IAEA-TECDOC-363

SELECTED TOPICS
IN
NUCLEAR ELECTRONICS

FOREWORD

For the last 25 years, the IAEA has been convening training courses on nuclear electronics; 23 courses have been held so far. During these years the courses have covered a specific approach of how to teach and train engineers, scientists and technicians in the sophisticated techniques of modern nuclear electronics has been developed. Thanks to several nuclear electronics experts who have been associated with the IAEA training courses for a number of years, this methodology today permits the transfer of specific and detailed knowledge of contemporary nuclear electronics to the participants in a training course in a relatively short period of time.

The IAEA training courses in this field heavily rely upon practical work in the laboratory. Accordingly, the IAEA has produced and published the Nuclear Electronics Laboratory Manual (TECDOC 309) where the experience on organizing the efficient practical training in nuclear electronics was compiled. The present publication is focused on the theoretical understanding of basic electronic circuits and is of particular importance to the attendees of the IAEA training course; the lecturing periods in these courses have been reduced to less than 40 percent of the available time and it is essential that the students have a suitable textbook to study the selected topics on their own.

The intention of the textbook is to find a reasonable balance between the phenomenological description of electronic circuits and the full and correct mathematical description of the phenomena. Any higher mathematical concepts are avoided, or are exposed only to the extent that they build a bridge for the student for further in-depth studies. Many practical tips are included in the desire to ease the road to design and repair of nuclear electronics instruments.

The present publication does not copy any available book on nuclear electronics and instrumentation. On purpose, it does not describe the elementary electronics circuits, as applied in nuclear instruments; they can be found in books. It starts the nuclear electronics study on the instruments level, continues to describe the technology and circuitries on the board level, and only in some exceptional cases investigates the circuits on the components level. It is believed that such an approach better reflects the advanced status of nuclear electronics and the philosophy of the modern design of nuclear instruments. For illustration, and wherever it appeared useful, some commercial instruments are described and analyzed.

EDITORIAL NOTE

CONTENTS

CHAPTER 1

FUNCTIONAL APPROACH

1   FUNCTIONAL APPROACH

Only radiation detectors suitable for energy measurements will be considered here: the most widely employed are the gas ionization chambers, the proportional counters, the scintillation detectors and the solid state detectors.

All these detectors deliver at their output a short current pulse whenever ionizing radiation interacts with their sensitive region. THE CHARGE CARRIED BY THE OUTPUT CURRENT PULSE OF THE DETECTOR IS LINEARLY RELATED TO THE ENERGY RELEASED BY THE RADIATION IN THE SENSITIVE REGION OF THE DETECTOR.

Therefore, an energy measurement requires the measurement of the charge carried by a short current pulse. It could be concluded, at a first glance, that this kind of measurement simply implies the time-integral of the current pulse. Unfortunately the solution is not so straightforward because several constraints are present.

First, it may occur that the charge delivered by the detector is small because the energy released is small or because the detector has a low sensitivity or because both reasons are present simultaneously. In these cases, the noise of the first amplification stages degrades the accuracy in the measurement.

Second, there is a possibility that the detector is exposed to a relatively high radiation intensity, in which case the time separation between single events is short. Other causes of accuracy degradation consequently appear. They are related to the reduced time which is available to perform the charge measurement, to the baseline fluctuations induced by the high event rates and to pulse-on-pulse pile-up effects.

The measurement of the energy released by a nuclear event in the sensitive region of the detector has to be considered performed, once the analog quantity which CARRIES THE INFORMATION OF SUCH AN ENERGY, OR THE CHARGE DELIVERED BY THE DETECTOR, IS CONVERTED BY AN ADC INTO A NUMBER AND THIS NUMBER IS STORED IN A DIGITAL MEMORY.

The current pulse delivered by the detector, before being presented to the ADC UNDERGOES SEVERAL OPERATIONS AIMING TO REDUCE THE ACCURACY DEGRADATION ARISING FROM THE PREVIOUSLY QUOTED EFFECTS AND TO DELIVER TO THE ADC A SIGNAL WITH SUITABLE CHARACTERISTICS TO GUARANTEE THE CORRECT ADC FUNCTION. These operations constitute the analog processing.

To explain what the purposes of analog signal processing are, let us start considering the time diagram of figure 1, which shows a hypothetical time-sequence of the current pulses delivered by a detector.

The pulses shown in Fig. 1.1 correspond to different energy releases. Their amplitudes are randomly distributed. Randomly distributed too is the time distance between subsequent pulses. It can be assumed, for the sake of simplicity, that all current pulses have the same width, are of rectangular shape and they are very short in comparison to the average time distance.

Several detectors are very far from respecting the first and second hypothesis, while in some particular experimental situations the third one
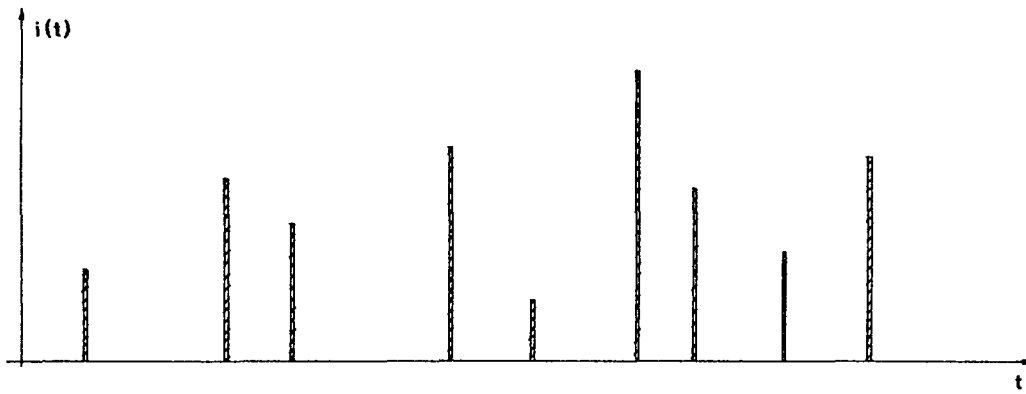
Fig. 1.1: Hypothetical time-sequence of current pulses delivered by
a detector
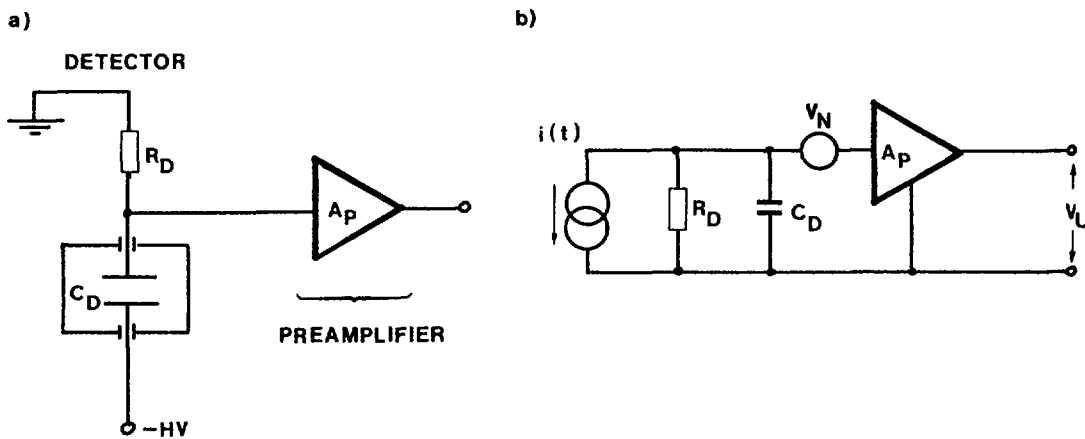
Fig. 1.2: Gas ionization chamber connected to a voltage-sensitive
preamplifier
a) real circuit showing the HV supply connection
b) equivalent circuit
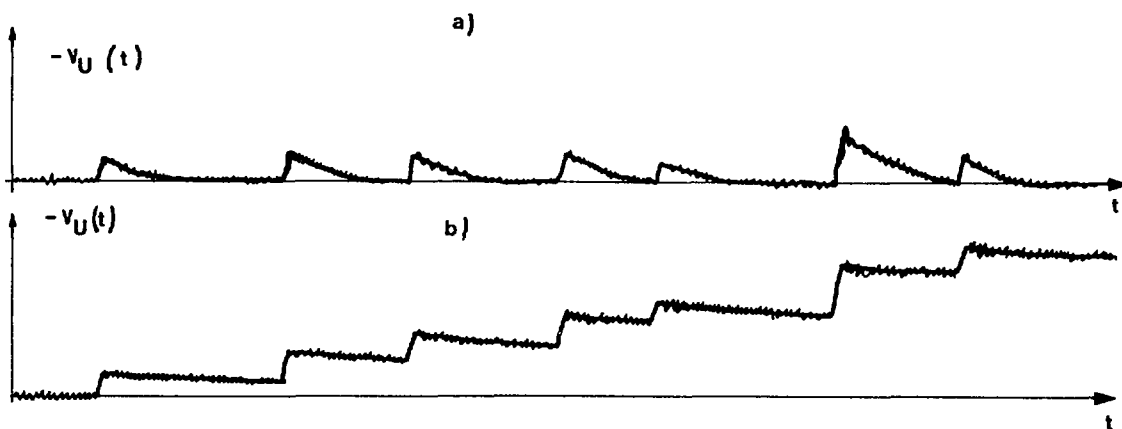
Fig. 1.3: Waveforms at the output of the preamplifier of Fig. 1.2
a) the time constant $R_D C_D$ is shorter than the average
time spacing between pulses

b) the time constant $R_D C_D$ is much longer than the
average time spacing between pulses.

may also be violated. They will, however, be maintained now, just to reach some basic conclusions in a very straightforward way.

REMEMBER THAT THE PARAMETER CARRYING THE ENERGY INFORMATION IS THE CHARGE ASSOCIATED WITH THE DETECTOR CURRENT PULSES.

The first idea which comes to the engineer's mind is to integrate the detector current pulse on the detector capacitance C_D and then to amplify the voltage step by simply using a voltage-sensitive preamplifier. The solution is shown in Fig. 1.2 in the case of a gas ionization chamber, for which this approach is really followed.

Assuming for i(t) the time-dependence of figure 1 and the sign accounted for by the arrow of figure 2b), which results from the fact that the electrode where the signal is taken in Fig. 1.2.a works on negative charge collection, the wave-forms of Fig. 1.3 would appear at the preamplifier output.

The parallel $R_D C_D$ circuit of Fig. 1.2 responds to a very short current pulse of area Q with a voltage signal

$$- \frac{Q}{C_D} \cdot e^{-t/R_D C_D}$$

Therefore the V_u(t) preamplifier output resulting from a single current pulse delivered by the detector would be:

$$V_u(t) = - A_P \cdot \frac{Q}{C_D} \cdot e^{-t/R_D C_D} + A_P \cdot V_N(t) \qquad \text{(Eq. 1.1)}$$

where A_P is the voltage gain of the preamplifier and V_N(t) is the noise, accounted for in Fig. 1.2.b by the generator V_N.

The behaviour of V_n(t) in the case of a sequence of current pulses like that of Fig. 1.1 is shown in Fig. 1.3. Fig. 1.3.a refers to the situation of a time constant $R_D C_D$ shorter than the average spacing between the pulses. Each voltage signal recovers to zero before the next current pulser arrives.

In Fig. 1.3.b it is assumed, instead, that $R_D C_D$ is much longer than the average time spacing between the current pulses. As a consequence the voltage pulses across the detectors as well as those at the preamplifier output pile-up on each other. In this case, a function called CLIPPING must be performed at the preamplifier output to reduce the width of the pulses.

CLIPPING IS THE OPERATION WHICH REDUCES THE TIME CONSTANT OF THE EXPONENTIALLY DECAYING PULSES AT THE PREAMPLIFIER OUTPUT, LEAVING THEIR AMPLITUDE UNCHANGED.

Referring to a single pulse at the preamplifier output the clipping operation gives the result shown in Fig. 1.4, where the noise has been neglected for the sake of clarity.

The clipping operation, applied to the staircase-rising wave-form of Fig. 1.3.b reproduce a sequence of non overlapping exponentially decaying pulses, as shown in Fig. 1.5.
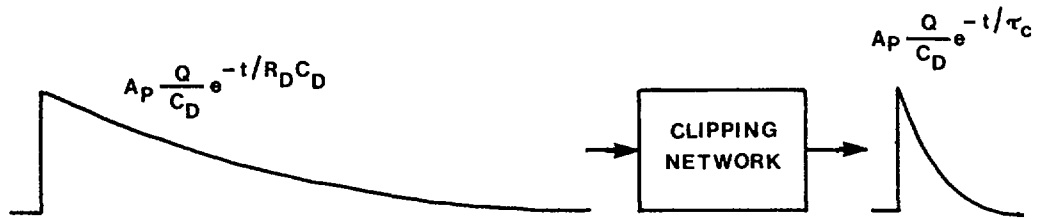
Fig. 1.4: Effect of clipping network on the slowly decaying
preamplifier pulses.

Fig. 1.5: Effect of clipping on the waveform arising from
overlapping pulses
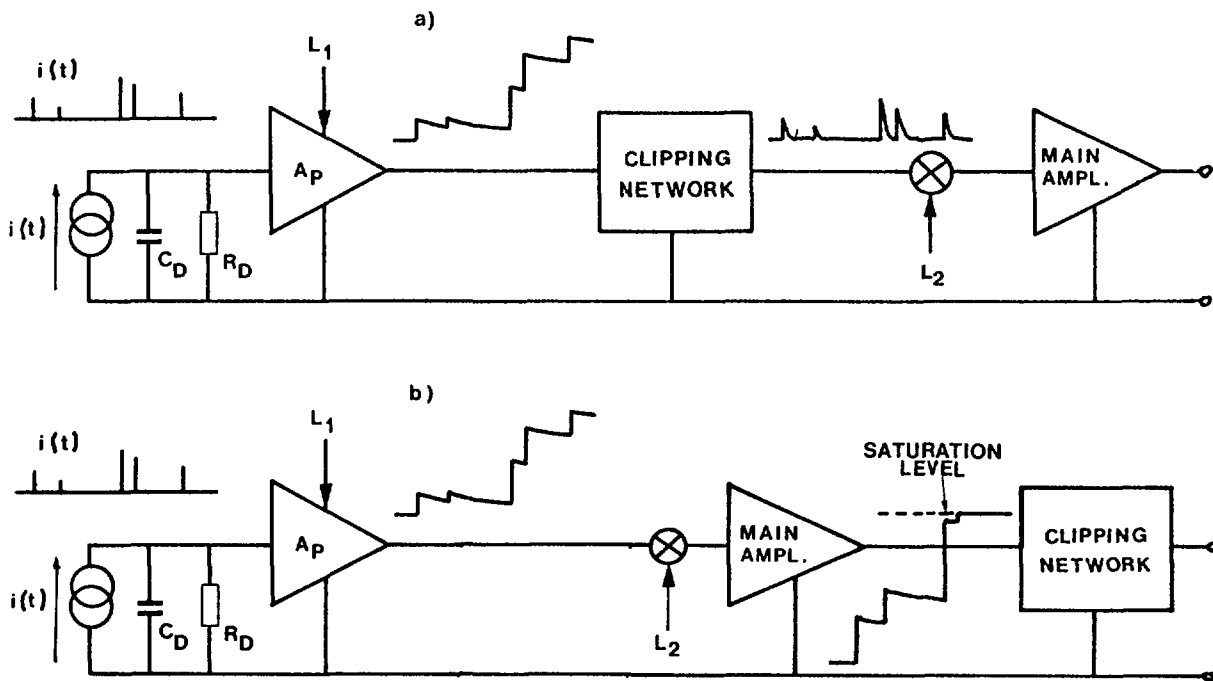
Fig. 1.6: Comparison between advantages and disadvantages of two
different positions of the clipping network in the analog
processing channel.

At the output of the clipping network of Fig. 1.5 the individual
pulses are again separated from each other.

It has to be pointed out that a CLIPPING NETWORK LEAVES UNCHANGED THE
ABRUPT CHANGES, FOR INSTANCE THE LEADING EDGE IN THE INPUT SIGNAL, BUT
STRONGLY ATTENUATES THE SLOWLY VARYING PARTS, LIKE THE LONG EXPONENTIAL
TAILS.

12

As to the effect of the clipping network on the noise, it has, therefore, to be expected that:

THE HIGH FREQUENCY COMPONENTS OF THE NOISE ARE LEFT UNCHANGED BY A CLIPPING NETWORK, WHILE THE LOW FREQUENCY COMPONENTS ARE ATTENUATED.

A clipping network is therefore a high-pass filter which is useful in rejecting LOW FREQUENCY noise contributions, like HUM and MICROPHONICS.

Considering this function of the clipping network, it might seem more reasonable to introduce it in a position closer to the end of the analog processing channel rather than immediately at the preamplifier output. Doing so, it may become effective in attenuating the low frequency noise which is generated AFTER the preamplifier.

For a better understanding of the advantages and disadvantages which may arise from a clipping network introduced immediatly after the preamplifier OR close TO THE end of the analog processing channel, the situation described in Fig. 1.6, a) and b) can be compared.

In Fig. 1.6 two sources of low frequency noise, L_1 and L_2 are considered, one inside the preamplifier and one at the input of the main amplifier, accounting for the noise pick-up in the connection between preamplifier and amplifier. In the case represented by Fig. 1.6.a the signals are clipped at the preamplifier output. The clipping network has no effect on the low frequency noise L_2, which appears unattenuated at the input of the main amplifier and then is amplified by the same factor as the signal. In Fig. 1.6.b the clipping network is at the end of the processing channel and therefore attenuates all the low frequency noise contributions. The main amplifier works on piled-up signals and if the amplification is large enough, the saturation level of the main amplifier may be reached. As shown in Fig. 1.6 the saturation level is reached after the fourth step in the piled-up signal. THE SIGNALS COMING AFTER THAN SATURATION IS REACHED ARE LOST. They obviously cannot be recovered by the clipping network.

