four bit counter. The outputs of this counter were fed to a DAC and the output of the converter buffered. Hence the required staircase waveform in synchronism with the scanning waveform was realised.

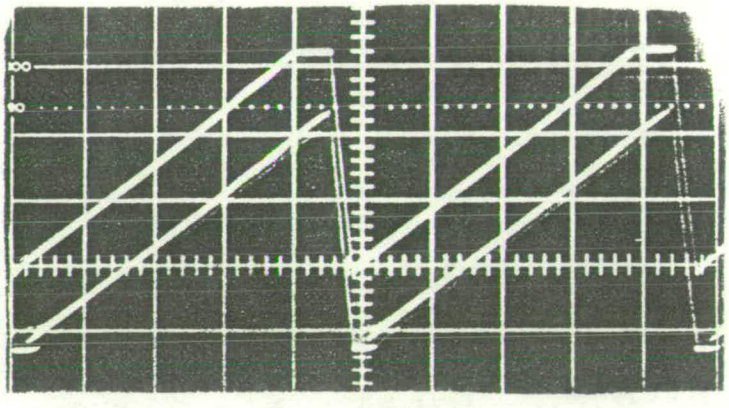
5.1.3) Selection and Summation.

As illustrated in Figure 5.1.7, the modulation and error signals had to be summed with the scanning bias waveform, prior to being applied to the retardation electrode. In addition, it was desirable to be able to control the amplitude of the ramp and modulation waveforms manually and independently. Since the signal generators used to provide the scanning and modulation waveforms each provided three waveforms, only one of which was necessary, selection was implemented by two manually controlled single pole triple throw analogue switches.

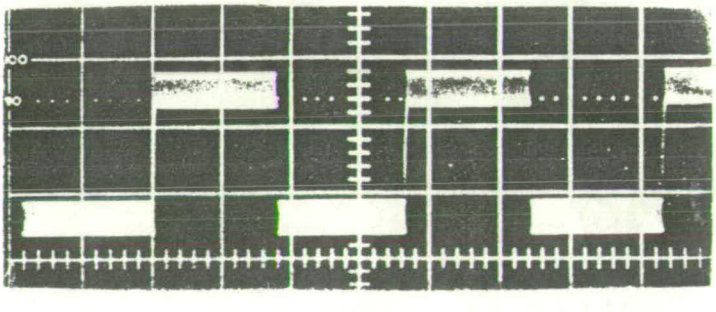
In summary, two signal channels, one high frequency and the other low frequency were summed. The addition was implemented with a high frequency operational amplifier configured in the standard



(a)
High frequency sine wave plus
low frequency ramp.



(b)
High frequency square wave plus
low frequency ramp.



(c)
High frequency triangle wave plus
low frequency square wave.

Figure 5.1.8. Examples of composite waveforms obtainable from the selection and summation stages. All traces: vertical scale: 0·5V per mm, horizontal scale: 1.25ms per mm.

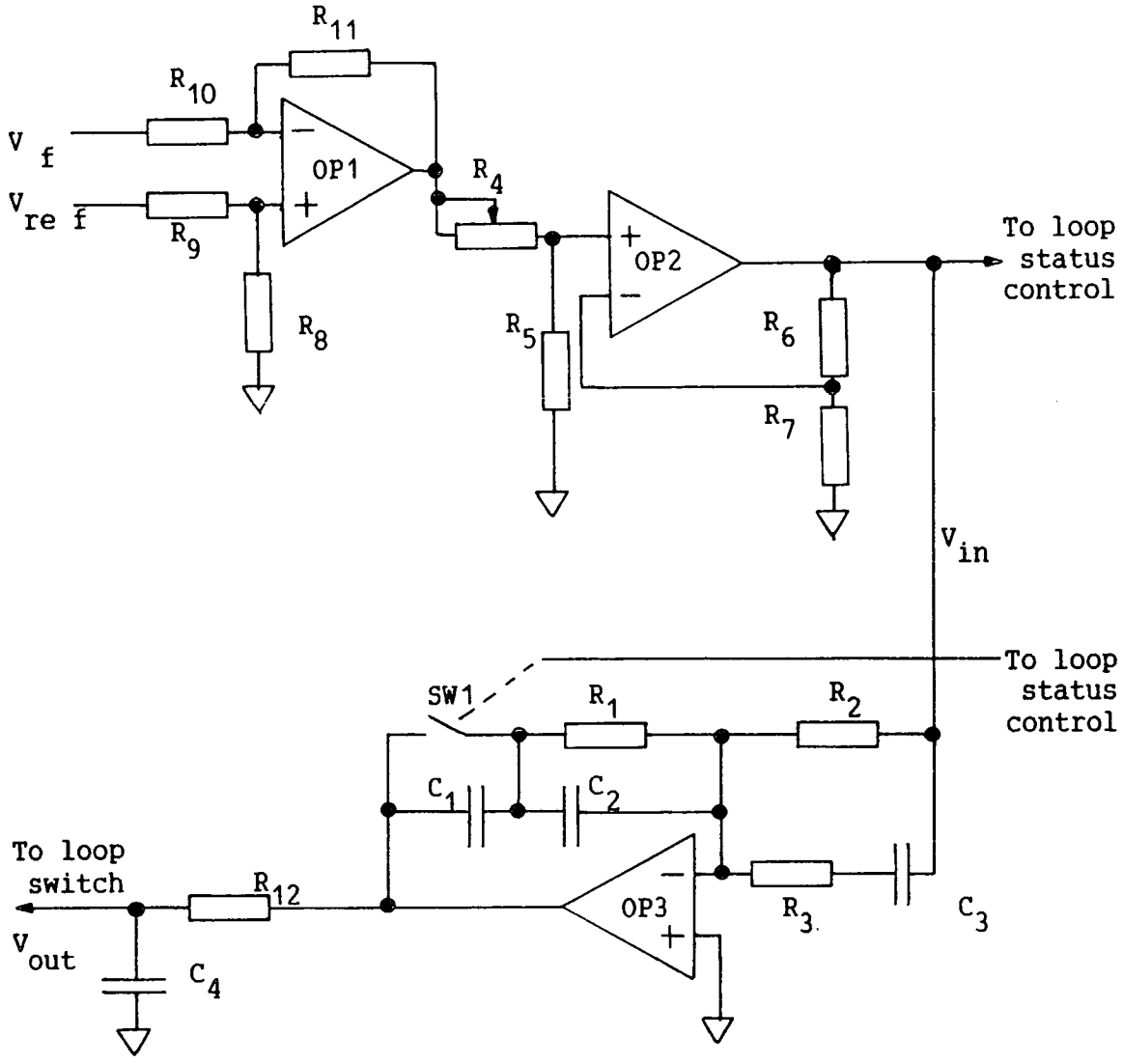
inverting summation circuit and the gain of each channel adjusted by varying the respective input resistor. Examples of the composite waveforms obtainable from the summation stages are shown in Figure 5.1.8.

5.1.4) Retardation Electrode Output Stages.

The composite waveform derived from the preceding summation circuits was then boosted in a voltage amplifier to produce a retardation waveform with a $40V_{pp}$ voltage range. This was necessary to enable entire "S"-curves to be reproduced and to accrue the measurement system a high dynamic range.

The normal power supply voltages of ±15V were insufficient to achieve the required voltage range and therefore a power supply was constructed solely for the output stages. This was a floating supply, incorporating a dual tracking regulator which allowed a maximum output swing of +-25V to be realised.

As well as providing large voltage swings at low frequency, the output stages also had to maintain a small signal response up to 1MHz, to reproduce a high frequency modulating square wave with good fidelity. The circuitry of the output stages consisted of an LM344 uncompensated operational amplifier which possessed a high slew rate(30V/μs) and large bandwidth(4MHz). These attributes would allow the modulation waveform to be reproduced at high frequency and amplitude. Although, the retardation electrode provided a purely capacitive load the operational amplifier's output was buffered with a class B push-pull complementary transistor stage and was fully protected. The output stage was compensated with a lead-lag network had a voltage gain of 6dB. A manually adjustable DC level control was



Transfer function:

$$\frac{V_{out}}{V_{in}} = -\frac{(SR_1(C_1+C_2)+1)(SC_3(R_3+R_2)+1)}{SR_2C_1(SR_1C_2+1)(SR_3C_3+1)(SR_{12}C_4+1)} \qquad (5.1.4)$$

If $C_1 \gg C_2$ and $R_2 \gg R_3$ then

$$\frac{V_{out}}{V_{in}} \simeq -\frac{(SR_1C_1+1)(SC_3R_2+1)}{SR_2C_1(SR_1C_2+1)(SR_3C_3+1)(SR_{12}C_4+1)} \qquad (5.1.5)$$

If SW1 is closed

$$\frac{V_{out}}{V_{in}} \simeq -\frac{R_1}{R_2}\frac{(SR_3C_3+1)}{(SR_1C_2+1)(SR_3C_3+1)(SR_{12}C_4+1)} \qquad (5.1.6)$$

Figure 5.1.9. Controller schematic and transfer function.

also included by varying the reference voltage of the operational amplifier circuit.

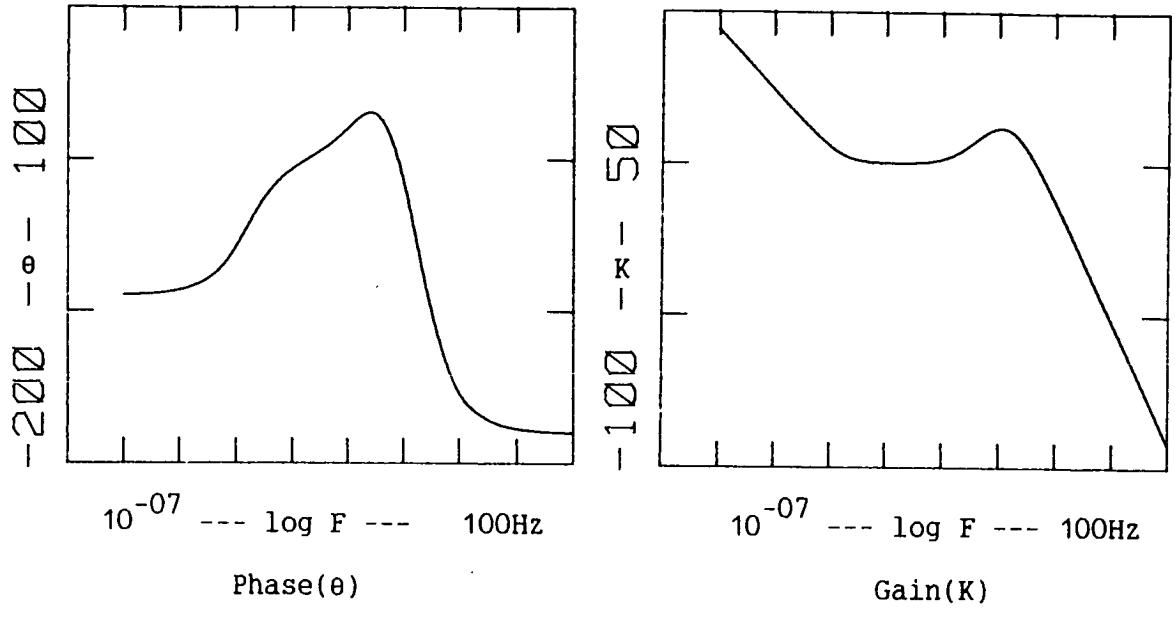
5.1.5) Controller.

The controller schematic and transfer function are presented in Figure 5.1.9 with the corresponding phase and frequency response in Figure 5.1.10(a). The compensation circuit incorporates proportional, derivative and integral control terms.

The loop error signal is generated by the difference circuit of OP1, while the forward loop gain or proportional term is provided by the circuit comprising OP2 and R_4,R_5 potential divider. OP2 is a 709 instrumentation amplifier capable of providing 40dB of gain up to 1MHz bandwidth. However it can only be stabilised for a fixed gain and hence the attenuator determines the forward loop gain.

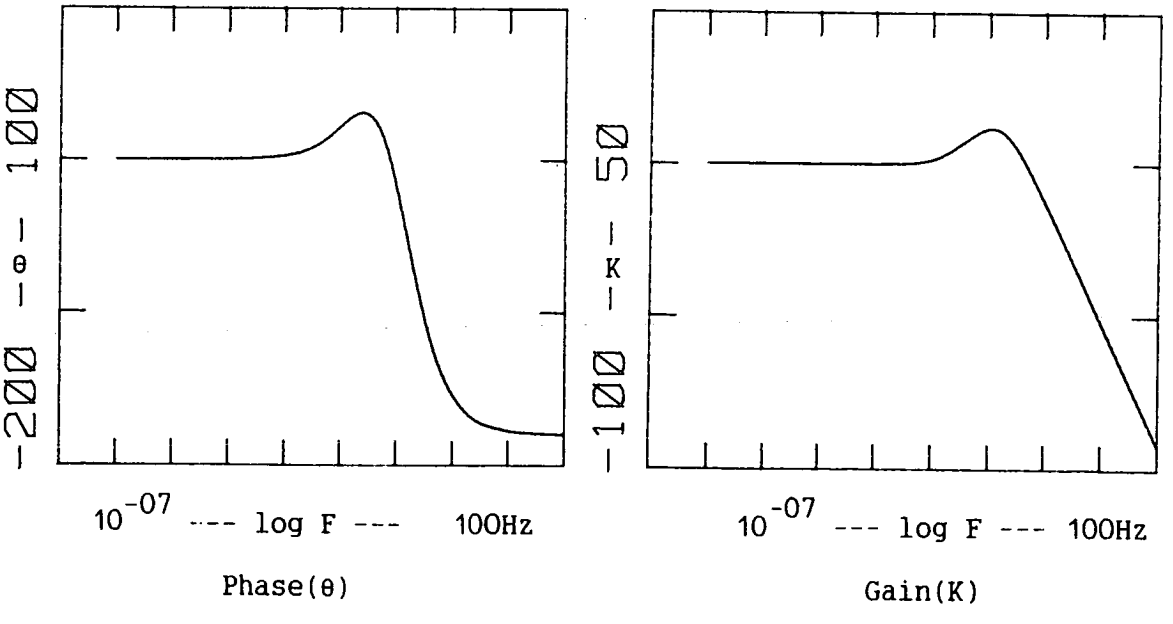
The transfer function shaping incorporating the proportional and derivative terms was implemented with OP3 and is adjacent circuitry. The integral term is included to eliminate any offset errors and to increase the DC accuracy of the servo-system. The break frequency is determined by the time constant of R_2 and C_1 and is made equal to the time constant of the lowest frequency zero.(R_1,C_1). This produces an in band gain of 1.

The phase advance or derivative term is utilised to introduce a positive phase shift at the unity gain frequency of the servo and thereby increase the bandwidth of the system. The frequency span of the derivative term is determined by the two time constants R_3, C_3 and R_2, C_3 and was fixed at one decade giving a maximum phase advance of 45^{o}. A 40dB per decade roll-off is achieved by making R_1, C_2



(a)

Closed loop



(b)

Open loop

Figure 5.1.10. Transfer function of controller for both modes of operation.

equal to R_{13}, C_4

Since the system was operated in open and closed loop modes, the controller's integral term gave rise to problems because it integrated its own offset and subsequently saturated the output. The effect on the system was to introduce a transient when the feedback loop was closed forcing the system into one of the out of lock regions of the operating characteristic. Hence, lock was impossible to attain. This problem was overcome by shorting out the integrating capacitor when the system was in the open loop mode and reducing the low frequency and DC gain of the amplifier to unity. This is illustrated in Figure 5.1.10(b).

The output of the measurement system in the closed loop mode would constitute the error signal as applied to the summation circuits. This signal was inverted and amplified by a factor equal to that of the retardation electrode output stages to give a 1:1 correspondence to the voltage applied to the specimen. In the open loop mode, the first and second derivatives were obtained immediately after the baseband filtering and amplified by an arbitrary amount and therefore are not accurately proportioned.

5.1.6) Automatic gain control.

Since the second derivative signal to be detected was a pure frequency component with no DC term, black level of the videohead amplifier could be suppressed to enable the higher amplification of the AC signal required, improving the SNR. Hence a higher photomultiplier tube gain could be realised before saturation occurred. When the servo-system is in lock, the amplitude of the fundamental term should not vary significantly with the surface voltage and therefore

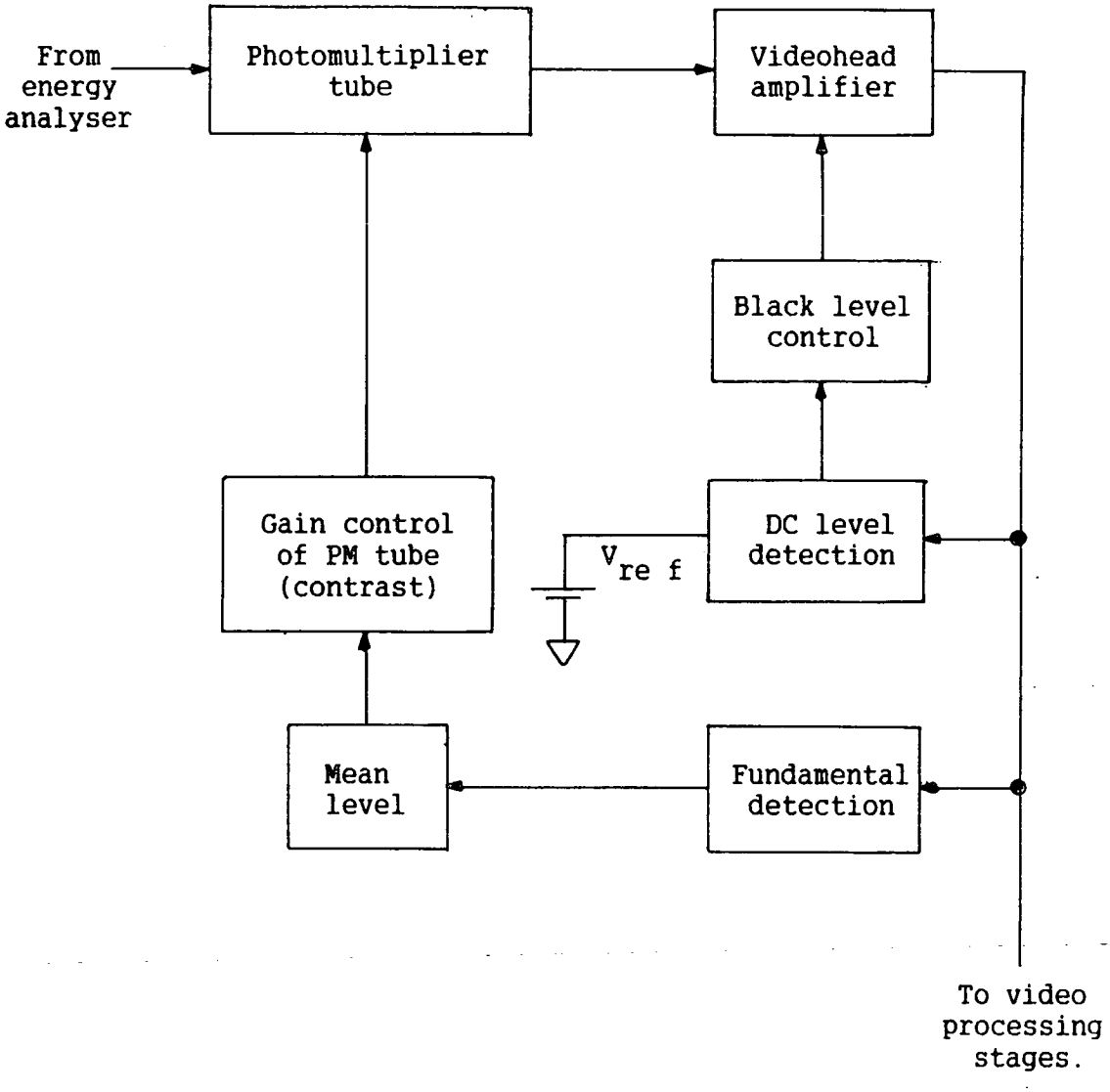


Figure 5.1.11. Block diagram of coherent automatic gain control circuits.

can be used as a source of coherent automatic gain control. Hence by demodulating the fundamental and comparing its RMS value to a reference, the variation of gain of the operating characteristic can be compensated.

The proposed system is shown in Figure 5.1.11, and maintains the videohead output DC level at half full-range. The gain of the photomultiplier tube is controlled by varying the dynode chain voltage and maintains an output fundamental amplitude of slightly less than half full range. When using a Cambridge S100 SEM, the solution is facilitated by the inclusion of a video DC level control on the front panel and the easy access of the photomultiplier gain control circuits.

The automatic gain control system can only be used when the system is in lock and before a measurement is made the "S"-curve should be normalised. To achieve this the gain control system can be operated in a sampled data mode, comparing the peak of the SED for each scan with a reference level.

5.2) Modulation System.

The modulation system performed the function of inducing, detecting and measuring the "S"-curve harmonic distortion and the major part of the noise reduction through filtering. A block diagram of the system is presented in Figure 5.2.1 and demonstrates the dual channel detection stages necessary to detect the first and second harmonics simultaneously. These two channels were identical in design and therefore only one channel will be discussed. The modulation system is transparent to the servo-system and can be represented as a

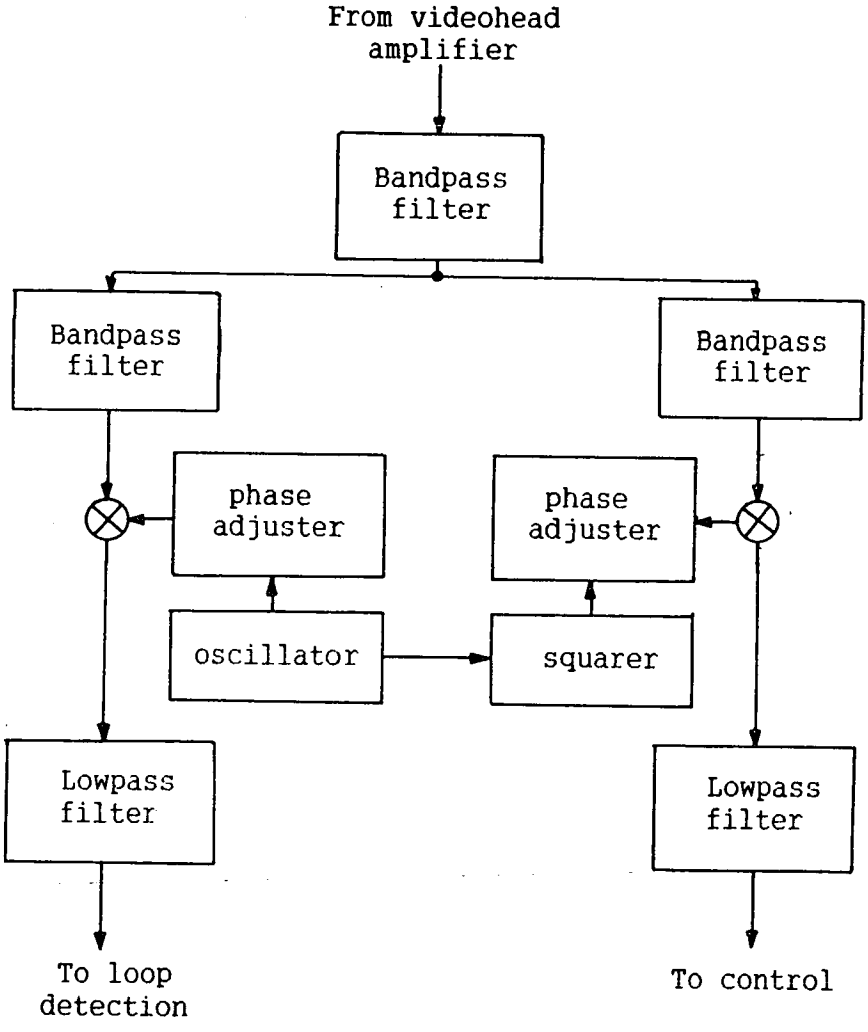


Figure 5.2.1. Block diagram of modulation system.

gain block. The input to the modulation system has an associated AC gain, while the output has a corresponding DC gain. The AC gain should be optimised in preference to the DC gain to maximise the output SNR.

The conventional method of measuring the harmonic distortion of linear systems is to carry out an INTERMODULATION test. Normally this consists of applying two sinewaves to the system and detecting the cross modulation product term. An alternative method is to apply a sinewave to the system and after nulling the original sinewave from the output, computing the TOTAL HARMONIC DISTORTION by taking the ratio of the sum of the mean square of all the distortion components to the mean square of the first harmonic. The former method requires two coherent low distortion oscillators and generates a series of higher harmonics, reducing the power in the cross product term. The latter method would reduce, in this case, to detecting the SECOND HARMONIC DISTORTION and require only one stable low distortion oscillator.

Since both ends of the communication channel are accessible simultaneously, the local oscillator frequency can be locked to the carrier frequency and in fact can be derived from the carrier frequency itself, eliminating the requirements for a stable oscillator. In addition, by minimising the phase discrepancy between the carrier and local oscillators, coherent detection can be performed, maximising the signal output from the detector stages. This can be demonstrated by considering two sinewave signals of the same frequency and slightly out of phase. The output from a linear detector can be shown to be

$$V_{out} = V_c \mathrm{Sin}(\omega_c t + \theta).V_o \mathrm{Sin}(\omega_o t) \qquad (5.2.1)$$

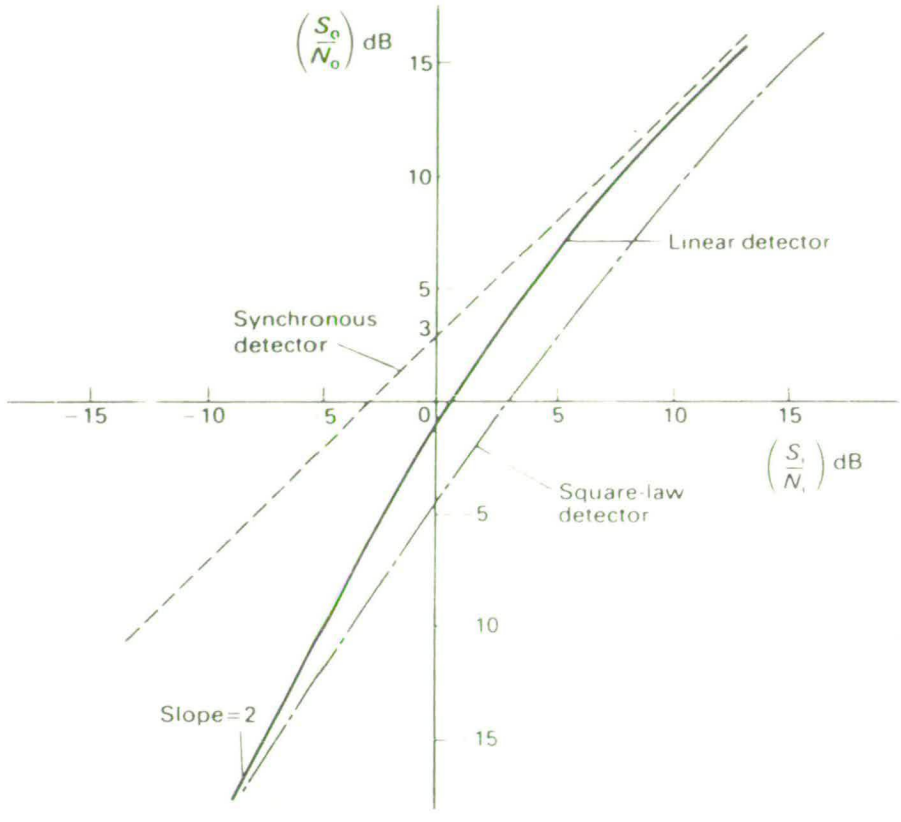
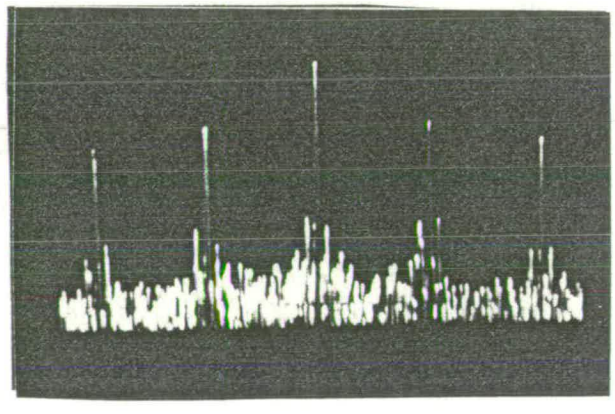


Figure 5.2.2. A comparison of the process gain of various detection techniques. (from Connor, 1982)



Power spectrum of output waveform when a 15kHz sine wave is applied to the specimen.(DC level centre of trace)

Figure 5.2.3. Modulation of surface voltage by retardation voltage.

which reduces to

$$V_{out} = \frac{V_c V_o \text{Cos}(\theta)}{2} + \frac{V_c V_o \text{Cos}(2\omega t + \theta)}{2} \qquad (5.2.2)$$

It can be seen from the latter equation that after lowpass filtering, the DC term is proportional to both the amplitudes of the carrier and local oscillator waveforms and the phase difference(θ) between them. Another distinct advantage coherent detection has over other methods of demodulation, linear and nonlinear, is that the output SNR is a linear function of input SNR. Therefore, this type of demodulation exhibits no "AM fading" at low input SNRs. Figure 5.2.2 illustrates the performance of different demodulation systems. Coherent demodulation also has a process gain of 2, which compensates for the signal power spreading inherent in the amplitude modulation process.

Shown in Figure 5.2.3. is the frequency power spectrum obtained when a sinewave is applied to the surface of the specimen and a high frequency triangular waveform is applied to the retardation electrode. From this it can be deduced that the output waveform is the result of multiplication in the frequency domain of the "S"-curve operating characteristic and the surface voltage. Extrapolating from this, if the retardation voltage is small in amplitude and a monotone, the output waveform is again the result of frequency multiplication of the surface voltage and the incremental "S"-curve, consisting of a linear and a square term. Therefore, the modulation is not strictly amplitude modulation, but a form of "pulse amplitude" modulation where the "pulse" in this case is the "S"-curve.

The baseband is governed by the Nyquist sampling criteria and is

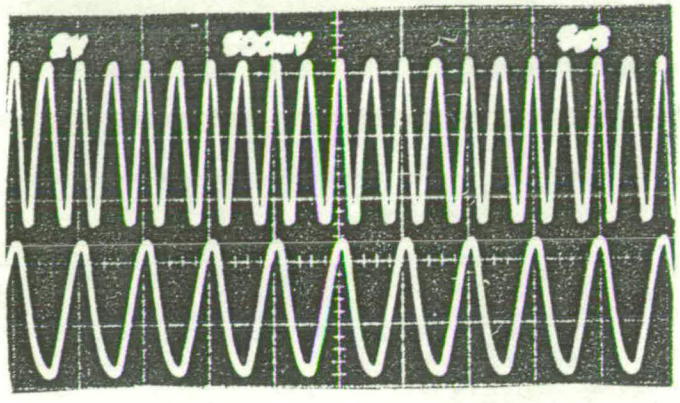


Figure 5.2.5. Upper trace: Output waveform of multipier 0·2V per mm, 0·5 μs per mm.
Lower trace: Input sine wave 0·1V per mm, 0·5 μs per mm.