To avoid saturation and to provide in the meantime some degree of rejection to all the contributions of low frequency noise it is a good practice to use two or more clipping networks of progressively shorter time constants, as shown in Fig. 1.7.
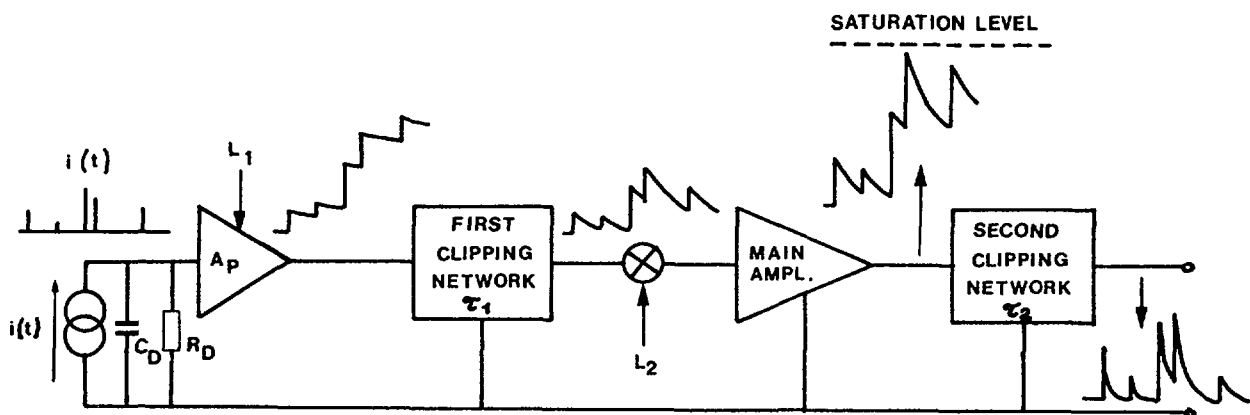

Fig. 1.7:   Use of two clipping networks to guarantee a suitable low frequency rejection without introducing problems of saturation in the main amplifier.

After this discussion it can be concluded that a situation like that shown in Fig. 1.3.a where the voltage signals across the detector have already a short decay time constant, so that a clipping network is unnecessary to shape the signal, is not satisfactory from the point of view of low frequency noise.

In such a situation both low frequency terms L_1 and L_2 would appear to the output of the processing channel without encountering any filter.

Therefore: TO AVOID SIGNAL DEGRADATION FROM LOW FREQUENCY NOISE, KEEP THE HIGHEST POSSIBLE $R_D C_D$ TIME CONSTANT, THAT IS, USE THE LARGEST POSSIBLE DETECTOR BIAS RESISTOR. USE TWO OR MORE CLIPPING NETWORKS OF PROGRESSIVELY SHORTER TIME CONSTANTS IN ORDER TO AVOID SATURATION IN THE AMPLIFIERS AND TO GUARANTEE IN THE MEANTIME A SUITABLE ATTENUATION OF THE LOW FREQUENCY NOISE THROUGHOUT THE ANALOG PROCESSING CHANNEL.

Pay attention to the following aspect.

A CLIPPING NETWORK IMPLEMENTS A TWOFOLD FUNCTION.
IT SHAPES THE SIGNAL TO A SHORTER WIDTH.
IT PERFORMS ON THE NOISE A HIGH PASS ACTION, FOR IT IS A HIGH PASS FILTER.

For this reason the words "shaping" or "filtering" are sometimes considered synonyms in nuclear electronics.

Once the signals are shaped to a suitably short duration so that overlapping no longer occurs and the individual signals can be recognized, it might be observed that the information about the charge Q is carried by the amplitude of the leading edge the exponentially decaying signal and that the energy measurement requires therefore the determination of the peak amplitude of the signal with respect to the baseline (see, for instance, Fig. 1.4).

If the measurement of Q were made in this way, however, it would not be accurate enough. The reason is that the processing discussed so far is either a large bandwidth one, like that implemented by preamplifier and main amplifier, or is of high-pass nature, like that implemented by the clipping networks. No high frequency limitations have yet been introduced and therefore the measurement would be badly affected by the high frequency noise in the preamplifier.

The measurement of Q requires a LOW PASS FILTERING to reduce the effect of the high frequency noise. The simplest way of implementing low pass filtering is through an approximate RC integrator (either passive or active) with a time constant equal to the shortest adopted in the clipping network.

A possible structure of a complete analog channel is shown in Fig. 1.7, where the main amplifier has been split into two parts.

In modern nuclear amplifiers the low pass filter may assume more sophisticated structures that have the advantage over the single integrator of Fig. 1.8 of a final signal more symmetric around the maximum than the signals 5 and 6. So, low pass filters employing several integrations to give SEMI-GAUSSIAN shapes are usual. The basic function is anyway, that of reducing the effects of high frequency noise, so that the single integrator can be considered now as reference low pass filter.
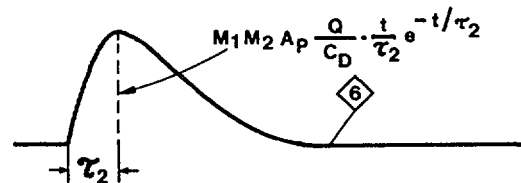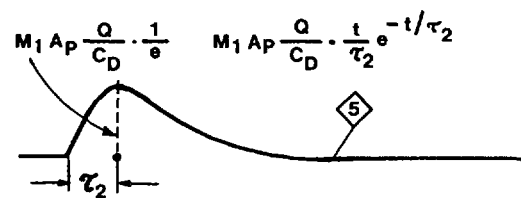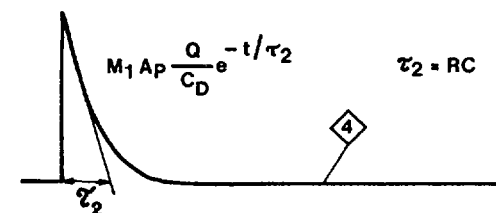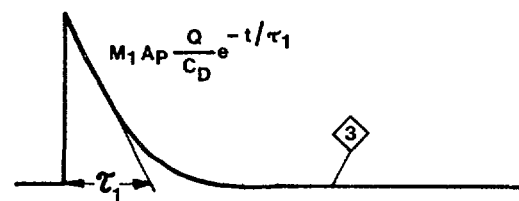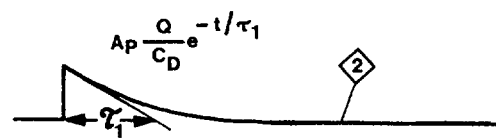
14
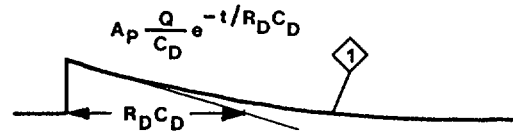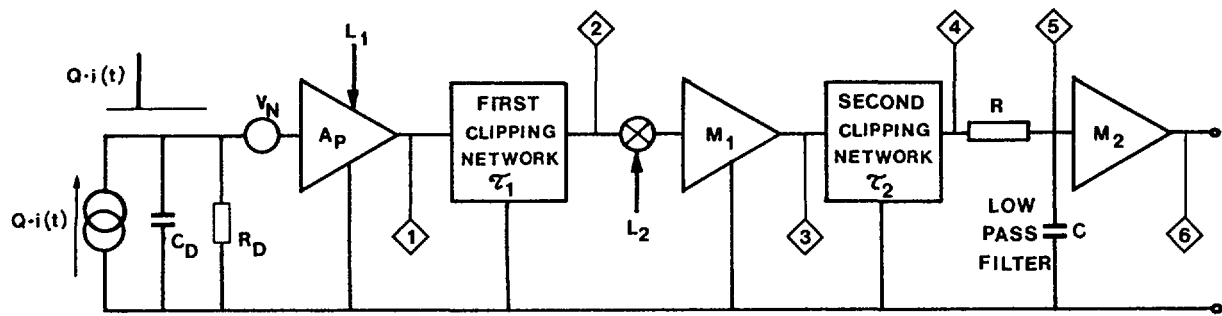
Fig. 1.8: Complete preamplifier-amplifier channel with shaping networks. The waveforms in the various points are shown.

The analog processing channel of Fig. 1.8 performs the following functions: GAIN, CLIPPING OR HIGH-PASS FILTERING AND LOW-PASS FILTERING.

An analog processing channel performs other functions in order to deliver a suitable signal to the analog-to-digital converter. A very important function is called BASELINE RESTORATION. So far, it has been assumed that the signals shown in Fig. 1.8 have a single polarity and that they return to zero without crossing the zero line. In practical cases, however, this does not occur. The signals sometimes have unwanted tails of opposite sign. These tails are due to badly designed CLIPPING NETWORK or to ac couplings inside the analog channel. The tails, for the individual signals may even be hardly noticeable for they are very small. Observing, for example, the signal b of Fig. 1.8 with a high sensitive oscilloscope dc coupled to the test point, the signal would, in a real case, look as in Fig. 1.9.
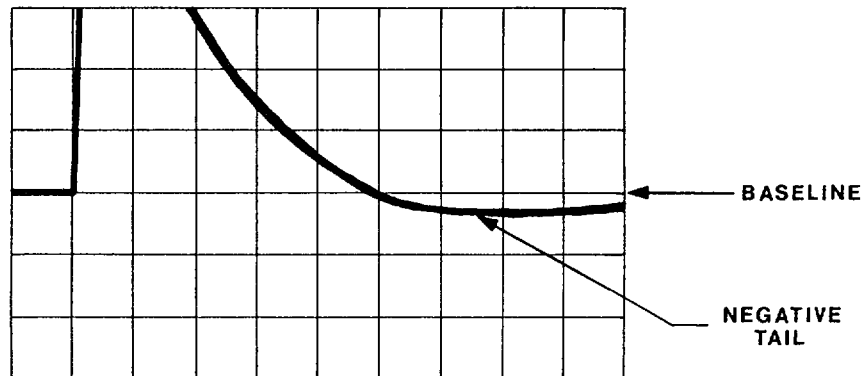
Fig. 1.9: Output signal from an improperly adjusted preamplifier-amplifier-shaping channel, observed with a high sensitivity oscilloscope.

The negative tail which follows the signal is put into evidence.

The negative tail shown in Fig. 1.9 is usually very small. However when the pulses occur at high rates, the tails pile-up thereby creating a negative average voltage by the mechanism described in Fig. 1.10.

As explained by Fig. 1.10, a transient process takes place upon the arrival of the first pulses. The first pulse is followed by a tail. It will be shown later that if the amplifier works in its linear region the area of the positive pulse is equal to the area of the negative tail. Whenever the tail comes from UNDESIRED effects, its amplitude V_T (see Fig. 1.10) is much smaller than the signal and consequently it lasts long. It is assumed in Fig. 1.10 that the tail has an exponential shape

$$V_T \cdot e^{-t/\tau_T}$$

The second pulse falls on the tail of the previous one, which makes it appear smaller if it is measured from the baseline. Besides, the tails of the first and second pulse add, thus further depressing the baseline. The third pulse induces a further depression and the process continues until a balance is reached.
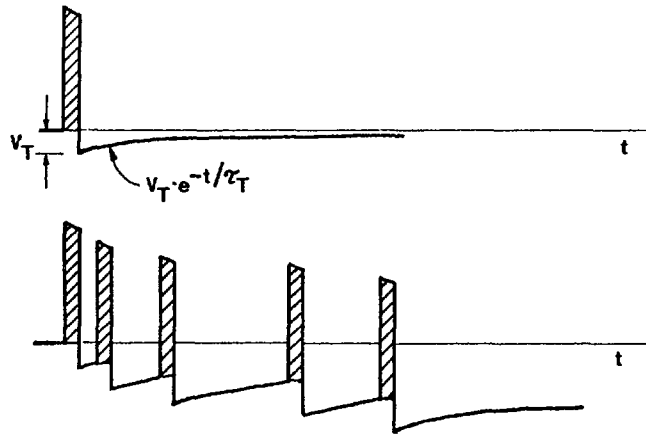
Fig. 1.10:   Baseline depression due to the superposition of negative
             tails.
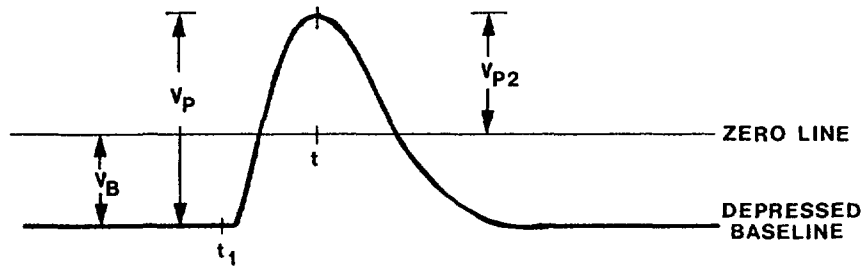
Fig. 1.11:   Amplifier output pulse referred to the depressed baseline

IN THE STEADY STATE, WITH A PULSE RATE OF N PULSES PER SECOND, THE
AVERAGE NEGATIVE VOLTAGE ON THE BASELINE WOULD BE:

$N \cdot V_T \cdot \tau_T$

THAT IS, PROPORTIONAL TO N.

So, a pulse which comes at a given instant to the output of an
amplifier working in a steady condition with an average rate of pulses will
not be referred to zero but to a negative voltage.

As pointed out by the waveform 6 of Fig. 1.9, the information about
the charge Q is carried by the peak amplitude of the signal at the amplifier
output.

If this signal is referred to a baseline which is not at zero volt,
precautions in the measurement of such a peak amplitude must be taken. If
the instrument which senses the peak amplitude PERFORMS THE MEASUREMENT
STARTING FROM THE ZERO LINE, IT WILL RECORD THE VALUE V_{P2}, WHICH IS WRONG.

In this case, it would be necessary to take a sample of the baseline
immediately before the arrival of the pulse ($t = t_1$), the correct peak
amplitude would be given by:

$V_{P2} - V_B$ (V_B is negative)

Alternatively, it is possible to use a circuit which removes the tail
of the pulse or strongly reduces its duration, as shown in Fig. 1.12. Such
a circuit is usually of a nonlinear nature and is called BASELINE RESTORER.

17
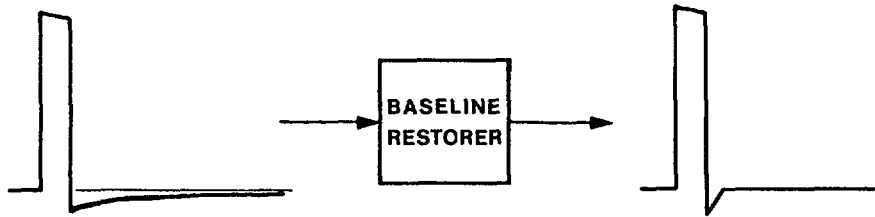
Fig. 1.12:   Operation of a baseline restorer on a single pulse. The
             duration of the negative tail is nonlinearly reduced.

Both methods are described, that is, one based upon baseline sampling
in the absence of pulse and then on subtraction of this value from the peak
amplitude and the other on employing a quick, nonlinear tail recovering to
zero perform the function called BASELINE RESTORATION.

Baseline restoration is an essential operation for all the analog
processing channels that have to work at high duty cycles. So, it is widely
employed in x-ray spectroscopy systems, in gamma-ray amplifiers, in
high-rate charged particles spectrometers and so on.

Some particularly advanced systems are designed to operate without
baseline restoration, in the sense that special care is take in order to
make sure that no tail is present after the signal. These systems, however,
are not of simple use, for they require careful adjustment of the
internal-clipping networks and a careful control of the shape of the output
signal. The reason why systems able to work without baseline restoration
are sometimes considered, lies in the fact that BASELINE RESTORATION IS A
POTENTIAL SOURCE OF DEGRADATION IN THE CHARGE Q MEASUREMENT.

To understand this point, consider the situation shown in Fig. 1.13.
The amplifier signal is smeared by noise and BASELINE RESTORATION is
implemented by sampling the baseline before the pulse arrival and by
subtracting this sample from the peak measurement. The noise is the high
frequency one; it changes quickly enough that the instantaneous values
BEFORE THE PULSE and THE PEAK OF THE PULSE are not correlated. From
Fig. 1.13, where the noiseless ideal pulse is also represented, it is clear
how the baseline restoration increases the error in the amplitude
measurement. This happens because the baseline sampling is affected by the
presence of noise.

If the noise samples at $t = t_1$ and at the peaking instant are
opposite signs, the error in the amplitude measurement would be

$$|\epsilon_p| \; + \; |\epsilon_B|$$

This leads to the conclusion that:

BASELINE RESTORATION SHOULD NOT BE EMPLOYED UNLESS STRICTLY
NECESSARY. THEREFORE, IN ENERGY MEASUREMENTS AT LOW COUNTING RATE NEVER
LEAVE THE BASELINE RESTORATION ON, AS THIS WOULD USELESSLY SACRIFICE THE
SIGNAL TO NOISE RATIO.

Another source of inaccuracy in the charge Q measurement is the
PULSE-ON-PULSE pile-up. Although the pulses are shaped by the clipping
networks and by the lowpass filter with the criterion that their basewidth
must be adequately shorter than the average distance between the detector
current pulses, owing to the fact that the events are randomly distributed

18
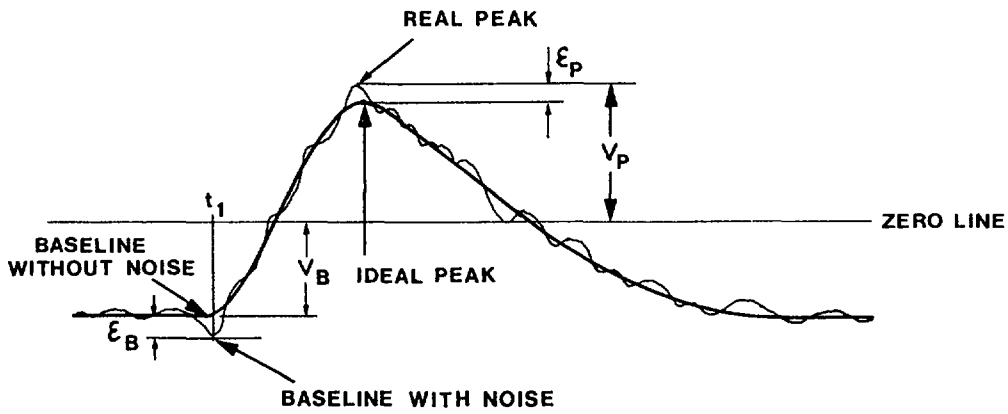
Fig. 1.13:   Effect of baseline restoration on high frequency noise
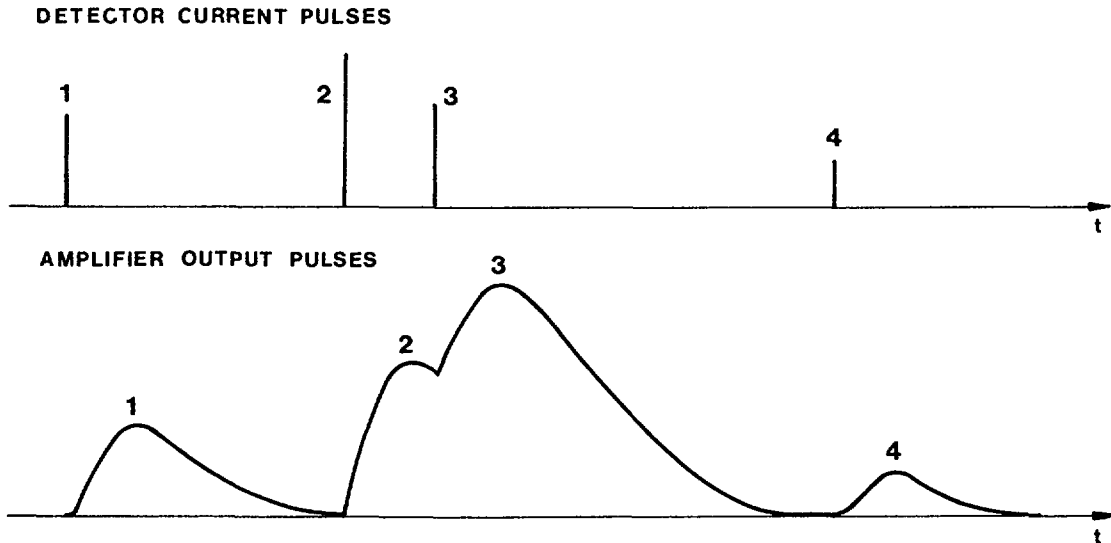
Fig. 1.14:   Pile-up of amplifier pulses.

in time, it may occur that two events superimpose on each other as in Fig. 1.14.