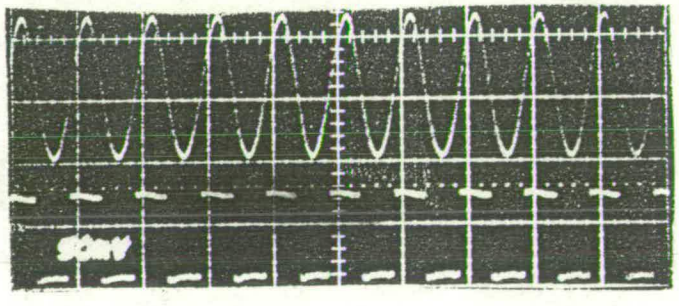


Figure 5.2.4. Output waveforms from High Frequency Channel
Upper trace: output sinewave 0·2V per mm 0·5 μs per mm.
Lower trace: output square wave 0·2V per mm, 0·5 μs per mm.

a maximum of half the modulation frequency to prevent aliasing errors. However, this would require very sharp roll-off filters and baseband prefiltering to maintain the baseband integrity. Therefore, since the baseband, in this case, could not be prefiltered, a large guard band was necessary and a rule-of-thumb that the highest baseband frequency was to be a factor of ten below the modulation frequency was adopted. A further constraint was necessary if the system was to be compatible with a primary beam sampling system. This was that the maximum modulation frequency should be a factor of ten, again below the sampling rate. Since, the maximum sampling rate envisaged was 10MHz, a modulation rate of 200kHz, was decided upon, because the second harmonic generated, would be at 400kHz. Theoretically, this would allow a baseband of 20kHz, which would enable a minimally clocked 5kHz waveform to be resolved.

An inherent advantage of a modulation system was that it totally eliminated all the 1/f flicker noise contributions from the detection system including mains interference and drift. This is due to the fact that the signal power is transferred through the videohead amplifier at a much higher frequency where only Gaussian white noise is present. However, modulation does not eliminate the effects of primary beam noise and interference coupled through the filament heater and high voltage power supplies.

5.2.1) Modulation.

The actual modulation was carried out by the energy analyser. That is by varying the energy filter's cut-off energy the collected output current was varied. The high frequency modulation signal generator was functionally identical to the low frequency signal genera-

tor presented in Figure 5.1.4. and again utilised the Intersil 8038 waveform generator i.c. Although all three waveforms were available for use as modulation functions only the sinewave was necessary. However, the square wave could be used to detect the operating characteristic itself, as described by Hamilton(1977). Therefore, it was decided to incorporate a means of selecting all the available waveforms into the design. Further, it was decided to use the output square wave as the digital system clock, it was converted to a TTL level by a comparator and also used as a synchronisation source.

As with the scanning function, the modulation function was synchronised to the mains with a phase locked loop identical to the one described in the previous section with the exception that a divide by 40,000 counter was included in the feedback loop to generate the mains frequency. An external synchronisation source was also made available to both loops. This would allow the system oscillators to be tied to the field synchronisation of an external scan generator and therefore allow VOLTAGE CODING and, with a sampling system, LOGIC STATE MAPPING.

The modulation waveform's amplitude was made manually variable to allow the output SNR to be maximised, allowing an output dynamic range of 60dB. An output sinewave obtained at a frequency of 200kHz, is shown in Figure 5.2.4. to illustrate the low level of distortion present.

To permit the coherent detection of the second harmonic term, a sinewave at twice the modulation frequency and in synchronism with it had to be generated. This was realised by using a 1546 linear multiplier, configured as a squarer, to exploit the trignometrical iden-

tity:

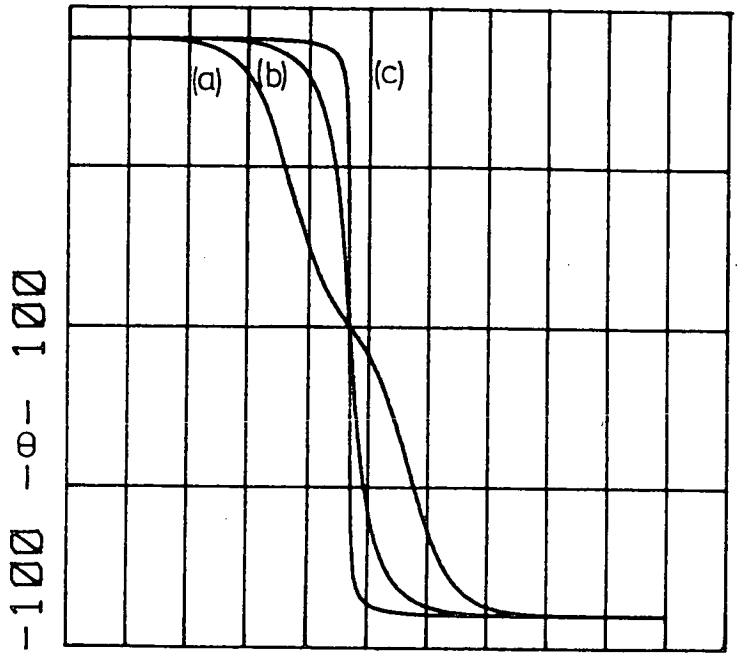
$$\mathrm{Sin}^2\omega t = \frac{(\mathrm{Cos}2\omega t + 1)}{2} \qquad (5.2.3)$$

The input and output waveforms are shown in Figure 5.2.5. The output was high pass filtered to remove the unwanted DC level, before being applied to the demodulation section.

5.2.2) AC Coupling.

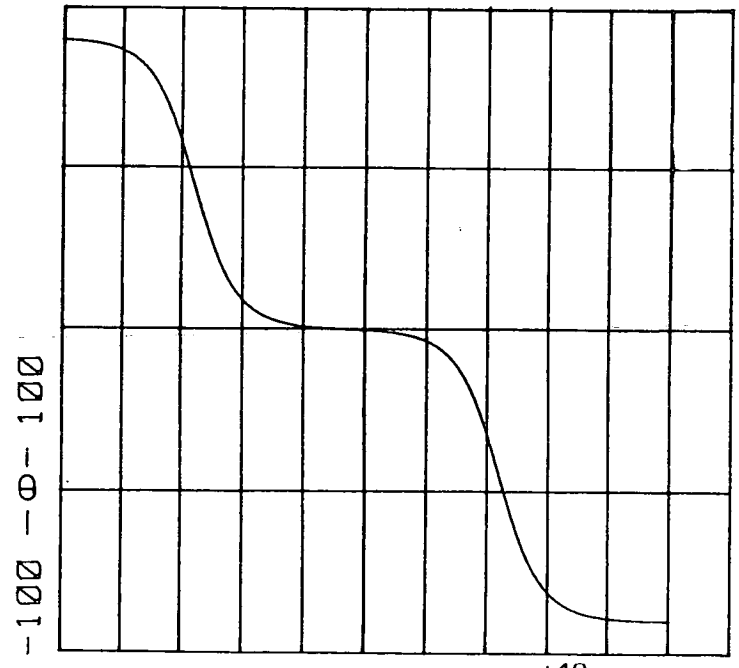
When the system operates in the open loop mode and the "S"-curve is generated at the output of the videohead amplifier, it was necessary to remove all the frequency components below the modulation frequency for two reasons. Firstly, the large "S"-curve voltage swing would saturate the high gain input stages of the detection system and secondly to maximise the SNR of the signal before further amplification of the first and second harmonic terms. This primary process gain can be large, since the output bandwidth of the videohead amplifier is at least 5MHz and the required signal bandwidth is only 300kHz. Therefore, the detection stages were AC coupled to the video head amplifier.

The detection stages were a high impedance balanced input amplifier with a gain of two. There followed two balanced single pole filters; one high pass and one low pass, before the signal was fed into a high gain differential amplifier to convert the signal to a single end. A variable gain two pole, one zero filter was used and in conjunction with the previous filter produced a second order low-pass Butterworth filter.



(a)

Phase(θ) response of a bandpass filter with various Q values, illustrating the large variation of phase around the centre frequency. curve (a) : Q = 0.1; curve (b) : Q = 1; curve (c) : Q = 10



(b)

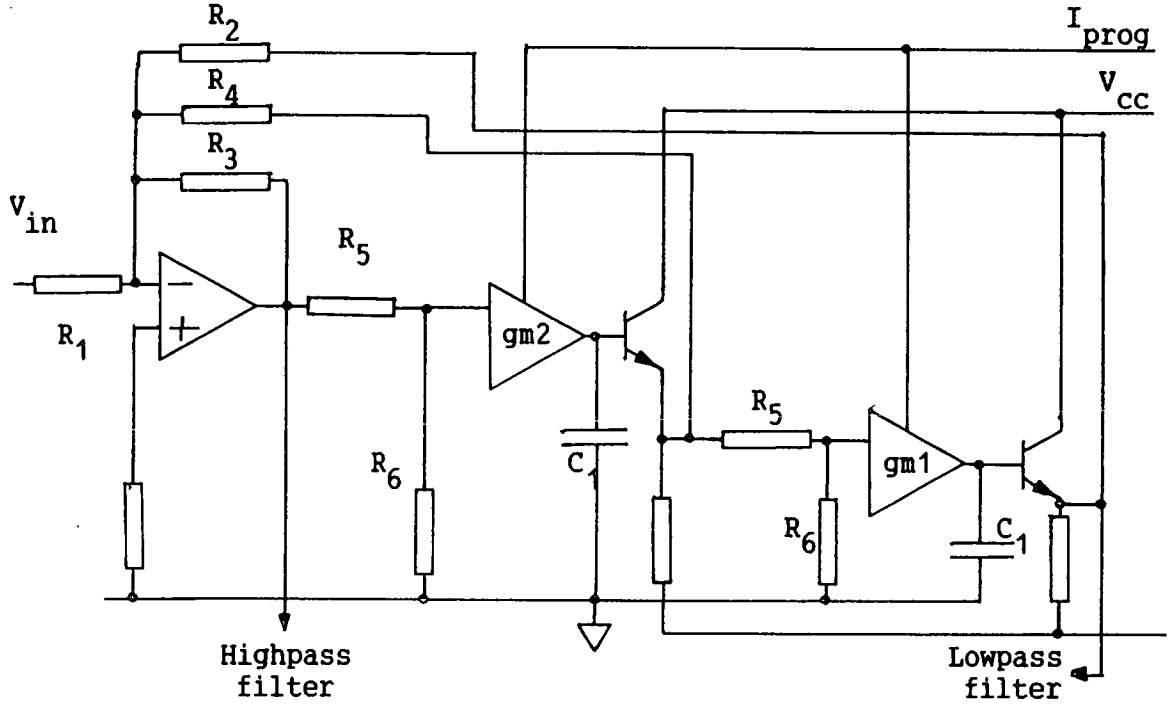
Phase(θ) response of a cascade highpass and lowpass filter, illustrating the small phase variation in the passband.

Figure 5.2.6. Possible channel filters' phase response.

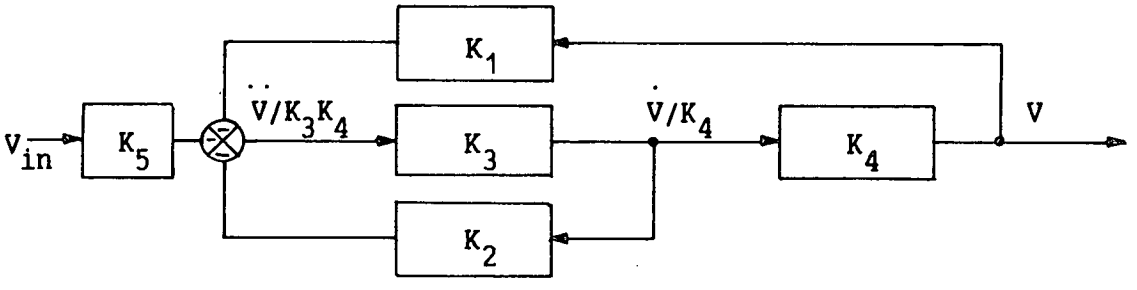
5.2.3) Channel Filtering.

Before the signal was demodulated, further filtering was introduced to reduce cross talk between channels and prevent saturation in the following stages. The signal channel was split into two; one for each harmonic and a series of three filters were used in each channel. These filters were one notch filter, one high pass and one low-pass.

A notch filter was used in each channel to attenuate any of the other channel's modulation component. Thus the notch filter in the fundamental channel was tuned to the first harmonic and vice-versa. This allowed the channel gain to be increased without saturation. High Q bandpass filters were not used in the filter section, but a high pass and low pass filter cascaded. The reason for this was that the phase response of a resonant filter is very strongly dependent on frequency, about the resonant frequency and that this dependency is directly proportional to the Q of the filter, as illustrated in Figure 5.2.6(a). Therefore, as the Q is increased, the gradient of the phase response increases and any drift in either the modulation frequency or the filter tuning would result in a large phase error and a consequent loss of coherence between the detected and local carrier. Hence the stability of the signal from the output stages may be compromised. The performance of the cascaded sections has a much less critical phase response with only a phase variation of 90^{o}, instead of the 180^{o} of a resonant filter, about the central frequency, as shown in Figure 5.2.6(b). An alternative method would be to use a homodyne method for frequency translation to a high frequency to allow the use of high Q, highly stable crystal filters.



(a)
Circuit schematic.



(b)
State variable diagram.

State equation

$$\ddot{V} = -(K_5 V_{in} + K_1 \dot{V} + \frac{K_2}{K_4} \dot{V}) \qquad (5.2.4)$$

Using the identity $SV = \dot{V}$ and choosing $\dot{V}$ as the state variable, the transfer function is

$$\frac{V_{out}}{V_{in}} = -\frac{K_5}{K_1} \frac{1}{\frac{s^2}{K_1 K_3 K_4} + \frac{SK_2}{K_1 K_4} + 1} \qquad (5.2.5)$$

Figure 5.2.7. Channel filter design.

The filters utilised were designed and constructed using a novel implementation of the State Variable approach after Kerwin et al.(1967) The design method is illustrated by Figure 5.2.7. and is capable of producing easily tuned highpass, lowpass and notch filters from the one design. As an example, the low pass filter transfer function is

$$\frac{V_{out}}{V_{in}} = -\frac{K_5}{\frac{S^2}{K_1K_3K_4} + \frac{S\,K_2}{K_1K_4} + 1} \qquad (5.2.5)$$

Giving a resonant frequency of

$$f_o = 2\pi\sqrt{K_1K_3K_4} \qquad (5.2.6)$$

and a Q factor

$$Q = \frac{1}{K_2}\sqrt{\frac{K_1K_4}{K_3}} \qquad (5.2.7)$$

In the normal implementation of these filters, tuning is achieved by varying either, two accurately matched resistors or capacitors, while by using two transconductance amplifiers as the integrators, this can be achieved by varying the transconductance(gm) which are inherently matched on a dual i.c. such as an LM13600 or the 3080. Therefore by supplying the bias current for the two transconductance amplifiers from the same current source, K_3 equals K_4. Further, since the summing amplifier is a voltage-controlled-voltage-source operational amplifier, K_1 can be easily made to equal unity and all three control parameters have been totally disassociated. The quality factor, gain and resonant frequency become respectively

$$f_o = K_3 = \frac{gm\ R_6}{(R_5 + R_6)} ; \quad Q = \frac{1}{K_2} = \frac{R_4}{R_3} ; \quad Gain = K_5 = \frac{R_3}{R_1} \quad (5.2.8)$$

The notch filter is obtained by subtracting the highpass and lowpass outputs. An additional advantage of this type of filter is its inherent stability and insensitivity to component tolerances, attained by having two integrators inside a feedback loop.

5.2.4) Demodulation.

Although linear filtering using linear multipiers would have been the optimum method of demodulation, eliminating any harmonic generation, switched multipliers were used for ease of use and economics. These devices have one linear large signal channel and one small signal linear channel or as it is more commonly used as, a switching channel. The outputs are balanced and a differential amplifier used to give a single ended output with a gain of two and to minimise carrier feed-through. From here the signal is buffered and fed to the final filtering stages. One advantage of the switched multipier is that it eliminates any errors in the output signal caused by variation in the amplitude of the demodulating waveform. For a normalised amplitude, equation 5.2.2 becomes

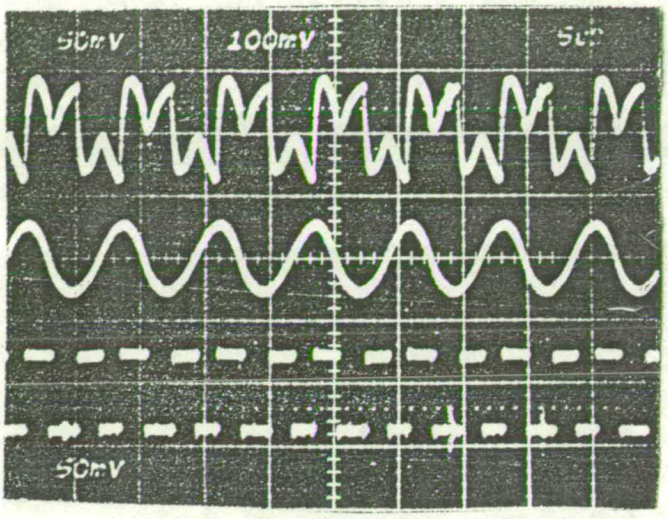
$$V_{out} = \frac{V_o Cos(\theta)}{2} \quad (5.2.9)$$

The effect of using a square wave as a demodulation signal can be shown by considering the first three terms of its Fourier series

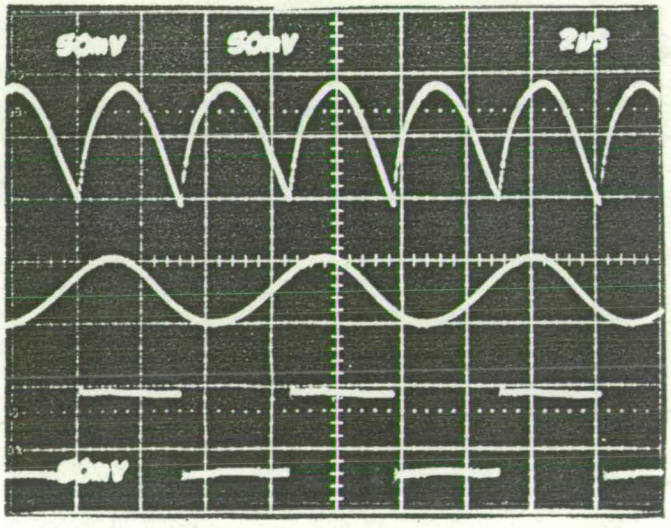
$$V_{sq} = \frac{4V}{\pi}\left[Sin(\omega_c t) - \frac{1}{3}Sin(3\omega_c t) + \frac{1}{5}Sin(5\omega_c t) \ \ldots \right] \quad (5.2.10)$$

For a unity amplitude square wave the output after filtering would be

$$V_{out} = \frac{2V_c}{\pi}Sin(\theta) \quad (5.2.11)$$



(a)
Second harmonic.
Upper trace: unfiltered output 0.05V per mm, 0.5 µs per mm.
Middle trace: Input sine wave 0.05V per mm, 0.5 µs per mm.
Lower trace: Switching input 0.1V per mm, 0.5 µs per mm.



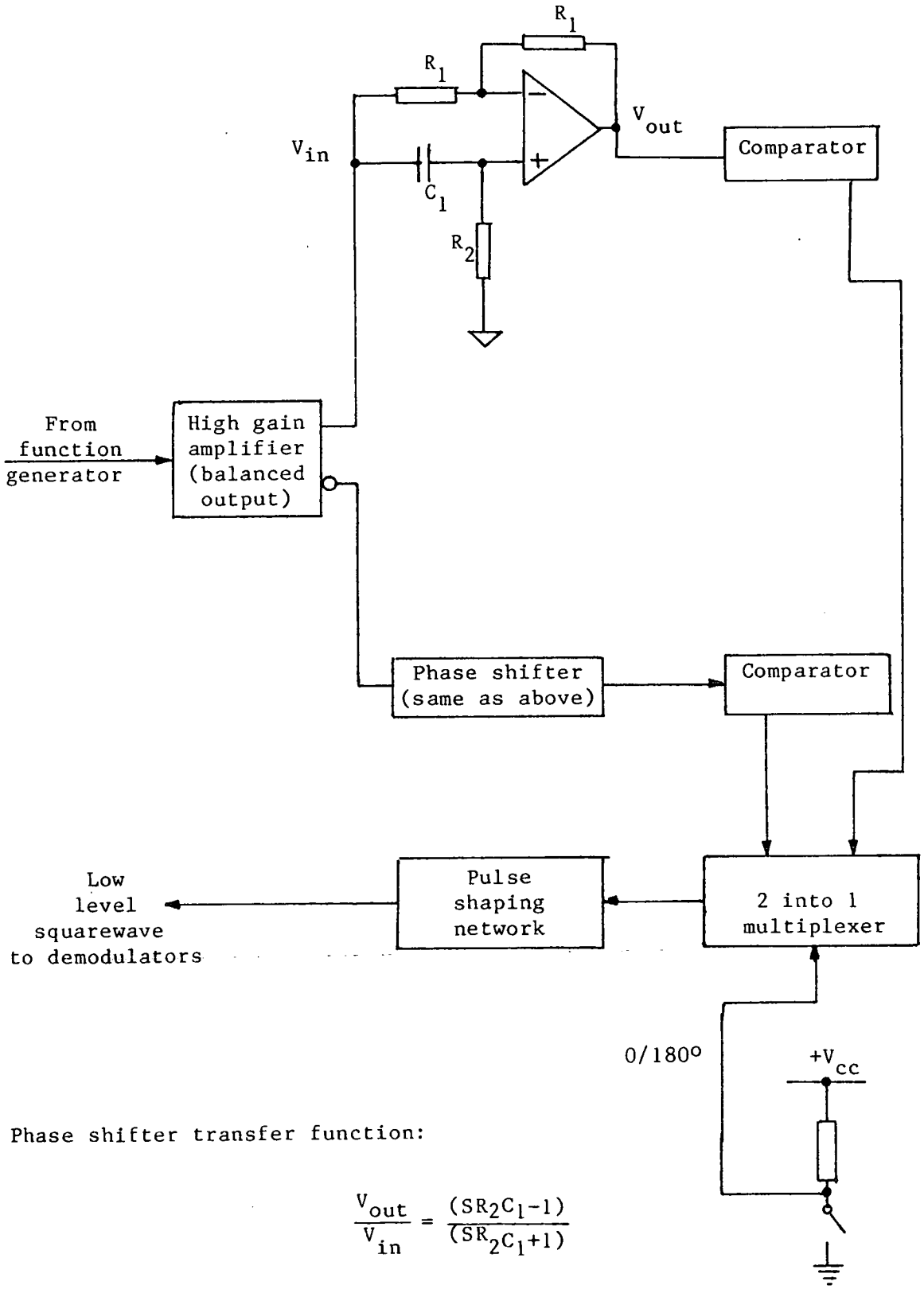
(b)
Fundamental.
Upper trace: unfiltered output 0.05V per mm, 0.2 µs per mm.
Middle trace: Input sine wave 0.05v per mm, 0.2 µs per mm.
Lower trace: Switching input 0.05V per mm, 0.2 µs per mm.

Figure 5.2.8. Test input and output waveforms of the demodulators.

If there is distortion present in the input signal, odd distortion harmonics will generate even output harmonics and vice versa. Note also, that any fundamental in the second harmonic channel and vice versa will not generate a DC output term and therefore compromise the detection accuracy. Only if appreciable components of the third harmonic and above are present in the input signal, will errors be generated in the output signal. Hence, the channel filters had to be of a sufficient bandwidth to exclude any third harmonic terms.

Therefore, the use of square wave demodulation does not affect the accuracy of the detection system, if the output signal is properly filtered. Indeed, the amplitude of the fundamental of a square wave is greater than the amplitude of a corresponding sine wave, giving an increased output signal. The output filters had to be able to attenuate the higher harmonics, generated in the detection process to prevent carrier feed-through. The lowest frequency component generated would be a first harmonic, i.e at the fundamental frequency. However, its baseband modulation would extend down to half the fundamental frequency and therefore, the output filters had to be designed to ensure that any frequency component above this frequency was attenuated to an insignificant level.

The input and unfiltered output waveforms of the switched multiplier for both channels are given in Figure 5.2.8 and demonstrate the device as a switched gain amplifier. These oscillographs were recorded using test signals obtained from the modulation oscillator and multiplier as inputs. This facility allowed the phase discrepancies accumulated from the system to be eliminated and thus when in operation, the phase of each demodulation signal requires only slight adjustment before maximum output SNR is obtained. Therefore only the



Phase shifter transfer function:

$$\frac{V_{out}}{V_{in}} = \frac{(SR_2C_1-1)}{(SR_2C_1+1)}$$

Figure 5.2.9. Phase shift circuit.

incoherence attributed to the videohead amplifier has to be negated and this should be a minimum at this frequency.

5.2.5) Phase Delay.

To ensure coherence between the carrier and the local oscillator signals, an all pass, constant time delay network, which could be manually adjusted to give a phase delay from $0^{o}$to 360^{o} in the frequency range of interest, was necessary. A block diagram of the phase delay unit and a schematic of the phase delay circuit is presented in Figure 5.2.9. The circuit transfer function can be formulated as

$$\frac{V_{out}}{V_{in}} = \frac{S\tau - 1}{1 + S\tau} \qquad (5.2.12)$$

Where τ is the time constant and equals $R_2 C_1$. This is a first order time delay approximation which gives a unity gain response and a phase response(θ) which can be expressed as:

$$\theta = -2\mathrm{Tan}^{-1}(\omega\tau) \qquad (5.2.13)$$

Differentiating this gives the group delay

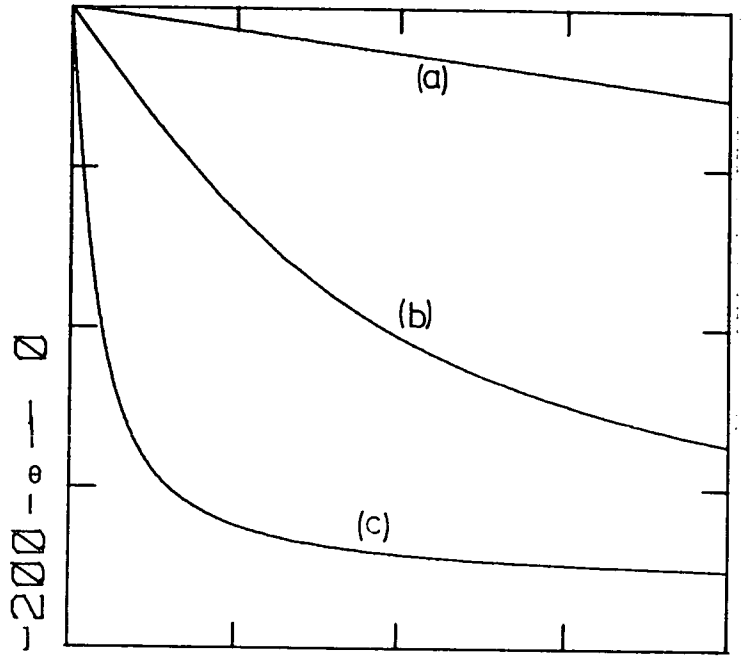
$$\frac{\delta\theta}{\delta\omega} = \frac{-2\tau}{1 + (\omega\tau)^2} \qquad (5.2.14)$$

Thus, for frequencies for which $(\omega\tau)^2 << 1$ is true, the group delay is constant

$$\frac{\delta\theta}{\delta\omega} \simeq -2\tau \qquad (5.2.15)$$

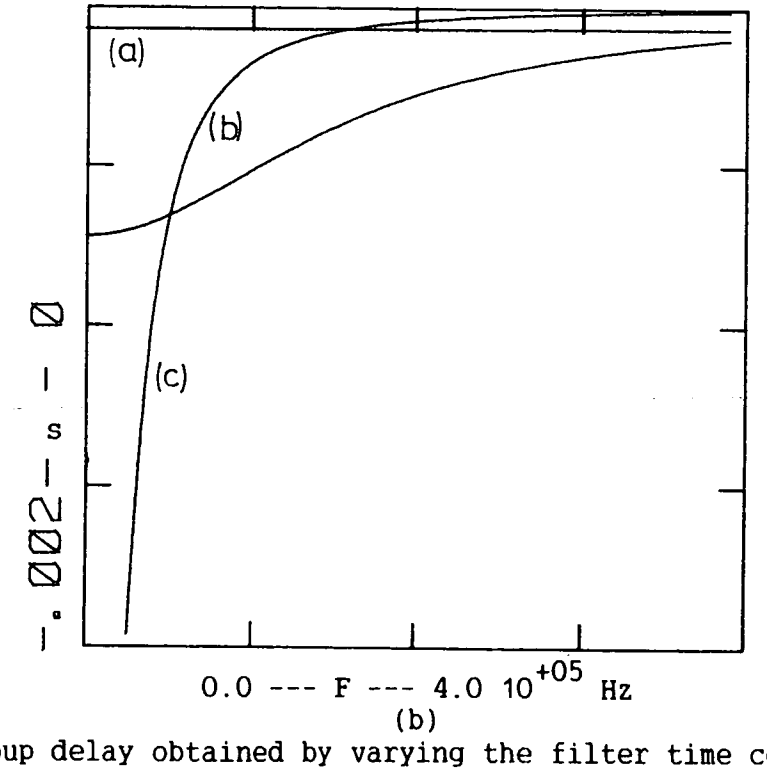
and hence the phase response linear

$$\theta \simeq -2\omega\tau \qquad (5.2.16)$$



(a)

Phase delay obtained by varying the filter time constant(τ)



(b)

Group delay obtained by varying the filter time constant

Figure 5.2.10. Characteristics of the phase delay filter with three different time constants: curve (a) : $\tau = 10^{-7}$; curve (b) $\tau = 10^{-6}$; curve (c) $\tau = 10^{-5}$

Therefore, the phase response and group delay of the network can be controlled by varying the filter time constant as shown in Figures 5.2.10(a) and (b) respectively. In Figure 5.2.10, the time constant is varied over two orders of magnitude and demonstrates that a 180^{o} phase delay can be achieved if the minimum natural frequency of the filter is an order of magnitude less that the maximum frequency of the input signal. However, for a constant time delay the natural frequency of the filter must be at least an order of magnitude greater than the maximum input frequency. For a similar order of filter, the frequency range for all pass constant time delay for this type of filter is twice that of the more widely used Bessel-Thomson(1949) approach, as discussed by Johnson(1970).

Since the circuit could only provide a maximum phase delay of 180^{o}., the full 360^{o} range was realised by inverting the original signal, delaying that also and switching the output to obtain the relevant waveform, as shown in Figure 5.2.11. The outputs of the phase delay circuits were amplified and applied to fast(10ns) comparators ,configured as zero-crossing detectors, with TTL compatible outputs. Thus the correct phase was selected with digital switching.

Finally the digital signals were applied to pulse interfacing circuits, which ensured that a constant amplitude, low level(≃ 100mV) signal was applied to the demodulators. This minimised carrier feed-through.

5.2.6) Baseband filters.

As has been stated, the baseband filters had to attenuate the harmonics generated in the demodulation process. They also added to the improvement of SNR by bandlimiting. However, their bandwidth

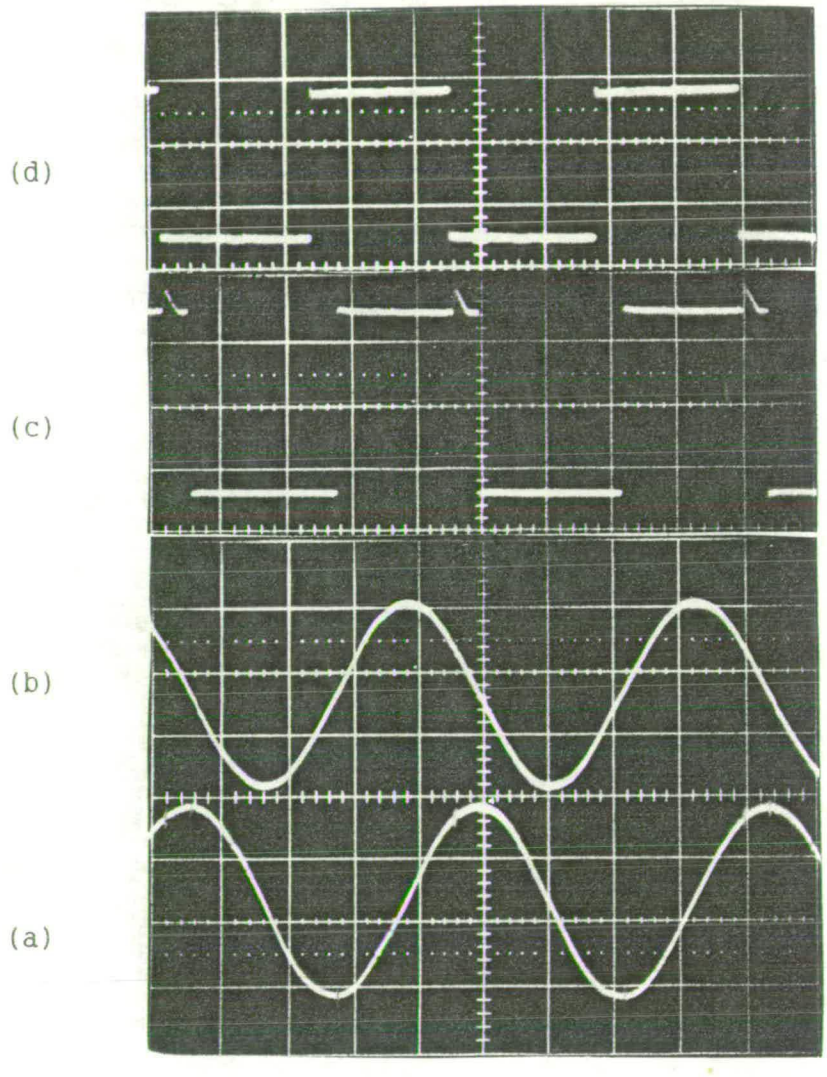


Figure 5.2.11. Phase delay circuit waveforms. All traces 0.5 μs per mm horizontal axis. (a) Input sinusoid 0.5 V per mm (b) Output sinusoid from phase delay filter 0.5 V per mm, (c) Output of comparator 100mV per mm, (d) Output from pulse shaping network 20mV per mm.

could not be so low that it altered the transfer function of the servo-system significantly and rendered it unstable. Therefore the 3dB frequency of the baseband filters had to be an order of magnitude greater than the servo-mechanism's bandwidth.

The baseband filters, again consisted of universal active filters designed using the state-variable approach. A large tuning range was required, from 1Hz to 100kHz, to permit the system to be effectively combined with a sampling system. Hence the implementation of the filters differed from those previously described, by utilising quad voltage-controlled-voltage-source operational amplifiers, which provided good parametric matching and thermal stability between all three amplifiers. In this case K_3 and K_4 of equation 5.2.6 were the time constants of the two integrators. To achieve 6 decades of tuning a combination of capacitor and resistor value switching was adopted. Two values of capacitors set a two decade range, either 1kHz to 100kHz or 1Hz to 100Hz. Variation of the resistor values allowed decade variation within these bands and finally a potentiometer allowed fine tuning. The coefficients of the all pole transfer function were set to give a second order Butterworth or maximally flat response with unity gain. The passive components were selected to give stability and accuracy of response using 2%, silvered mica, for the low value capacitor, polycarbonate capacitors for the high capacitor value and 1% metal film resistors.

5.3) Practical Design.

With the voltage measurement system comprising 12 half-size eurocards and the utility system 8 eurocards certain construction practices had to be met to ensure satisfactory performance of the

system as a whole.

5.3.1) Power Supples.

The power supplies utilised were linear ±15V for analogue circuits and +5V for digital circuits. Several low cost units were incorporated to provide the relevant unregulated supplies to each rack. All were regulated on the rack prior to distribution to the circuit cards.

Each card was decoupled with a 10μF tantalum electrolytic and a 100nF disc ceramic in parallel. Further on card regulation was provided by monolithic regulators on the analogue supplies for the fast switching components such as the demodulators. Additional decoupling was utilised were necessary in the form of in-line 10ohm resistors and 0.01μF capacitors.

5.3.2) Earthing.

Four earth returns were required in the system: a) low current analogue for virtual earths, b) high current analogue for high frequency and current buffers, c) low current digital for low frequency digital circuits such as control lines for analogue switches and d) high current digital for clocked logic circuits. All the earth lines were star connected with low impedance cable. The chassis of the rack system was earthed to the star point with a copper braid and all coaxial feed-throughs were isolated from the chassis.

5.3.3) Signal Levels.

Whenever possible high slew rate waveforms were limited to $1V_{pp}$ to prevent electromagnet coupling between signal lines. Where neces-

sary high input impedance operational amplifiers were screened with earthed metal shields. On board switching was utilised and cable runs to control panels minimised.

5.3.4) Design Aids.

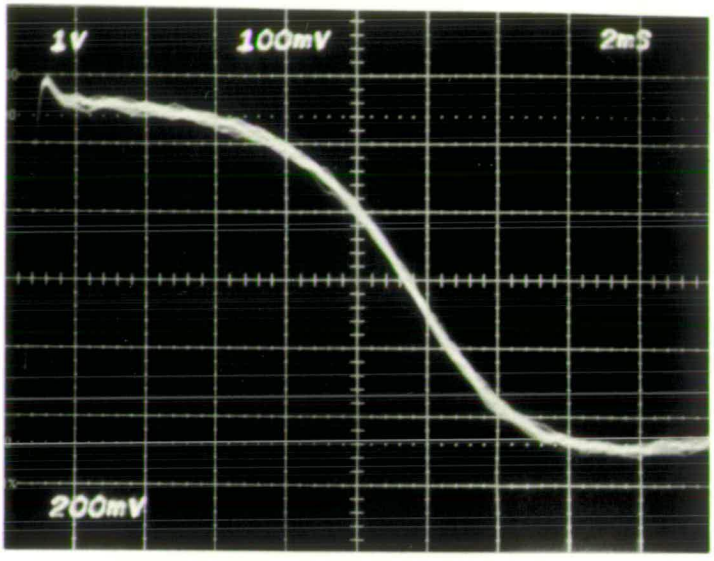
A computer program was written to generate Bode plots of system magnitude and phase response. The program also provided a Root-Loci diagram and a system analysis, which included the break frequencies of all the poles and zeros and the system gain and phase margin. Figures 5.1.10 and 5.2.10 were generated using this software. The program is discussed in detail in Appendix A.

5.4) Results.

A series of experiments was implemented to investigate the performance of the voltage measurement system. A Cambridge Instruments S150 Mk1, SEM with a 2kV accelerating voltage and a stationary primary beam was used throughout. The specimen was a gold bond pad 0.4mm by 0.1mm, utilised to eliminate any affects caused by local fields. Again the energy analyser described by Ranasinghe and Khursheed(1983) was employed.

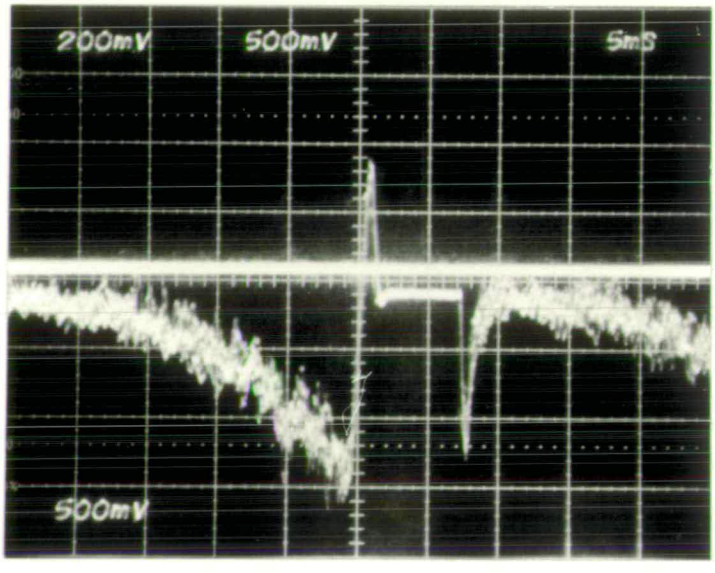
5.4.1) Derivatives.

Prior to operating the system in a closed loop mode, the derivatives of the "S"-curve were recorded and compared to the theory of chapter 4. With the specimen earthed the "S"-curve shown in Figure 5.4.1. was obtained. Since the slew rate of the retardation waveform was 663V/s, the width of the "S"-curve as a function of the retardation potential is approximately 10V.



$+V_r$ $-V_r$

Figure 5.4.1. Characteristic "S"-curve, 0.2V per mm, 0.02s per mm. S_r = 663V per s.



$-V_r$ $+V_r$

Figure 5.4.2. Second derivative of "S"-curve, 0.5V per mm, 0.05s per mm. S_r = 454V per s.

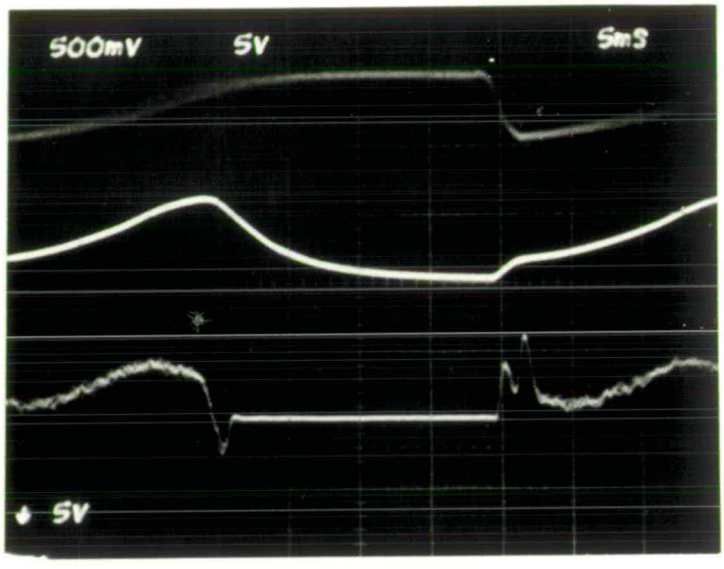
The second derivative is presented in Figure 5.4.2. and with a retardation voltage slew rate of 454V/s, the width of the linear region of the operating characteristic is 1.3V. The amplitude of the modulation sinewave was 0.5V and the horizontal trace signifies 0V. The asymmetry observed in the theoretical operating characteristic of chapter 4 is not present. However, this may be attributed to some degree by the smoothing effect of the modulation system output filters. The amplitudes of the "S"-curve and the second derivative have not been normalised and therefore, no inference concerning the relative transconductances can be made.

5.4.2) Static Voltage Measurements.

Preliminary voltage measurements were made in the open loop mode to verify that the 1:1 correspondence between the applied surface voltage and the translation of the SED in the energy domain, could be reproduced by the instrumentation system.

The "S"-curve, the SED and the operating characteristic, presented in Figure 5.4.3(a) were obtained for an earthed specimen. The modulation system output filters had cut-off frequencies of 100Hz and 1kHz for the first and second derivatives respectively, while the "S"-curve was monitored at the output of the videohead amplifier. The modulation signal is apparent on the "S"-curve about the point of inflection. The point of inflection of the "S"-curve, the peak of the SED and the zero-crossing of the operating characteristic all coincide at the same retardation voltage and therefore, at the same SE energy as predicted.

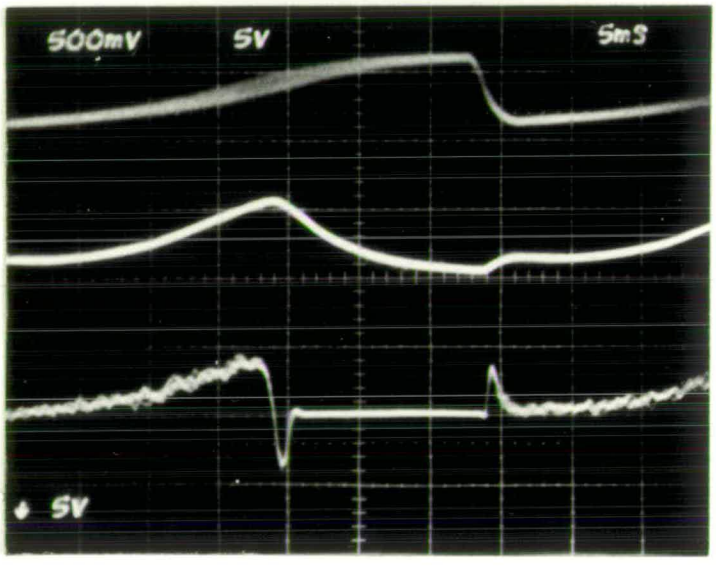
A +2.5V potential was applied to the specimen and the "S"-curve, SED and operating characteristic recorded as shown in Figure



$-V_r$ $+V_r$

(a)

V_s = 0V.



$-V_r$ $+V_r$

(b)

V_s = +2.5V.

Figure 5.4.3. DC voltage measurement.

Upper trace: "S"-curve, 0.5V per mm.
Middle trace: First derivative, 0·5 V per mm.
Lower trace: Second derivative(inverted), 0·5 V per mm.
0.5ms per mm. S_r = 428V per s.

5.4.3(b). The application of a positive surface voltage causes a translation of the three curves to the right or low SE energy regime, since the retardation voltage is scanning negative to positive. The change in surface voltage can be expressed as

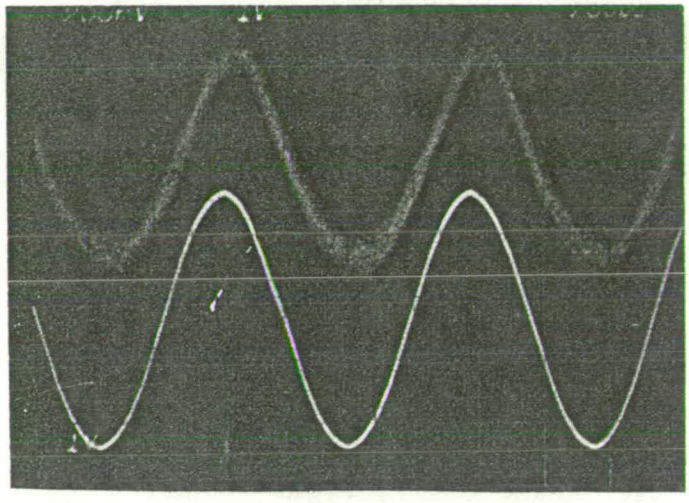
$$V_s = S_r(T_m - T_o) \qquad (5.4.1)$$

Where S_r is the slew rate of the retardation potential in V/s and T_o is the time delay, in seconds, between the commencement of the scan period and the zero-crossing of the second derivative when the specimen is earthed. T_m is the time delay between the commencement of the scan period and the zero-crossing of the translated operating characteristic, corresponding to the applied surface voltage, in seconds.

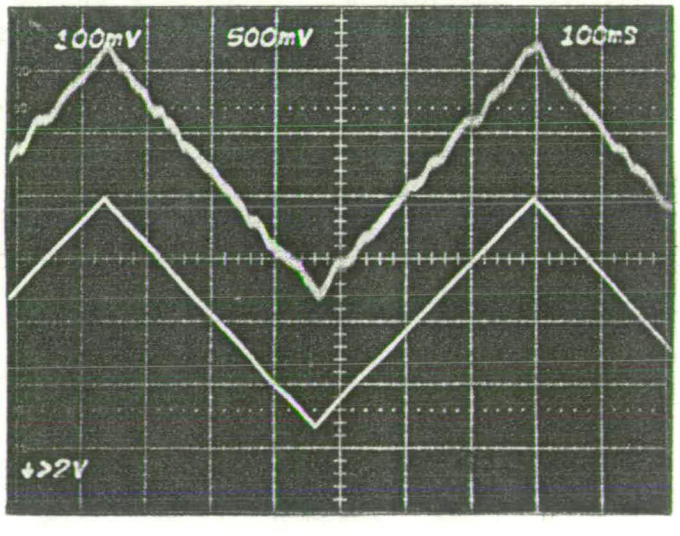
With a retardation waveform slew rate of 455V/s and an estimated time difference between the two operating characteristics in Figure 5.4.3, of 5.5ms, the required 1:1 correspondence can be seen to be achieved.

5.4.3) Dynamic Voltage Measurements.

In order to make dynamic measurements and demonstrate the linearity of the system a $8V_{pp}$ triangular waveform was applied to the specimen and the error voltage monitored, as shown in Figure 5.4.4(a). Unlike contemporary voltage measurement systems, the correction voltage applied to the retardation grid cannot be taken as the output of the system, since the modulation waveform is superimposed on it. Hence the output waveform was taken prior to the summation and output stages and since the output stages had a gain of two, there is a scalar factor of two between the recorded surface voltage



(b)
Upper trace: Measured voltage, 0·10V per mm.
Lower trace: Applied voltage, 0·1 V per mm, 20μs per mm.



(a)
Lower trace: Applied voltage, 0·2 V per mm,
Upper trace: Measured voltage, 0·1 V per mm, 10ms per mm.

Figure 5.4.4. Dynamic voltage measurement.

and the applied surface voltage, as observed in Figure 5.4.4. Although the modulation output filter bandwidth was limited to 10Hz, there is still low frequency noise in the measured waveform. This is attributed to mains hum and possibly noise coupling into the system at the loop differencing node.

To estimate the bandwidth of the measurement system, a sinusoidal waveform of $5V_{pp}$. was applied to the specimen and the error voltage monitored. The surface voltage and error signal are shown in Figure 5.4.3.(b) and demonstrate a 1:1 correspondence in phase and magnitude. The frequency of the surface voltage was increased until a phase discrepancy in the two waveforms appeared. Therefore, since the phase discrepancy was noticeable on the oscilloscope at a frequency of 1kHz, giving a phase error of about 2%, the 3dB bandwidth would occur at a frequency of about 5 times this.

However as the frequency of the surface voltage was increased the voltage range within which the loop remained in lock reduced accordingly. Also distortion became apparent in the measured waveform. These two effects can be explained in terms of the reduction of loop gain at high frequencies, giving rise to large error voltages and hence movement of the operating point onto the nonlinear region of the operating characteristic or in the worst case into the out of lock region. The measurement accuracy can be estimated from the SNR of the sinusoidal surface voltage and can be found to be about 15dB or 20% at a bandwidth of approximately 5kHz.

5.5) Discussion of Results.

It has been shown that it is possible to make quantitative voltage contrast measurements based on the principle of tracking the

most linear part of the "S"-curve.

However, these results have to be qualified in that not all the goals were fulfilled. Total automatic gain control over the photomultiplier tube was not attained and hence the lock and no-lock regions of the operating characteristic had to be determined manually. This in turn led to difficulty in achieving automatic lock of the feedback loop and therefore the feedback loop was closed manually for the measurements made.

Another difficulty encountered, occurred after the operating region had been detected and the feedback loop closed. Owing to the high forward loop gain, the error amplifier remained saturated for a time after the operating region had been detected and when the loop was closed forced the correction signal into an out of lock region of the operating characteristic.

If the system is to rendered fully automatic, the method of detecting the operating characteristic must be made independent of the emitted SE current, possibly by detecting the extrema of the 2nd derivative. However, this signal should only be used to arm the loop switch, which should only be closed when the amplified loop error corresponds to a SE energy between the peak and either of the inflection points of the SED.

CHAPTER 6.

DYNAMIC VOLTAGE CONTRAST.

The design and construction of the necessary digital hardware for the realisation of dynamic voltage contrast measurements on an i.c. will be presented. A block diagram of the digital instrumentation is presented in Figure 6.1.1. and illustrates the internal architecture of the system. In addition, the software developed to control, both these circuits and the data-acquisition system will be described. A detailed discussion of the IEEE488 General Purpose Interface Bus(GPIB) and its protocols will not be given, since these are widely known and can be found in standard texts. (IEEE, 1978, Summers, 1980, Poe and Goodwin, 1980).

6.1) Computer Interface.

The computer was equipped with an IEEE488 instrumentation bus and therefore it was decided to control the digital equipment and gather data using this bus. The interface design was based on the Fairchild 94LS488 i.c., (Fairchild Inc., 1980) which decoded all the bus protocols, handshakes and commands. The device is fabricated using bipolar low power Schottky technology and consumes 0.25W of power. The status of the interface is determined by programming a four bit mode word input to the device. In doing this, the interface can be set to Talker, Listener or both, single or dual address and low or high speed. It is capable of operating the bus handshakes at full 1MHz rate and only requires one external single phase clock. It also provides an automatic reset on power-up. A block diagram of the interface is given in Figure 6.1.2.

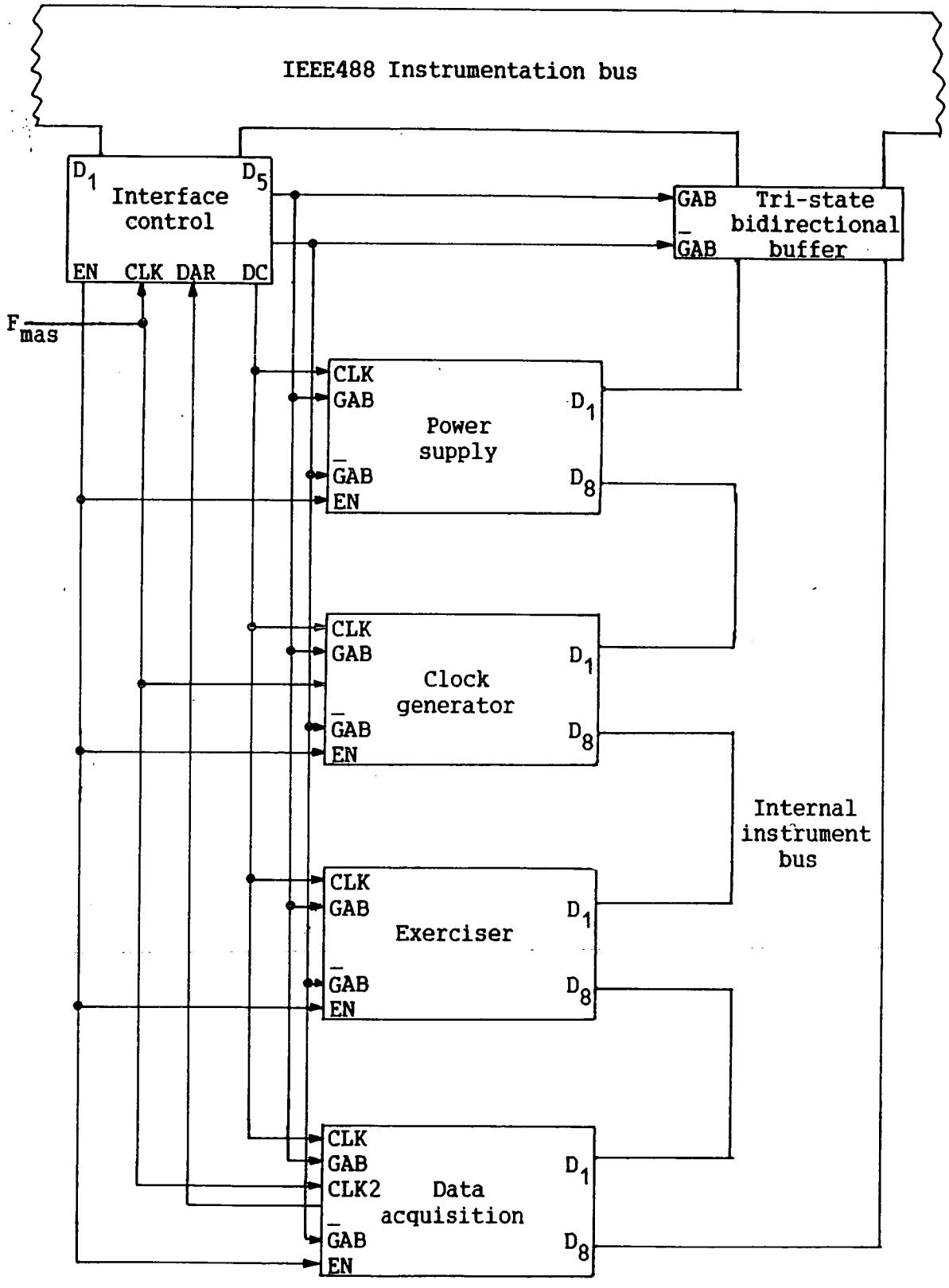


Figure 6.1.1. Architecture of digital system hardware.

6.1.1) Address Mode.

The device's operating mode was set to Talker and Listener, since data was to flow in both directions between the controller, in this case the Superbrain microcomputer (Intertec Data Systems, 1980), and the instrument. As a number of registers within the instrument were to be accessed, it was decided to operate in the Extended Address Mode, using the secondary address as an internal instrument address. Since other instruments such as the digital scan generator and specimen stage position controller were to be controlled via the bus simultaneously, each instrument needed a unique address to enable the controller to discriminate between them. This was realised as the primary address which remained peculiar to each instrument on the IEEE488 bus. However, the method adopted to implement the instrument address possessed an inherent problem in that the interface had to recognise a variable secondary address as well as a constant primary address to enable data transfer. The interface was set to the extended address mode by selecting the correct code for the four bit address mode byte, (M_0 - M_3 in Figure 6.1.1) by means of four single pole double throw DIL switchs.

The instrument addresses are presented to the interface as two 5 bit words(D_1-D_5) on the data bus when the $\overline{ATN}$ line of the GPIB is low. The device compares these bytes with the primary and secondary address bit internally and if both comparisons are true , enables the data transfer, the direction of which is dictated by the controller via the GPIB handshake lines: not ready for data($\overline{NRFD}$), data available($\overline{DAV}$) and no data accepted($\overline{NDAC}$). As shown on the timing diagram presented in Figure 6.1.3, the interface device provided an address select line $\overline{ASEL}$ to allow primary and secondary addresses to

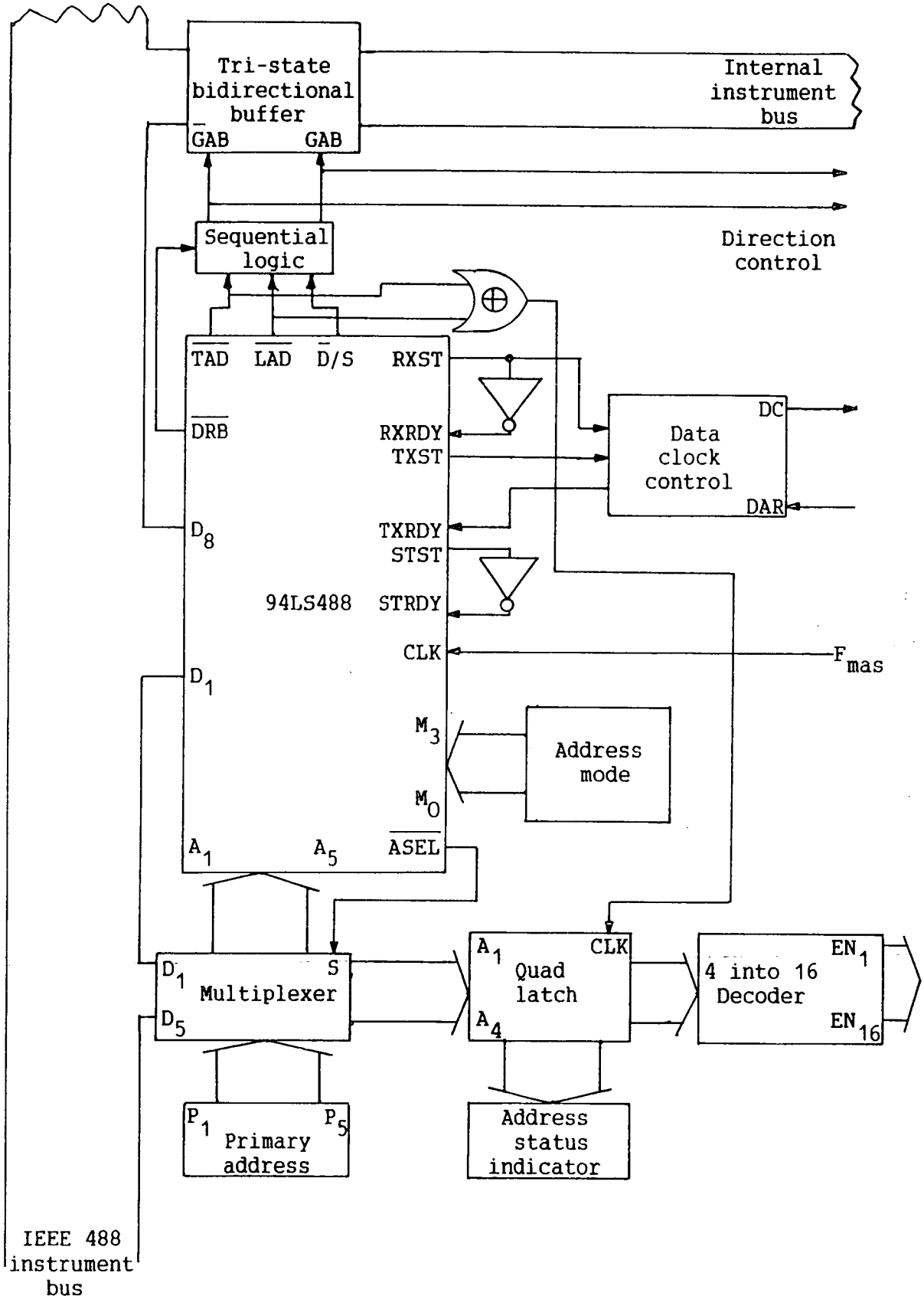


Figure 6.1.2. IEEE 488 bus interface and address control.

be multiplexed into its address input A_1-A_5. To achieve a variable secondary address, the relevant five data lines D_1-D_5 of the GPIB were buffered and fed into the second channel of the multiplexer. Hence, the problem associated with a variable secondary address was solved.

The secondary address had to be clocked into a latch during the data valid time on the instrumentation bus to enable the relevant internal instrument register. This was achieved by generating a clock pulse by exclusive OR-ing the device's Listen addressed($\overline{\text{LAD}}$) and Talker Addressed($\overline{\text{TAD}}$) outputs since one of these always changed state when the interface was enabled, as shown by the timing diagram of Figure 6.1.3. After the four MSBs of the secondary address had been loaded into the nibble(4 bit) latch, they were decoded into 16 address lines. The LSB of the secondary address could be used for future expansion. Figure 6.1.4. tabulates the 16 internal instrument addresses. The nomenclature will be explained as the discussion proceeds. The addresses are split into data and control lines.

Summarising the interface enable routine, the bus controller initially pulls $\overline{\text{ATN}}$ low and places the primary address on the data bus. The 94LS488 interface device pulls $\overline{\text{ASEL}}$ low, to enable the primary address input of the multiplexer. Subsequently, all instruments on the bus carry out an internal comparison with their own primary address. The 94LS488 interface, takes $\overline{\text{ASEL}}$ high and the process is repeated for the secondary address, which in this case, reduces to a comparison with itself via the multiplexer. If both comparisons are true, the interface device decodes the direction of data transfer from the GPIB handshake lines and sets the appropriate output flag($\overline{\text{TAD}}$ or $\overline{\text{LAD}}$)low. The secondary address is then clocked into a

latch on the falling edge of this data direction flag, after which the byte is decoded by the quad decoder and the relevant register enabled. Finally, the controller takes $\overline{ATN}$ high to enable data transfer. The GPIB handshake operates throughout this set-up procedure, as shown by the timing diagram in Figure 6.1.3. However, this has been omitted in the discussion in the interest of clarity.

6.1.2 Direction Control and Data Transfer.

After the primary and secondary addresses have been recognised by the device, the Talk Addressed ($\overline{TAD}$) and Listen Addressed ($\overline{LAD}$) direction latches are set to the relevant levels. These were decoded using combinational logic to set the tri-state octal transceiver, used for interfacing the instrument to the bus, to the correct state, as shown in Figure 6.1.1. The combination logic included Data/$\overline{Status}$ (D/$\overline{S}$) and Drive bus ($\overline{DRB}$) outputs, since these were utilised to set the interface transceiver into the high impedance state whenever a status or command byte was placed onto the bus by the interface i.e. during a parallel poll sequence or an untalk command. In addition to setting the interface transceiver to the correct state, any addressed internal transceiver also had to be programmed accordingly. This was achieved by enabling the data direction command signals(GAB, $\overline{GAB}$) with the internal instrument address enable.

Since no status byte information was to be used the interrupt request was tied to line 7 of the data bus, which concurred with the device specifications for answering a parallel poll procedure.