This shows two ionizing events that hit the detector within a time interval shorter than the basewidth of the shaped amplifier pulses. Therefore, the pulses 2 and 3 pile-up and a charge measurement performed on them would lead to a wrong result. For instance, by measuring the peak of the resulting signal it may be inferred that an energy higher than that actually released was left in the detector. In the case of Fig. 1.14, the peak of pulse 2 is reached before the arrival of pulse 3; the information carried by 2 is correct, while that carried by 3 is inaccurate and pulse 3 has to be rejected.

Should, instead, pulse 3 arrive before the peak of 2, then both pulses 2 and 3 have to be rejected.

Rejection of overlapping pulses is realized with an auxiliary unit called PILE UP INSPECTOR. The logic diagram of a pile-up inspector is shown in Fig. 1.15.

As shown in Fig. 1.15, a fast signal is obtained by taking the preamplifier output and passing it through an ideal differentiation. The signals at the output of the differentiation reproduce the amplified

19
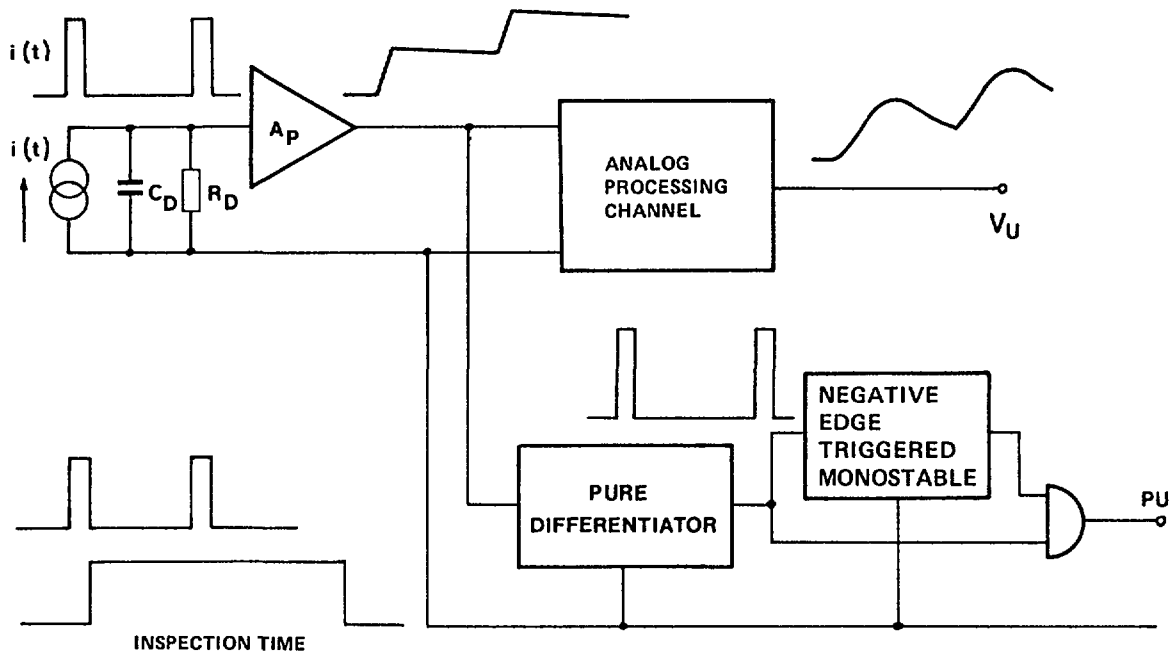
Fig. 1.15: Pile-up inspector

detector current pulses. They are sent to a negative edge triggered
monostable multivibrator which generates a rectangular signal equal in width
to the duration of the shaped pulses at the output of the analog processing
channel. The width of the monostable pulse differentiation is also sent to
one of the inputs of an AND circuit, whose other input is controlled by the
monostable.

Therefore a signal appears at the AND output whenever the second of
the detector current pulses follows the first one within a time period
shorter than the inspection time.

As the monostable is negative edge triggered, a detector pulse cannot
fall within the inspection time generated by itself.

The PILE-UP signal which appears at the AND output can be used as
INHIBIT signal to discard or simply to mark events that have arrived while
the previous one had not yet recovered to zero. A little more elaborated
logic would permit rejection of both pulses when the second one arrives
before the peak of the previous pulse.

If the monostable of Fig. 1.15 is RETRIGGERABLE TYPE, the system
described may be employed to detect multiple PILE-UP. Two more functions
have to be implemented before presenting the signal to the analog-to-digital
converter.

First, it has to be pointed out that not all the detector pulses, in
practical cases, have to be measured. Sometimes, the experimenter has some
constraints to be met for the events that can be accepted. The previous
discussion about pile-up, for instance, showed that only the events that are
not affected by pile-up have to be accepted and the others have to be
discarded. For these reasons the analog processing channel should be
provided with a linear gate whose state is controlled by either external
logic requirement or by the pile-up inspector itself or by both.

Second, as already pointed out, the measurement of the energy
released in the detector is reduced to a charge measurement and because of

20

DETECTOR

A_P

FIRST
CLIPPING
NETWORK
τ_1

M_1

SECOND
CLIPPING
NETWORK
τ_2

R

LOWPASS
FILTER $\quad$ C

M_2

BASELINE
RESTORER

O

C_D $\quad$ R_D

PILE-UP INSPECTOR

PURE
DIFFERENTIATOR

NEGATIVE
EDGE
TRIGGERED
MONOSTABLE

EXTERNAL LOGIC
SIGNAL

LOGIC
INTERFACE

GATE
CONTROL

DELAY

LINEAR
GATE
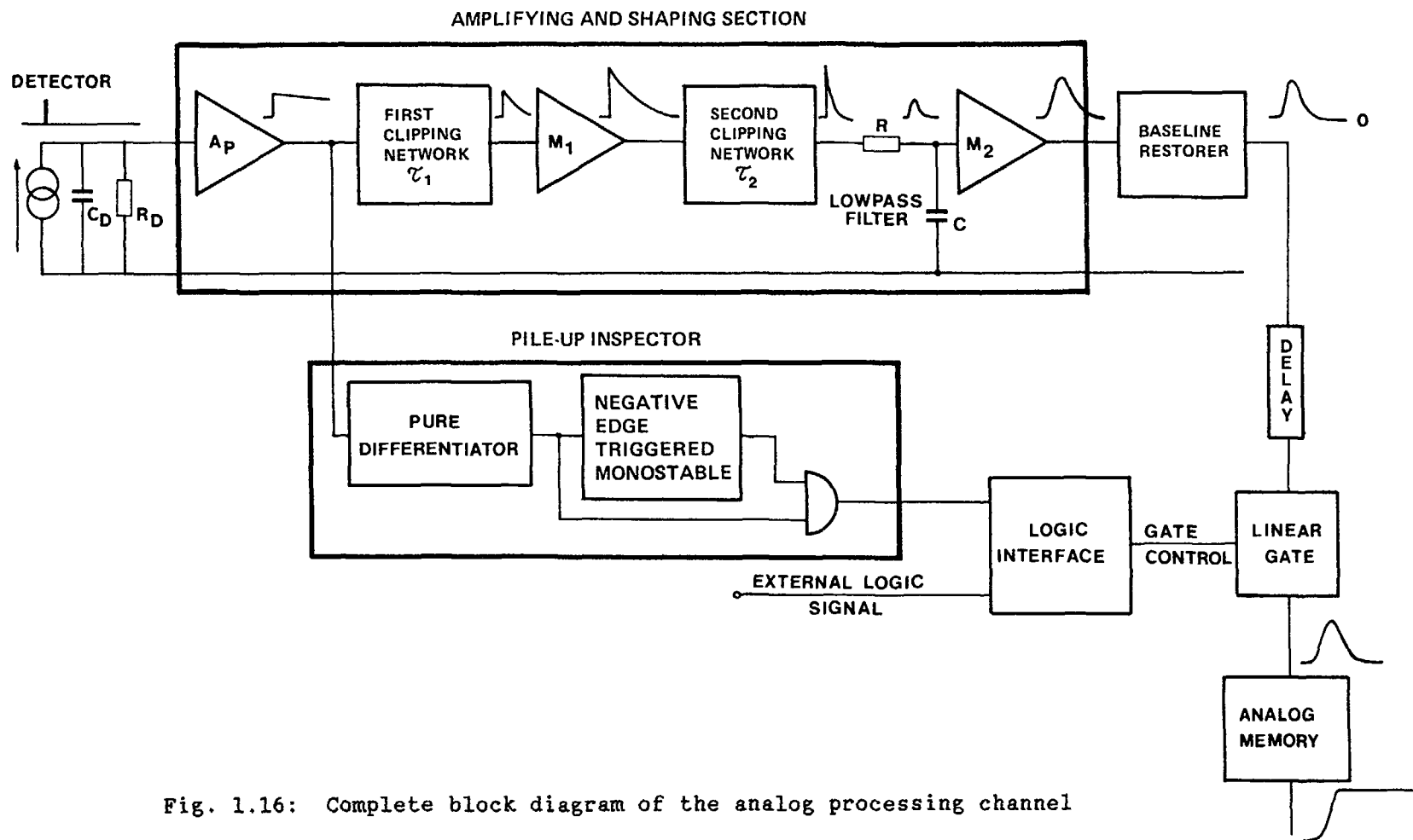
ANALOG
MEMORY

Fig. 1.16: Complete block diagram of the analog processing channel

the particular processing employed, to the measurement of the peak amplitude
of the pulse at the output of the analog processing channel. To convert the
peak amplitude into a number it is necessary in most cases to store it into
an analog memory, as current multichannel analyzers usually require either a
flat-topped pulse or work on the principle of linearly discharging a
capacitor where the peak amplitude of the pulse has been stored. Therefore
an analog memory of the peak-sensing type must be added to the analog
processing channel to interface the output shaped signal to the
analog-to-digital converter.

Summarizing into a single block diagram all the functions discussed
so far, the structure of Fig. 1.16 is obtained. A delay has been introduced
between baseline restorer and linear gate. The delay serves the purpose of
adjusting the mutual time relationship between gate control command, which
is determined by the pile-up inspector and possibly by external logic
conditions and the analog output of the amplifying and shaping section.

The block diagram of Fig. 1.16 reproduces the real structure of a
modern analog channel for energy measurements. The actual system usually
has a more elaborate low pass filter. The disadvantage of the simple
shaping method employed here is that the pulse at the output of the
amplifying and shaping section is not symmetric around the peak and has a
relatively long recovery to the baseline. This requires a long inspection
time in the pile-up circuit and lead to an increased probability of event
rejection becaue of pile-up effects. The best amplifying and shaping
systems use Gaussian or triangular pulses that feature a faster recovery to
the baseline.

CHAPTER 2

CIRCUIT DESIGN

2   CIRCUIT DESIGN

2.1   INTRODUCTORY REMARKS ON CIRCUIT TECHNOLOGIES

The purpose of the previous chapter was to discuss the functions implemented by an analog processing channel, on pulses from nuclear radiation detectors in energy measurements. Those functions were introduced at a block-diagram level. The present chapter, aims at explaining how the different blocks must be implemented from a circuit point of view. Therefore, detailed circuit diagrams will be discussed for preamplifiers, clipping networks, amplifiers, low-pass active filters and so forth. Nowadays, part of the circuits for nuclear pulse processing in energy measurements is realized with monolithic components. Two types of active devices will be employed in this chapter, the OPERATIONAL AMPLIFIER and the ANALOG COMPARATOR. Although the monolithic technology has substantially improved over the last few years, leading to operational amplifiers that suit several requirements in a sophisticated design, such as that of nuclear instrumentation, there are parts in the amplifier and filter section where monolithic components are unsuitable, and discrete design is still employed. For instance, monolithic operational amplifiers are still too noisy to be employed as preamplifiers in low noise applications. Consequently PREAMPLIFIERS FOR IONIZATION CHAMBERS AND SOLID-STATE DETECTORS ARE DISCRETE CIRCUITS. Monolithic operational amplifiers are even too noisy to be employed as first amplification stages, like M_1 in Fig. 1.8. It has to be pointed out, moreover, still referring to Fig. 1.8 that the signal at the input of M_1 may have a rather fast leading edge, because the smoothing action due to the low-pass filter has not yet taken place. It may happen that the speed requirements fixed for M_1 are not met by monolithic operational amplifiers and that, also for this reason, a descrete design is advisable for M_1.

As to M_2, baseline restorer, linear gate and analog memory, it has to be pointed out that they work on relatively slow pulses, at least in most of conventional spectrometry applications. Common values of the low-pass filter time constant RC and of the second clipping time-constant $(T_2 = RC)$ lie in the $1\mu s \div 10$ μs range. In these cases all the circuits following the low-pass filter, with the only possible exception of the linear gate can be implemented using monolithic operational amplifiers as basic building blocks.

The design of linear and nonlinear circuits based upon the operational amplifier, offers several advantages over the discrete approach:

FIRST   The monolithic operational amplifier offers characteristics of dc accuracy (low offset and low offset drift) that can hardly be reached in a discrete design. This is a consequence of the fact that the "CLOSE PROXIMITY" at which the components of the input stage are diffused on the monolithic substrate guarantees a high degree of MATCHING for these components, (which results in a low offset) and in a high degree of TRACKING with temperature (which results in a low offset drift).

SECOND   The design of instruments employing monolithic operational amplifiers as basic active elements has obvious advantages of compactness and reliability, for the number of connections is greatly reduced in comparison to the discrete design.

THIRD   Troubleshooting and maintenance work are highly simplified and a
        defective active element can easily be recognized and replaced.

FOURTH Analog processing can be realized by using a few basic linear and
        nonlinear circuits that employ monolithic operational amplifiers and
        few external elements.

The monolithic process, however, has some limitations that must be
considered in the design of nuclear instrumentation. Please refer to
Fig. 2.1 to understand one of these limitations.
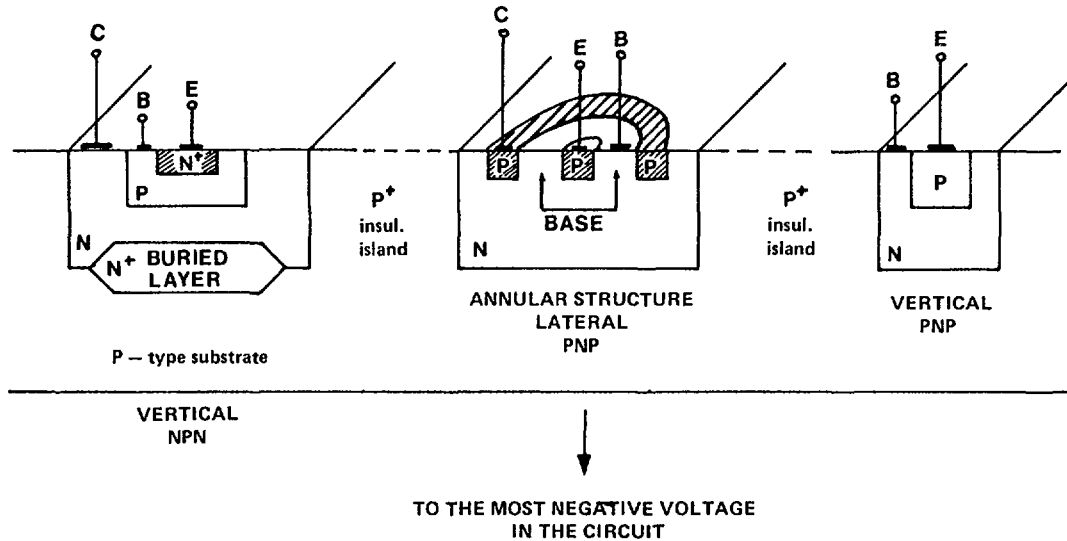
Fig. 2.1: Realization of three different types of transistors on a
           monolithic substrate

The figure shows how the transistors are realized in a monolithic process.
The vertical NPN transistors have the same characteristics as good discrete
elements, this is, typically, at a collector current of a few hundred μA,
the β would be between 100 and 1000 and the f_T cut-off frequency larger
than 500 MHz.

The PNP transistor, on the other hand, is a critical element in the
monolithic process. The lateral structure is shown in Fig. 2.1. Because of
the technological difficulties in realising a thin base for such structure
it usually has, at collector currents of a few tens of μA, a current gain
lower than 50 and an f_T cutoff frequency of only 5 MHz. Both parameters
would drop as the collector current is increased. The vertical PNP of
Fig. 2.1 has better characteristics than the lateral one. However, because
its collector is identical to the substrate, the vertical PNP transistors
can only be employed as a common collector element, that is, as an ouput
stage.