The interface device transferred the bus handshake to the instrument side of the interface in the form of three sets of two handshake lines; one each for the device Talking(TXST, TXRDY),

Listening(RXRDY, RXST) and Status(STST, STRDY) modes. For fast operation without a handshake, the strobe output(ST), of each handshake was inverted and fed into the ready(RDY) input of each of the three channels. Whenever a strobe line was pulled high and therefore the ready line pushed low, there would be a delay, determined by the IEEE488 bus handshake and the master clock frequency, of at least one master clock period before the interface would acknowledge the request by pulling the strobe line low again. Hence, the data was clocked into or out of the internal registers with the rising edge of strobe pulse. Since the data transfer was asynchronous, the instrument registers had to possess a set-up time much less than one master clock cycle, otherwise not only would the data be corrupted but the controller would be unaware of this fact. Therefore all internal registers were 74LS components for speed and compatibility.

Since the instrument was required to both talk and listen, the two clocks one each from the Talk(TXST) and Listen(RXST) channels were multiplexed onto a system data clock(DC) by using the direction control signals(GAB, $\overline{GAB}$). A handshake was necessary for the data-acquisition system, since the ADC conversion period was much longer than one master clock cycle and therefore , the interface handshake was inhibited until the conversion was complete. This was implemented by "Anding" the transmit ready signal(TXRDY) with a data ready signal(DAR) issued by the ADC when a conversion had been completed and the data placed in a latch. The DAR line was taken low on the rising edge of the data clock, when the ADC was addressed, and initiated the conversion. Thereafter the talker strobe was held low until the conversion was finished and DAR went high.

6.2) Exerciser.

As stated in chapter 4, the philosophy of the design was that the exerciser would execute exclusively functional dynamic testing, leaving the electron beam test system to perform any parametric, AC or DC testing required. Therefore, the exerciser hand to be able to detect functional faults on a 1's and 0's basis. The microcomputer would be able to control the rate at which the device under test would be tested, the cycle time and the cycle length of the test.

6.2.1) Exerciser Architecture.

As shown in Figure 6.2.1, the exerciser was divided into 5 memory segments, a 10 bit programmable counter, output buffer stages and clock control logic. Since the majority of i.c.s to be tested were signal processing devices with a bus orientated architecture, a minimum of 8 bits wide, it was decided to subdivide the memory segments into multiples of 4. Thus, there were 2, 8 bit bytes and 4, 4 bit nibbles of memory giving 6 memory partitions, and 32 data channels.

Although all the data lines should have been independently bidirectional for maximum flexibility, it was felt that owing to the architecture of the i.c.'s to be tested, the economics of this flexibility could not be justified. Therefore, a compromise was made, in that the memory was organised into direction programmable bytes and nibbles. This allowed the number of input and output lines to be varied to suit a particular device.

This was implemented by storing an output direction bit pattern in the remaining memory segment(MEMSEG5) and clocking this into the

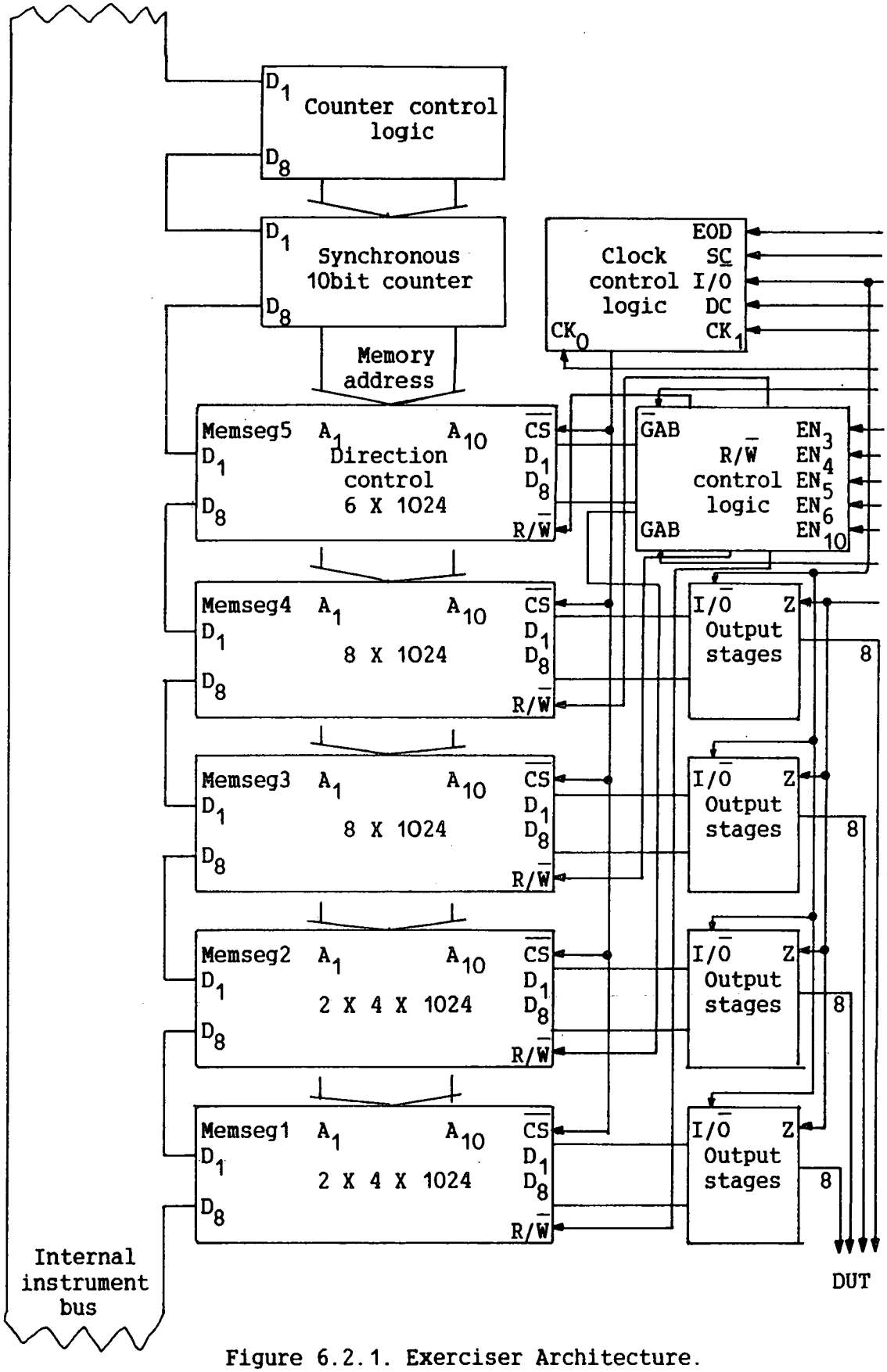


Figure 6.2.1. Exerciser Architecture.

output buffers in synchronism with the system clock. Hence the data direction for each data channel could be altered cycle by cycle. One byte wide memory segment was divided into 6 bits for direction control of the other 4 memory segments and 2 bits for special chip command functions such as reset, enable and hold, giving a maximum of 34 data lines. Since the maximum pin-out of the devices to be tested at this stage was 40 and 3 of these would be committed to power supply, biasing and grounding and a further 2 for clock inputs, only 35 data lines were necessary. Hence, one chip command line could be made bidirectional if necessary.

The control logic(Figure 6.2.1) gave over-riding command to the microcomputer and automatically set the output stages of a memory segment to high impedance, if addressed to either talk or listen. The functions of bad-level detection and automatic hardware comparison were both sacrificed for reducing the cost. The input/output comparison could be done by software in the microcomputer, while bad-level detection could be executed by the electron beam test system. Since the exerciser was to be used in conjunction with the electron beam test system as a diagnostic test instrument, it was felt that maximum clock rate testing would not be necessary and therefore, NMOS static RAMs(access time 150ns) were used to construct the instrument's memory. The exerciser's memory was made 1Kbit in depth, but this could be easily extended.

6.2.2) Counter Control.

A 10 bit synchronous binary counter, was utilised to generate the memory address word, and as such had to provide flexible operation to fulfill all the requirements of the envisaged modes of opera-

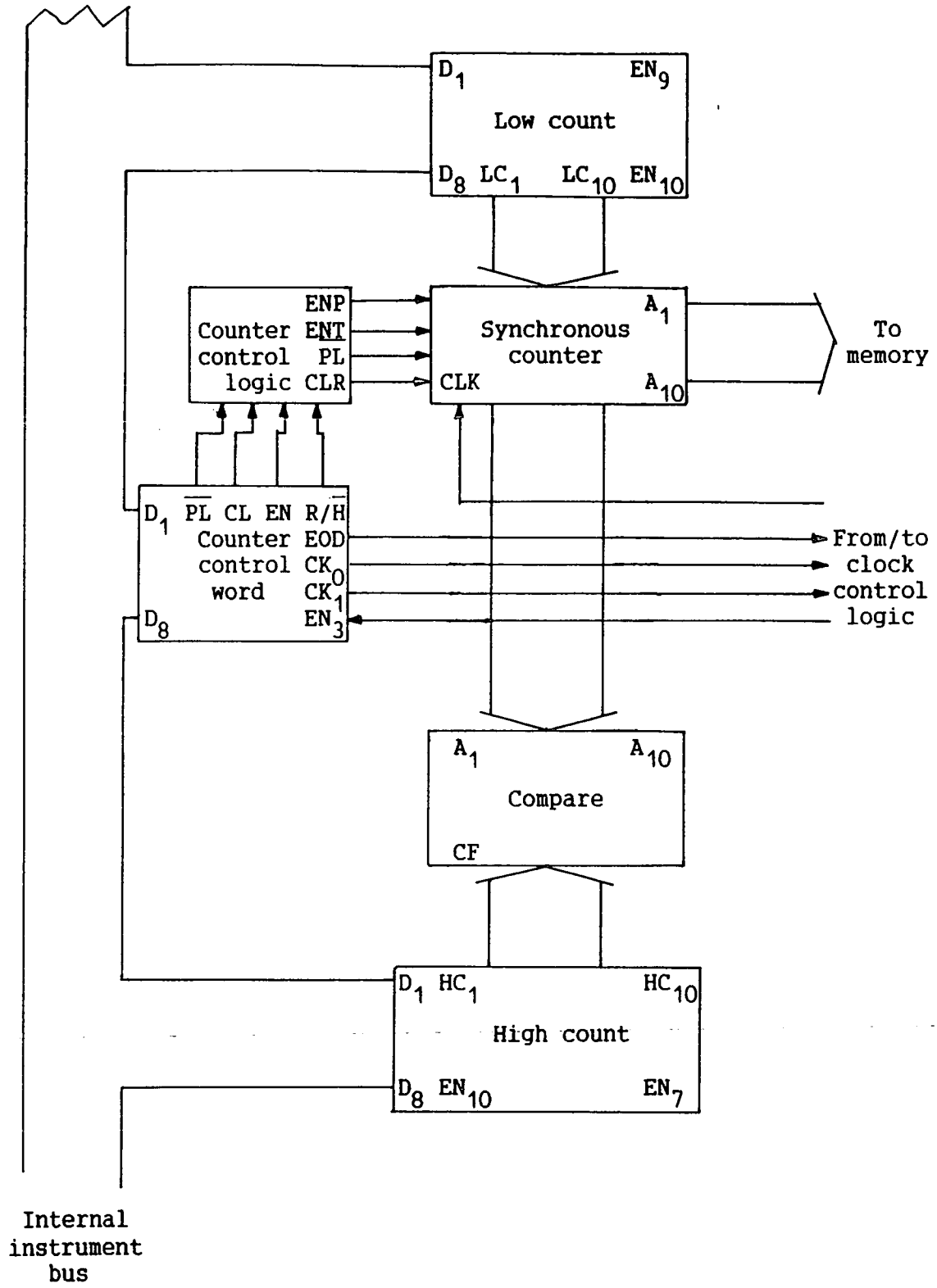


Figure 6.2.2. Memory address logic.

tion.

It had to cycle between any two memory location addresses in one direction only and also preset and reset to any number. As shown in the block diagram of the counter in Figure 6.2.2, this was achieved by utilising programmable maximum and minimum count latches and comparing the counter output word to the stored maximum count word. When true, a compare flag(CF) was issued and a reset to the minimum count word was executed, depending on the mode of the counter.

By placing both combinational and sequential logic within the counter feedback path, several modes of operation were possible. These included a STEP mode, a RUN mode, a CYCLE mode and a RESET mode. The STEP mode allows the operator to manually step through a sequence of inputs by applying a single pulse to the counter clock input(CP) in synchronism with the input/$\overline{\text{output}}$ clock(I/$\overline{\text{O}}$). A CYCLE mode was included to enable a single pass test to be carried out, in that the counter will count from the minimum count word to the maximum count word, once and issue a count complete flag. The count complete flag could be used to initiate an interrupt request, resulting in an automatic read of the memory.

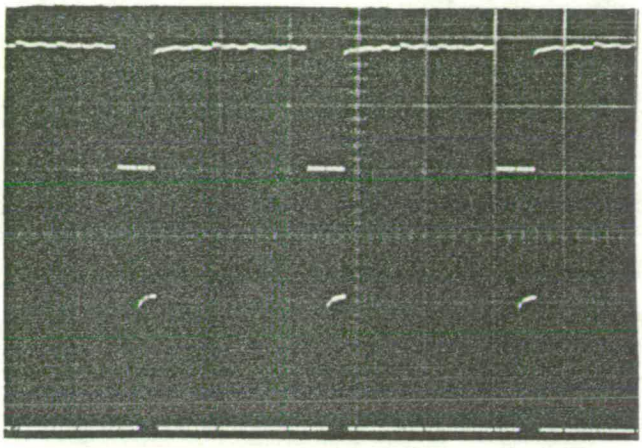
The RUN mode puts the counter into a continuous count from the minimum count word to the maximum count word until instructed to stop. i.e. it inverts the compare flag and applies it to the parallel load($\overline{\text{PL}}$) input of the three binary counters. This mode of operation is essential if either logic state mapping or sampling is to be implemented by the electron beam test system. Finally, the RESET mode implements an effective parallel load of the minimum count word.

Both the minimum and maximum count word and the operation mode are programmable via the GPIB bus, as illustrated by Figure 6.2.2. The counter control word consists of several flags which, when selected in certain combinations produce the given set of commands, as will explained in the software section. All of the command modes discussed were designed to operate in synchronism with whichever clock was selected for the counter by the clock control logic.

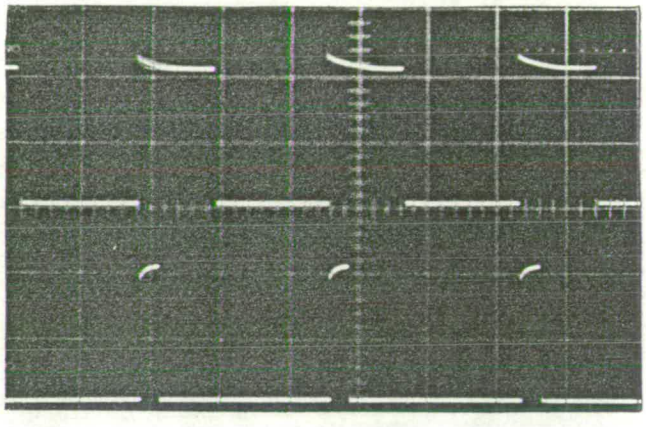
6.2.3) Clock Control.

The clock control logic provided the clock signals for the exerciser memory, the counter and the output stages, all in synchronism. In addition, these clocks had either to be synchronised with the asynchronous interface handshake or the input/$\overline{\text{output}}$ clock(I/$\overline{\text{O}}$), which was synchronised to the θ_1 phase of the two phase clock generator. The clock control was essentially a multiplexer which issued the system clock, as the chip select and counter clock and the I/$\overline{\text{O}}$ as the output stage clock, to the exerciser unless addressed by the microcomputer, in which case the data clock(DC), was issued as the chip select signal and $\overline{\text{DC}}$ as the counter clock. Since the output stages were disabled throughout the period of communication between the exerciser and the microcomputer no output stage clock was necessary.

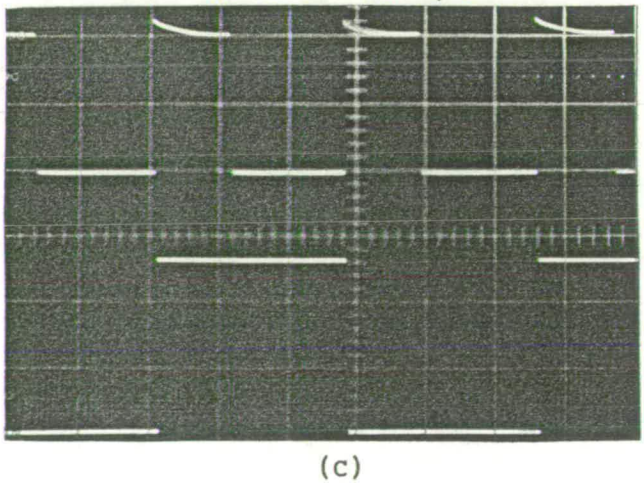
The memory timing cycle for the microcomputer accessing the exerciser commenced with the counter being initialised to a given address. The counter had to be disabled during this write sequence since the data clock was delayed by the multiplexer long enough for the clock to increment on this instruction. Thus instead of being initialised to an address, N, it was initialised to N+1. Therefore, a



(a)
Upper trace: SC, counter is incremented on rising edge.
Lower trace: I/O, output stage buffers latch on rising edge.



(b)
Upper trace: θ_1 of two phase clock generator.
Lower trace: I/O.



(c)
Upper trace: θ_1.
Lower trace: Output from LSB when a binary count is stored in the exerciser.

Figure 6.2.3. Exerciser clock waveforms.

hardwire disable was included to disable the counter, while the counter control latch was being addressed.

The data clock was inverted before it was connected to the counter but was applied directly to the memory chip select line. After the counter initialised, the memory was remained enabled and since valid data must be placed on the bus before the data clock is pulled high, the rising edge of the data clock, clocked the data into the memory. Subsequently, the counter was incremented and the memory enabled on the falling edge of the data clock, when the data became invalid. When writing the last memory address in a data transfer sequence, an End Of Data ($\overline{EOD}$) flag was necessary to disable the counter on the last count and prevent the data clock going false. If this was not implemented, the memory would return to its active quiescent state and capture the last data byte on the bus before the remainder of the handshake was executed, in the last plus one memory address. Hence, the data in this memory location would be corrupted. Again this was implemented in hardware, with the MSB of the counter control word being the $\overline{EOD}$ flag.

The memory timing cycle for the system clock was similar and is shown in Figure 6.2.3.(a) The system clock(SC) was applied to both the memory chip select and the counter. Therefore, the memory was enabled one master clock period before the $I/\overline{O}$ clock enabled the output stages, which clocked the contents of the memory into the output latches or vice-versa. One master clock period after the rising edge of the $I/\overline{O}$ clock, the rising edge of SC increments the counter. As shown in Figure 6.2.2(b) the $I/\overline{O}$ clock is in synchronism with θ_1 of the two phase clock generator which is applied to the DUT.

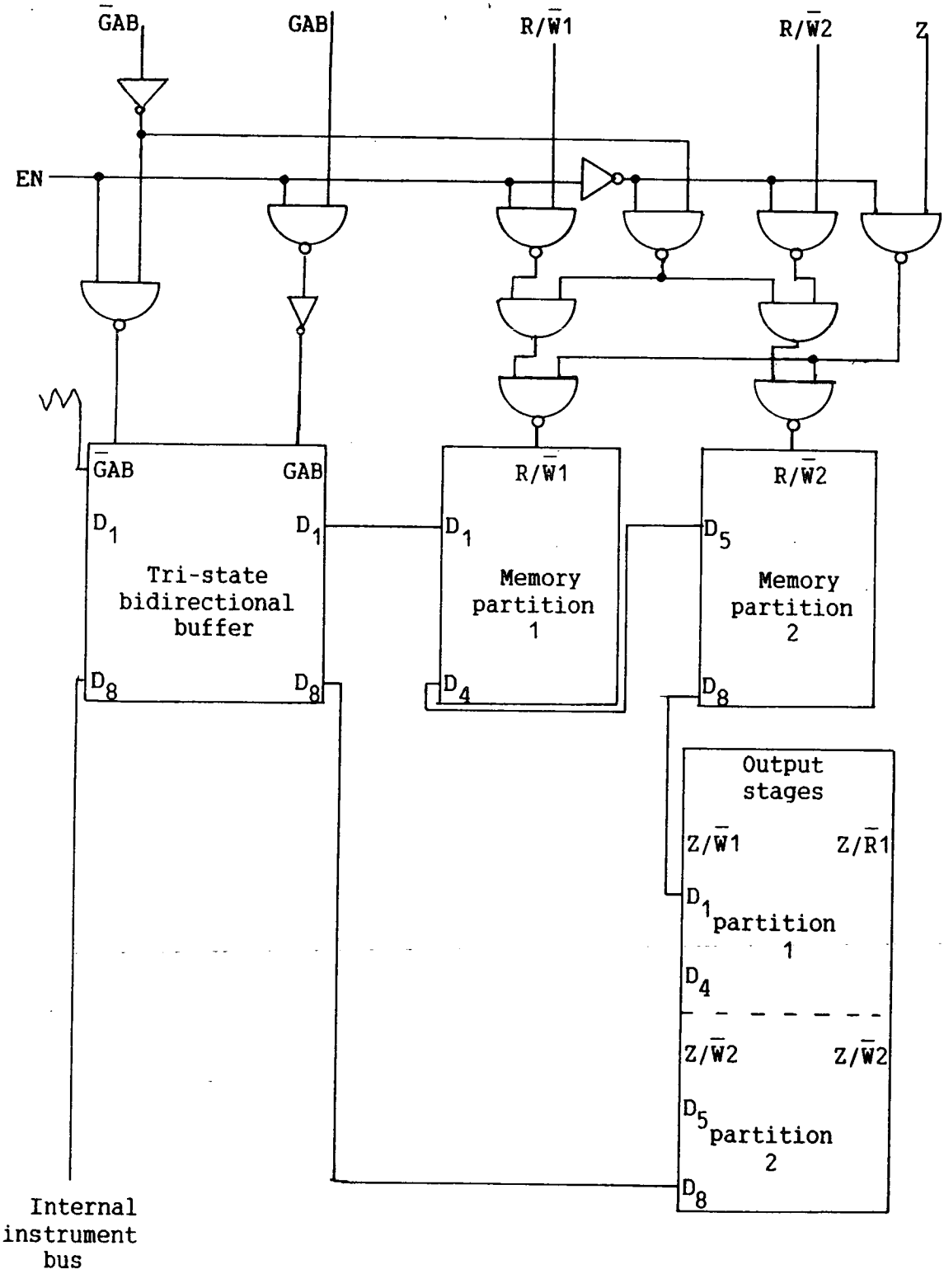


Figure 6.2.4. R/$\overline{W}$ control logic.

Both these timing cycles were designed to minimise the latent memory corruption by, in the former case inhibiting the memory immediately after reading or writing and incrementing the counter and in the latter by minimising the percentage of system clock period that the memory was enabled.

6.2.4) Read/$\overline{\text{write}}$ Control.

The read/$\overline{\text{write}}$ control logic was necessary to supervise the flow of data to and from the exerciser memory and either the DUT or the microcomputer. When communicating with the microcomputer, the direction of data flow was decoded from the bus transceiver control lines(GAB,$\overline{\text{GAB}}$) and applied to the memory read/$\overline{\text{write}}$ input. Simultaneously the output stages were placed in a high impedance state and therefore, isolated from both the exerciser memory and the DUT.

For communication between the DUT and the exerciser memory, the direction control bits contained in MEMSEG5 were issued to the relevant memory partition as well as the output stages. A schematic of the control logic is shown in Figure 6.2.4. and illustrates that the control logic is essentially a multiplexer between the direction control vector of the microcomputer and MEMSEG5. The quiescent state of the memory was $\overline{\text{write}}$ and it could only read when forced to do so. In this way the likeliehood of memory corruption was minimised.

6.2.5) Output stages.

The output stages of the exerciser had to possess the capability of driving NMOS 5V logic over about 0.5m of cable and therefore, the outputs had to be not only buffered but converted from TTL logic levels to NMOS levels. However, the incoming signals did not have to be

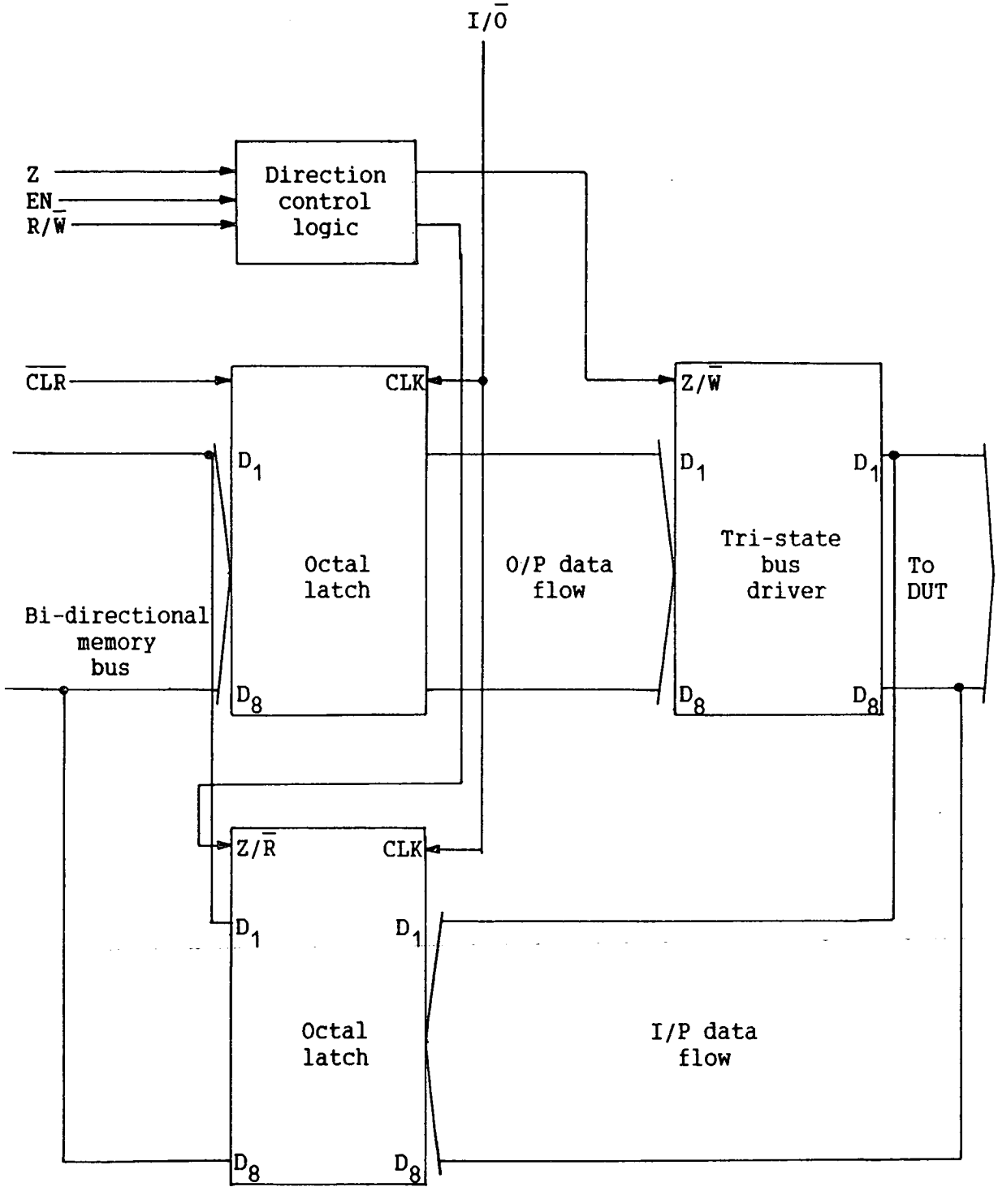


Figure 6.2.5. Output stages.

reconverted to TTL levels, since the memory i.c.s were NMOS compatible. This interfacing was facilitated by using high voltage output TTL devices, such as the 74LS364, octal latch to drive a 74C series octal bus buffer. This selection of output buffers ensured easy upgrading when HC and HCT series buffers became more widely available, which would increase the drive capability of the output stages and hence the maximum frequency of operation.

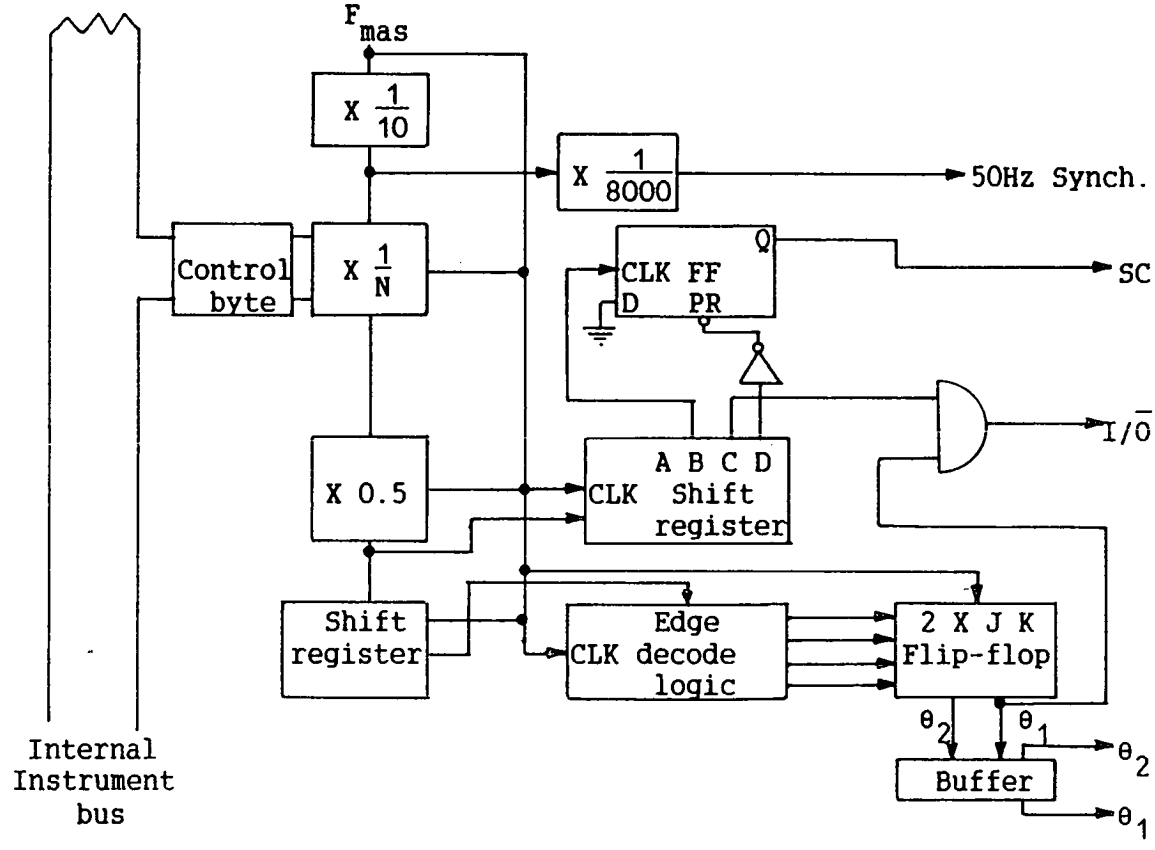
The output stage consisted of two channels, as shown in Figure 6.2.5. Both input and output data was latched on the rising edge of the system clock and the data direction control was derived from the $R/\overline{W}$ bit contained in MEMSEG5. The latches could be cleared by taking bit 4 of the clock control word low, which would also set the outputs to high impedance. When the instrument memory was being accessed by the controller, the outputs were automatically set to high impedance.

6.3) Utilities.