The poor characteristics of the lateral PNP transistors explain why
there are so few monolithic operational amplifiers suitable for the fast
applications met in nuclear electronics. The relatively fast monolithic
operational amplifiers either employ only NPN transistors (in this case
their output signal range is limited to a few volts) or employ special
technologies to realise PNP transistors of high characteristics (then they
are expensive).

There are other limitations in the monolithic process that the
designer of nuclear electronics equipment must know. Resistors have large

absolute tolerances, though the relative tolerances are good (.1% or better). High-valued resistors are difficult to make and consume too much space on the substrate. For these reasons active loads and current sources replace the resistors to a large extent.

Only small capacitors can be integrated on the monolithic substrate: values above 100 pF can be hardly realized.

Monolithic operational amplifiers exist with both bipolar transistor and field-effect transistor input stage. In neither case, however, can the input stage be optimized for low-noise operation.

It has to be remembered, that the output current capability of a monolithic operational amplifier is usually limited to 10 ÷ 20 mA and that the output stage does not transmit equally well negative-going and positive-going pulses. Therefore, in the design of the output stage for a nuclear amplifier intended for bipolar operation, with a cable matched at both ends (if $Z_o = 93ohm$, $V_{OUT} = \pm 10V$, $I_{OUT} = \pm 55mA$) it is advisable to increase the output current capability of the operational amplifier with a discrete power booster, as in Fig. 2.2.
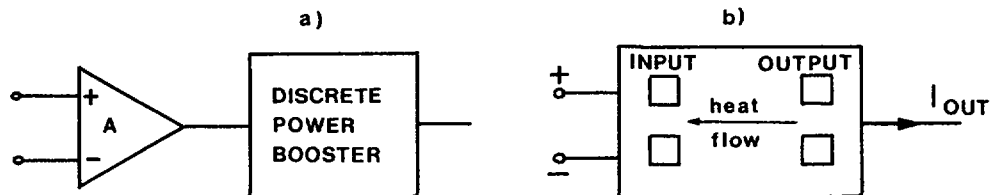
Fig. 2.2: a) Design of output stage with power booster
           b) Thermal feedback on a monolithic substrate

Fig. 2.2 also points out that the use of a separate power booster eliminates the thermal feedback, which otherwise would be present on the monolithic substrate.

The thermal feedback takes place when a large current is drained from a monolithic operational amplifier. The power dissipated by the output stage would increase, a thermal gradient would appear across the monolithic substrate. The resulting heat flow changes the temperature of the input stage, thereby impairing the accuracy of the amplifier.

Most of the discussed limitations of monolithic technology are eliminated with the hybrid one. Hybrid circuits use an insulating substrate and the resistors are realized with either a thin film or a thick film technology. As the hybrid process presents much fewer constraints than the monolithic one, it can be employed also in the realisation of critical circuits, like low noise preamplifiers and wideband operational amplifiers that are out of the reach of monolithic technology.

As to the ANALOG COMPARATOR, it has to be said that very fast units – for instance comparators with a delay of a few nanoseconds, emitter-coupled logic (ECL) compatible, are nowadays available in monolithic technology. The point is that the analog comparator is not required to have a large bidirectional output swing like the operational amplifier and therefore can be made with NPN transistors only.

Below, the basic circuits employing operational amplifiers will be described. For introductory purposes the static characteristics of the operational amplifier will be discussed along with that of the comparator.

The electric symbols of the operational amplifier and of the analog comparator are shown in Fig. 2.3: both circuits have in general a positive and a negative supply, an inverting and a non-inverting input and a single-ended output.
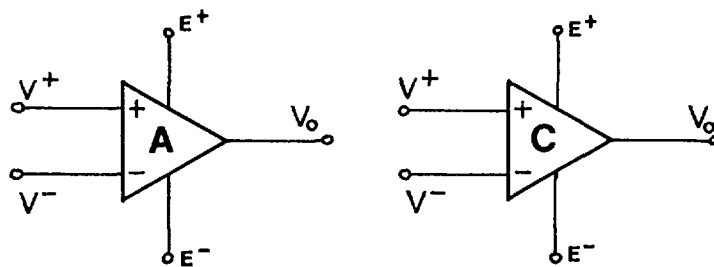
Fig. 2.3: Electric symbols of operational amplifier (A) and analog comparator (C)

OPERATIONAL AMPLIFIER   COMPARATOR
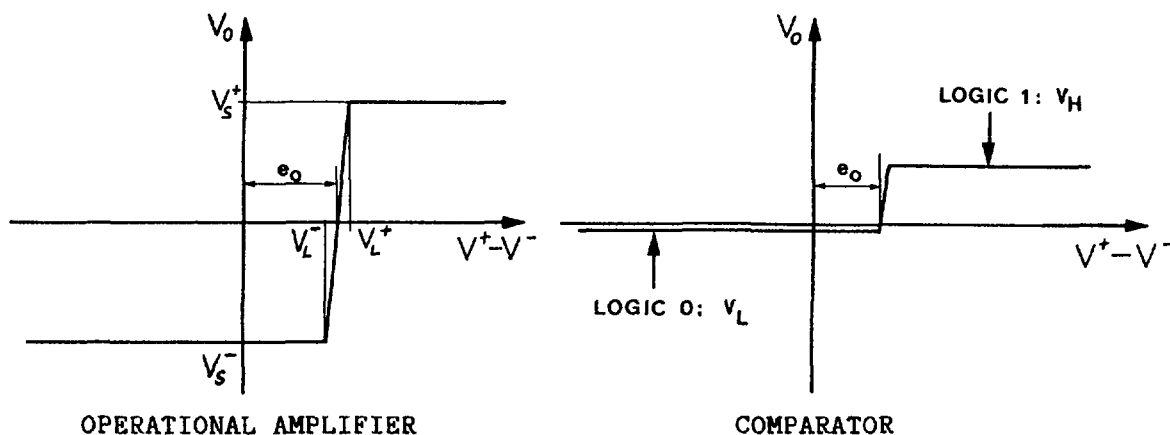
Fig. 2.4: Static input-output characteristics of operational amplifier and analog comparator

If the input signals are referred to 0V, the E^+, E^- supplies of the operational amplifier are symmetric with respect to zero, while those of the comparator are not. Further differences appear in their static input-to-output characteristics, obtained by plotting the output voltage V_O as a function of the input difference $V^+- V^-$. These characteristics are shown for the operational amplifier and for the comparator in Fig. 2.3. Both characteristics present two saturation levels, and a linear region in between. Both characteristics cross the horizontal axis in a point; the distance of this point to the origin is called OFFSET VOLTAGE and represents a deficiency of the real devices. In other words, for an ideal operational amplifier and for an ideal comparator the output voltage should be zero when the input voltage difference is zero.

The offset voltage is less than 1 mV for a good monolithic operational amplifier and is of the order of a few mV for a good comparator. Some special operational amplifiers feature less than 100 μV offset voltage, for others offset adjustment can be obtained through a potentiometric network connected to externally available pins.

As can be deduced from Fig. 2.4 the linear output range for the operational amplifier is symmetric around the zero value. For a well designed circuit, such a range, should differ from the supply interval E^+ E^- by not more than 4V.

For the analog comparator, the two saturation levels must be compatible with the logic levels of the family the comparator is intended to work with. The saturation levels of the comparator of Fig. 2.4 reproduce those of the well known L711, TTL compatible.

The upper and lower saturation levels of the analog comparator must be stable enough to cope with the tolerances of a specific logic family.

Stability of the saturation levels, is less important in the operational amplifiers.

It is important to point out the different static condition in the operational amplifier and in the comparator and the different connections employed to obtain it.

THE STANDING WORKING POINT FOR THE OPERATIONAL AMPLIFIER ALWAYS LIES IN THE LINEAR REGION. THIS IS ACHIEVED BY STABILIZING THE WORKING POINT THROUGH NEGATIVE FEEDBACK.

THE STANDING WORKING POINT OF THE ANALOG COMPARATOR ALWAYS LIES ON EITHER SATURATION LEVEL AND THE LINEAR REGION IS CROSSED ONLY DURING THE TRANSITIONS. THE COMPARATOR IS BIASED IN THE OPEN-LOOP CONFIGURATION OR WITH A POSITIVE FEEDBACK CONNECTION.

THE OPERATIONAL AMPLIFIER IS GENERALLY EMPLOYED FOR LINEAR OR NON LINEAR, NON REGENERATIVE APPLICATIONS.

THE COMPARATOR IS EMPLOYED FOR TRIGGERING AND NON LINEAR, REGENERATIVE APPLICATIONS.

2.2 STANDING CONDITIONS FOR THE OPERATIONAL AMPLIFIERS

From now on attention will be concentrated on the operational amplifier, and the first step will be the external bias circuit. As can be seen from Fig. 2.4, the slope of the characteristics in the linear region,

$$\frac{V_S^+ - V_S^-}{V_L^+ - V_L^-}$$

represents the voltage gain. For monolithic commercial units such a voltage gain ranges between 10^4 and 10^6.

As V_S^- and V_S^+ are nearly equal to the supply voltages, say $\pm$ 10V (for $E^- = -12V$, $E^+ = + 12V$) it turns out that the input linear range V_L^-, V_L^+ is very narrow, between 20μV and 2 mV. It is therefore impossible to fix the working point of the operational amplifier in the linear region without making use of a stabilizing network.

The bias network commonly employed in operational amplifiers is shown in Fig. 2.5. Negative feedback is applied from the output to the inverting input via the voltage divider R_2, R_1. For the evaluation of the working point in the circuit of Fig. 2.5, the effect of the input bias currents, I^+, I^- will be taken into account. A simplifying assumption will be made: as shown in Fig. 2.5 the gain of the operational amplifier is considered infinite and the linear region approximated by a vertical straight line. This means that, irrespective of the working point in the linear region,

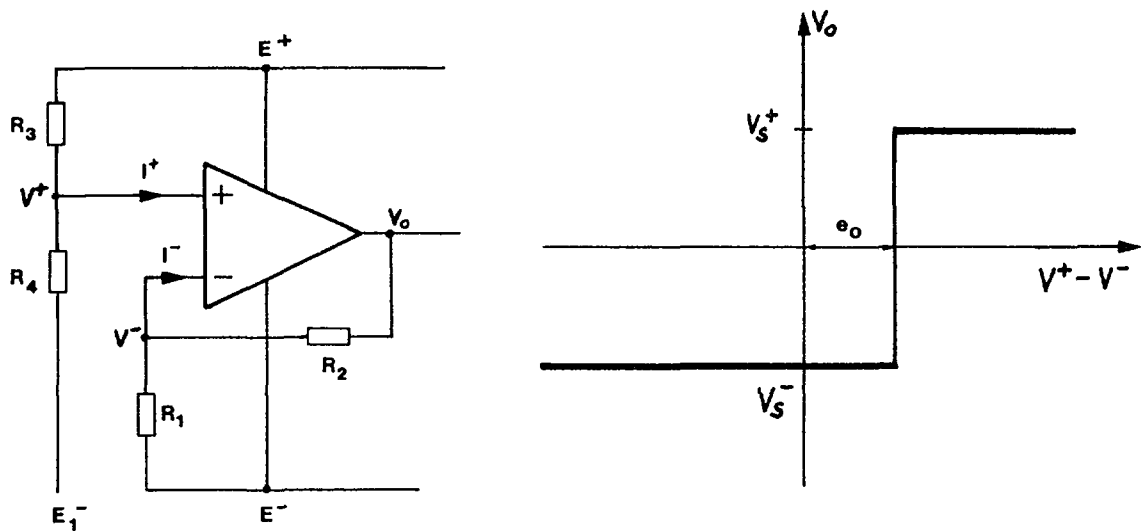$$V^+ - V^- = e_o \qquad\qquad\qquad \text{(Eq. 2.1)}$$

Fig. 2.5: Biasing the operational amplifier

From the analysis of the circuit of Fig. 2.5, the following equations can be written:

A)   Current balance at node V^+ is

$$\frac{E^+ - V^+}{R_3} = \frac{V^+ - E_1^-}{R_4} + I^+ \qquad \text{or} \qquad \frac{E^+}{R_3} + \frac{E_1^-}{R_4} = V^+ \cdot \frac{R_3 + R_4}{R_3 \, R_4} + I^+$$

Voltage can be calculated to be:

$$V^+ = \frac{R_4}{R_3 + R_4} \cdot E^+ + E_1^- \cdot \frac{R_3}{R_3 + R_4} - I^+ \cdot \frac{R_3 \, R_4}{R_3 + R_4} \qquad \text{(Eq. 2.2)}$$

2)   Current balance at node V^- is

$$\frac{V_o - V^-}{R_2} = \frac{V^- - E^-}{R_1} \quad E + I^- \qquad \text{or} \qquad \frac{V_o}{R_2} + \frac{E^-}{R_1} = V^- \frac{R_1 + R_1}{R_1 \, R_2} + I^-$$

For V^- we obtain:

$$V^- = \left(\frac{R_2}{R_1 + R_2} \right) \cdot E^- + \left(\frac{R_1}{R_1 + R_2} \right) \cdot V_o - \left(\frac{R_1 \, R_2}{R_1 + R_2} \right) \cdot I^- \qquad \text{(Eq. 2.3)}$$
$$(2)$$

The difference $V^{\pm}$ V^- can now be taken and it can be made use of the fact that $V^+ - V^- = e_o$. Subtracting results of Eq. 2.3 from Eq. 2.2 we get:

$$V^+ - V^- = e_o = \left(\frac{R_4}{R_3 + R_4} \right) \cdot E^+ + E_1^- \cdot \left(\frac{R_3}{R_3 + R_4} \right) - \left(\frac{R_2}{R_1 + R_2} \right) \cdot E^- - \left(\frac{R_3 \, R_4}{R_3 + R_4} \right) I^+ +$$

$$+ \left(\frac{R_1 R_2}{R_1 + R_2} \right) \cdot I^- - \left(\frac{R_1}{R_1 + R_2} \right) \cdot V_o \qquad (3) \qquad \qquad \text{(Eq. 2.4)}$$

The unknown of the problem is V_o. Solving Eq. 2.4 for V_o the following result is obtained:

$$V_o = \left(\frac{R_4}{R_3 + R_4} \cdot \frac{R_1 + R_2}{R_1} \right) \cdot E^+ + \left(\frac{R_3}{R_3 + R_4} \cdot \frac{R_1 + R_2}{R_1} \right) \cdot E_1^- - \frac{R_2}{R_1} E^- -$$

$$- \left(\frac{R_3 R_4}{R_3 + R_4} \cdot \frac{R_1 + R_2}{R_1} \right) \cdot I^+ + I^- R_2 - \left(\frac{R_1 + R_2}{R_1} \right) \cdot e_o$$

<div align="right">(Eq. 2.5)</div>

The expression now found for V_o contains a remarkable amount of information and, as will be shown later, can be employed to analyze the behaviour of the operational amplifier from several points of view, including that of thermal drift and small signal analysis.

The first analysis concerns the accuracy in the working point V_o. As shown by Eq. 2.5 there are error terms determined by the input bias currents I^+, I^-, and by the offset voltage e_o. The error in the working point is given by:

$$- \left(\frac{R_3 R_4}{R_3 + R_4} \cdot \frac{R_1 + R_2}{R_1} \right) \cdot I^+ + I^- \cdot R_2 - \left(\frac{R_1 + R_2}{R_1} \right) e_o$$

and it can be reduced if appropriate measures are introduced. Use, whenever possible, operational amplifiers with a FET input stage. REMEMBER THAT AT ROOM TEMPERATURE AN OPERATIONAL AMPLIFIER WITH JFET INPUT HAS INPUT BIAS CURRENTS IN THE $10^{-11} \div 10^{-10}$ A RANGE, WHILE AN OPERATIONAL AMPLIFIER WITH AN ORDINARY BIPOLAR INPUT STAGE HAS INPUT BIAS CURRENTS IN THE $10^{-8} \div 10^{-6}$ A RANGE.

A remarkable exception is represented by the operational amplifiers employing super β bipolar transistors as input elements, that is, transistors made by a special process leading to values of β of the order of 5.10^3. An example of amplifier of this type is the LM 108 (National Semiconductor) which features an input bias current of less than 1nA at 25°C.

If the input bias currents I^+, I^- cannot be made low by choosing a suitable operational amplifier, keep the values of $R_3 R_4/(R_3+R_4)$ and the resistance R_2 small. It is worth pointing out that with a balanced resistor configuration, that is, with

$$\frac{R_3 R_4}{R_3 + R_4} = \frac{R_1 R_2}{R_1 + R_2}$$

the error term due to the bias currents would become $(I^- - I^+) R_2$, that is, it would be determined by the difference I^+-I^- rather than by the individual values of the bias currents. The term I^+-I^-, called OFFSET CURRENT in a well designed operational amplifier may be for a factor 10 lower than the individual values of I^+, I^-. Therefore, biasing the operational amplifier with balanced resistors gives always a reduction in the error term due to the input currents.

The following conclusions can be drawn:

IN ORDER TO REDUCE THE ERROR IN THE OUTPUT STANDING VOLTAGE DUE TO THE INPUT BIAS CURRENTS, ONCE THE OPERATIONAL AMPLIFIER IS CHOSEN, USE BALANCED RESISTOR NETWORKS, THAT IS, SATISFY THE CONDITION:

$$\frac{R_1 R_2}{R_1 + R_2} = \frac{R_3 R_4}{R_3 + R_4}$$

AND KEEP THE VALUE OF THE FEEDBACK RESISTOR R_2 SMALL. DOING SO, THE ERROR ON V_0 BECOMES:

$$(I^- - I^+) \; R_2 \; - \; \left(\frac{R_1 + R_2}{R_1} \right) e_o$$

As to the error term determined by the offset voltage, there is nothing to be done but chose an operational amplifier with a low offset voltage. It will be shown soon indeed that the factor $(R_1+R_2)/R_1$ which multiplies the offset e_o, determines also the signal gain of the operational amplifier.

Referring again to Eq. 2.5, the thermal drift of the output voltage can be determined by differentiating with respect to temperature T. Temperature dependent parameters are the resistors, the input bias currents and the offset voltage. We shall, however, neglect the temperature dependence of the resistors because suitably stable resistors can be employed in the design of low drift active circuits. The analysis will therefore be restricted to I^+, I^-, e_o. From Eq. 2.5 the following result is obtained:

$$\frac{dV_o}{dT} = - \frac{dI^+}{dT} \cdot \left(\frac{R_3 \; R_4}{R_3 + R_4} \cdot \frac{R_1 + R_2}{R_1} \right) + \frac{dI^-}{dT} \cdot R_2 - \frac{de_o}{dT} \cdot \left(\frac{R_1 + R_2}{R_1} \right) \qquad (\text{Eq. 2.6})$$

Attention must be paid to the signs.

dI^+/dT and dI^-/dt have equal signs either positive or negative. It is worth pointing out that for a given operational amplifier the signs of dI^+/dT and dI^-/dT are known for they depend on the type of input circuit. On the other hand, the sign of de_o/dT is not specified

It is worth now discussing the behaviour of dI^+/dT and dI^-/dT for the two common input stages, the bipolar and the JFET long-tailed pairs, (see Fig. 2.6.)
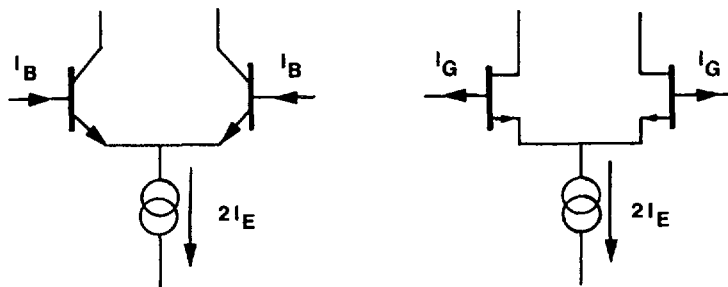
Fig. 2.6: Input bias currents for a bipolar transistor and for a JFET input stage.

For the bipolar transistor long-tailed pair the input current is the base current which has the direction shown by the arrows (in the case of NPN transistors). For a transistor working at an emitter current I_E, the base current I_B is given by:

$$I_B = - \frac{I_E}{1 + \beta} - I_{co}$$

where β is the current gain and I_{co} is the reverse saturation current of the collector-base junction. At fixed I_E, β increases with temperature:

$$\frac{1}{\beta} \cdot \frac{d\beta}{dT} \simeq 10^{-2} \; K^{-1}$$

I_{CO} increases with temperature. In silicon, I_{CO} doubles with every 10K temperature variation.

I_{CO} in a monolithic circuit can be kept quite small, less than 100 pA. In a well designed circuit, where the term $I_B/(1+\beta)$ is two or three orders of magnitude higher than I_{CO} it can be concluded that:

THE INPUT BIAS CURRENT DECREASES AS T INCREASES.

For a junction field-effect input stage, instead, I_G is the reverse current of the gate-to-channel junction and therefore it has the same behaviour as I_{CO} in the bipolar transistor that is I_G doubles every 10K increase in the temperature.

It has to be emphasized that the JFET is a device with low input bias currents at room temperature. At high temperatures a good bipolar transistor operational amplifier may become competitive. To give an example, the LM 108 is based upon super β technology and employs an input circuit connection with transistors working at low base-to-collector voltage. This has the advantage of a reduced I_{CO}, resulting in the behaviour of input bias currents and input offset current as shown in Fig. 2.7.a.

a) **BIPOLAR**
   LM 108

Input Currents

b) **JFET INPUT**
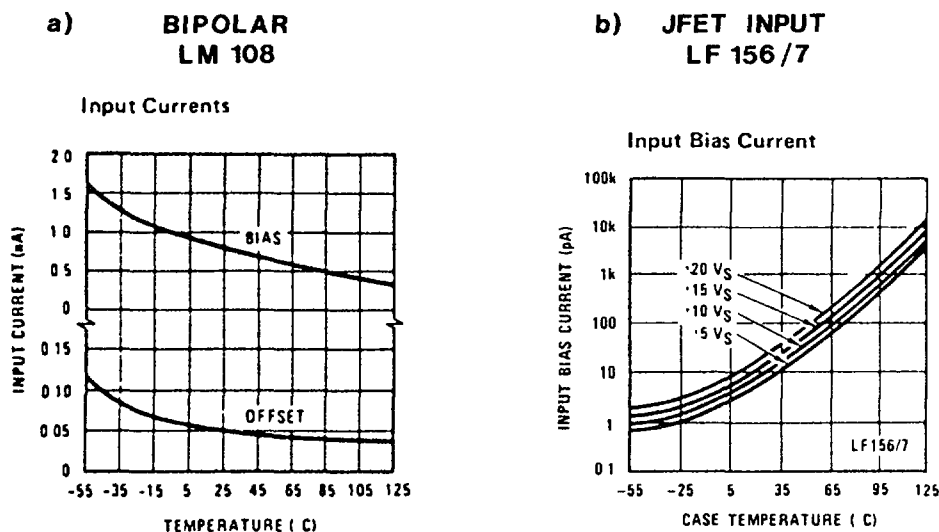   **LF 156/7**

Input Bias Current

Fig. 2.7: Temperature dependence of input bias currents for a bipolar and a JFET input stage.

Fig. 2.7.b shows how the input bias current increases with temperature for a JFET input stage: at 125°C the input bias current is larger than 10nA which is an easily achievable value with a conventional bipolar technology. As a matter of fact, the super β transistors start behaving better than the JFET of Fig. 2.7.b, at a temperature around 65°C. Going back to Eq. 2.6, the thermal drift in the output voltage in the case of balanced resistive networks, (i.e. when the condition $R_3\ R_4/(R_3+R_4) = R_1\ R_2/(R_1+R_2)$ is fulfilled) becomes

$$\frac{dVo}{dT} = R_2 \cdot \left(\frac{d[-(I^- - I^+)]}{dT} \right) - \frac{de_o}{dT} \cdot \left(\frac{R_1 + R_2}{R_1} \right)$$

which shows that the output voltage drift is expressed as a linear combination of offset current drift and offset voltage drift.

The operational amplifier has been considered so far as a black box and the input-to-output static characteristic of Fig. 2.4 has been assumed to describe its behaviour.   Eq. 2.5 which gives the output voltage as a function of the supply voltages and of the undesired contributions (e_0, I^+, I^-) has been deduced from the knowledge of the black-box model of the operational amplifier.   From Eq. 2.5 the behaviour of the fundamental connections of the operational amplifier can easily be obtained, at least for slowly varying signals, as will be shown in Section 2.4.   It can be said, therefore, that the black-box approach of the operational amplifier provides a good amount of information.   Such an approach, as a matter of act, is adequate in most of the conventional applicatons of operational amplifier.   It is not sufficient, however, for the applications of operational amplifiers in the nuclear field where signals with duration in the 1 to 10 μs region have to be dealt with, and where high linearity and stability have to be guaranteed.

It is for these reasons that a diversion in the main line of the presentation is being made now.   This diversion aims at describing the internal design of a real wideband, operational amplifier.   Such a diversion has a twofold purpose.   The first purpose is to make the users interested in applications to the nuclear field, and aware of the limitations of monolithic operational amplifier.   To achieve this goal, a detailed circuit diagram will be investigated.   The circuit assumed as an example can be easily made also in discrete form.   The second purpose is to enable the designers of nuclear instrumentation to realise their own discrete operational amplifiers and therefore to carry on instrumentation development, overcoming the difficulties related with high cost or with nonavailability of suitable operational amplifiers.

Let us take a monolithic wideband operational amplifier.   The internal block diagram is shown in Fig. 2.8.
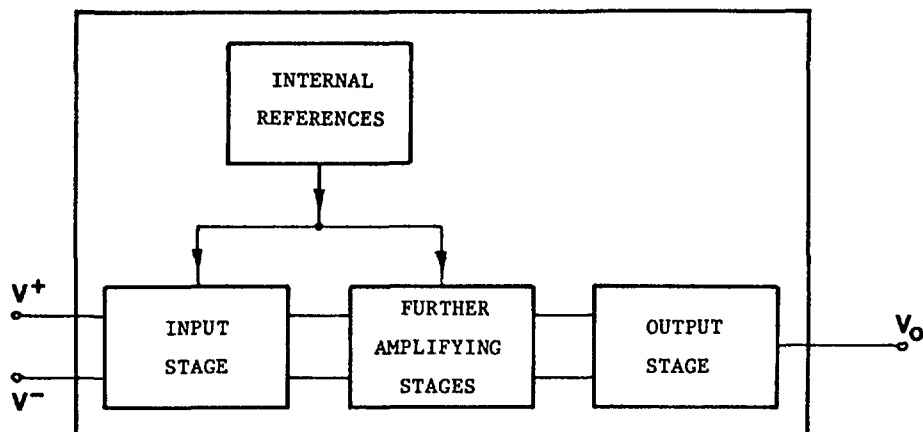
Fig. 2.8:   Block diagram of a monolithic operational amplifier

Voltage and current references required for the amplifying blocks are generated on the monolithic substrate and then distributed to the various amplifying stages.   The signal path goes through the input stage, through further amplifying stages and finally to the output stage.

A simple and yet good operational amplifier configuration is shown in Fig. 2.9.   All the current sources are intended to be derived from a single internal reference not shown in the figure.

34

Fig. 2.9:   Detailed circuit diagram of a monolithic operational
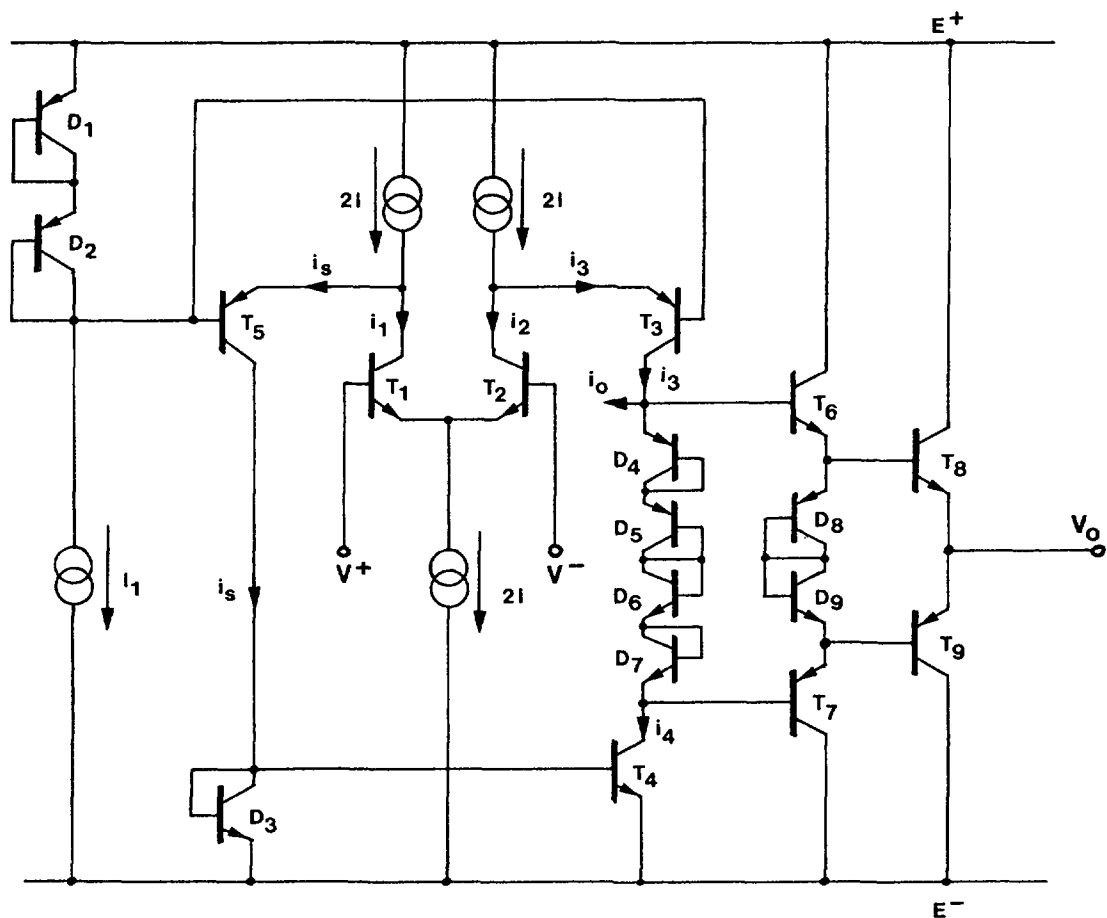amplifier

In the circuit of Fig. 2.9 the transistors, T_1, T_2, T_3, T_4,
T_5 all work at the same current I.   (T_1, T_2) constitute the input
long-tailed pair, which is fed at a constant current 2I by the current source
on the common emitter.   The input long-tailed pair is assumed to be realised
with perfectly matched transistors.   Perfectly matched are also the
diode-connected transistor D_3 and transistor T_4.   If the inputs V^+, V^-
are connected to the same voltage, for instance to ground, the structure T_1,
T_2, T_3, T_5, D_3, T_4 will be balanced with all the elements working at
the same current I.

The D_3, T_4 combination is a current mirror, which has the purpose
of transforming the current flowing in the collector of T_5 into an equal
current absorbed by the collector of T_4.

The output stage is a high impedance-isolation one, consisting of two
complementary emitter followers in cascade (T_6, T_7, and T_8, T_9).   The
two transistors T_8, T_9 are biased by the forward voltage drops in the
diode-connected transistors D_8, D_9.   The transistors T_6, T_7 and the
diodes D_8, D_9 are biased by the forward voltage drop in the chain of four
diode-connected transistors D_4, D_5, D_6, D_7.   If all the PNP
transistors are perfectly matched and so are the NPN transitors, it can be
concluded that (T_6, T_7) and (T_8, T_9) work at the same current I as the
other circuit elements.

The transistors T_3 and T_5 are common base amplifiers that work as
current amplifiers on the collector current signals of T_1 and T_2.   The
base voltage of T_5, T_3 is fixed by the forward drops across the diodes
D_1, D_2 fed by the current source I_1.

35

The behaviour of the operational amplifier on the signals can be analyzed in the following way. A difference signal $v^+ - v^- = e$, $e > 0$ increases the collector current of T_1 and decreases the collector current of T_2. Let $\Delta I = i$ be the current variation in T_1. The variation in the collector current of T_2 will be $-i$, as the sum of the currents in T_1, T_2 is constant. Also constant are the sum of the currents in the emitter of T_3 and in the collector of T_2, and the sum of the collector current of T_1 and of the emitter current of T_5. That is (small letters represent the signals):

$$i_1 + i_2 = 0, \qquad\qquad i_1 = -i_2$$
$$i_1 = i, \qquad\qquad\qquad i_2 = -i$$

Furthermore, we see that
$$i_5 + i_1 = 0 \qquad\qquad i_3 + i_2 = 0$$
$$i_5 = -i_1 \qquad\qquad\quad i_3 = -i_2$$
$$i_5 = -i \qquad\qquad\qquad i_3 = i$$
$$i_4 = i_5 = -i$$

The total current variation at the collectors of T_3 and T_4 is $i_0 = +i_3 - i_4$ flowing out of the collector node. That is: $i_0 = 2i$.

Calling Z_0 the impedance seen between the collector of T_3 and ground, the output voltage signal is

$$v_0 = 2iZ_0. \qquad\qquad\qquad\qquad\qquad\qquad \text{(Eq. 2.7)}$$

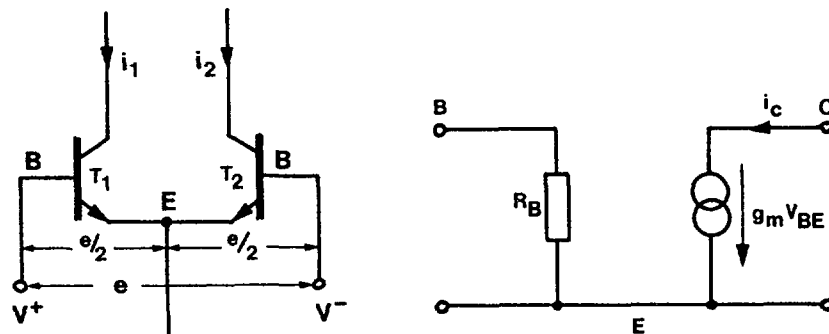Next, i and Z_0 have to be related to the circuit parameters.

Fig. 2.10: Evaluation of the transconductance of a long-tailed pair

The relationship between the input signal e and the collector currents i_1, i_2 in the long-tailed pair can be easily found with reference to Fig. 2.10, where the simplified equivalent circuit of the bipolar transistor is also shown. Owing to the symmetry in the long-tailed pair it can be written:

$$v^+ - v_E = \frac{e}{2}$$

$$v_E - v^- = \frac{e}{2}$$

Therefore T_1 and T_2 behave as grounded-emitter transistors with a base-to-emitter signal equal respectively to $+e/2$ and $-e/2$.

Employing the equivalent circuit of Fig. 2.10, we obtain:

$$i_1 = g_m \cdot \frac{e}{2} \qquad\qquad\qquad i_2 = -g_m \cdot \frac{e}{2}$$

where $g_m = \dfrac{qI}{kT} = \dfrac{I}{V_T}$, and $V_T = 25$ mV at T = 300K

Let us recall:

k = Boltzmann's constant, $1.38 . 10^{-23}$ Joule/K
q = electron charge $1.6 . 10^{-19}$ Coulomb
T = absolute temperature

From Eq. 2.7 it follows:

$$v_O = g_m \cdot Z_O \cdot e$$

and now Z_O has to be evaluated. As shown by Fig. 2.9., Z_O is the impedance seen looking into the collectors of the two elementary amplifiers connected as in Fig. 2.11.
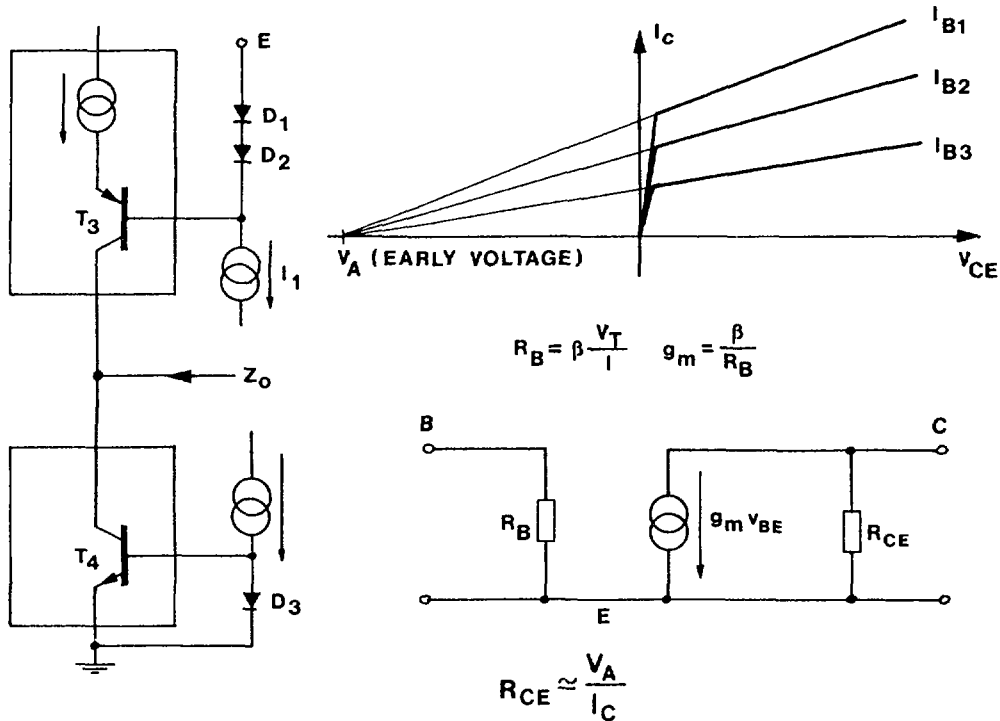

$$R_B = \beta \frac{V_T}{I} \qquad g_m = \frac{\beta}{R_B}$$

$$R_{CE} \cong \frac{V_A}{I_C}$$

Fig. 2.11: Evaluation of the impedance Z_O.