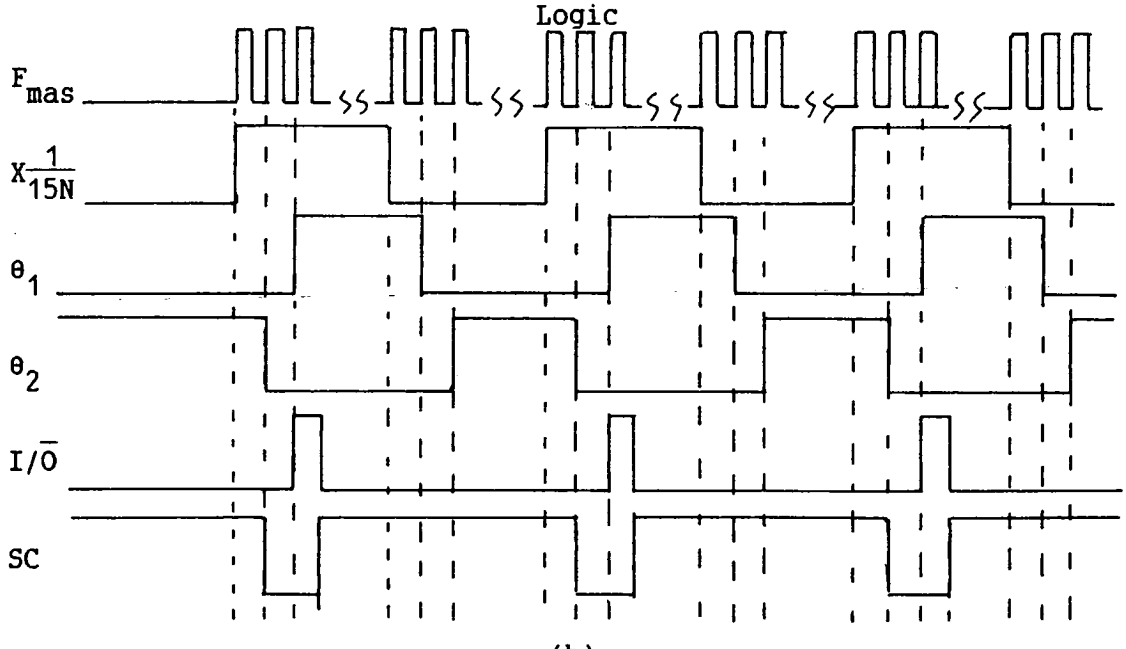
The utility sub-system comprised all the miscellaneous functions which were necessary to make the DUT function such as the clock generator, the power supplies and the exerciser/device interface.

6.3.1) Clock Generator.

The clock generator was necessary to generate the two phase, non-overlapping clock required for the operation of 5V dynamic NMOS logic, as discussed in chapter 1. It was decided to derive the clock from the analogue modulation frequency(200kHz), although any external source could easily be substituted. As shown in the block diagram of the clock generator circuitry presented in Figure 6.3.1(a), the master clock frequency(F_{mas}) was divided by a minimum of 10 before being



(a)
Logic



(b)
Timing diagram.

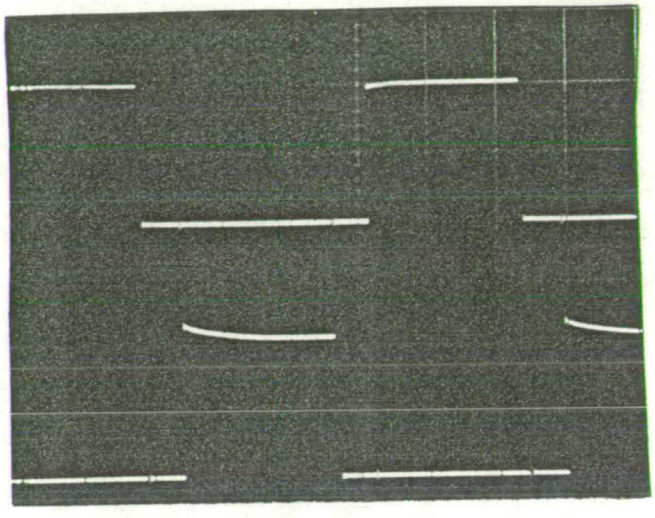
Figure 6.3.1. Clock generator circuit and timing waveforms.

applied to the DUT. The maximum device operating frequency of 20kHz was determined by the highest frequency component the servo-system could theoretically track. However, if a sampling system was employed the modulation frequency could be phase locked to it and the maximum operating frequency determined by the memory access time(150ns). The scalar division of 10 was realised in two stages: a divide by 5 followed by a divide by 2, interspersed with a divide by N. The divide by N was implemented with an 8 bit programmable counter which allowed the system frequency to be varied according to the expression:

$$F_{sys} = \frac{F_{mas}}{10.N} \qquad 1 \leq N \leq 255 \qquad (6.3.1)$$

Where F_{sys} is the system frequency and N is determined by the contents of an octal latch. All the counters were binary and synchronous with the master clock. Hence with a master clock frequency of 200kHz, the input/$\overline{\text{output}}$ clock frequency could range from 80Hz to 20kHz.

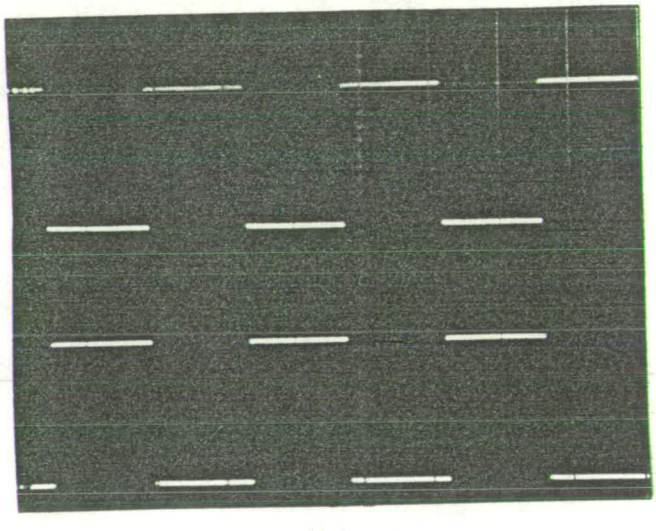
The divide by 2 counter was incorporated to generate a square wave with a 50% M:S ratio, which was required to produce the two phase clock. The resultant square wave of the preceding operation was applied to a 5 bit serial to parallel shift register and the rising and falling edges detected and used to force the inputs of two J-K synchronous flip-flops in such a way as to ensure that the two $\overline{Q}$ outputs were non-overlapping by one master clock period, as illustrated in Figure 6.3.1(b). The Q outputs were then buffered with inverter memory drivers to give the desired two phase clocks. Figures 6.3.2(a) and (b) show the two phase output clocks at high and low frequencies respectively. In Figure 6.3.2(a) the simultaneous "off" time is an appreciable fraction of the clock period, but not



(a)

$$F_{I/\overline{O}} \simeq \frac{F_{mas}}{10}$$

Upper trace: θ_1
Lower trace: θ_2



(b)

$$F_{I/\overline{O}} \ll F_{mas}$$

Upper trace: θ_1
Lower trace: θ_2

Figure 6.3.2. Two phase clock waveforms.

long enough to corrupt the data within the device since its period is much smaller than that of the time constant of a circuit node. In (b) the "off" time delay is not an appreciable fraction of the clock period and therefore, is not evident in the traces.

The $I/\overline{O}$ clock had to be generated in synchronism with the two phase clock to ensure that the data was presented to the inputs of the DUT in synchronism with the two phase clock. This was implemented by tapping the clock prior to the divide by two counter and applying it to a second shift register. From the timing diagram of Figure 6.3.1(b), it can be seen that after two clock delays, the $I/\overline{O}$ clock is in synchronism with the θ_1 phase of the two phase clock. The $I/\overline{O}$ clock was generated by "Anding" these two signals. The system clock signal was achieved by firstly taking the outputs from the second shift register, one delay prior and one delay after the $I/\overline{O}$ signal and applying them to the clock and preset inputs of a D-type flip-flop, respectively. The input to the flip-flop was tied to earth. The $\overline{Q}$ output was taken as the system clock.

In addition to generating the two phase clock, the 50Hz synchronising signal had to be derived from the master clock to enable mains locking of the master oscillator. This was achieved by tapping off the master clock divided by 5 and further dividing by 10, 10 and 8, making a total of 4000, giving a 50Hz output frequency. Again this was achieved by synchronous binary counters.

6.3.2) Power Supplies.

Two power supplies were necessary in order to enable an NMOS 5V integrated circuit function. These are a +5V ±5% and a -2.5V ±5% for substrate biasing, although the latter may be provided via an on-chip

current pump in some instances. In order to provide some flexibility and allow power supply dependent failures to be analysed, both power supplies were made programmable. Using a normal linear power supply regulator configuration with voltage feedback and a transistor pass element, the voltage reference of the voltage loop was made programmable, using a DAC. A current limit was provided in the output stages and a current monitor which allowed the i.c.'s power consumption to be estimated, since the current drain was estimated by measuring the differential voltage across the current limiting accurate resistor. This differential voltage was filtered using a lowpass balanced Butterworth filter and converted to a single-ended output.

6.4) Data-Acquisition System.

The data-acquisition system was used for logging the outputs of the voltage measurement system in both the open and closed loop modes. In addition, it could be used to monitor circuit functions such as power supply voltage and current consumption. Finally, it gave the system the ability to monitor the outputs of the DUT directly over an entire clock period, unlike the exerciser which only records outputs at clock transitions. It is based on a 16 channel analogue multiplexer and an 8 bit ADC, as shown in Figure 6.4.1.

6.4.1.) Analogue-to-Digital Converter.

Owing to financial constraints during the design stages, a monolithic ADC was not utilised. Instead a hybrid converter, comprising several standard i.c. parts and based on the popular successive approximation technique, as explained by Taub and Schilling(1977), was employed. This method accrued the advantage that the

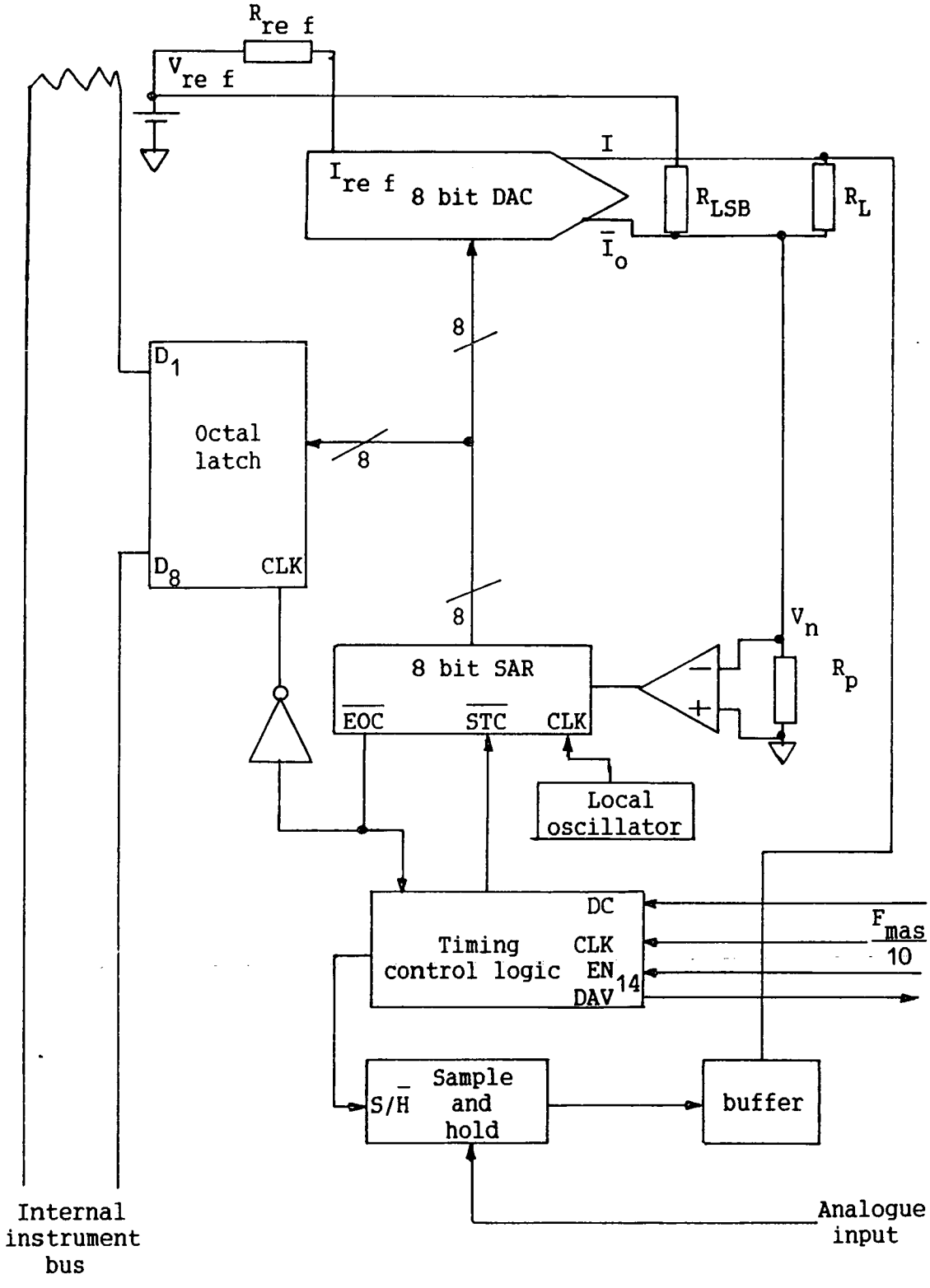


Figure 6.4.1. Analogue-to-digital converter.

achievable sampling rate was much greater than that of a cost comparable monolithic ADC. Although only eight bit accuracy was used to make the converter compatible with the instrument bus, this was felt adequate when used in conjunction with the variable input gain range. The system was incorporated a discrete feedback loop consisting of a DAC, a successive approximation register(SAR) and a comparator.

A conversion commences when the successive approximation register puts the MSB high, giving an output voltage from the DAC equal to half the full scale value. The DAC output is compared with the input voltage by means of a passive summing node and the resultant value is applied to a zero crossing configured comparator. If V_{in} is greater than V_{dac} a '1' is entered into the successive approximation register; if not, a '0' is entered. This process is repeated for eight cycles of the conversion clock, corresponding to eight shifts in the SAR, implementing a binary search. When the conversion is complete the word in the successive approximation register is a digital approximation to the analogue input.

The voltage comparison was actually implemented using currents after Soderquiste and Schoeff(1981). Since by using a current mode DAC, the usual operational amplifier could be omitted, permitting faster conversion times. The equation for the node voltage is

$$V_n = \left[\frac{R_p R_1}{R_1+R_p}\right]\left[\frac{V_{in}}{R_1} + \frac{V_{ref}}{R_{lsb}} - \bar{I}_o\right] \qquad (6.4.1)$$

The pull-down resistor, R_p reduces the voltage swing at the input to the comparator to ±1.67V maximum and therefore to within the comparator's common mode voltage range. It also reduces the load impedance of the DAC which results in a faster settling time. The

minimum voltage swing is ±6.5mV and therefore, the comparator must be chosen to have as small a propagation delay as possible with this amount of overdrive. The resistor R_{lsb} is introduced to compensate for the comparator's offset and to provide a 0.5 LSB of current into the node as required to reduce the quantisation error to a maximum of 0.5LSB.

The minimum conversion time possible is determined by the delay around the feedback loop for an accuracy of 0.5 of a LSB. Hence the conversion time(T_{con}) consists of

$$T_{con} = T_{dac} + T_{comp} + T_{sar} \qquad (6.4.2)$$

Where T_{dac} is the settling time of the DAC, T_{comp} is the propagation delay of the comparator and T_{sar} is the set-up time of the SAR. The settling time of a DAC08 converter is $6.2R_pC_o$ where C_o is the output capacitance of the DAC and, in this case, is specified as 15pF. This is the time taken for the output current to settle to ±0.5LSB. Hence for this configuration the settling time is 90ns. The comparator used was the LM361 with a propagation delay off 12ns with 5mV overdrive and the SAR was fabricated in low power Schottky with a set-up time of 10ns, giving a theoretical minimum conversion time of about 112ns per bit and therefore a minimum conversion clock frequency of ~9MHz. The conversion clock frequency utilised was 2MHz, which gave a factor of 5 margin.

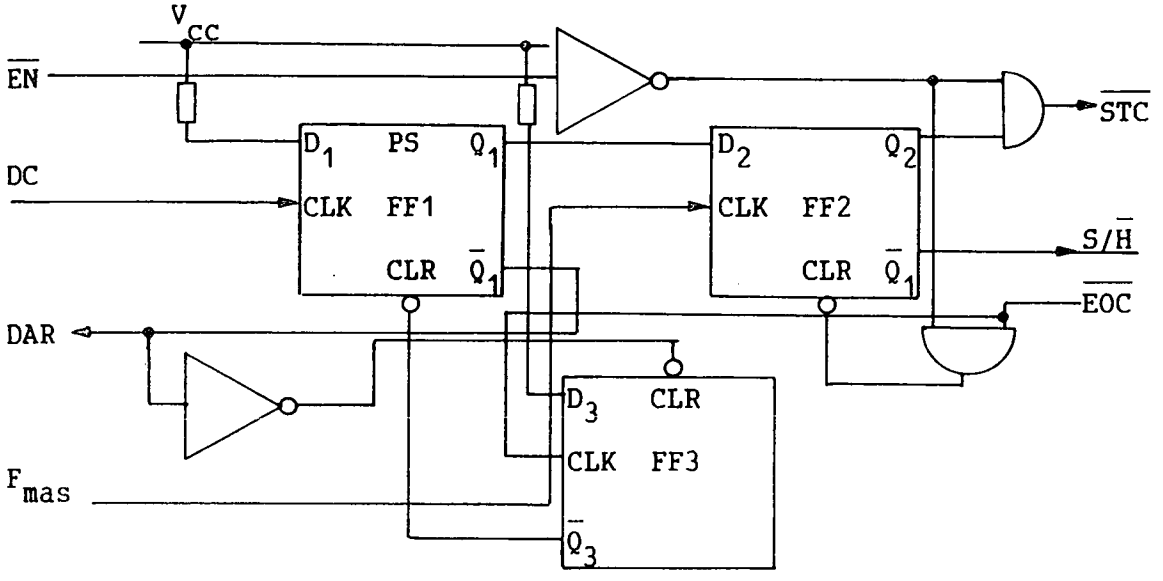
The input signal was required to provide a constant current of 3.6mA while the DAC output changed at a frequency of 2MHz. In addition, the input voltage had to be maintained at a constant level over the entire conversion period(4μs) to ensure optimum accuracy. This was achieved by using a monolithic sample and hold in cascade with a

wideband low output impedance unity gain follower(LM310). The input signal was sampled synchronously at the master clock frequency (200kHz) giving a sampling rate a factor of ten above the maximum device clock frequency. The sample and hold had a droop rate of 0.02V/s when used with a 1nF low leakage capacitor(NPO monolithic ceramic) which would result in an error of 8nV over the conversion period. i.e. it is insignificant since $V_{droop} \ll 0.5LSB$. The settling time was 0.8μs to within 1mV which corresponded to 4LSBs

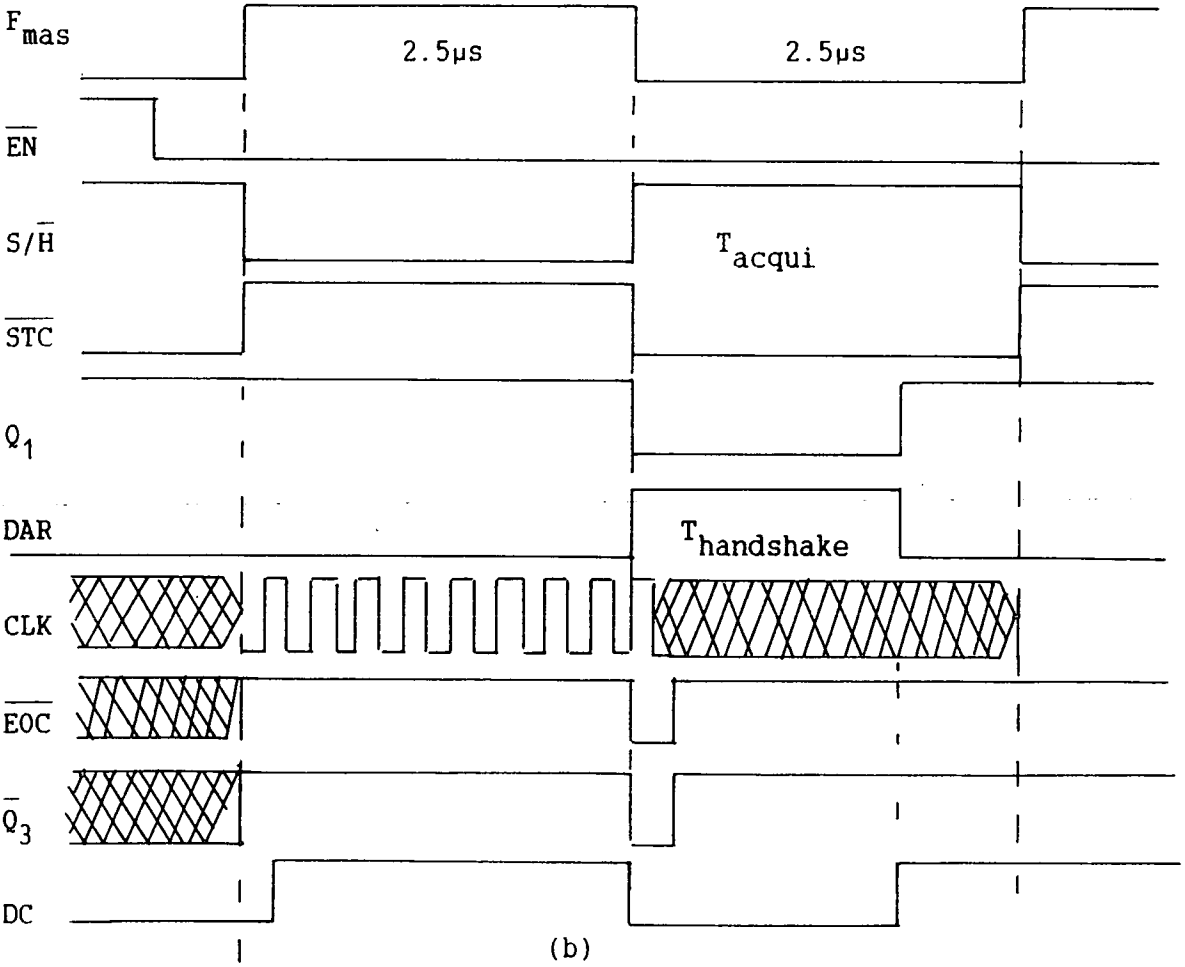
6.4.2.) Timing.

In order to make the sampling synchronous with the specimen voltage, the sampling rate had to be an integral multiple of the DUT clock. However, the conversion itself was asynchronous and performed at a greater rate than the master clock. In addition since the data was to be gathered by the microcomputer, the conversion was initiated by the computer and thus the converter was disabled when not addressed.

The interface timing functions were implemented using the circuit shown in Figure 6.4.2.(a) with the timing diagram presented in Figure 6.4.2.(b) In the quiescent state FF1 is preset and FF2 is cleared resulting in the input signal being continuously sampled and the SAR disabled, since the start convert($\overline{STC}$) is held low. When the converter is addressed by the microcomputer $\overline{EN}$ is taken low and the $\overline{SC}$ and Hold($\overline{H}$) functions are enabled. Since Q_1 is high, Q_2 will go high to initiate a conversion when FF2 is clocked with the system clock. After eight cycles of the asynchronous conversion clock, the end-of-convert($\overline{EOC}$) line is taken low by the SAR. This signal also clocks the outputs of the SAR into the octal output latch on the fal-



(a)
ADC timing control circuit.



(b)
ADC timing diagram.

Figure 6.4.2. ADC converter timing.

ling edge. The data ready line(DAR) is taken high to indicate that there is valid data on the bus. Simultaneously, FF1 is cleared and any further conversion is inhibited.

The handshake from the instrument interface is returned by taking the data clock high, loading FF1, enabling conversion and taking DAR low. However, the analogue input is not resampled and the conversion not initiated until the system clock loads FF2 and pulls Q_2 high. This process is repeated until the converter is disabled and returned to its quiescent state.

Inherent in the design of the ADC handshake is the fact that the conversion time is dependent on the IEEE488 handshake and the master clock frequency. Hence the interface must operate with a frequency of 1MHz, at least, if the maximum sampling rate is to be achieved. However, with the system configuration as discussed the interface was not operated at its full rate and thus the sampling rate was compromised to 20kHz.

6.4.3) Multiplexer and Switched Gain Amplifier.

The sixteen channel multiplexer consists of 2, 8 into 2 analogue multiplexers followed by 2, 2 into 1, in cascade with a single 2 into 1 stage, giving a four bit channel select control word. The switches were constructed by a combination of monolithic CMOS analogue multiplexers and double pole single throw analogue switches.

Following the multiplexer, the selected input channel was fed into a unity gain bandlimiting amplifier which also had a switched input offset to condition bipolar waveforms for the monopolar ADC. Hence the transfer function of the amplifier was

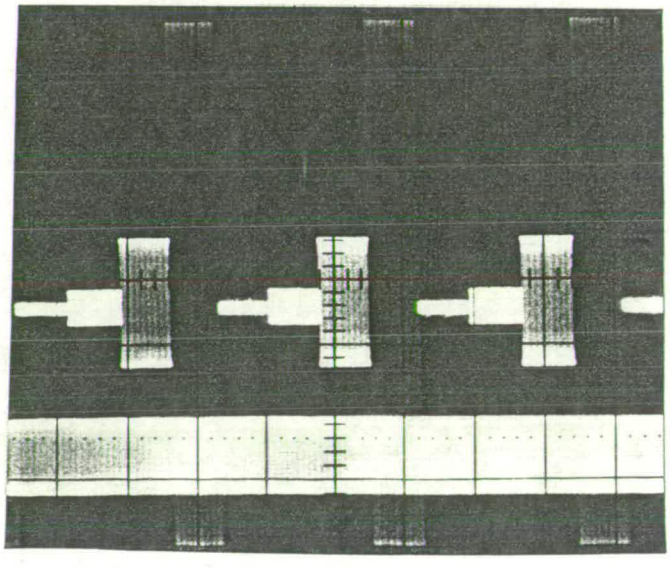


Figure 6.4.3. Datacquisition system input gain ranges.

Upper trace: Output of programmable gain amplifier, when the gain adjust command is put into an infinite incrementing loop. Regions (a), (b), (c) and (d) correspond to gains of 0.13, 0.49, 1.95 and 7.71 respectively. Lower trace: Input sine wave.

$$V_{out} = -\frac{V_{in}}{(1 + \omega\tau)} - N.5 \qquad N = 0,1 \qquad (6.4.3)$$

With a sampling rate of 200kHz, the 3dB frequency of the single pole filter was 10kHz to give 40dB attenuation at the Nyquist frequency in order to limit aliasing. The DC offset was contolled by one bit from the data-acquisition control byte.

After the bandlimiting amplifier, the input signal was applied to the switched gain amplifier which was realised by placing a multiplying 8 bit DAC in the feedback loop of a standard operational amplifier inverting circuit. The DAC performed the function of a digitally controlled floating resistor and its value(R_f) can be expressed as

$$R_f = R_{dac} \cdot \frac{256}{N} \qquad 1 \leq N \leq 255 \qquad (6.4.4)$$

Where R_{dac} is the resistance of the R-2R ladder of the DAC. Hence the transfer function of the amplifier becomes:

$$\frac{V_o}{V_{in}} = -\frac{256}{\alpha\ N} \qquad 1 \leq N \leq 255 \qquad (6.4.5)$$

Where α is the quotient of R_{dac} and the value of the amplifier input resistor, R_{in}. Figure 6.4.4. shows the output of the digitally programmable amplifier, when the gain range is incremented and looped. This clearly demonstrates the four gain ranges available: x 0.13, x 0.5, x 2.0 and x 7.8.

Taken together the ADC and the switched gain amplifier constitute a primitive floating point converter with the exponent derived from the latter and the mantissa from the former. With α = 10, the gain range varied from 0.1 to 25.6. This was reduced further to 0.1

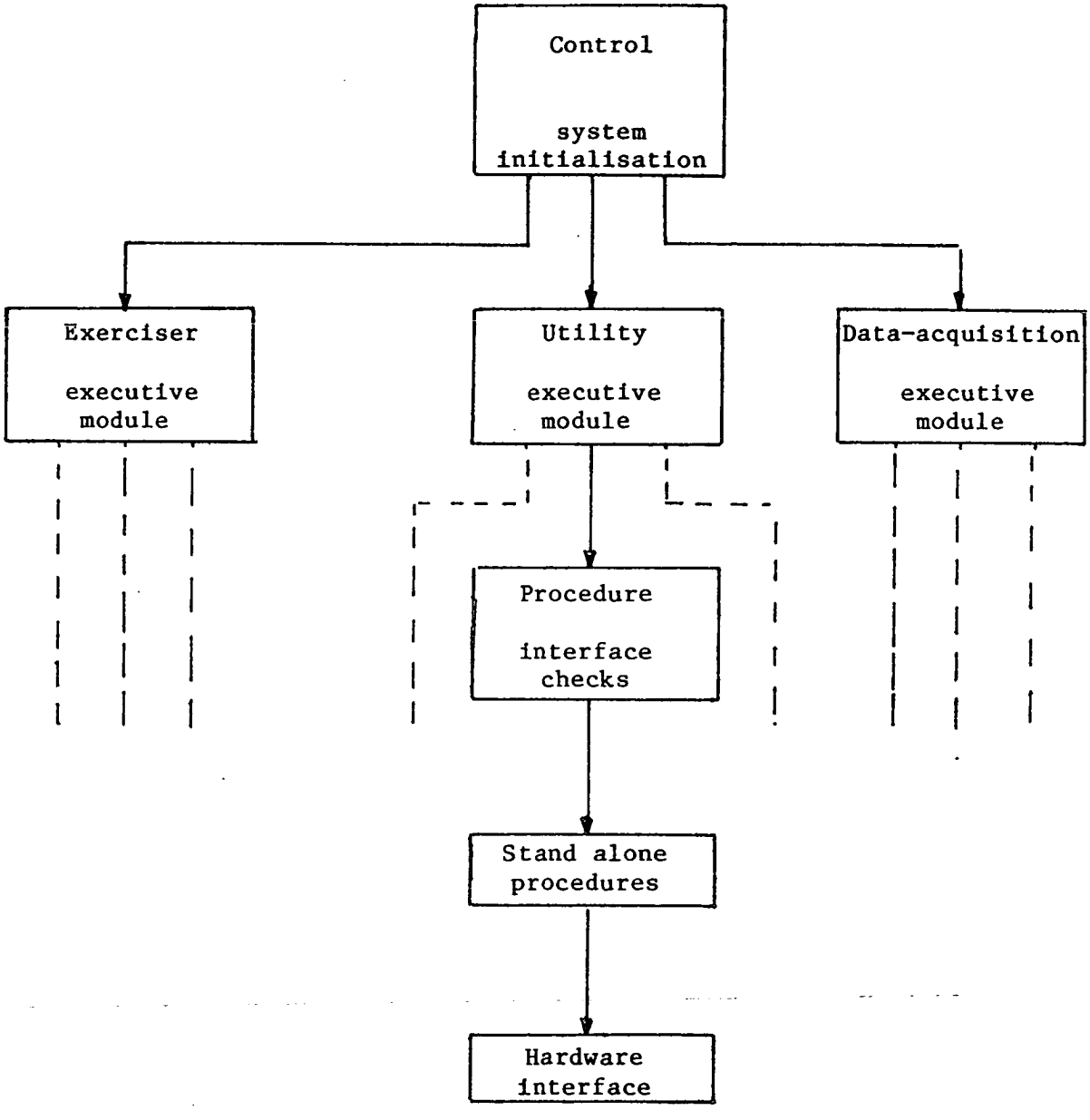


Figure 6.5.1. Software architecture.

to 7.8 because only three control bits remained from the data-acquisition control word. However, even with this limited variation of gain, a dynamic range of 75dB compared with 48dB for an 8 bit ADC with a 10V reference was achieved.