It is shown in Fig. 2.11 that the impedance Z_O is the parallel combination of the output impedance of the PNP grounded base amplifier T_3 and the output impedance of the current mirror D_3, T_4. To evalute both output impedances the equivalent circuit of Fig. 2.11 will be used where the collector-to-emitter dynamic resistance R_{CE} has been included. The expression of R_{CE} is written in Fig. 2.11, it is the ratio between the EARLY VOLTAGE V_A and the standing collector current I_C. In a conventional monolithic process for operational amplifiers

$V_A \cong 100$ V   for NPN transistors
$V_A \cong 60$ V   for PNP transistors

For instance, at a collector current of 100 µA, R_{CE} would be 1 Mohm for an NPN transistor and 0.6 Mohm for the PNP transistors.

The output impedances of the grounded-base amplifier and of the current mirror can now be evaluated.
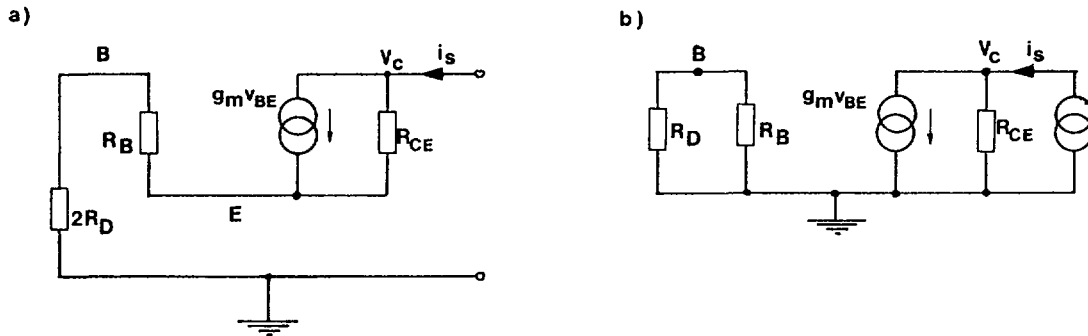
Fig. 2.12: Equivalent circuits for impedance evaluation.

Evaluation can be based on the equivalent circuits of Fig. 2.12 a) for the grounded base amplifier and Fig. 2.12.b for the grounded emitter amplifier. In both circuits the resistance of the biasing diodes has been accounted for. Thus in Fig. 2.12.a, $2R_D$ is the resistance of diodes D_1 and D_2 in series (see Fig. 2.9 and R_D in Fig. 2.12.b is the resistance of diode D_3.

From Fig. 2.12.b it is immediately concluded that the OUTPUT IMPEDANCE OF THE GROUNDED EMITTER AMPLIFIER IS R_{CE}.

The calculation of the input impedance of the circuit (Fig. 2.12.a) is less straightforward. The probe current i_s injected to measure the output impedance, develops a voltage v_C, which can be expressed as:

$$v_C = v_C - v_E + v_E \qquad \text{(Eq. 2.8)}$$

The current i_s flows entirely through the series combination $R_B + 2R_D$. Therefore: $v_E = i_s (R_B + 2 R_D)$ and $v_{BE} = - R_B i_s$.

Moreover, writing the current balance at node V_C:

$$i_s = (v_C - v_E) \cdot \frac{1}{R_{CE}} + g_m \cdot (-R_B i_s)$$

whence: $V_C - V_E = R_{CE} \cdot i_s + g_m R_{CE} R_B i_s$ $\qquad$ (Eq.2.9)

Remembering now that $g_m R_B = \beta$ and substituting Eq. 2.9 into Eq. 2.8 the following result will be obtained:

$$v_C = (1 + \beta) \cdot R_{CE} \cdot i_s + (R_B + 2 R_D) \cdot i_s$$

As ordinarily $R_{CE} > R_B$, $R_{CE} > R_D$, it can simply be stated that the output impedance of the grounded base amplifier

$$\frac{v_c}{i_s} \text{ is } \beta R_{CE}$$

THE OUTPUT IMPEDANCE OF THE COMMON BASE AMPLIFIER IS ABOUT β TIMES HIGHER THAN THAT OF THE GROUNDED EMITTER AMPLIFIER.

Returning now to the evalution of Z_0, from Fig. 2.11 it can be said that Z_0 is the parallel combination of:

$$\beta_{PNP} \frac{V_{A\,PNP}}{I} \qquad \text{(grounded base)}$$

and

$$\frac{V_{A\,NPN}}{I} \qquad \text{(grounded emmitter)}$$

The voltage gain of the operational amplifier, v_o/e is given, as already pointed out, by:

$$g_m \, Z_o.$$

Z_o can be approximated as:

$$\frac{V_{A\ NPN}}{I}$$

Remembering the expression of g_m, the $g_m \, Z_o$ product will be written as:

$$g_m \cdot Z_o = \frac{I}{V_T} \cdot \frac{V_{A\ NPN}}{I} = \frac{V_{A\ NPN}}{V_T}$$

The result above is important, as it states that the VOLTAGE GAIN OF THE OPERATIONAL AMPLIFIER OF Fig. 2.9 IS INDEPENDENT OF THE CURRENT AT WHICH THE TRANSISTORS WORK, AND IS SIMPLY THE RATIO BETWEEN THE EARLY VOLTAGE OF THE NPN TRANSISTOR AND THE THERMAL VOLTAGE V_T.

For normal values of V_A, the gain at room temperature is around 40.000.

Obviously the above calculation refers to the unloaded operational amplifier. Loading the output gives a decrease in gain.

The practical circuit of operational amplifier discussed so far is an example of a circuit which looks attractive because of its simplicity. As shown, it allows a reasonably high gain. If implemented with discrete elements, so that the PNP transistors T_3 and T_4 do not introduce appreciable bandwidth limitations, it can be fast enough to suit most of the requirements of nuclear pulse processing. Gain-bandwidth products between 30 MHz and 100 MHz are easily achievable.

If designed in discrete form, the circuit of Fig. 2.9 should be modified by replacing the diodes D_1 and D_2 with a 4-to-6V reference diode, in order to leave a larger voltage margin for the 2I current sources that feed the emitters of T_3 and T_5. The bases of T_3, T_5 should be filtered by a 0.1 µF capacitor to ground.

2.4   SIGNAL BEHAVIOUR OF BASIC OPERATIONAL AMPLIFIER CONNECTIONS

Now that the fundamentals about operational amplifiers have been introduced and that a practical design of a circuit which may represent a monolithic realisation as well as a suggestion for a discrete implementation has been discussed, the signal behaviour of operational amplifiers will be analyzed.

The basic signal connections of operational amplifiers are shown in Fig. 2.13.

For the signal analysis the same symbols employed to obtain Eq. 2.5 will be adopted, but the signals will be represented by small letters. So, v^+ and v^- are the voltage signals at the noninverting and inverting input respectively, v_o is the output voltage and e^+, e^-, i^- are the voltage and current input signals.

The signal input-to-output relationships are obtained by differentiating Eq. 2.5 with respect to E^+, E^-, I^-. It has to be pointed out, that the signal connections of Fig. 2.9 can be obtained as
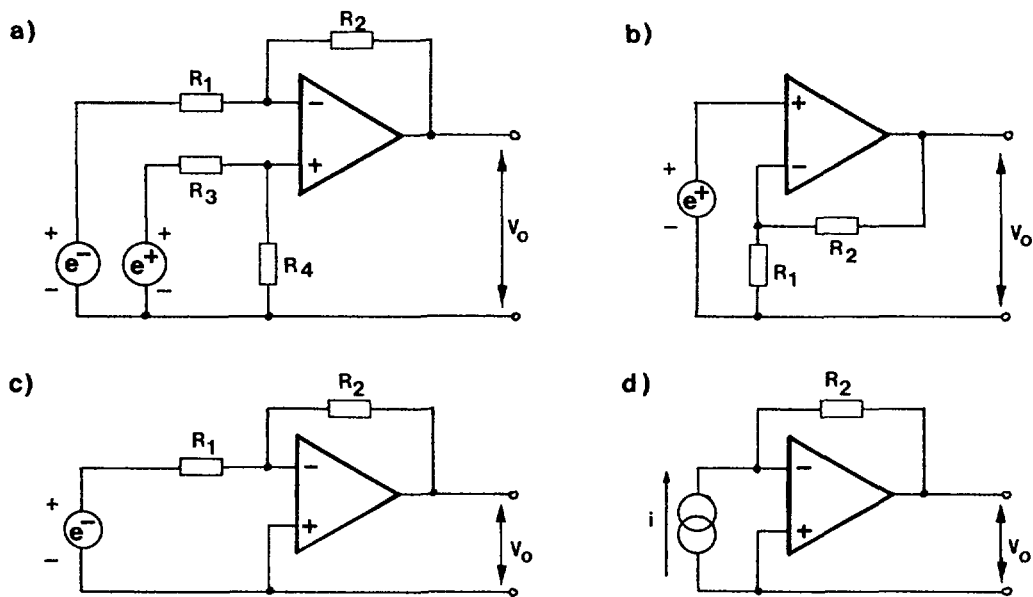
Fig. 2.13:   Different connections of signal generators
            a)   two generators configuration
            b)   one generator non inverting connection
            c)   one generator inverting connection
            d)   current source driving

particular cases of the general connection of Fig. 2.5 and that the signal input-to-output relationships for the four circuits a), b), c), d) of Fig. 2.13 can be deduced from the general Eq. 2.5, as is shown below.

By comparing the circuit of Fig.2.13.a with that of figure Fig. 2.5 can be seen that the signal generator e^+ of the former is in the same circuit position as E^+ in the latter, that the lower terminal of resistor R_4 in Fig. 2.13.a is connected to ground and that the generator e^- of Fig. 2.13 is in the same circuit position as E^- in Fig. 2.5. This means that e^+ and e^- can be considered as the differentials of E^+ and E^-:

$$e^+ = \delta E^+, \qquad e^- = \delta E^-$$

Moreover I^-, I^+, e_0 have to be treated as constants. Differentiating Eq. 2.5 and remembering that $v_0 = \delta V_0$, the following result is obtained:

$$v_0 = \left(\frac{R_4}{R_3 + R_4} \cdot \frac{R_1 + R_2}{R_1} \right) \cdot e^+ - \frac{R_2}{R_1} \cdot e^-$$

If, in particular, $R_1 = R_3$, $R_2 = R_4$, then

$$v_0 = \frac{R_2}{R_1} \cdot (e^+ - e^-) \qquad\qquad \text{(Eq. 2.10)}$$

which shows that the circuit of Fig. 2.13.a, with a balanced resistor configuration realizes THE DIFFERENCE BETWEEN THE TWO INPUT SIGNALS WITH GREAT ACCURACY. It is also worth pointing out that by differentiating Eq. 2.1 the following important result is obtained for the signals at the inverting and noninverting input:

$$v^+ - v^- = 0 \qquad \text{or} \qquad v^+ = v^- \qquad\qquad \text{(Eq. 2.11)}$$

HAVING ASSUMED INFINITE GAIN FOR THE OPERATIONAL AMPLIFIER, THE SIGNALS AT THE INVERTING AND NONINVERTING INPUTS ARE EQUAL. NO VOLTAGE DIFFERENCE APPEARS BETWEEN THE TWO INPUTS.

ONE GENERATOR, NONINVERTING CONNECTION, Fig. 2.13.b. The signal input-to-output relationship for this circuit can be obtained as a particular case of Eq. 2.8 simply by putting $e^- = 0$, $R_3 = 0$, $R_4 \rightarrow \infty$. Doing so, Eq. 2.8 gives:

$$v_o = \frac{R_1 + R_2}{R_1} \cdot e^+ \qquad \text{(Eq. 2.12)}$$

which shows that the noninverting configuration gives a gain which is always larger than 1 and becomes equal to 1 in the case of the buffer or voltage-follower circuit ($R_2 = 0$) shown in Fig. 2.14.
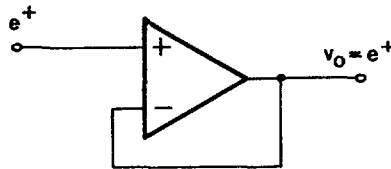

Fig. 2.14: Gain one noninverting configuration (output stage)

A buffer realized with the connection of Fig. 2.14 and with an operational amplifier of finite voltage gain A has a gain which is slightly less than one. Its value is $A/(1 + A)$. The circuit of Fig. 2.14 is a very good impedance translator. Remembering the signal relationship $v^+ = v^-$, it turns out that no signal current can be absorbed from the generator because the two input terminals of the operational amplifier are at the same voltage. This means that the input impedance of the circuit of Fig. 2.14 would be infinite in the ideal case. The total negative feedback ($e^- = v_o$), as a second advantage, keeps the output impedance very low, to a fraction of an ohm.

ONE GENERATOR, INVERTING CONNECTION, (Fig. 2.13.c)
The signal input-to-output relationship for the circuit of Fig. 2.13.c can be obtained as a particular case of Eq. 2.18, by putting $e^+ = 0$. The following result is obtained:

$$v_o = - \frac{R_2}{R_1} \cdot e^- \qquad \text{(Eq. 2.13)}$$

which shows that the inverting configuration allows values of the gain whose magnitude can be smaller or larger than 1. This remark is important in view of the analog switching applications of the operational amplifier.

Another important remark concerns the voltage at the inverting input. The relationship given in Eq. 2.13 applied to the present case in which $e^+ = 0$ and therefore $v^+ = 0$ gives

$$v^- = 0 \qquad \text{(Eq. 2.14)}$$

The Eq. 2.14 is the key to the understanding of all the inverting configurations of operational amplifiers in both linear and nonlinear applications. It states that WHEN THE NONINVERTING INPUT IS KEPT AT ZERO SIGNAL VOLTAGE AND A SIGNAL IS APPLIED THROUGH THE SERIES RESISTOR R_1 IN THE CONFIGURATION OF FIG. 2.13.c, NO SIGNAL APPEARS BETWEEN THE INVERTING INPUT

AND GROUND. SUCH A PROPERTY OF THE INVERTING CONFIGURATION IS EXPRESSED BY SAYING THAT THE INVERTING INPUT BEHAVES AS A VIRTUAL GROUND.

This property will be extensively applied in the next sections.

From Eq. 2.13 it can be seen that the condition $R_1 = R_2$ gives $v_0 = -e^-$. The circuit behaves in this case as an inverting buffer, Fig. 2.15.
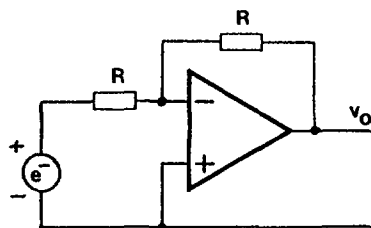

Fig. 2.15: Gain one inverting configuration (output stage)

CURRENT DRIVEN CONFIGURATION, Fig. 2.13.d. The previously discussed property of the virtual ground suggests that it can be employed as an ideal current input. For this reason, a current signal can be injected into the virtual ground and a transresistance amplifier implemented according to the circuit diagram in Fig. 2.13.d. The relationship between input current signal and output voltage signal can be obtained through differentiation from Eq. 2.4 observing that: E^+, $E_{\bar{1}}$, E^-, I^+, e_0 must be kept constant and $I^- = -i$. Therefore:

$$v_0 = -i . R_2 \qquad\qquad (Eq. 2.15)$$

The inverting configuration and the transresistance-type connection are very frequently employed to realize simple active filters. The resistors R_1, R_2 are for this purpose, replaced by impedances obtained connecting L, R, C elements. Active filters based on operational amplifiers usually employ only R,C components because real inductors differ too much from their ideal model. When inductors are required, they can be simulated with active networks employing only R,C elements. Here attention will be restricted to the filters that are of interest in nuclear electronics applications.
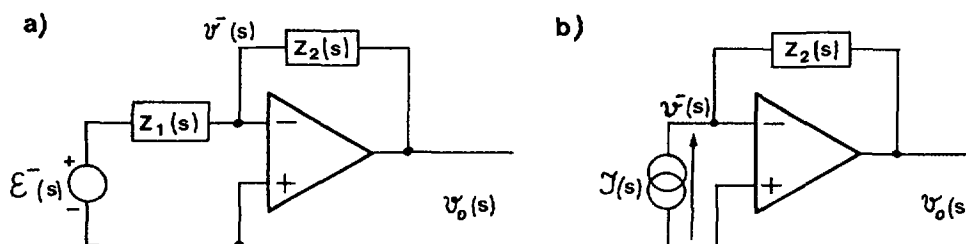

Fig. 2.16: Operational configurations with nonresistive external elements

The circuits of Fig. 2.16 can be analyzed by means of Laplace transform. According to this method the input time-dependent generators e(t), i(t) are replaced by their Laplace transform generators E(s), I(s), where s is the complex Laplace variable. The Laplace-Transform methods assignes to each resistor R a resistance of value R, to each capacitor C an equivalent resistance of value 1/sC and to each inductor an equivalent

42

resistance of value sL. The correspondence is summarized below and it is a particular case of the following general rules: MULTIPLICATION BY A CONSTANT IN THE TIME DOMAIN IS MULTIPLICATION BY A CONSTANT IN THE LAPLACE TRANSFORM CASE: DIFFERENTIATION IN THE TIME DOMAIN BECOMES MULTIPLICATION BY s IN THE LAPLACE TRANSFORM CASE AND INTEGRATION IN THE TIME DOMAIN BECOMES A DIVISION BY s IN THE LAPLACE TRANSFORM CASE.
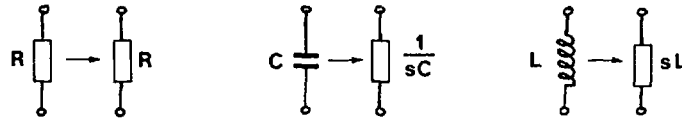
Fig. 2.17: Correspondence between real circuit element and equivalent symbolic resistances according to Laplacetransform method.

Correspondence between some time functions of frequent use and their Laplace transforms is given in Table I.

TABLE I

TIME DOMAIN	LAPLACE TRANSFORM DOMAIN
MULTIPLICATION BY A CONSTANT, K	MULTIPLICATION BY A CONSTANT, K
TIME DIFFERENTIATION	MULTIPLICATION BY S
TIME INTEGRATION	DIVISION BY S
FUNCTION OF TIME	LAPLACE TRANSFORM
δ IMPULSE	1
STEP 1(t)	$\dfrac{1}{S}$
RAMP	$\dfrac{1}{S^2}$
EXPONENTIAL $e^{-\frac{t}{\tau}}$	$\dfrac{1}{S+\dfrac{1}{\tau}}$

The circuits of the type shown in Fig. 2.16 can be analyzed in the following way. The inverting input of the operational amplifier is a virtual ground, therefore $V^-(s) = 0$. The current balance at the inverting input, expressed in the Laplace transform domain gives:

For circuit in Fig. 2.16.a

$$\frac{E^-(s) - 0}{Z_1(s)} = \frac{0 - V_o(s)}{Z_2(s)}$$

whence
$$\frac{V_0(s)}{E^-(S)} = - \frac{Z_2(s)}{Z_1(s)}$$ (Eq. 2.16)

43

For circuit 2.16.b

$$I(s) = -\frac{0 - V_o(s)}{Z_2(s)}$$

whence

$$\frac{V_o(s)}{I(s)} = -Z_2(s) \qquad\qquad (Eq. \ 2.17)$$

The following elementary active filters can now be introduced (see Fig. 2.18).
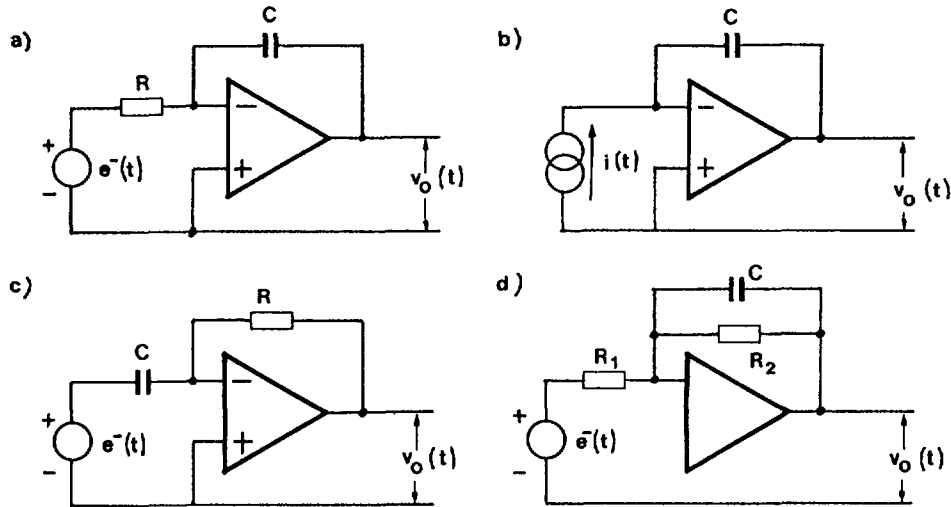

Fig. 2.18: Four different types of elementary active filters of frequent use in nuclear electronics
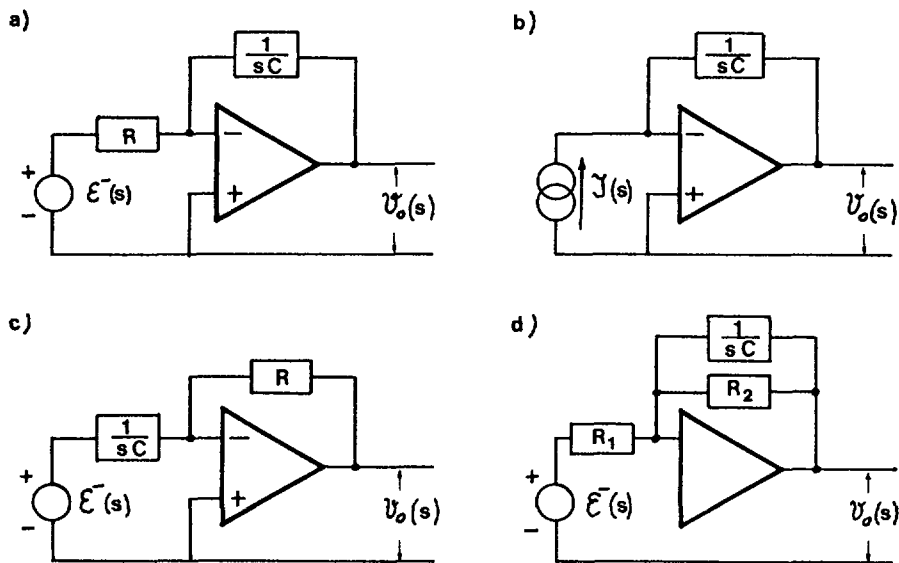

Fig. 2.19: Laplace transform symbolic circuits corresponding to those of circuit in Fig. 2.19.a.

In order to analyze them, the Laplace transform method will be employed. The corresponding symbolic circuits, ready for the application of such a method and obtained from the real circuits of Fig. 2.18 with the correspondence rules outlined in Fig. 2.17 are shown in Fig. 2.19. E(s) in Fig. 2.19 is the Laplace transform of e(t) in Fig. 2.18, so I(s) and V_o

44

(s) are the Laplace transforms of i(t) and v_o(t) respectively. The circuits a), c) and d) of Fig. 2.19 belong to the general type of Fig. 2.16, while b) is of the current drive type.

Applying the Eq. 2.16 we have:

$$\frac{V_o(s)}{E^-(s)} = - \frac{\frac{1}{sC}}{R} = - \frac{1}{sRC} \quad \text{or} \quad V_o(s) = - \frac{1}{sRC} \cdot E^-(s)$$

The last relationship shows that, in the Laplace transform domain V_o(s) is obtained from E^-(s) by multiplying this last by 1/s which corresponds to integration in the time domain and then by multiplying again by the constant 1/RC. As multiplication by a constant in the Laplace transform domain remains multiplication by a constant in the time-domain, the final conclusion is:

$$v_o(t) = - \frac{1}{RC} \cdot \int_o^t e^-(t) \, dt$$

that is, the circuit in Fig. 2.19.a is an integrator for voltage signals.

Circuit in Fig. 2.19.b. As this circuit is current driven, Eq. 2.17 must be used. This gives:

$$\frac{V_o(s)}{I(s)} = \frac{-1}{sC} \quad \text{or} \quad V_o(s) = \frac{-1}{sC} I(s)$$

which shows that the transform of the output voltage is obtained by multiplying I(s) by 1/s

and then by a constant 1/C

So, in the time domain:

$$v_o(t) = \frac{-1}{C} \int_o^t i(t) \, dt$$

that is, the circuit as shown in Fig. 2.19.b is a current integrator.

Circuit in Fig. 2.19.c. The Eq. 2.16 applied to this case gives:

$$\frac{V_o(s)}{E^-(s)} = - \frac{R}{\frac{1}{sC}} \quad \text{or} \quad V_o(s) = - sRC \cdot E^-(s)$$

So, V_o(s) is obtained from E^-(s) through multiplication by s which corresponds to differentiation in the time domain and by multiplication by the constant -RC. Therefore the circuit is a differentiator:

$$v_o(t) = - RC \frac{de^-}{dt}$$

Circuit in Fig. 2.19.d. The parallel combination of R_2 and 1/sC be obtained first. Then:

$$Z_2(s) = \frac{R_2 \cdot \frac{1}{sC}}{R_2 + \frac{1}{sC}} = \frac{R_2}{1 + sR_2C}$$

The Eq. 2.16 can now be applied and gives:

$$\frac{V_o(s)}{E^-(s)} = \frac{\dfrac{R_2}{1 + sR_2C}}{R_1} = -\frac{R_2}{R_1} \cdot \frac{1}{1 + sR_2C}$$

The circuit behaves therefore as an approximate integrator in cascade with an amplifier of gain $-R_2/R_1$.

The active filter of Fig. 2.18.d is commonly used to implement the low pass filtering in nuclear amplifiers.

2.5 OPERATIONAL AMPLIFIER REALIZATON OF THE ENTIRE LINEAR PART OF A NUCLEAR SPECTROMETRY SYSTEM

The principles of operational amplifier applications described so far and the fundamentals of Laplace transform method can be utilized to describe the design of the entire linear part of a nuclear spectrometry system. In other words, it is easy now to analyze the actual structure of the different blocks that perform the operations graphically illustrated by the waveforms of Fig. 8 in the Section 1, devoted to the functional approach of nuclear pulse processing.

In modern spectrometry systems the design is based upon an extensive use of the operational amplifier connections: it has to be emphasized once more that THE PREAMPLIFIER IS USUALLY SUBJECT TO SEVERE REQUIREMENTS ON NOISE REDUCTION AND IMPLEMENTED IN MOST CASES IN DISCRETE FORM. NEVERTHELESS IT IS NOTHING BUT A SPECIAL, LOW NOISE, OPERATIONAL AMPLIFIER CIRCUIT. Most of the commercial low-noise preamplifiers are housed in a small metal box, which performs a shielding function for the external disturbances. The preamplifier box is located as close as possible to the detector to keep the signal connection between detector and preamplifier as short as possible, as dictated by the need of reducing the noise pick-up. It is a normal practice to put inside the preamplifier box also the first amplification stage. In this way the SIGNAL AT THE OUTPUT OF THE PREAMPLIFIER BOX IS NOT TOO SMALL AND THEREFORE IT IS LESS AFFECTED BY THE NOISE PICK-UP IN THE CONNECTION CABLE, WHICH MAY BE UP TO SEVERAL TENS OF METERS IN LENGTH, BETWEEN PREAMPLIFIER AND AMPLIFIER.

Remembering the discussion made in the section "FUNCTIONAL APPROACH" about pile-up and clipping networks it is advisable to put THE FIRST CLIPPING NETWORK INSIDE THE PREAMPLIFIER BOX.

Therefore, the circuitry housed inside the preamplifier box is that shown in the block diagram of Fig. 2.20.
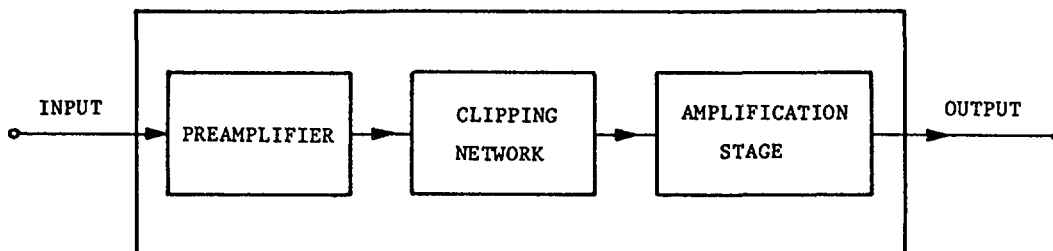

Fig. 2.20: Block diagram of the circuitry housed in the preamplifier box

The actual diagram of the circuitry housed in the preamplifier box is shown in Fig. 2.21, where the detector has also been represented with the equivalent circuitry already discussed in the Section 1, "functional approach".
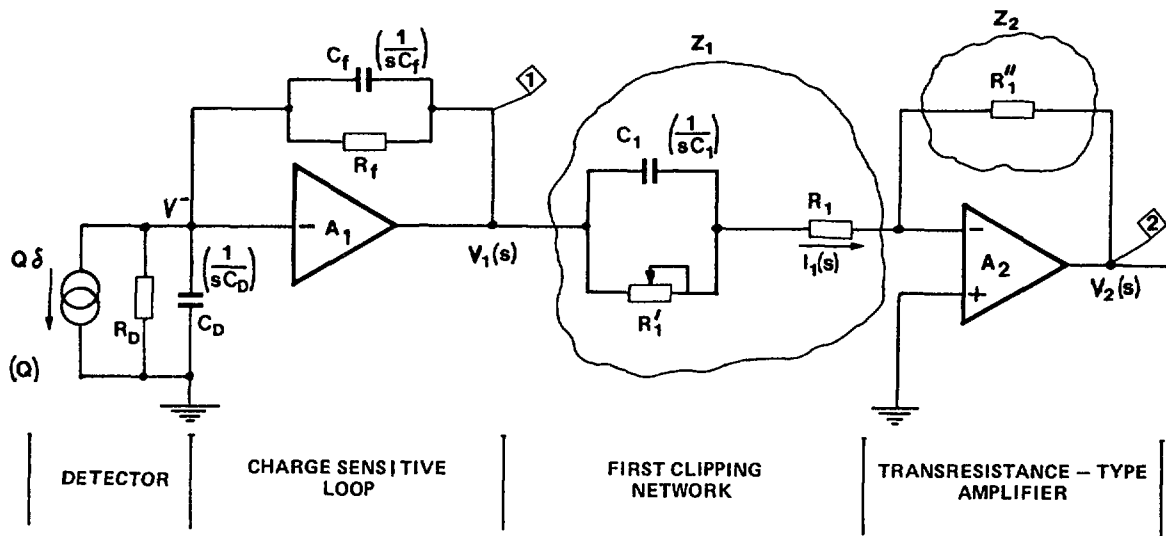

Fig. 2.21: Preamplifier box circuitry using operational amplifier.

In the diagram of Fig. 2.21 the detector signal has been assumed to be an impulse carrying a charge Q. In the same diagram are the quantities that have to be introduced in the Laplace transform analysis shown along with the corresponding time domain parameters. Q is the Laplace transform of the detector current signal and $1/sC_D$, $1/sC_f$, $1/sC_1$ are the equivalent resistances corresponding to C_D, C_f, C_1.

Pay, moreover, attention to the following point. It is not by chance that the operational amplifier in the charge-sensitive loop has been drawn with the INVERTING INPUT ONLY. THE PREAMPLIFER IS BASED ON THE OPERATIONAL AMPLIFIER A_1. A DIFFERENTIAL INPUT CONFIGURATION BEING ALWAYS MORE NOISY THAN A SINGLE-ENDED ONE WOULD NOT BE ADVISABLE IN THE PREAMPLIFIER. Bear in mind that to implement the operational feedback the only STRICTLY ESSENTIAL input is the INVERTING ONE. The NON INVERTING INPUT is useful, but NOT ESSENTIAL.

It is important, moreover, to observe that the actual preamplifier of Fig. 2.21 differs from the simplified structure discussed in the Section 1. The functional approach aimed at introducing basic principles, while the present section discusses the real circuit implementation. Most of MODERN RADIATION DETECTOR PREAMPLIFIERS ARE OF THE CHARGE-SENSITIVE TYPE. THE CHARGE-SENSITIVE PREAMPLIFIER IS AN OPERATIONAL INTEGRATOR.

Observe the following point: THE INPUT VOLTAGE V⁻ IS ZERO BECAUSE THE INVERTING INPUT OF THE OPERATIONAL AMPLIFIER A_1 IS A VIRTUAL GROUND. THEREFORE THE DETECTOR CURRENT SIGNAL CANNOT FLOW ACROSS EITHER R_D OR C_D, BECAUSE THESE TWO CIRCUIT ELEMENTS WORK AT ZERO VOLTAGE. Therefore the detector current signal HAS NO OTHER WAY TO GO THROUGH, BUT THE FEEDBACK NETWORK, PARALLEL COMBINATION OF R_f AND C_f.

Applying this conclusion in the Laplace transform domain, that is, writing the current balance at node V^-:

$$\frac{V_1(s) - V^-(s)}{\dfrac{R_f \cdot \dfrac{1}{sC_f}}{R_f + \dfrac{1}{sC_f}}} = Q \qquad \text{(Eq. 2.18)}$$

Remembering that $V^-(s) = 0$ and rearranging the denominator:

$$V_1(s) = \frac{QR_f}{1 + sR_fC_f} \qquad \text{(Eq. 2.19)}$$

In order to apply the Laplace transform table of this chapter, it is better to write $V_1(s)$ as:

$$V_1(s) = \frac{Q}{C_f} \cdot \frac{1}{s + \dfrac{1}{R_fC_f}} \qquad \text{(Eq. 2.20)}$$

Looking now at TABLE I it is recognized that $V_1(s)$, apart from the constant factor Q/C_f, has the same expression as the Laplace transform of the exponential pulse. Therefore:

$$v_1(t) = \frac{Q}{C_f} \cdot e^{-t/R_fC_f} \qquad \text{(Eq. 2.21)}$$

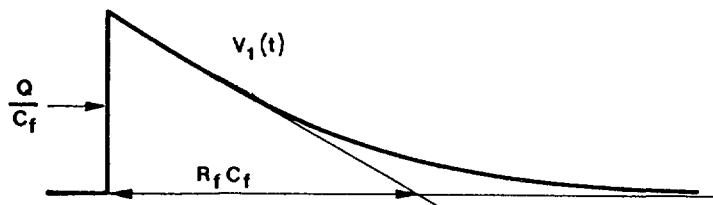The signal $v_1(t)$ is represented in Fig. 2.22.

Fig. 2.22: Signal at the output of the charge-sensitive loop.

It is interesting to compare the signal $v_1(t)$ at the output of the charge-sensitive loop with the signal 1 of Fig. 8 in Section 1. The initial amplitude of $v_1(t)$ is Q/C_f; it is proportional to the charge carried by the detector signal, like in the signal 1 of Fig. 8. However, there is an important difference between the actual case and that analyzed in Section 1. In the relationship between Q and the amplitude of $v_1(t)$, the detector capacitance C_D does not appear. This is the basic advantage of a charge-sensitive preamplifier over a voltage sensitive one: in the charge-sensitive loop the detector capacitance C_D is shunted by a virtual ground, as a consequence the detector current signal cannot be integrated on C_D, it is integrated on the EXTERNALLY ADDED CAPACITANCE C_f. Remember that in semiconductor detectors C_D is not very reliable; it depends on the bias voltage, on the temperature and its value may also be slightly affected by radiation damage effects. Removing the dependence on C_D is a very important property of the charge-sensitive preamplifier.