The converter could be improved by incorporating an auto-zeroing function, since as it stands the switched gain amplifier assumes an apriori knowledge of the amplitude of the input waveform.

6.5) Software Control

A master control program was written in UCSD Pascal to supervise all the functions of the Utility system. Using a top down decomposition approach, the program was divided into three executive modules and one overall control module as shown in Figure 6.5.1. The three executive modules correspond to the control of the three sections of the hardware. i.e. the exerciser, the utilities and the data-acquisition system.

The separate compilation option of UCSD Pascal(Softech Microsystems, 1981) allowed each executive module to be compiled separately as units, which could be called from the control program. Each executive module consisted of one or more sub-modules, each of which contained a single procedure. Each procedure was divided into a check part and an executable part. The check part of the procedure ensured that no run time errors would occur and that no illegal writes to the hardware could be made. i.e. a maximum count word which was smaller than the minimum count word. The executable parts were written as procedures themselves and contained units to read and write from the IEEE488 bus. These read and write units were written by Nye(1983) and

```
IM HANGING ON THE BUS
```

(Error message)

```
ROUTINE: C(lock, D(ata_acq, A(ddress, E(xerciser,P(ower, Q(uit
```

(Control prompt)

There are a number of subroutines for setting up the excerciser, these are:

```
              CLOCK              ADDRESS
          EXCERCISER         POWER SUPPLY
     DATA ACQUISITION
Please select:
```

(Control menu)

```
The master clock frequency is: 200kHz.
The circuit clock frequency is derived as a decimal
fraction of this frequency, i.e. 0.010masterclock =  2.00kHz.
Select clock frequency in the range 0.0000 to 0.1000 : 0.masterclock

        THE CIRCUIT CLOCK FREQUENCY IS:  1.50kHz
```

(Utility procedure, clock control prompt)

```
POWERSUP: P(ositive, N(egative, Q(uit
```

There are two Power Supplies contained within the test-rig; one positive and one negative. Their voltage ranges are + 9.95 and - 9.95 respectively. Their resolution is 0.04

Value of supply voltage is :

POSITIVE SUPPLY = + 4.49 Volts.
NEGATIVE SUPPLY = - 0.00 Volts.

(Utility procedure, power supply prompt)

Figure 6.5.2. An example of the user friendly prompts and menus.

controlled the IEEE488 interface handshake.

Since all the executable parts were written as self-contained procedures, it is possible to place them in a system library and allow faster programs to be written without any checks. The alternative, which would result in a slower but safer program is to run the microcomputer's Monitor program while executing the master program. The Monitor program copies all the answers made to the master program's prompts, from the console into a data file which can be used as an input for the master program.

The master program was designed to be self-explanatory and to allow an operator to control the instrument from the console without fear of causing a fatal fault in the hardware. Hence in each executive module and the control section, there is a "user-friendly" interface. Full use of the microcomputer's screen operation codes or procedures (Intertec Data Systems, 1980) was made to generate menu type displays. An example is shown in Figure 6.5.2. Each set of statements separated by a dash indicates a new page on the console. In this sequence, the program passes through control into the utility section to set-up the power supplies and the clock. The initial prompt is an error message issued within the initialisation process indicating that the instruments on the IEEE488 bus are not completing the handshake.

6.5.1) Exerciser Software.

The exerciser software performed five functions: READ, WRITE, COMPARE, VIEW and PATTERN. The data structure of the software was such that the five hardware memory segments were echoed within the microcomputer's memory and the two memories were interconnected

	MEMVIEW															
	MEMSEG2								MEMSEG3							
CLOCK CYCLE	D7	D6	D5	D4	D3	D2	D1	D0	D7	D6	D5	D4	D3	D2	D1	D0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1
2	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1	0
3	0	0	0	0	0	0	1	1	0	0	0	0	0	0	1	1
4	0	0	0	0	0	1	0	0	0	0	0	0	0	1	0	0
5	0	0	0	0	0	1	0	1	0	0	0	0	0	1	0	1
6	0	0	0	0	0	1	1	0	0	0	0	0	0	1	1	0
7	0	0	0	0	0	1	1	1	0	0	0	0	0	1	1	1
8	0	0	0	0	1	0	0	0	0	0	0	0	1	0	0	0
9	0	0	0	0	1	0	0	1	0	0	0	0	1	0	0	1
10	0	0	0	0	1	0	1	0	0	0	0	0	1	0	1	0
11	0	0	0	0	1	0	1	1	0	0	0	0	1	0	1	1
12	0	0	0	0	1	1	0	0	0	0	0	0	1	1	0	0
13	0	0	0	0	1	1	0	1	0	0	0	0	1	1	0	1
14	0	0	0	0	1	1	1	0	0	0	0	0	1	1	1	0
15	0	0	0	0	1	1	1	1	0	0	0	0	1	1	1	1
16	0	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0
17	0	0	0	1	0	0	0	1	0	0	0	1	0	0	0	1
18	0	0	0	1	0	0	1	0	0	0	0	1	0	0	1	0
19	0	0	0	1	0	0	1	1	0	0	0	1	0	0	1	1
20	0	0	0	1	0	1	0	0	0	0	0	1	0	1	0	0
21	0	0	0	1	0	1	0	1	0	0	0	1	0	1	0	1
22	0	0	0	1	0	1	1	0	0	0	0	1	0	1	1	0
23	0	0	0	1	0	1	1	1	0	0	0	1	0	1	1	1
24	0	0	0	1	1	0	0	0	0	0	0	1	1	0	0	0
25	0	0	0	1	1	0	0	1	0	0	0	1	1	0	0	1
26	0	0	0	1	1	0	1	0	0	0	0	1	1	0	1	0
27	0	0	0	1	1	0	1	1	0	0	0	1	1	0	1	1
28	0	0	0	1	1	1	0	0	0	0	0	1	1	1	0	0
29	0	0	0	1	1	1	0	1	0	0	0	1	1	1	0	1
30	0	0	0	1	1	1	1	0	0	0	0	1	1	1	1	0
31	0	0	0	1	1	1	1	1	0	0	0	1	1	1	1	1
32	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0
33	0	0	1	0	0	0	0	1	0	0	1	0	0	0	0	1
34	0	0	1	0	0	0	1	0	0	0	1	0	0	0	1	0
35	0	0	1	0	0	0	1	1	0	0	1	0	0	0	1	1
36	0	0	1	0	0	1	0	0	0	0	1	0	0	1	0	0
37	0	0	1	0	0	1	0	1	0	0	1	0	0	1	0	1
38	0	0	1	0	0	1	1	0	0	0	1	0	0	1	1	0
39	0	0	1	0	0	1	1	1	0	0	1	0	0	1	1	1
40	0	0	1	0	1	0	0	0	0	0	1	0	1	0	0	0

Figure 6.5.3. Pattern generation

The contents of two memory segments after a binary count pattern has been written to both and observed using the VIEW command

through a software buffer. The program had access to the microcomputer's memory segments only for internal calculations such as comparisons. This was to minimise the number of hardware read/writes performed to carry out an instruction. The commands were primitives in that more complex commands could be constructed with them and all were built around the executable procedures. Hence, to implement a comparison, the two hardware memory segments to be compared, were read into the software memory using the READ command. The COMPARE command would then be used to compare the two software memory segments and store the result in a sixth software memory segment. Finally, a VIEW command would be necessary to inspect the result.

The PATTERN command initiated a software pattern generator within the master program. Several patterns can be generated including walking ones, checkerboard, alternate columns and a binary count. As well as being useful for testing memory i.c.s, these patterns proved useful for debugging the exerciser hardware. Access to an external file while the program was running was essential, since externally generated test vectors had to be read into the exerciser's memory. This facility was provided within the PATTERN subroutine and the only constraint was that the data format of the file was integer decimal with a sub-range of 0 to 255. The packed array variable type peculiar to UCSD Pascal(Clarke and Koehler, 1982) was exploited to allow the external file name to be assigned while the program was running, avoiding the pre-execution file handing of standard Pascal.

An example of the PATTERN routine is shown in Figure 6.5.3. and demonstrates the binary count pattern in memory segments three and four, locations 0 to 40. This also illustrates the VIEW command, since the printout is identical to that obtained on the console

MEMVIEW																	
	MEMSEG2								MEMSEG6								
CLOCK_CYCLE	D7	D6	D5	D4	D3	D2	D1	D0	D7	D6	D5	D4	D3	D2	D1	D0	
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	
2	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	
3	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	
4	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	
5	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	
6	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	
7	0	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0	
8	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	
9	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	
10	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	
11	0	0	0	0	1	0	1	1	0	0	0	0	0	0	0	0	
12	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	
13	0	0	0	0	1	1	0	1	0	0	0	0	0	0	0	0	
14	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0	
15	0	0	0	0	1	1	1	1	0	0	0	0	0	0	0	0	
16	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	
17	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	
18	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	
19	0	0	0	1	0	0	1	1	0	0	0	0	0	0	0	0	
20	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	
21	0	0	0	1	0	1	0	1	0	0	0	0	0	0	0	0	
22	0	0	0	1	0	1	1	0	0	0	0	0	0	0	0	0	
23	0	0	0	1	0	1	1	1	0	0	0	0	0	0	0	0	
24	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	
25	0	0	0	1	1	0	0	1	0	0	0	0	0	0	0	0	
26	0	0	0	1	1	0	1	0	0	0	0	0	0	0	0	0	
27	0	0	0	1	1	0	1	1	0	0	0	0	0	0	0	0	
28	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	
29	0	0	0	1	1	1	0	1	0	0	0	0	0	0	0	0	
30	0	0	0	1	1	1	1	0	0	0	0	0	0	0	0	0	
31	0	0	0	1	1	1	1	1	0	0	0	0	0	0	0	0	
32	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	
33	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	
34	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	
35	0	0	1	0	0	0	1	1	0	0	0	0	0	0	0	0	
36	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	
37	0	0	1	0	0	1	0	1	0	0	0	0	0	0	0	0	
38	0	0	1	0	0	1	1	0	0	0	0	0	0	0	0	0	
39	0	0	1	0	0	1	1	1	0	0	0	0	0	0	0	0	
40	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	

Figure 6.5.4. COMPARE command

The two memory segments shown in the previous figure have undergone a software compare process and the result stored in a software buffer, MEMSEG6, within which 1 = different and 0 = same.

screen. Figure 6.5.4. gives the result of a COMPARE command executed on the previous memory segments. There are no errors as indicated by the 0 return of all the comparisons.

6.5.2) Utility Software.

The utility software consists of CLOCK, POWERSUP and COUNTER commands which control the clock generator, the power supplies and the counter. The COUNTER command will be taken as an example of a procedure.

A flow diagram is shown in Figure 6.5.5. of the UFCOUNTER procedure.i.e. a user friendly counter. The LIMEM subroutine is global in that it is used by several modules and, as show by its flow diagram in Figure 6.5.6., is used to check the operator's inputs for the counter maximum count and minimum value. This is to ensure that both values within the variable sub-range and that the maximum is greater than the minimum. After the memory subrange has been set-up, the counter's mode of operation is selected. The modes have been discussed in the hardware section. If the Q(uit mode is selected, the program exits from the UFCOUNT module. If not, the executable procedure COUNTER is executed. This is implemented within the module by the call

```
Counter(countmax, countmin, mode);
```

A flow diagram of the procedure is given in Figure 6.5.7. and consists of writing the memory bounds to the appropriate registers in the hardware and then carrying out a multi-way selection process, according to the mode selected. The counter control word corresponding to the mode of operation is then written to the counter control

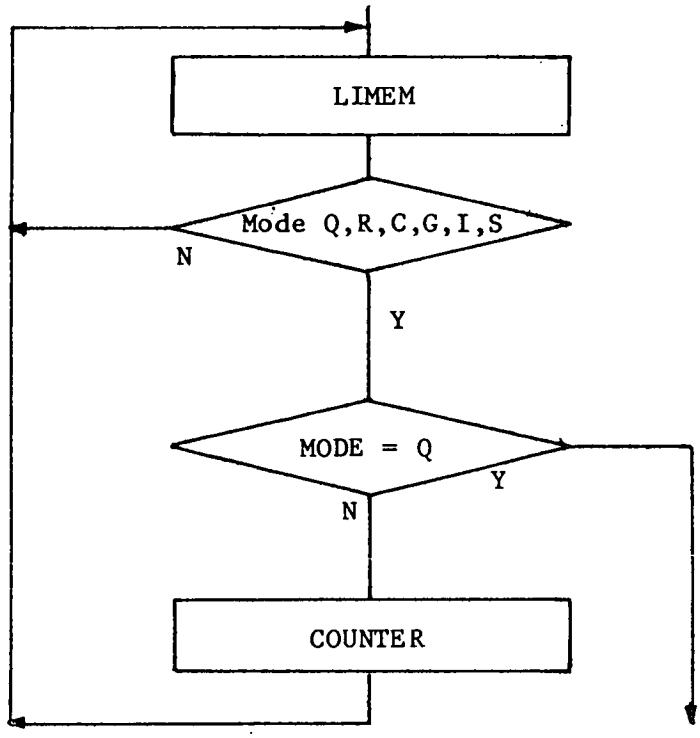


Figure 6.5.5. Procedure UFCOUNTER

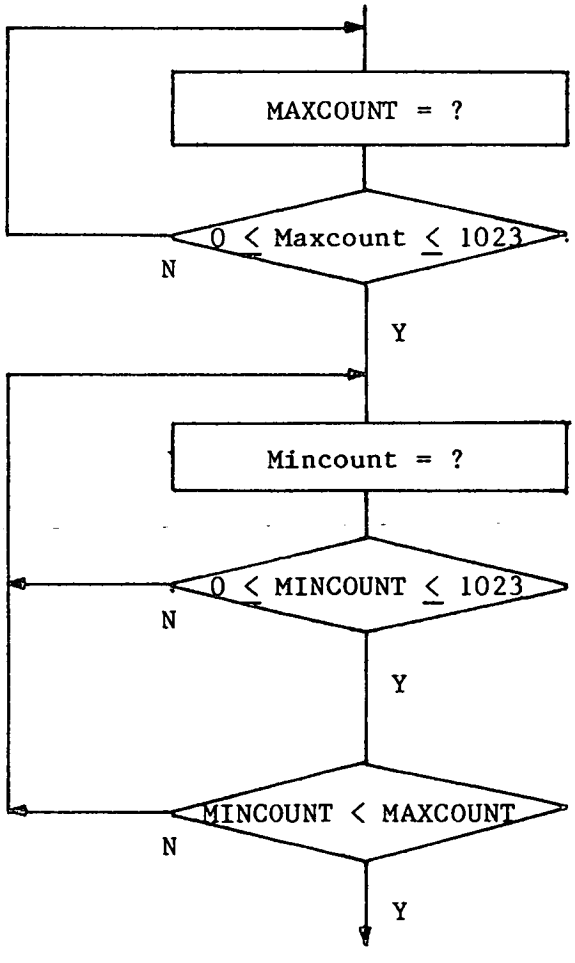


Figure 6.5.6. Procedure LIMEM.

register. In actual fact the mode of operation usually takes two write commands after which an interface clear command is issued. This clears the instrument address register. Note that the counter control word is not returned to a quiescent state and that not all the commands are synchronous.

Appendix B contains a full list of all the executable procedures and the required variable declarations. These procedures taken in conjunction with computer axioms of conditional testing and universal jumps could form the basis of a high level test language for the system, facilitating its use, since the operator does not have to program specific bytes into specific registers. Thus with all the executable procedures in a system library, test programs could be written in Pascal.

6.5.4) Data-acquisition Software.

The software control of the data-acquisition system can be split into two sections: that needed to set-up the data-acquisition system and that required to read from it. Within the master program the commands are CHANNEL, RANGE, POLARITY and NOSAMPS. CHANNEL selects the channel to be sampled(0-15), RANGE selects one of four input gain ranges and POLARITY, either a monopolar or bipolar input. NOSAMPS determines how many samples are required in the range 0 to 1023.

The data is read from the data-acquisition system into one of five pages of memory, 1kByte in depth. This can then be transferred to the console or the line printer. Unfortunately, time did not permit a graph plotter to be interfaced to the system.

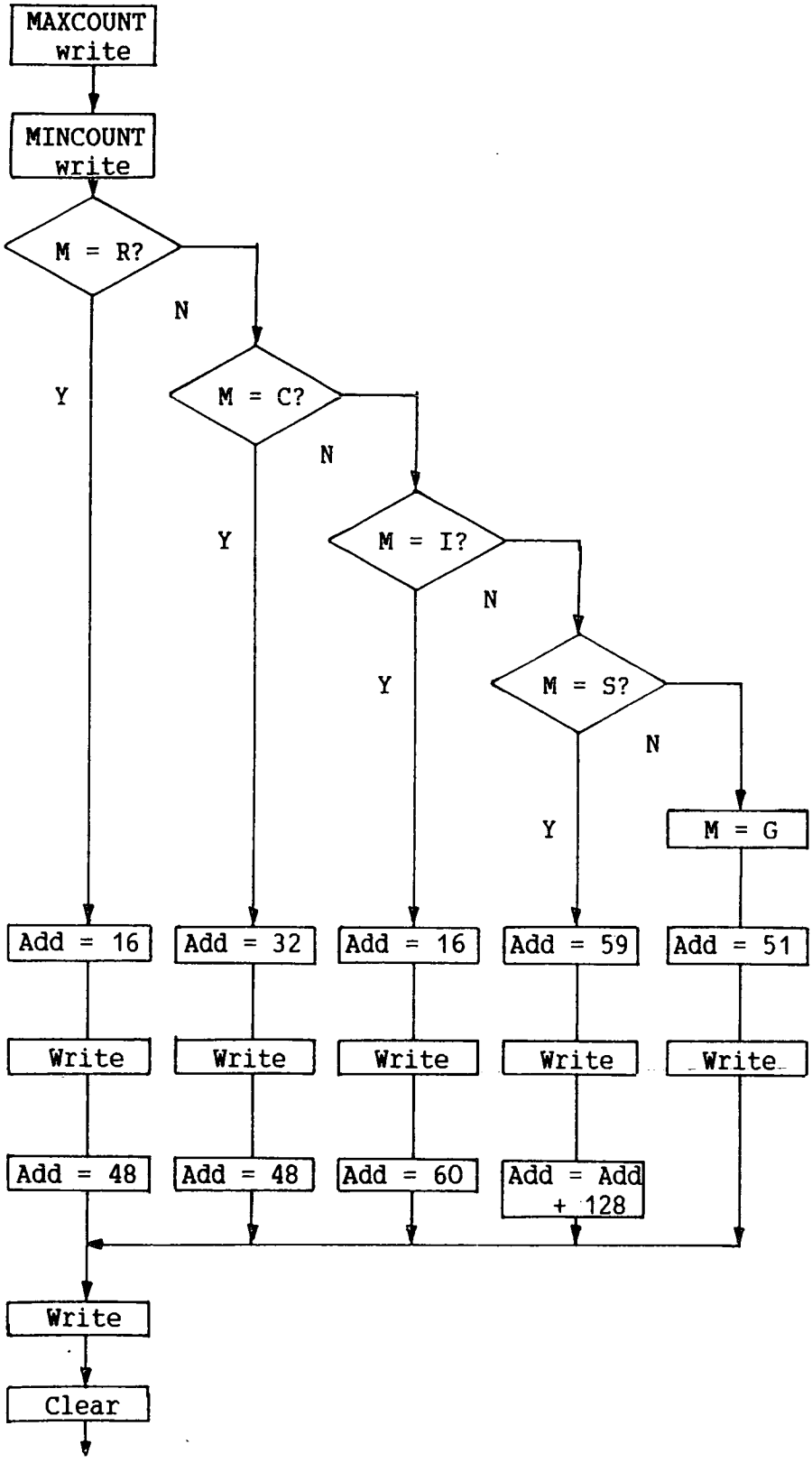


Figure 6.5.7. Procedure COUNTER.

CHAPTER 7.

CONCLUSIONS.

7.1) Conclusions.

- An i.c. test and measurement system has been designed and constructed based on the phenomena of voltage contrast. A bandwidth of 2kHz and a voltage resolution of 0.1V were realised without the aid of a sampling system. It has been shown that voltage measurements are feasible using the technique of tracking the most linear region of the "S"-curve and therefore, the peak of the SED, by inducing and suppressing second harmonic distortion, although more sophisticated signal processing techniques than currently employed are necessary. Further, it was demonstrated that the "S"-curve characteristic could be differentiated by lineshape analysis which is independent of frequency as opposed to the usual electronic differentiation.

The effect of local fields on the performance of the energy analyser was investigated and it was demonstrated that under specific conditions, the energy analyser characteristics, over the voltage range 0 to +5V, were unperturbed for a 5μm wide conductor track with similar adjacent electrodes with an intertrack spacing of similar dimensions.

A computer controlled low cost functional test system was designed, constructed and interfaced with an SEM. Although limited by the modulation frequency of the measurement system to 20kHz, the test system possessed the capability of operating up to 1MHz and in conjunction with a sampling system, allowing logic state mapping. A control program was written to facilitate the continuing use of the test system, by reducing all the set-up procedures to self contained

subroutines, which could reside in a system library and form the basis of a high level test language.

A suite of numerical analysis programs was written in UCSD Pascal to perform integration, differentiation and convolution. In addition, a program which generates Bode magnitude and phase plots for an unfactorised polynomial transfer function was realised. Further, the program also displays a root-loci diagram for varying forward loop gain of a servo-system. and calculates all the frequency break points and the system gain and phase margins.

7.2) Future Work.

Future work, continuing from this thesis would be to carry out a study to compare the measurement technique discussed and the more conventional "s"-curve feedback loop, to quantify the advantages and disadvantages of each system.

Local fields could be further investigated by the design of a test point which could be included within signal paths of i.c.s, the characteristics of which could be accurately simulated and therefore, accounted for in measurements. Hence, the variable test point geometry and local field effects could be eliminated for very accurate results.

An important aspect of future electron energy analyser design must be the ability to collect all or a fixed fraction of the emitted secondary electrons, and thus enable real-time normalisation of the "S"-curve characteristic.

The measurement system could be rendered fully automatic with the inclusion of a hybrid electron beam and specimen stage

positioning system for accurate location of test points. A test point of known geometry would facilitate beam positioning, since it would provide a constant feature within the field of view of the energy analyser, with which pattern recognition techniques could be used to give accurate alignment. Specimen stage movement could be used for rough test point alignment.

APPENDIX A : Root Locus Program.

A computer program was written to facilitate the design and stabilisation of the servo-system. As well as producing a Bode plot of phase and gain the Root Locus was obtained for a given gain range. A system analysis was also derived including the gain and phase margins and all the break frequencies of the system. An example of the graphics output is presented in Figure A.1 in which a second order system has been analysed.

The input to the program was the coefficients of a polynomial quotient approximation to the transfer function(T(s)), where S is the Laplacian operator.

$$T(s) = \frac{Z(s)}{P(s)} = \frac{\sum_{n=0}^{n=r} N_n S^{r-n}}{\sum_{n=0}^{n=r} D_n S^{r-n}} \qquad (A.1)$$

Where r is the order of the polynomial. Both the denominator and the numerator can be considered as the product of several first order and second order irreducible quadratic equations. The solutions of the numerator polynomial are termed zeros since the transfer function is reduced to zero a these values. Correspondingly, solutions of the denominator are termed poles since the transfer function becomes infinite at these values.

In order to plot the Root Locus the transfer function should be made dependent on the forward loop gain(A_o) and equation A.1 becomes for unity gain negative feedback:

$$T(s) = \frac{\sum_{n=0}^{n=r} A_o N_n S^{r-n}}{\sum_{n=0}^{n=r} S^{r-n}(D_n + A_o N_n)} \qquad (A.2)$$

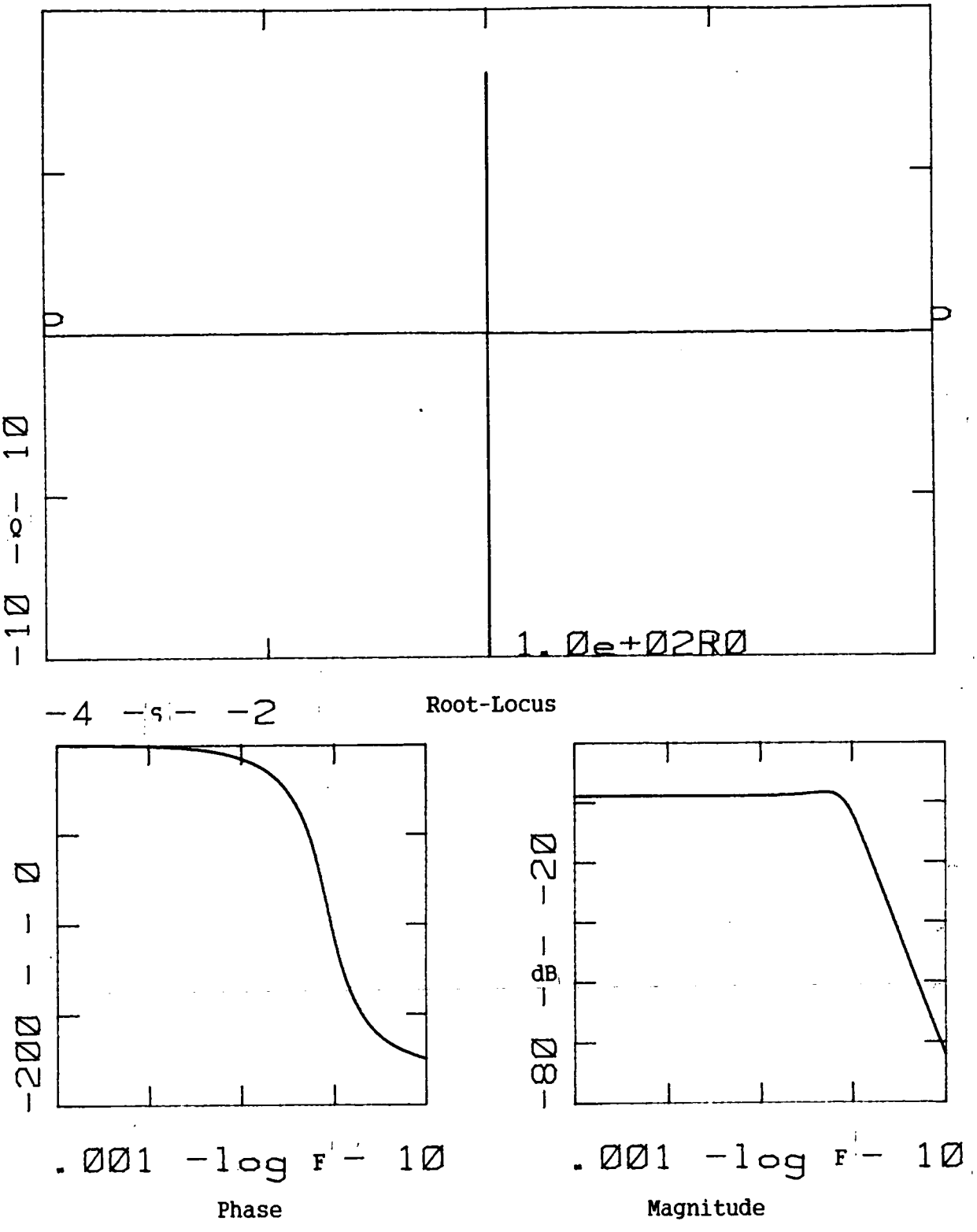


Figure A.1(a). Root-Locus and Bode Plot of second order system.

By determining solutions to both the numerator and denominator polynomials as a function of gain the Root Locus can be plotted and the Bode plot for a single gain value. Hence the program was written to solve this equation. i.e. For T(s) = 0 and T(s) = inf For first and irreducible second order polynomials standard analytic solutions exist and were exploited.i.e

$$(S\omega_o + 1) = 0 \quad \Rightarrow \quad S = \frac{-1}{\omega_o} \tag{A.3}$$

and

$$S^2 + S\frac{\omega_o}{2Q} + \omega^2 = 0 \quad \Rightarrow \quad S_{1,2} = -\frac{\omega_o}{2Q} \pm \sqrt{\left(\frac{\omega_o}{2Q}\right)^2 - \omega_o^2} \tag{A.4}$$

When the polynomial was of a higher order than two numerical methods had to be used. Laguerre iteration in conjunction with a Horner scheme were utilised as explained by Henrici(1974). The iteration function(F(s)) was

$$F(s) = \frac{V.G(s)}{G'(s) + \sqrt{(V-1)^2G'(s)^2 - V(V-1)G(s)G''(s)}} \tag{A.5}$$

Where the argument of the root is chosen to differ by less than $\frac{\pi}{2}$ from the argument $(V - 1)G'(s)$ and is constant and chosen so that $V \neq 0, 1$ This is a third order iteration function for solving $G(s) = 0$ and has the advantage that it will branch out into the complex plane automatically if no real roots are found.

The above equation not only requires the value of the function but also it's first and second derivatives. These plus the coefficients of a deflated polynomial can be found using the Horner Algorithm which can be formulised in the following reoccurrence relationships:

System Analysis:

SYSTEM ANALYSIS

System Gain	:	2.000000e+01
Gain crossover frequency	:	0.000000e+00Hz.
Phase margin	:	0.000000e+00degs.
Phase crossover frequency	:	0.000000e+00Hz.
Gain margin	:	0.000000e+00Hz.

ROOTNO	BREAK FREQUENCY(Hz)	DAMPING FACTOR
1	8.406643e-01	5.679615e-01
2	0.000000e+00	0.000000e+00

Transfer Function:

$$\frac{V_o}{V_{in}} = \frac{A_o}{S^2 + 6S + 8 + A_o} \qquad (A.1)$$

Figure A.1(b). System Analysis and Transfer Function for second order system.

$$B_n^{-1} = D_n \quad ; \qquad n = 0,1,.....r \qquad (A.6)$$

$$B_o^m := D_o \quad ; \qquad m = 0,1,.....r \qquad (A.7)$$

$$B_n^m := B_n^{m-1} + S.D_{n-1}^m \quad ; \quad m = 0,1...r; n = 1,...r-m \qquad (A.8)$$

Where D_n n = 0..r are the coefficients of the polynomial, r is the number of coefficients and S is the value of the input variable at which the Horner scheme is to be calculated. Hence for a third order polynomial the Horner scheme is

D_o	D_1	D_2	D_3
$B_0^{(0)}$	$B_1^{(0)}$	$B_2^{(0)}$	$B_3^{(0)}$
$B_0^{(1)}$	$B_1^{(1)}$	$B_2^{(1)}$	
$B_0^{(2)}$	$B_1^{(2)}$		
$B_0^{(3)}$			

Another characteristic of the Horner Algorithm is the relationships:

$$B_{r-1} = P'(s) \quad ; \quad 2.B_{r-2}^r = P''(s) \quad ; \quad B_r^{(0)} = P(s) \qquad (A.9)$$

Also by computing the Horner scheme for a polynomial at a root point the coefficients of the deflated polynomial are given by:

$$B_n^{(1)} \qquad n = 0,1.....r-1 \qquad (A.10)$$

i.e. the second row of the above Horner scheme.

Hence higher order irreducible polynomials were solved by taking an initial value of -1, computing the Horner scheme for this value and executing Laguerre iteration. When a root was found the second row of the current Horner scheme was taken as the deflated polynomial and the iteration procedure repeated. If the root was complex, the

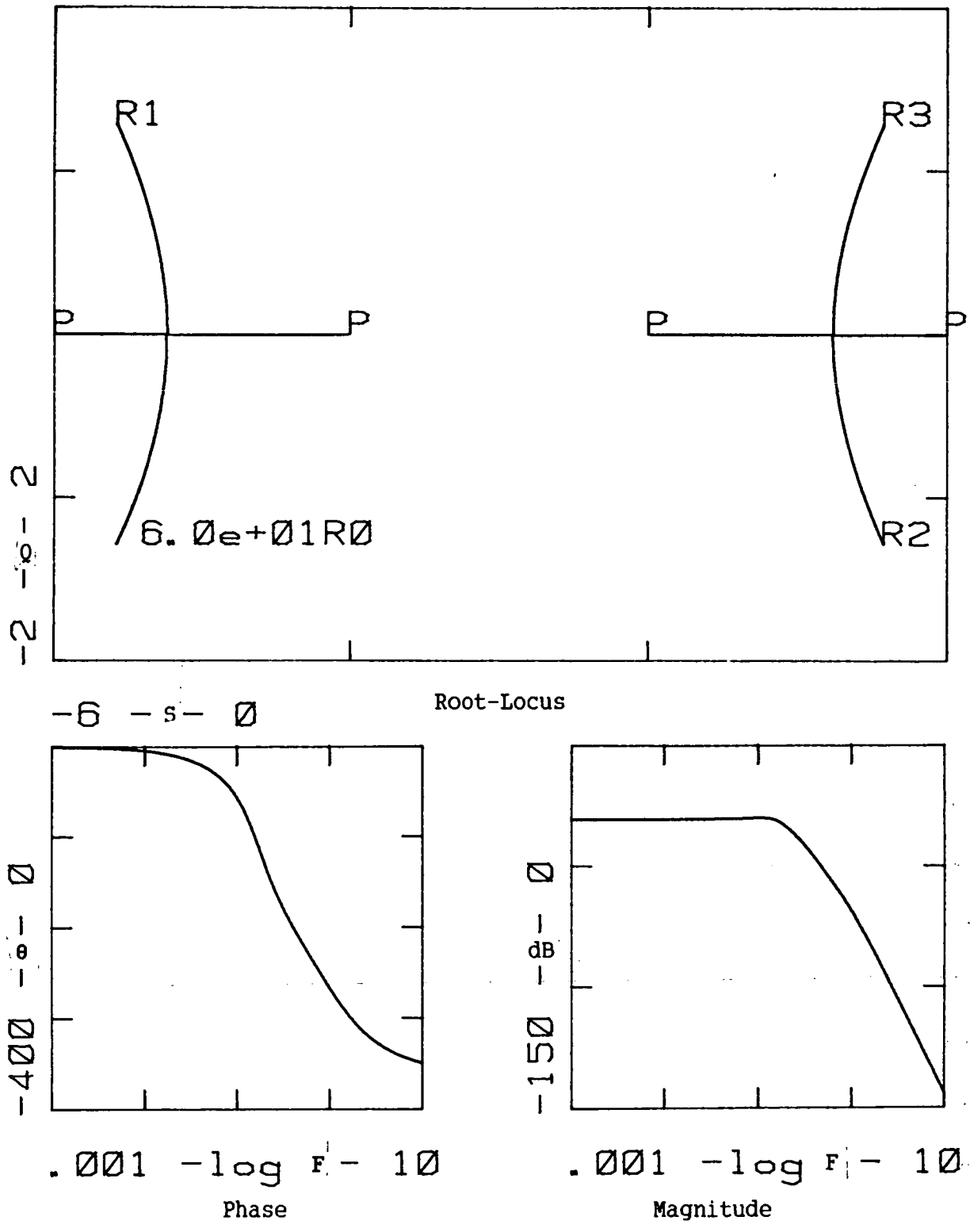


Figure A.2(a). Root-Locus and Bode Plot of fourth order system.

polynomial was immediately deflated again with the complex conjugate of the root without further iteration. This process was repeated until all the roots were found. The present roots were used together with linear extrapolation as an estimate of the new root after the gain was incremented. In this way the Root Locus was derived.

The Bode plots were generated at a gain specified in the input parameters together with the number of decades to be plotted. At this gain the Root Locus part of the program would be interrupted and the current coefficients fed into an internal file. After the Root Locus was completed, the coefficients were used to compute gain and phase contributions in terms of first and second order irreducible quadratic components. These contributions were then added logarithmically to give the Bode plots. A further example of the program is given in Figure A.2

System Analysis:

SYSTEM ANALYSIS

System Gain	:	3.500000e+01
Gain crossover frequency	:	0.000000e+00Hz.
Phase margin	:	0.000000e+00degs.
Phase crossover frequency	:	3.183374e-01Hz.
Gain margin	:	4.194608e+01Hz.

ROOTNO	BREAK FREQUENCY(Hz)	DAMPING FACTOR
1	8.730577e-01	9.864012e-01
2	0.000000e+00	0.000000e+00
3	1.713998e-01	5.469343e-01
4	0.000000e+00	0.000000e+00

Transfer Function:

$$\frac{V_o}{V_{in}} = \frac{A_o}{S^4 + 12S^3 + 44S^2 + 48S + A_o} \qquad (A.2)$$

Figure A.2(b). System Analysis and Transfer Function for fourth order system.

APPENDIX B : Executable Procedures.

The executable procedures can be split into hardware and software routines. Hardware procedures set-up registers within the instrument while software procedures perform functions on software memory segments.

a) Hardware

1) COUNTER(Hi,Lo,Mode)

Hi, Lo : 0..1023; Hi > Lo; Mode := R,C,G,Q

This routine sets the memory bounds of the counter(Hi, Lo) and the mode of operation of the counter. i.e. RESET, GO, STEP, CYCLE etc.

2) CLOCKFREQ(circlk, ckfreq, clock)

circlk : master clock frequency; constant

ckfreq : desired clock frequency 1..255.

clock : returned variable; closest approximation

The clock frequency at which the DUT functions is determined by this routine.

3) POWSUPP(Vmax,value,P,retvalue);

Vmax: maximum supply voltage; constant;

value: desired value;

P: polarity, positive or negative; P,N

retvalue: returned closest approximation.

The positive and negative power supplies are independently adjusted using this procedure.

4) MEMBOUND(Hi,Lo);

Hi, Lo: 0..1023, Hi > Lo;

This procedure resets the memory bounds of the counter independent of the counter's mode of operation.

5) MEMTALK(Lo,Hi,seg,mode)

Lo, Hi: 0..1023 Hi > Lo;

seg, memory segment: 1..5;

mode of operation: R,W

The microcomputer communicates with the hardware using this procedure, transferring the contents of the software buffer of a given memory segment to the exerciser or vice-versa. The number of bytes transferred is determined by the memory bounds.

6) DATA-ACQUI(CHAN,POLARITY,RANGE,NOSAMPS,PAGE);

chan : channel;0..15 0 = earth;

Polarity : -5 to +5 or 0 to +10 : P,N;

Range : input gain range : A, B, C, D;

Nosamps : number of samples : 0..1023;

Page : memory page number : 1..5;

This routine selects the input data channel to be sampled, places a DC offset in the signal channel if the input waveform is AC, adjusts the gain of the switched gain amplifier to give maximum accuracy and takes a predetermined number of data samples. These samples are then stored in a five page memory buffer.

b) Software.

1) SEEMEM(S1,S2,high,low,file);

S1,S2 : memory segments : 1..5;

high, low : memory bounds : 0..1023 high > low;

file : output file : string 7 letters, CONSOLE, PRINTER.

This routine allows the operator to view the contents of the software memory buffers either on the console screen or the line printer. In addition it can store them in an external file. To view the contents of the exerciser memory, a hardware read has to be executed(MEMTALK) prior to a software view(SEEMEM).

2) COMMEM(mem1,mem2,Hi,Lo);

mem1,mem2 : memory segment : 1..5;

Hi, Lo : memory bound : 0..1023 Hi > Lo;

A software comparison between two software memory segments is performed by this procedure and the result is placed in a software buffer(MEMSEG6).

3) PATTERN(mode, file, seg);

mode : generating mode : K, C, A, Z, W, B, F

file : name of external file : string 7 letters.

seg : memory segment.

This routine generated a test pattern determined by the mode and stored in in the given memory segment. Alternatively, it could take a test pattern from an external file and place it in the given memory segment. The patterns available are: K: binary count repeated after 255, C : checkerboard, A : alternating ones and zeros in columns, Z : all zeros, O : all ones, W : walking ones, B : either one column ones and the remainder zeros or vice-versa and finally, F : find test pattern in external file.

REFERENCES.

NB: References given as: IITRI SEM, "year" are to:

Scanning Electron Microscopy ("year")

Proceedings of the annual scanning electron symposium,

Editor: O. Johari,

SEM, INC.,

P.O. Box 66507,

AMF O'Hare,

IL 60666, USA.

Allison L.:" Electronic Integrated Circuits", McGraw-Hill Book Company(U.K.) Ltd., 1975.

Ardezzone F.J.:"Probe Parameters and Considerations", Solid State Technology, 17, 1974, March pp. 51-54.

Aris F.C., Davis P.M. and Lewis T.J., J. Phys. C.: Solid St. Phys., 9 , 1976, pp. 797-808.

Balk L.J., Feuerbaum H.P., Kubalek E. and Menzel E.:"Quantitative Voltage Contrast at High Frequencies in the Scanning Electron Microscope", SEM, IITRI, 1976, pp. 615-624.

Banbury J.R. and Nixon W.C.:"Voltage Measurement in the Scanning Electron Microscope", SEM, IITRI, 1970, pp. 473-480.

Banbury J.R. and Nixon W.C.:"A High-Contrast Directional Detector for the SEM", J. Sci. Inst, (J. Phys. E), series 2, 2 , 1969, pp. 1055-1059.

Beaulieu R.P., Cox C.D. and Black T.M.:"A SEM Voltage Measurement

Technique", Proc. 10th Annual IEEE Rel. Phys. Symp., 1972, pp. 32-35.

Black H.S.:"Stabilised Feedback Amplifiers", Electrical Engineering (Trans. AIEE), 53 , Jan. 1934, pp. 114-120.

Bruining H.:"Physics and Applications of the Secondary Emission Process", Pergamon Press, London, 1954.

Burlison P.:"The L.S.I. Manufacturers Testing", Tutorial LSI Testing, IEE, 2ed., IEEE cat. no. 122-2, 1978, pp. 6-8.

Clarke R. and Koehler S.:"The UCSD Pascal Handbook", Prentice-Hall, USA, 1982, pp. 71-75.

Connor F.R.:"Modulation", 2ed, Edward Arnold, London, 1982.

Conru H.W. and Laberge P.C.:"Oil Contamination with the SEM operated in the Spot Scanning Mode", J. Phys. E., 8 , 1975, pp. 136-138.

Crawford R.H.:"MOSFET in Circuit Design", McGraw-Hill, New York, 1967.

Crichton G., Fazekas P. and Wolfgang E.:"Electron Beam Testing of Microprocessors", IEEE Test Conf., 1980, pp. 444-449.

Dinnis A.R., Khursheed A. And Nye P.D.:"Quantitative Voltage Contrast in the SEM", EMAG 1981, Inst. Phys., 1982, pp. 527-530.

Driver M.C.:"A Spherically Symmetrical Detector for the Scanning Electron Microscope", SEM, IITRI, 1969, pp. 403-413.

Emmerson N.:"An Introduction to NMOS Fabrication", 2nd Microfabrication School, Solid State Devices Group, SERC, 1982.

Eichelberger E.B. and Williams T.W.:"A logic design structure for LSI testing", Proc. 14th Design Automation Conf. IEEE, June 1977, pp. 462-468.

Eichelberger E.B. and Williams T.W.:"A logic design for LSI testability", J. Design-Automation Fault Tolerant Comput., 2 ,(2), May 1978, pp. 165-178.

Everhart T.E.:"Reflections on Scanning Electron Microscopy", SEM, 1968, pp. 1-12.

Everhart T.E. and Hoff P.H.:"Determination of Kilovolt Electron Energy Dissipation vs Penetration Distance in Solid Materials", J. Appl. Phys., 42 , 1971, pp. 5837-5846.

Everhart T.E. and Thornley R.F.M.:"Wide-band Detector for Micro-microampere Low-energy Electron Currents", J.Sci.Instrum., 37 , 1960, pp. 246-248.

Everhart T.E., Wells O.C. and Matta R.K.:"Evaluation of Passivated Integrated Circuits Using the Scanning Electron Microscope", J.Electrochemical Society 3, (8), 1964, pp. 929-935.

Fairchild Camera and Instrument Corporation.:"Fairchild General Purpose Interface Bus(GPIB) Circuit", Preliminary data sheet, 1980.

Fentem P.J. and Gopinath A.: "Voltage Contrast Linearisation with a Hemispherical Retarding Analyser", J.Phys.E., 7, 1974, pp. 930-933.

Feuerbaum H.P.:"VLSI Testing Using the Electron Probe", Scanning electron microscopy (1), 1979, pp. 285-296.

Feuerbaum H.P., Kantz D., Wolfgang E. and Kubalek E.: "Quantitative

Measurment with High Time Resolution of Internal Waveforms on MOS RAM's Using a Modified Scanning Electron Microscope", IEEE J.Solid-State Circuits SC-13 (3), june 1978, pp. 319-325.

Fleming J.P.:"A Technique for Accurate Measurement and Display of Applied Potential Distributions Using the Scanning Electron Microscope", SEM, IITRI, 1970, pp. 467-472.

Fleming J.P. and Ward E.W.:"Improved Technique for Voltage Measurement in the Scanning Electron Microscope", Electronics Letters, 5 ,(5), 1969, pp. 435-436.

Flemming J.P. and Ward E.W.:"Improved Technique for Voltage Measurement Using the SEM", Electronics Letters, 6 , (1), 1970, pp. 7-9.

Fleming J.P. and Ward E.W.:"A Technique for Accurate Measurement and Display of Applied Potential Distributions Using the SEM", SEM, IITRI, 1970, pp. 465-472.

Fuji S., Kenji N., Tohru F., Shozo S., Haruki T., Takeshi T. and Osamu O.: "A Low-power Sub 100ns 256K Bit Dynamic RAM", IEEE, J-SC, Oct 1983, pp. 441-446.

Fujioka H., Hosokawa T., Kanda Y and Ura K.:"Submicron Electron Beam Probe To Measure Signal Waveform At Arbitrarily Specified Positions On MHz IC", SEM, IITRI, 1 , 1978, pp. 755-762.

Fujioka H., Nakame K., Takaoka H. and Ura K:"Function Testing of Bipolar and MOS LSI Circuits with a Combined Stroboscopic SEM-Microcomputer System", Trans. IECE of Japan, 64 , (5), 1981.

Fujioka H., Nakame K. and Ura K.:"Local Field Effects on Voltage Measurement using a Retarding Field Analyser in the SEM", SEM, IITRI,

1 , 1981, pp. 323-332.

Fujishima K., Hideyuki O., Hideshi M., Shigeo U., Masao N., Kazunori S., Kazuhiro S. and Hisao O.:"A 256K Dynamic RAM With Page-Nibble Mode", IEEE, J-SC, Oct. 1983, pp. 470-478.

Gibbons D.J.:"Secondary Electron Emission", Handbook of Vacuum Physics, 2 , Beck A.H. Ed., Pergamon Press, London 1966.

Glasser A.B. and Subak-Sharpe G.E.:"Integrated Circuit Engineering", Addison Wesley Publishing Company Ltd, London, 1979.

Gopinath A.:"Estimate of Minimum Measurable Voltage Change in the SEM", J. Sci. Instrum. (J. Phys. E), 10 , 1977, pp. 911-13.

Gopinath A., Gopinathan K.G. and Thomas P.R.:" Voltage Contrast: A Review", SEM, IITRI, 1978, (1), pp. 375-380.

Gopinath A. and Hill M.S.:"A Technique for the Study of Gunn Devices at 9.1GHz. Using a Scanning Electron Microscope", IEEE Trans. Electron Devices ED-20, 1973, pp. 610-612.

Gopinath A. and Hill M.S.:"Some Aspects of the Stroboscopic Mode: A Review", SEM, IITRI, 1974, pp. 235-242.

Gopinath A. and Hill M.S.:"Deflection Beam Chopping", J. Phys. E., 10 , 1977, pp. 229-236.

Gopinath A. and Sanger C.C.:"A Method of Isolating Voltage Contrast in the Scanning Electron Microscope", J.Phys.E., 4 1971, pp. 610-611.

Gopinath A. and Sanger C.C.:"A Technique for the Linearisation of Voltage Contrast in the Scanning Electron Microscope", SEM, IITRI, 1970, pp. 473-480.

Gopinath A and Tee W.T.:"A New Electron Energy Analyser for Voltage Measurement in the SEM", J. Phys. E., 10 1977, pp. 660-663.

Green D., Sander J.E., O'Keefe T.W. and Matta R.K.:"Reversible Changes in Transistor Characteristics Caused by Scanning Electron Microscope", J.Appl.Lett, 6, 1965, pp. 3-4.

Haas G.A. and Thomas R.E.:"Electron Beam Scanning Technique for Measuring Surface Work Function Variations", Surface Science, North Holland Publishing Co., Amsterdam, 4 , 1966, pp.64-68.

Hanisch J.R., Hughes G.P. and Merrill J.R.:"Theoretical Limitations and Extensions of a Modulation Technique for Lineshape Analysis", Rev. Sci. Instrum, 46 , (9), 1975, pp. 1262-1266.

Hamilton T.S.:"Handbook of Linear Integrated Electronics for Research", McGraw-Hill Book Company(UK) Ltd., 1st ed., 1977.

Hannah J.M.:"Scanning Electron Microscope Applications to Integrated Circuit Testing", Ph.D Thesis University of Edinburgh, 1974.

Hardy W.S., Behera S. and Cavan D.:"A Voltage Contrast Detector for the SEM", J.Phys.E., 2, 1975, pp. 789-793.

Hearle J.W.S, Sparrow J.T. and Cross P.M.:"The Use of the Scanning Electron Microscope", Pergamon Press, London 1st ed., 1972.

Henrici, P.:"Complex Analysis Volume 1", John Wiley and Sons, London, 1974, pp. 433-552.

Hosokawa T., Fujioka H. and Ura K.:"Generation and Measurement of Sub- picosecond Electron Beam Pulses", Rev. Sci. Instrum., 49 , (5), 1978, pp. 624-628.

IEEE Std. 488-1978.:"IEEE Standard Interface for Programmable Instrumentation", IEEE, 1978.

Intertec Data Sytems.:"Superbrain QD", Columbia, USA, 1980.

Johnson D.E.:"Introduction to Filter Theory", Prentice-Hall Inc., Englewoods Cliffs, N.J., 1970.

Jonker J.L.H.:"The Angular Distribution of the Secondary Electrons of Nickel", Philips Res. Repts., 6 ,1912, pp. 372-

Joy D.C. and Titchmarsh J.M.: "The Mechanism of Voltage Contrast In the SEM", 7th Int. Cong. of Electron Microscopy, Grenoble 1970, pp. 221-222.

Keery W.J., Leedy K.O. and Galloway K.F.:"Electron Beam Effects on Microelectronic Devices", SEM, IITRI, (1), 1976, pp. 507-514.

Kerwin W.J., Huelsman L.P. and Necombe R.W.:"State-Variable Synthesis for Insensitive Circuit Transfer Functions", IEEE J.SC-2, 1967, pp. 7-92.

Kimoto S. and Hashimoto H.: "Voltage Contrast in Scanning Electron Microscopy" Proc. 4th European Conf. on Electron Microscopy, Rome 1968, pp. 83-84.

Khursheed A.:"The Computer Aided Design of Electron Detectors for the Scanning Electron Microscope", Ph.D. Thesis University of Edinburgh 1983.

Kotorman L.:"Non-Charging Electron Beam Pulse Prober on FET Wafers", SEM, IITRI, 1980, (4), pp. 77-84.

Lefferts P.:"60Hz Synchroniser Offers Stability, Safety", EDN, 1982, pp. 188.

Lin Y.C. and Everhart T.E.:"Study On Voltage Contrast In SEM", J.Vac.Sci. Technol., 16 , (6), 1979, pp.1856-1860.

Lischke B., Frosiere J. and Schmitt, R.:"Electron-optical Instrumentation for Dynamic E-Beam Testing of Integrated Circuits", Microcircuit Engineering Academic Press, London 1983, pp. 465-474.

Lukianoff G.V. and Touw T.R.:"Voltage Coding Temporal Versus Spatial Frequencies", SEM, IITRI, 1975, pp. 465-471.

MacDonald N.C. and Everhart, T.E.:"Selective Electron-Beam Irradiation of Metal-Oxide-Semiconductor Structures", J.App. .Phys., 39, (5), April 1968, pp. 2433-2447.

McKay. K.G.:"Secondary Electron Emission", Advances in Electronics Academic Press, New York 1948, pp. 66-120.

Menzel E. Ph.D. Thesis, University of Duisburg, West Germany, 1983.

Menzel E. and Kubalek E.:"Electron Beam Chopping Systems in the SEM", Scanning Electron Microscopy 1, 1979, pp. 305-318.

Menzel E., Kubalek.:"Secondary Electron Detection Systems for Quantitative Voltage Measurements", Scanning, 5 , 1983, pp. 151-167.

Miyoshi M., Ishikawa M. and Okumura. T.:"Effects of Electron Beam Testing on the Short Channel Metal Oxide Semiconductor Characteristics", SEM, IITRI, 4 , 1982, pp. 1507-1514.

Muehldorf E.I. and Sakar A.D.:"LSI Logic Testing - An Overview", IEEE Trans Computers, C-30 , (1), 1981, pp. 1-17.

Nakamae K., Fujioka H. and Ura K.:"Measurements of Deep Penetration of Low-Energy Electrons into Metal-Oxide Semiconductor Structures", J.App. .Phys., 52, (3), March 1981, pp. 1306-1308.

Nicolas D.P.:"Role of SEM in Microcircuit Failure Analysis", SEM, 4 , 1974, pp. 955-962.

Oatley C.W.:"The Scanning Electron Microscope Part 1: The Instrument", Cambridge University Press, 1972.

Oatley C.W. and Everhart, T.E.:"The Examination of P-N Junctions with the Scanning Electron Microscope", J.Electron 2, 1957, pp. 568-570.

Ostrow M., Menzel E. and Kubalek E.:"Real Time Logic State Analysis", Microcircuit Engineering 1981, pp. 514-521.

Plows G.S.:"Stroboscopic Scanning Electron Microscopy and the Observation of Microcircuit Surface Voltages", Ph.D. Dissertation, Cambridge University, 1969.

Plows G.S. and Nixon W.C.:"Stroboscopic Scanning Electron Microscopy", J. Sci. Instrum.(J. Phys.E), 1 , 1968, pp. 595-560.

Plows G.S. and Nixon W.C.:"Operational Testing of LSI Arrays by Stroboscopic Scanning Electron Microscopy", Microelectronics and Reliability, 10 , 1971, pp. 317-323.

Poe E.C. and Goodwin J.C.:"The S100 and Other Micro Buses", 2nd ed., Howard W. Sams and Co. Inc., Indianapolis, Indiana, USA, 1980, pp. 149-160.

Ramseyer R.R., Lee D.T. and Kinney L.L.:"Strategy for testing VHSIC chips", Proc. IEEE Int. Test Conf., 1982, pp. 15-518.

Ranasinghe D. and Khursheed A.:"A High Efficiency Energy Detector", Proc. Microscopy, of Semiconducting Materials Inst. Phys., 1983, pp. 433-438.

Roth J.P.:"Diagnosis of Automata Failures: A Calculus and a Method", IBM Journal, July 1966, pp. 78-291.

Seiler H. and Stark M.:"Determination of Mean Free Path of Secondary Electrons in Polyermised Films", Z. Phys., 183 , 1965, pp. 527-531.

Soderquist D. and Schoeff J.:" A Low-cost High Speed Analogue-to-Digital Conversion with the DAC-08", Linear Integrated Handbook, Precision Monolithics Inc, 1981, AN-16, pp. 15/52-58

Softec Microsystems Inc.:"UCSD Pascal User's Manual Version IV.0", , Softec Microsystems Inc., San Diego, USA, 2nd Ed., 1981.

Speth A.J. and Fang F.F.:"Effects of Low-energy Electron Irradiation on Si Insulated Gate FETS", App. ied Physics Letters, 17 , (6), 1965, pp. 145-146.

Stewart J.H.:"App. ication of Scan/Set Error Detection and Diagnosis", Proc. IEEE Int. Test Conf., 1978.

Sumney L.W.:"VLSI with a vengeance", IEEE Spectrum, 17, 1980, pp. 24-27.

Taub H. and Schilling D.:"Digital Integrated Electronics", McGraw-Hill Book Company, 1977, pp. 522-526.

Taylor D.M.:"The Effect of Passivation on the Observation of Voltage Contrast in the Scanning Electron Microscope", J.Phys.D., 11, 1978, pp. 2442-2454.

Taylor N.J.:"Resolution and Sensitivity Considerations of an Auger Electron Spectrometer Based on LEED Optics", Rev. Sci. Inst., 40, (6), June 1969, pp. 792-804.

Tee W.J. and Gopinath A.:"A Voltage Measurement Scheme for the SEM Using a Hemispherical Retarding Analyser", SEM, IITRI, 1, 1976, pp. 595-602.

Tee W.J. and Gopinath A.:"Improved Voltage Measurement System Using the Scanning Electron Microscope", Rev. Sci. Instrum., 48 (3), 1977, pp. 50-355.

Tee W.J., Gopinath A. and Farquhar S.G.:"Voltage Measurement with an Electron Beam Probe", Microciruit Engineering, 1981, pp. 455-464.

Thomson W.E.:"Delay Networks Having Maximally Flat Frequency Characteristics", Proc. IEE, 96 , (3), 1949, pp. 85-490.

Thornton P.R., Davis I.G., Shaw D.A., Sulway D.V. and Wayte R.C.:"Device Failure Analysis By Scanning Electron Microscopy", Microelectron. Rel., 8 , 1969, pp. 33-53.

Torrero E.A.:"A.T.E.: Not so easy", IEEE Spectrum 14, 4, 1977, pp. 29-34.

Touw T.R., Herman P.A. and Lukianoff G.V.:'Practical Techniques for App. ication of Voltage Contrast to Diagnosis of Integrated Circuits", SEM, IITRI, 1977, pp. 177-182.

Ura K., Fujioka H. and Yokobayashi T.:"Calculation of Local Field Effects on Voltage Contrast of SEM", Proc. 7th European Cong. on Electron Microscopy 1 , pp. 330-331, 1980.

Walton A.J.:"Process Control and Evaluation", 2nd Microfabrication School, Solid State Devices Group, SERC, 1982.

Wells O.C.:"Scanning Electron Microscopy", McGraw-Hill Inc., London 1974.

Wells O.C.:"Method for Measuring Voltages in the Scanning Electron Microscope", SEM, IITRI, 2 1969, pp. 397-401.

Wells O.C. and Bremner C.G.:"Voltage Measurement in the Scanning Electron Microscope", J.Phys.E., 1, (2), 1968, pp. 902-906.

Wells O.C. and Bremner C.G.:"Improved Energy Analyser for the Scanning Electron Microscope", J.Phys.E., 2, (2), 1969, pp. 1120-1121.

Wolfgang E., Lindner R., Fazekas P and Feuerbaum H.P.: "Electron-Beam Testing of VLSI Circuits", IEEE J.Solid-State Circuits, SC-14 (2), April 1979, pp. 71-481.

Yakowitz H "The Clyindrical Secondary Electron Detector as a Voltage Measuring Device in the Scanning Electron Microscope", SEM, IITRI, 1972, pp. 33-40.

A STUDY OF LOCAL FIELD EFFECTS ON VOLTAGE CONTRAST MEASUREMENTS

B. Gilhooley and A. Khursheed

Department of Electrical Engineering
University of Edinburgh, Edinburgh
EH9 3JL, Scotland

ABSTRACT

The effect of local surface fields on quantitative voltage measurements in the scanning electron microscope has been investigated. By utilising a specially fabricated test structure, with dimensions similar to those of current semiconductor devices, the way in which energy analyser characteristics are compromised by local fields is demonstrated.

1. INTRODUCTION

Most closed loop quantitative voltage contrast measurement systems, implicity assume that not only the total collected current, but also the shape of the secondary electron distribution remains constant with variations in the applied surface voltage(Flemming and Ward(1970), Gopinath and Sanger(1971)). It has been shown by Rau and Spivak (1979) and Ura et al.(1980) that local field effects can constitute large sources of error in quantitative voltage contrast measurement systems and that these effects will be accentuated with smaller geometry integrated circuits.

Local field effects are caused by the introduction of transverse electric fields between adjacent electrodes, at different potentials on the surface of a specimen. The effect of these local fields is to distort the initial secondary electron distribution and hence, the "S"-curve characteristic of a retarding field, highpass type energy filter. This in turn causes the feedback loop to sense an apparent change in the operating point of the system and to generate a corresponding correction voltage.

The local fields can be partially suppressed by the use of high extraction electric fields. However, Fujioka et al. (1981) showed that the effects of local fields on an aluminium /silicon test structure, with dimensions greater

than those of a typical integrated circuit by a factor of two, cannot be suppressed to an insignificant level by increasing the extraction voltage alone.

This paper will investigate conditions on MOS integrated circuit structures for which reliable measurements can be made, using a test structure with comparable dimensions of present day integrated circuits.

2. INSTRUMENTATION

In order to demonstrate any distortion present in the "S" curves of the energy analyser caused by local field effects, a series of linearly displaced "S"-curves, covering the dynamic range of the measurement system was generated for each measurement point. A typical display is shown in figure 3, in which the upper trace is the applied specimen voltage, going vertically negative to positive and maps onto the horizontal axis of the lower trace which displays the displaced "S"-curves, going left to right.

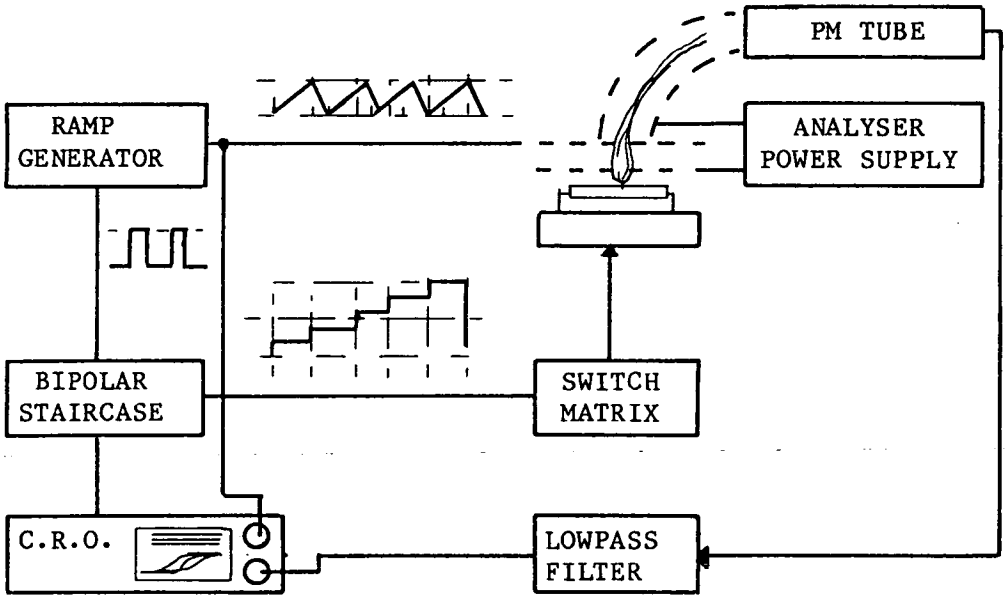


Fig.1. Instrumentation set-up used to generate characteristic "S"-curves.

The instrumentation set-up used to obtain the analyser characteristics is shown in figure 1. The waveform generator provided a 40V peak-to-peak linear ramp waveform, which was applied to the retardation grid of the analyser and a fly-back pulse, which was used to trigger and blank a dual

channel oscilloscope. In addition, the flyback pulse was used to synchronise the applied specimen voltage, which was a bipolar staircase waveform, in such a way that the period of the retarding waveform equaled the period of one step. Both the number of steps and the step magnitude could be altered manually. The output from the video head amplifier was lowpass filtered and applied to the dual channel oscilloscope, while the applied surface voltage was connected to the second channel. The time base was adjusted until all the "S"-curves appeared superimposed as in figure 3.

A retarding field, highpass energy analyser as described by Ranasinghe and Khursheed(1983) was utilised throughout the series of experiments and although, results from the analyser have been previously published by both Ranasinghe and Khursheed(1983) and Gilhooley and Dinnis(1983), the results presented here differ since an integrated test structure was used and not a copper stub. A Cambridge S100 scanning electron microscope(sem) was used throughout the experiments with a stationary primary beam and an accelerating voltage of 1kV. The sem's working distance was 27mm.

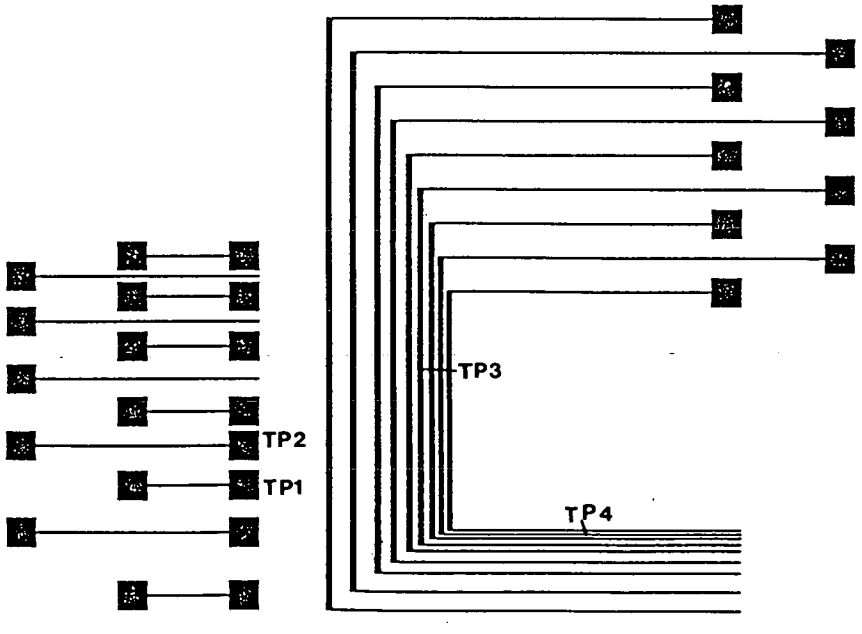


Fig.2. Floor plan of Test Structure.

3. TEST STRUCTURE

A special test structure(figure 2), was fabricated to allow the investigation of field effects along parallel aluminium tracks with varying track and inter-track dimensions. The inter-track spacing varies from 3um to 50um, while the track width varies from 3um to 25um. In addition, there are two arrays of 100um by 100um test pads with inter-pad spacings varying from 10um to 100um. In the upper array of pads a 10um wide electrode bisects the intergap space. The test structure has silicon dioxide deposited beneath the aluminium tracks to reconstruct typical device conditions. Each test integrated circuit contains two identical test structures; one passivated with silicon dioxide, the other unpassivated.

4. TEST PADS

Initially, a gold test pad 0.4mm by 1mm was probed to ensure that linearly displaced "S"-curves could be obtained with the analyser without any apparent local field effects and therefore, any subsequent distortion in the "S"-curves could be attributed to local fields and not the analyser. The series of "S"-curves obtained is shown in figure 3 for a surface voltage range of -10V to +10V and demonstrates that no apparent distortion is present. An aluminium test pad 100um by 100um (TP1) was probed and a series of "S"-curves produced. A positive potential of 30V applied to the adjacent pad, resulted in a 50% reduction of the amplitude of the "S"-curve corresponding to a surface voltage of +5V. However, with an applied potential of +20V no observable distortion was present. The inter-pad spacing was 25um. Both these experiments were carried out with the primary beam positioned in the centre of the test pad.

To highlight the effect of local fields on "S"-curves an aluminium pad (TP2) was probed close to the edge bordering an adjacent pad with an interpad spacing of 10um. Both the voltage applied to the adjacent and the probed pad was varied in the range -6V to +5V. The series of "S"-curves shown in figure 4(b) was obtained with 0V applied to the adjacent pad and demonstrates that "S"-curves corresponding to positive surface voltages are affected predominantly, reducing in amplitude and giving rise to an apparent shift from their original position (figure 3). The "S"-curves corresponding to negative surface voltages also reduce in amplitude, but the 0V curve remains unaffected. These results are as expected, since the magnitude of the surface

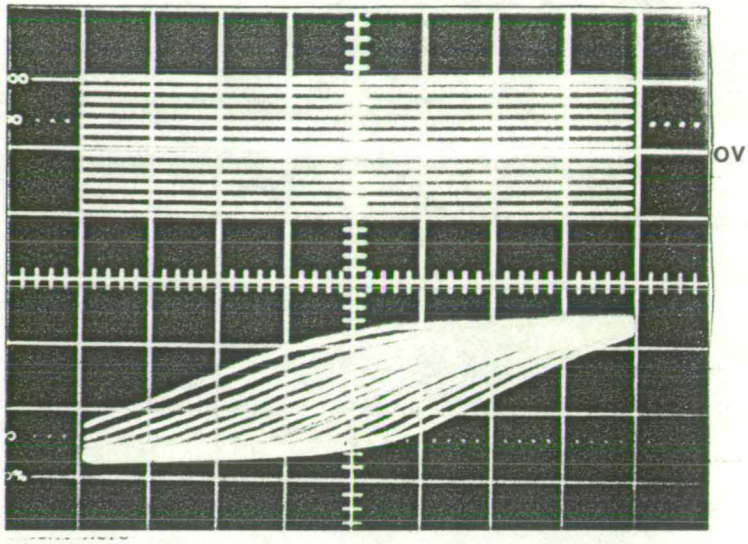


Fig.3. Gold test pad. Extraction voltage +100V. Upper trace: Applied surface voltage to gold test pad. 1V per mm. Lower: "S"-curve series generated.

fields is dependent on the potential difference between the two pads. Hence distortions are present for positive and negative potential differences, but absent when the pads are maintained at the same potential. As shown in figures 4(b) to (d), the degree of distortion increases with the magnitude of the potential difference of the test pads and is greater for positive potential differences. Figure 4(c) demonstrates the effect produced when +5V is applied to the adjacent pad. The "S"-curves corresponding to surface voltages around +5V are seen to increase in amplitude and become more regular in shape, while the remaining "S"-curves reduce in amplitude. This is as expected since the relative voltage between the two pads for a surface voltage around +5V is very small and hence the effect of the surface fields is greatly reduced for these conditions. Figure 4(d) shows "S"-curves obtained for -6V applied to the adjacent pad and again, as expected the "S"-curves for surface voltages around -6V become more regular, while the other "S"-curves greatly decrease in amplitude.

Figure 4(e) demonstrates that as the electron beam is moved away from the pad edge towards the centre the effect of the local fields on the "S"-curve decreases. Although the effect of surface fields can be reduced by increasing the extraction voltage by a factor of four, it could not be eliminated.

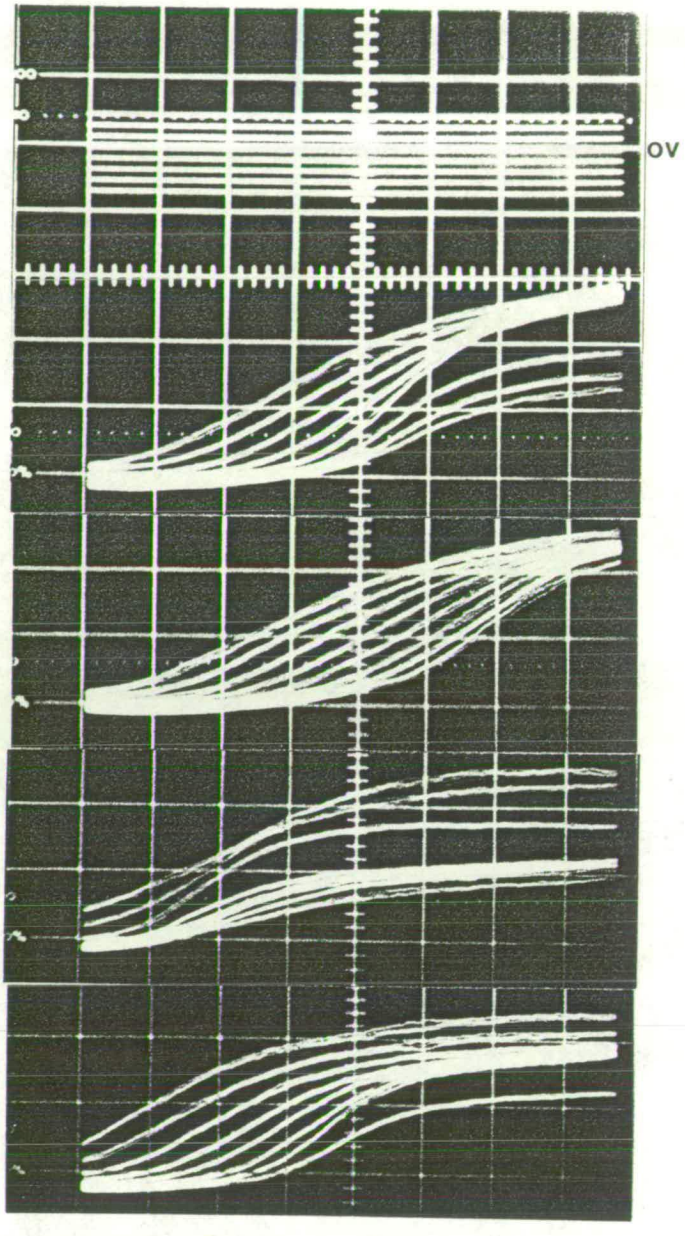


Fig.4. Pad edge probe. Extraction voltage: +100V. (a) Applied surface voltage 1V per mm (b) Adjacent pad voltage 0V (c) Adjacent pad voltage +5V (d) Adjacent pad voltage -6V (e) Adjacent pad voltage 0V, middle of pad.

5. PARALLEL TRACKS: 5um WIDE, 3um SPACING

Initially 10um tracks with intertrack spacing of 25um were probed at test point three (TP3). However, the "S"-curves obtained were grossly distorted and could not be the basis of a measurement system. This was thought to be caused by the inter-electrode oxide charging and producing complex

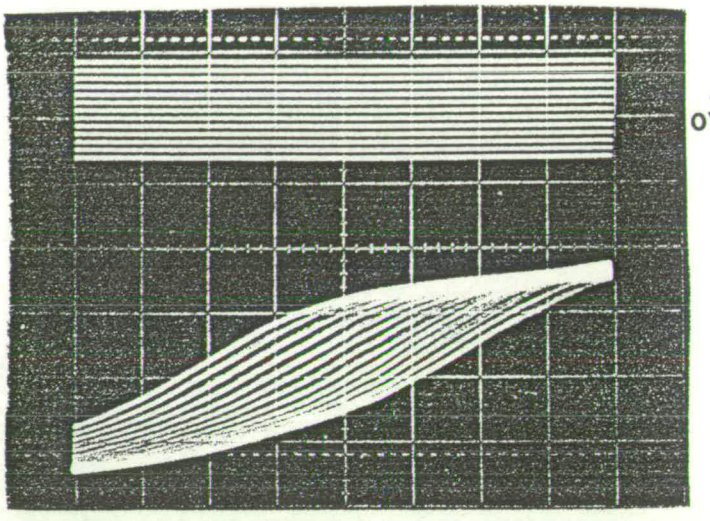


Fig.5(a). Surface voltage applied to all electrodes. Extraction voltage +200V. Upper trace: Applied surface voltage. 1V per mm Lower trace: "S"-curve group.

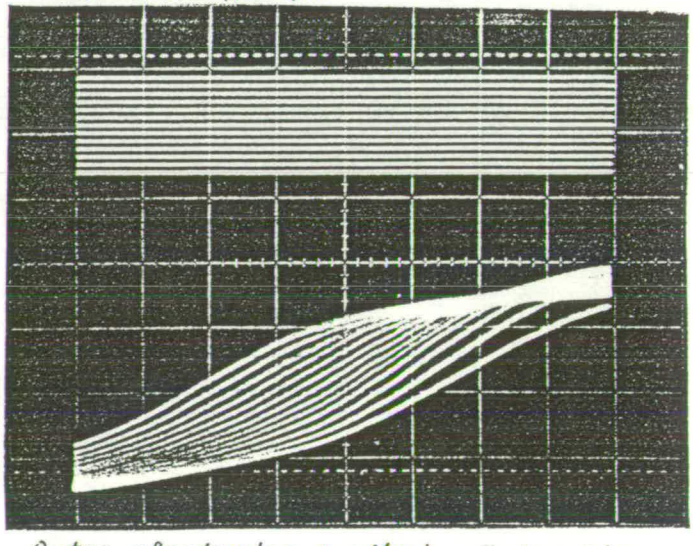


Fig.(b). Outer electrodes earthed. Extraction voltage +200V. Upper trace: Applied surface voltage 1V per mm. Lower: Surface voltage 1V per mm.

surface fields.

Shown in figures 5(a) and 5(b) are the results of an experiment in which three parallel conductors, 5um wide and with an inter-track spacing of 3um were utilised (TP4). A bipolar staircase waveform was applied to all three tracks and the central track probed, as shown in figure 5(a). A linear set of undistorted "S"-curves was generated. When the outer two tracks were grounded the distortion apparent in the second display (figure 5(b)) was observed. At positive surface voltages of about +10V and -10V a nonlinear translation of the "S"-curve occurred. In addition, the amplitude of the secondary electron signal was reduced owing to the presence of large transverse surface fields which became greater as the potential difference between the central track and its neighbouring track was increased. However, the surface voltage/energy displacement relationship remains linear in the region of interest. i.e. from 0V to +5V and thus it is possible under these conditions to carry out quantitative measurements.

6. CONCLUSIONS

It has been shown empirically that the effect of surface fields is most dominant when the probed conductor is positive with respect to close adjacent conductors and that when probing MOS structures the electron beam should be positioned in the middle of the widest track, in the area of least oxide. Also it has been demonstrated that 5um conductor tracks can be probed, but only under special conditions.

7. REFERENCES

Flemming, J.P. and Ward, E.W. (1970) Proc. 3rd. Ann. SEM Symp., IIT Research Institute, Chicago, ill., 465-472.

Fujioka, H., Nakamae, K. and Ura, K. (1981) Scanning Electron Microscopy (Chicago) 1 , 323-332.

Gilhooley, B. and Dinnis, A.R. (1983) Proc. Microscopy of Semiconducting Materials, (Cambridge) To be Published.

Gopinath, A. and Sanger, C.C. (1971) J.Sci. Inst. (J. Phys. E.), 4 , 334-336.

Ranasinghe, D.W. and Khursheed, A. (1983) Proc. Microscopy of Semiconducting Materials (Cambridge) To be Published.

Rau, E.I., and Spivak, G.V. (1979) Scanning Electron Microscopy, 1 , 325-331.

Ura, K., Fujioka, H. and Yokobayashi, T. (1980) Proc. 7th European Cong. on Electron Microscopy 1 , 330-331.

Inst. Phys. Conf. Ser. No. 67: Section 9
Paper presented at Microsc. Semicond. Mater. Conf., Oxford, 21–23 March 1983

Quantitative voltage contrast: instrumentation and signal processing

B Gilhooley and A R Dinnis

Department of Electrical Engineering, University of Edinburgh, The King's Buildings, Mayfield Road, Edinburgh. EH9 3JL., Scotland.

Abstract The design and operating principles of an instrumentation system for real-time testing of L.S.I. integrated circuits, using the electron beam probe in the voltage contrast mode, are described. The system utilises coherent detection and second harmonic suppression to realise a linear relationship between the surface voltage and the error signal in the feedback loop. The preliminary evaluation results are presented.

1. Introduction

The electron microprobe utilised in the voltage contrast mode has established itself in semiconductor research laboratories as a practical and precise instrument for the detection and characterisation of timing faults in the early design stages of integrated circuits (Wolfgang, 1982, Lukianoff, 1981, Feuerbaum, 1979). With the advent of high density, high impedance V.L.S.I. circuits and the resultant economic penalty of high pin count packages, it will become necessary to implement internal nodal testing at an early stage of production to verify functional devices. If functional testing is to be implemented by the electron beam probe, a fully automated test system, with a real-time 3dB bandwidth of approaching five times the minimum clock frequency of the process will have to be realised. High voltage and temporal resolution will not constitute major design constraints, as in the parametric testers, but rather the ability to discriminate between "1"s, "0"s and bad levels. High speed open loop systems have been developed (Ostrow, 1981, Macari, 1982), but suffer from being both qualitative and requiring user set-up time.

To this end an electron beam test system is being developed to test dynamic nmos integrated circuits, fabricated using a 5μm polysilicon gate process at Edinburgh. The system is being designed to complement rather than replace conventional test systems and will be able to implement both functional and parametric testing. The system consists of IEEE 488 controlled, 2 phase clock generator, plus and minus power supplies, test vector generator and a data acquisition system. The stimuli are interfaced to the device-under-test via a patch board and are controlled by a microcomputer using the Pascal language. This paper addresses the problem of deriving quantitative voltage measurements from the surface of integrated circuits, while maintaining a system bandwidth capable of resolving minimally clocked circuits.

2. Operating Principles

The system is based on the surface voltage modulation of the secondary

electron distribution and incorporates an energy/velocity filter (Ranasinghe, 1983). For signal processing purposes, the output of the video head amplifier can be considered as a function of the integral of the product of the energy filter's impulse response (H(E)) and the secondary electron distribution (N(E)):

$$Vo = Zm.T.\int_{0}^{50eV} \frac{H(E-e)N(E-f(Vs))dE.Isec}{N(E)dE}$$

where Zm is the transimpedance of the video head amplifier, T the transport efficiency of the filter, e the filter cut-off energy, Vs the surface voltage and Isec the secondary electron current. Although bandpass energy filters have been used (Hannah, 1974), the signal-to-noise ratio and convolution errors involved have been such that they have been superceded by highpass, retarding grid filters. This type of filter was used throughout these experiments. If the energy filter's cut-off is made time dependent, the characteristic "S"-curve can be generated by convolving the impulse response of the filter with the secondary electron distribution. Ideally, the energy filter's impulse response should be a step function and independent of the secondary electron angle of emission.

In order to obtain a linear relationship between the indicated value and the actual voltage on the specimen, Flemming(1970) stated the now common practice of applying a feedback voltage to the detector. The effectiveness of this depends on the surface conditions and the electron optical performance of the detector. Since feedback to the retarding grid was utilised, an operating point had to be selected and to this end the characteristics of the secondary electron distribution were assessed. The "S"-curve is a monotonic function of energy and is the result of a stationary statistical process: i.e. secondary electron distribution. Owing to these facts, the peak of the distribution was selected, as this should remain fixed independent of beam current and time. Inherently, the effects of surface fields and the energy filter itself are assumed to be constant for each measurement point, these being a function of energy filter design and not signal processing.

Employing a bandpass energy filter, Hannah(1974) used differentiation followed by zero-crossing detection to estimate the point of inflection on the "S"-curve. However, this method degrades the signal-to-noise ratio and is not suited to a feedback system because of the required scanning. The method employed was spectral differentiation. That is, if a Taylor expansion of an estimating polynomial of the"S"-curve is taken, it can be shown (Taylor, 1969) that under the correct conditions, the second term in the expansion can be made proportional to the first derivative of the "S"-curve and the third component to the second derivative (figure 1). Since the system had to be compatible with the sampling parametric tester, coherent detection was employed to eliminate the detection system's 1/f noise and improve the signal-to-noise ratio, at low system bandwidths. By applying the modulation function to the retarding grid and detecting the first and second harmonics the first and second derivatives could be estimated and thus, the point of inflection in the "S"-curve. The feedback loop was constructed to null out the second harmonic by applying a correction voltage to the analyser's retarding grid and therefore, to track the zero point of the second derivative. By employing a modulating function, the system bandwidth is subject to the limitations set by the Sampling Theorem and the

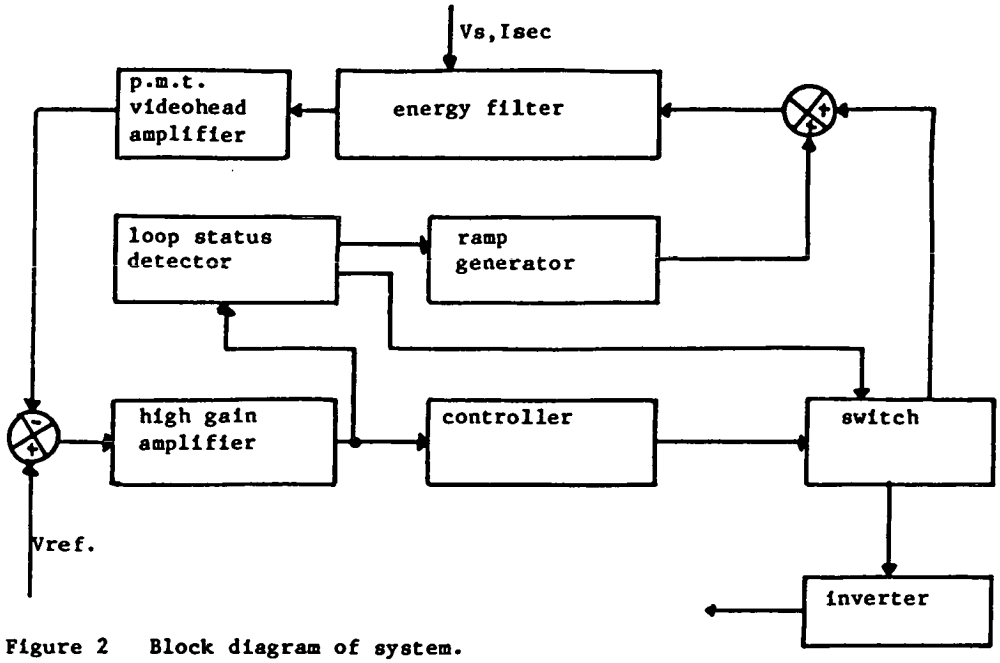


Figure 2 Block diagram of system.

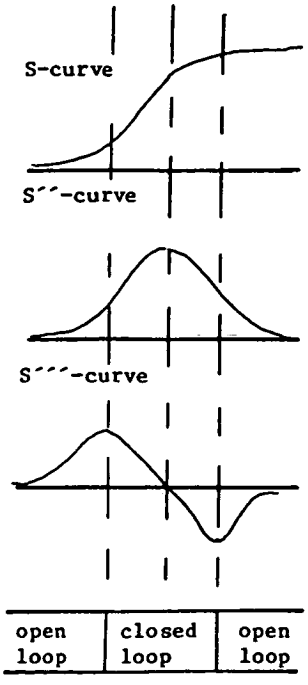


Figure 1 System operating characteristics.

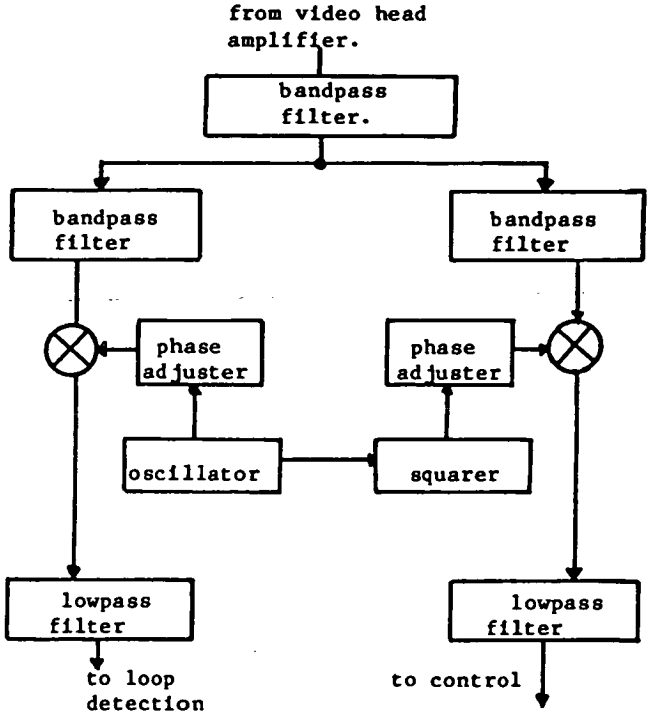


Figure 3 Modulation system block diagram.

signal-to-noise ratio is effectively halved for large bandwidths.

3. System

A block diagram of the system is presented in figure 2. Initially, the system operates in open loop and a 200kHz sine wave superimposed on a slow slewing ramp(10Hz) is applied to the retarding grid of the energy filter. The first and second harmonics are detected and when the peak of the first derivative is encountered, a level detector circuit samples the ramp voltage and closes the loop. The feedback loop pulls the error signal into the null point and tracks the surface voltage. The system is a.c. coupled to the video head amplifier with a bandlimiting predetection filter (figure 3), to prevent saturation in the detection circuits. Further filtering is employed to prevent cross-talk between the slow first and fast second harmonic channels. To maximise the output signal-to-noise ratio these filters have to be matched closely with the signal band. The phase response of these filters is critical, since drift in the modulating frequency can cause loss in coherence. Since the feedback loop gives no improvement in signal-to-noise ratio and the video head amplifier has a bandwidth of 5MHz, the resolution of the system is determined by the amount of bandlimiting employed. This is controlled by the bandwidth of the lowpass filters employed at the output of the modulation system, which are second order Butterworths, tunable from 1-100Hz in decades. With a signal band of 20kHz, 5MHz video bandwidth and white noise, the process gain is about 30dB. The accuracy with which the system can track a surface voltage is dependent on the Loop Gain of the system, which is predominantly controlled by the photomultiplier tube gain.

4. Results

Preliminary experiments to evaluate the system's performance were carried out using a copper stub and a beam accelerating voltage of 2kV on a Cambridge Instrument's S150 scanning electron microscope. With the system operating in the open loop mode, the "S"-curve, first and second derivatives were obtained as shown by figure 4.a. Filter cut-off frequencies of 100Hz and 1kHz were used to achieve the first and second derivatives, respectively, while the "S"-curve itself was obtained directly from the video-head amplifier output. When +1.5V level was applied to the specimen, a corresponding shift in the energy domain occurred. This is demonstrated by the right lateral shift of the inflection point, shown by the curves in figure 4.b. as compared with its position in figure 4.a. Since a positive slewing 15V peak-to-peak ramp waveform, with a period of 35mS was used to scan the energy domain, a 1:1 correspondence between the applied surface voltage and the correction voltage can be inferred. While in the closed loop, tracking mode a 8V peak-to-peak triangular waveform was applied to the specimen to demonstrate the linearity of the system. Although a 10Hz cut-off output filter was used to detect the correction voltage, the output waveform was distorted by low frequency noise, as shown by figure 4.c However, a linear system can be extrapolated from the results. Due to the beam current noise and the gain of the system, the maximum closed loop tracking bandwidth achieved at this stage was about 1-2kHz, as shown by the sine wave response given in figure 4.d. Although there is only slight or no phase distortion of the output sinewave, above a cut-off frequency of 1kHz the system would pull out and in of lock unpredictably, before the system began to oscillate. These are preliminary results, without either the bandlimiting filters or the modulation depth having been optimised: i.e. the predetection filter had a 400kHz bandwidth, while the signal band

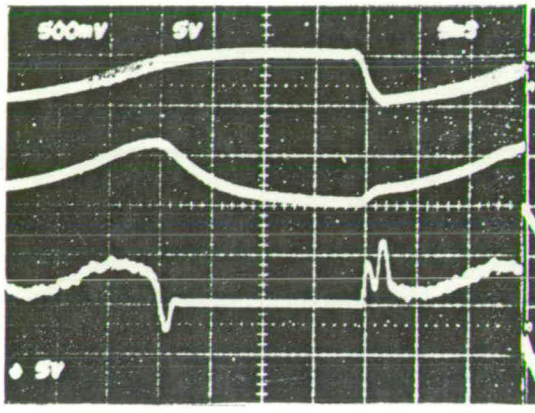


Figure 4 a from top to bottom: "S"-curve, 1st derivative, 2nd derivative. Surface voltage = 0v.

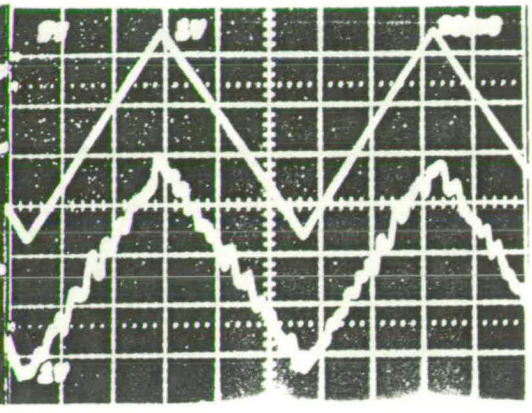


Figure 4 c Upper 8Vpp triangle surface voltage. Lower output votage / two.

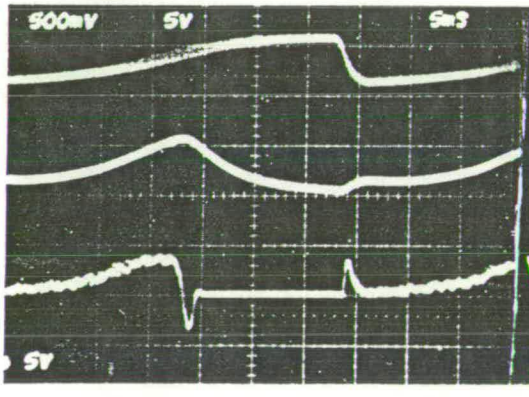


Figure 4 b from top to bottom: "S"-curve, 1st derivative, 2nd derivative, Surface voltage = +1.5V.

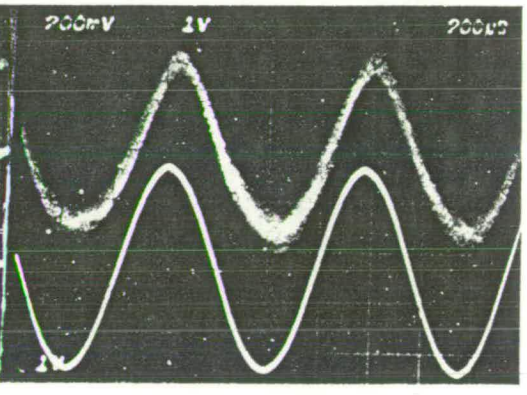


Figure 4 d Upper sinewave output voltage. Lower sinewave surface voltage.

was 2kHz and the modulation is barely visible on the "S"-curves of figures 4.a and b. Therefore, it is believed that by accurately aligning the filters and optimising the modulation depth, the system will eventually be able to track the minimum process clock frequency of 5kHz.

5. Conclusions

A voltage contrast instrumentation system has been developed which can track surface voltages using second harmonic suppression. The preliminary results are encouraging with a system bandwidth of 1-2kHz which is adequate for a parametric sampling system. With accurate alignment of the system filters and optimisation of modulation depth an improvement in signal-to-noise ratio and system bandwidth should be attained.

References

Feuerbaum H P 1979 Proc. 12th SEM Symposium IITRI, Chicago 1 285
Flemming J P 1970 Proc. 3rd SEM Symposium IITRI, Chicago 467
Hannah J M 1974 Ph.D. Thesis University Of Edinburgh
Lukianoff G W, Wolcott S J and Morrissey J M 1981 Digest of Papers IEEE Int. Test Conf. 68
Macari M, Thangamuthu K and Cohens S 1982 Proc. Int. Phys. and Rel. Symp. 163
Ostrow M, Menzel E and Kubalek 1981 Microcircuit Engineering 514
Ranasinghe D and Khursheed A 1983 Proc. Microscopy of Semiconducting Materials, to be published
Taylor N J 1969 Rev. Sci. Instrum. 40 6 792
Wolfgang E, Fazekas P, Otto J and Crichton G 1982 Hardware and Software Concepts in VLSI (Amsterdam: Van Nostrand)