It is useful to point out that the charge-sensitive preamplifier belongs to the general configuration of Fig. 2.16.b.

48

Going now to the first clipping network and to the output trans-resistance amplifier of Fig. 2.21, it is clear that the relationship between $v_1(s)$ and $v_2(s)$ can be determined as a special case of the Eq. 2.16 applied to the circuit in Fig. 2.16.a, where

$$Z_1(s) = R_1 + \frac{R_1'}{1 + s\,R_1'\,C_1} \qquad\qquad Z_2(s) = R_1''$$

Using Eq. 2.16, the following relationship is obtained

$$\frac{V_2(s)}{V_1(s)} = \frac{R_1''}{R_1 + \dfrac{R_1'}{1 + sR_1'\,C_1}} =$$

$$= \frac{-R_1''\,(1 + SR_1'\,C_1)}{R_1 + R_1' + s\,R_1\,R_1'\,C_1} \qquad\qquad (Eq. 2.22)$$

Eq. 2.22 can also be written as:

$$\frac{V_2(s)}{V_1(s)} = -\frac{R_1''}{R_1 + R_1'} \cdot \frac{1 + s\,R_1'\,C_1}{1 + s\,\dfrac{R_1 R_1'}{R_1 + R_1'} \cdot C_1} \qquad\qquad (Eq. 2.23)$$

$V_2(s)$ can be determined by introducing into Eq. 2.23 the expression of $V_1(s)$ given by Eq. 2.19.

$$V_2(s) = -\frac{R_1''}{R_1 + R_1'} \cdot \frac{1 + sR_1'\,C_1}{1 + s\,\dfrac{R_1 R_1'}{R_1 + R_1'}C_1} \cdot \frac{QR_f}{1 + sR_f C_f} \qquad\qquad (Eq. 2.24)$$

Acting upon R_1', which for the purpose has been shown as a potentiometer in Fig. 2.21, adjust it so as to make

$$R_f\,C_f = R_1'\,C_1$$

Doing so, you implement the pole-zero cancellation, in other words, as shown by Eq. 2.24, the term $1 + sR_1'C_1$ becomes equal to $1 + sR_fC_f$ and they cancel each other. Eq. 2.24 becomes:

$$V_2(s) = -\frac{R_1''}{R_1 + R_1'} \cdot \frac{Q\,R_f}{1 + s\,\dfrac{R_1 R_1'}{R_1 + R_1'}\,C_1} \qquad\qquad (Eq. 2.25)$$

Remember now that $R_f C_f = R_1'\,C_1$, that is,

$$R_f = \frac{R_1'\,C_1}{C_f}$$

and introduce this value of R_f into Eq. 2.25.

$$V_2(s) = - \frac{R_1''}{R_1 + R_1'} \cdot \frac{Q \dfrac{R_1' \, C_1}{C_f}}{1 + s \dfrac{R_1 \, R_1'}{R_1 + R_1'} \, C_1}$$

$V_2(s)$ can therefore be put into the following form:

$$V_2(s) = - \frac{R_1''}{R_1} \cdot \frac{R_1 R_1'}{R_1 + R'_1} \, C_1 \cdot \frac{1}{1 + s \dfrac{R_1 \, R_1'}{R_1 + R_1'} \, C_1} \cdot \frac{Q}{C_f} =$$

$$= - \frac{Q}{C_f} \cdot \frac{R_1''}{R_1} \cdot \frac{1}{s + \dfrac{R_1 + R_1'}{R_1 \, R_1'} \cdot \dfrac{1}{C_1}} \qquad \text{(Eq. 2.26)}$$

Compare now $V_2(s)$ given by Eq. 2.26 with $V_1(s)$ given by Eq. 2.20. Both terms have the same type of dependence on s. Therefore, everything goes AS THOUGH THE SIGNAL WERE AMPLIFIED BY THE FACTOR (R''_1/R_1) AND CLIPPED, BUT LEFT UNCHANGED IN SHAPE. CLIPPING IS ACCOUNTED FOR BY THE FACT THAT THE NEW TIME CONSTANT T_1 REPLACES $R_f C_f$ IN PASSING FROM Eq. 2.20 TO Eq. 2.26. REMEMBER THAT $R_f C_f = R'_1 C_1$. T_1 CAN BE WRITTEN IN A MORE SIGNIFICANT WAY:

$$T_1 = \frac{R_1 \, R'_1}{R_1 + R'_1} \cdot C_1 = R_1' \cdot C_1 \cdot \frac{1}{1 + \dfrac{R_1'}{R_1}} = \frac{R_f \, C_f}{1 + \dfrac{R_1'}{R_1}}$$

The signals $v_1(t)$ and $v_2(t)$ are compared in Fig. 2.23, where the sign inversion on $v_2(t)$ is not accounted for.
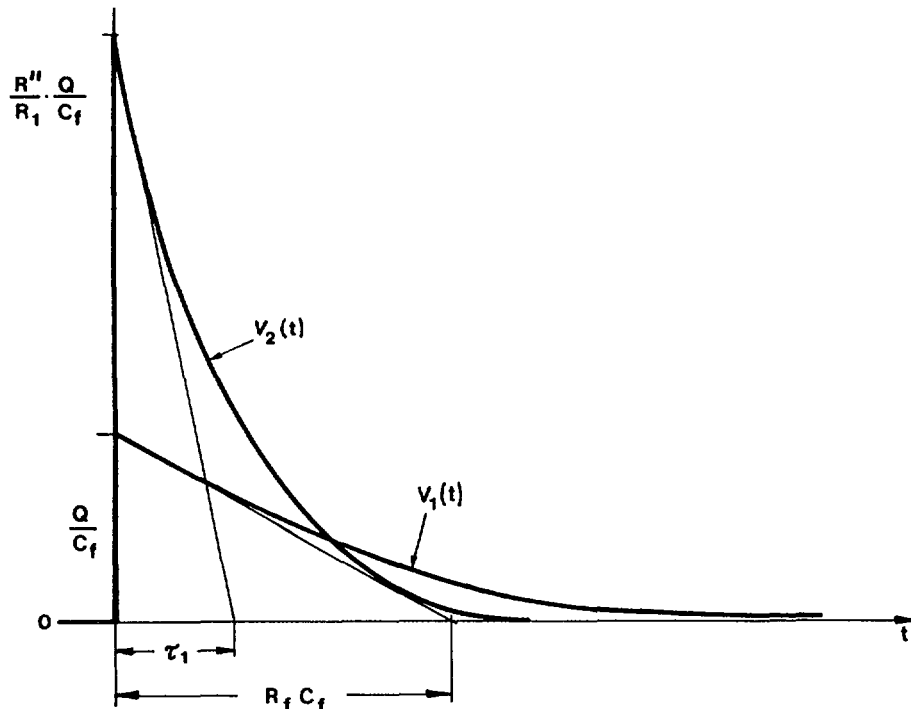

Fig. 2.23: Comparison between the signals $V_1(t)$ and $V_2(t)$

50

Remember: POLE-ZERO CANCELLATION IS ACHIEVED BY MAKING
$R_f C_f = R_1 {}' C_1$; THIS ADJUSTMENT HAS TO BE DONE VERY ACCURATELY.
Remember, moreover, that THE SIGNAL GAIN BETWEEN $v_1(t)$ AND $v_2(t)$ IS
DETERMINED BY THE RATIO R''/R_1.

The signal $V_2(t)$ is sent through a coaxial cable to the further
amplifying unit which is known as "SPECTROSCOPY AMPLIFIER" or "RESEARCH
AMPLIFIER". The coaxial cable MUST BE MATCHED AT BOTH ENDS AS SOON AS ITS
LENGTH EXCEEDS TWO-THREE METERS. Fig. 2.24 shows the proper connection of
the coaxial cable, based upon the assumption that the output dc voltage of
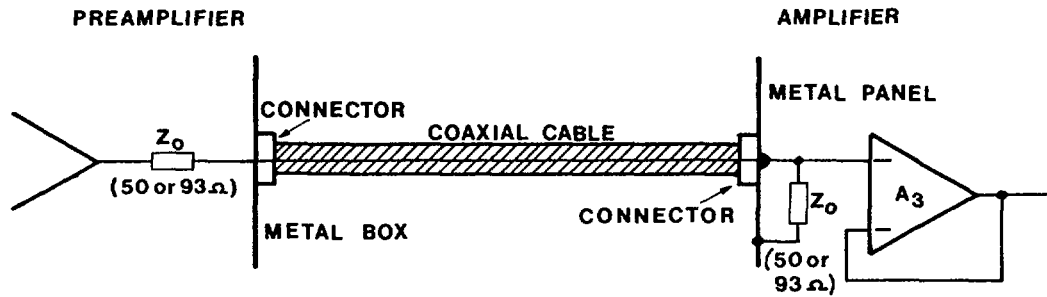the preamplifier is close to zero.

Fig. 2.24:   Coaxial cable connection between preamplifier and
             amplifier.

In Fig. 2.24, Zo is the characteristic impedance of the coaxial
cable: either 50 or 93ohm depending on the type of cable. Fig. 2.24 also
shows the input buffer of the amplifier, which serves the purpose of
impedance isolation at the amplifier input. The same operational amplifier
can be connected like in Fig. 2.13.b if some additional gain is required and
therefore adjusted to work at a value of gain 2 ÷ 3 rather than 1.
Immediately after A_3 it is advisable to put the second clipping network
and some gain. These two functions are implemented by a network which is
identical to the one performing the same function inside the preamplifier
box, Fig. 2.25. Remember that now the condition for correct POLE-ZERO
CANCELLATION IS $R_2 {}' C_2 = T_1$. The signal at the output V_4 is
therefore clipped to a time constant T_2, where

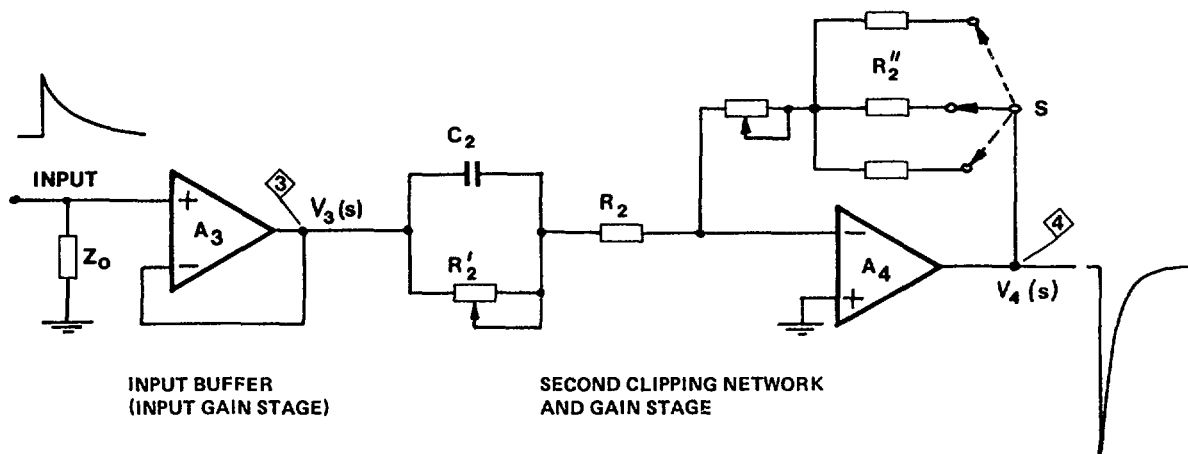$$T_2 = \frac{R_2 R'_2}{R_2 + R'_2} \cdot C_2$$

Fig. 2.25:   Input circuit, second clipping network and gain stage.

51

The gain on the leading edge of the signal is determined, exactly as in the previous case, by the ratio (R''_2/R_2), where R'' can assume different values, selectable through the switch S. Therefore the gain can be set with a coarse control and a fine control. The shaping is completed by sending the signal $v_4(t)$ to a low-pass filter. Usually the simple, passive RC integrator of Fig. 8 in the Section 1 on functional approach is replaced by active filters of the approximate integrator type as shown in Fig. 2.18.d. In order to give the final signal an almost Gaussian shape, which has good performances with respect to signal-to-noise ratio and counting rates, at least two integrators are used at the output of A_4, as shown in Fig. 2.26.
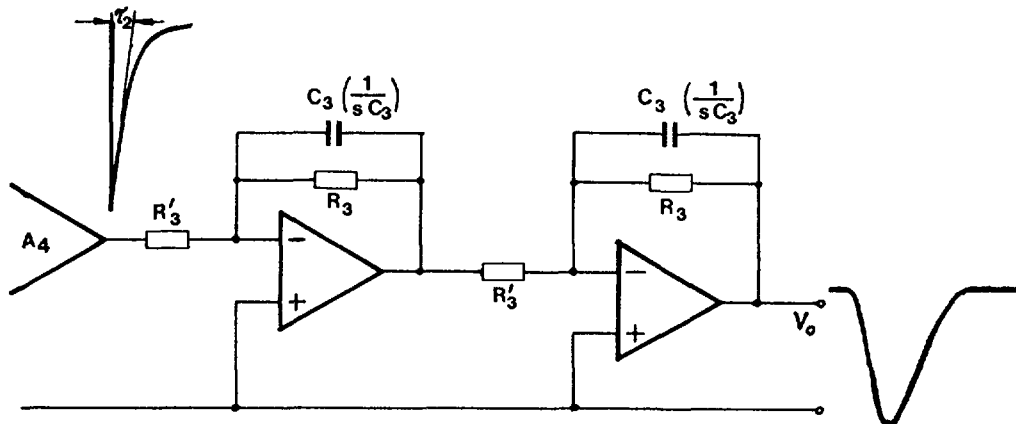

Fig. 2.26: Cascaded active low-pass filter to obtain quasi Gaussian shaping.

To obtain the semigaussian shaping, the time constants R_3C_3 of the approximate integrators of Fig. 2.26 must be made equal to T_2, that is, to the decay time constant of the exponential signal at the output of A4:

$$R_3C_3 = T_2$$

The signal $v_0(t)$ has an almost Gaussian shape, and this signal is ready for pulse amplitude analysis. It has to be pointed out that the low pass filters used in some modern shaping amplifiers may actually differ from those shown in Fig. 2.26. The low pass filters of Fig. 2.26, however, are sufficiently accurate, though simple and can be assumed as an adequate solution to the problem of nuclear pulse shaping.

CHAPTER 3

NOISE AND RESOLUTION

3   NOISE AND RESOLUTION

3.1   INTRODUCTION

Noise gives rise to a degradation of the signal delivered by a nuclear radiation detector, thus limiting the accuracy in the information carried by such a signal.

Noise, in a nuclear pulse processing system consists of several contributions. According to their nature, we can sort them into three groups.

Belonging to the first group are the noise terms directly related with the detection mechanism itself. An example are the fluctuations in the number of electron-ion pairs created by a monoenergetic radiation in an ionization chamber or in a semiconductor detector. To the same group belong the fluctuations in the charge at the output of a scintillation detector when a monoenergetic radiation interacts with it. These fluctuations are due to the statistics of the scintillation process and of the phototube multiplication mechanism.

Within the second group are the purely stochastic noise in circuit devices; to this group belong all the terms associated with passive and active components, like resistor noise, diode noise, bipolar and field-effect transistor noise and so on.

The third group contains the noise contributions due to the influence of the surrounding world, for instance electromagnetic disturbances, ground loop noise, power supply ripple and similar. All these types of noise fall under the denomination of interference noise.

There is a basic difference between the noise groups 1, 2 and group 3. As a matter of fact, the electronic equipment can be, at least conceptually, shielded in a perfect way, and the power supplies filtered so as to make the third type of noise negligible. Instead, noise of types 1 and 2 is directly connected with the physical nature of radiation detectors and electron devices and THE DESIGNER CANNOT AVOID IT.

3.2   HOW NOISE AFFECTS ENERGY MEASUREMENTS IN NUCLEAR SPECTROSCOPY

Let us consider a strictly monoenergetic beam of nuclear particles or photons interacting with the sensitive volume of a nuclear detector. The electronic setup of the spectrometer used in such measurements, is indicated in Fig. 3.1.

Every interaction event between incoming radiation and detector gives rise to a short current pulse. The charge carried by such a pulse is proportional to the energy delivered to the detector. Therefore, energy measurements in nuclear spectrometry imply the evaluation of a transient charge. This is done by integrating the detector current pulse on a capacitor C_{INT} of known value, as shown in Fig. 3.2, and measuring the amplitude of the voltage step which appears across it.

C_{INT} may either be the detector capacitance or an articifical capacitance as in the charge-sensitive preamplifiers. If these voltage steps are amplified and sorted out by a multi-channel analyser as shown in Fig. 3.3, we will not find an indefinitely narrow spectral line, as we would
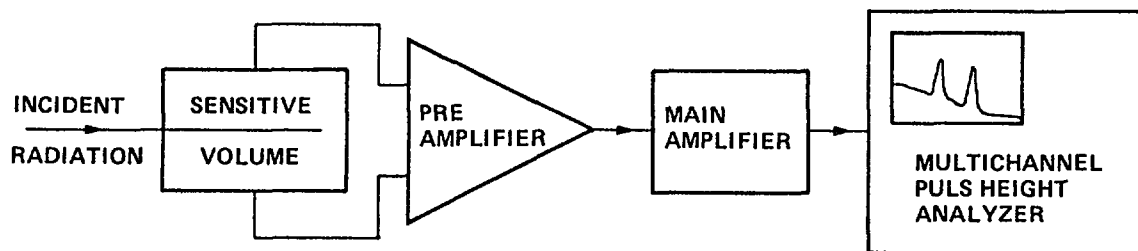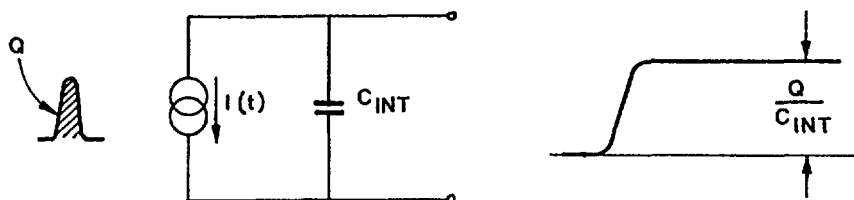
Fig. 3.1: Spectrometric set up

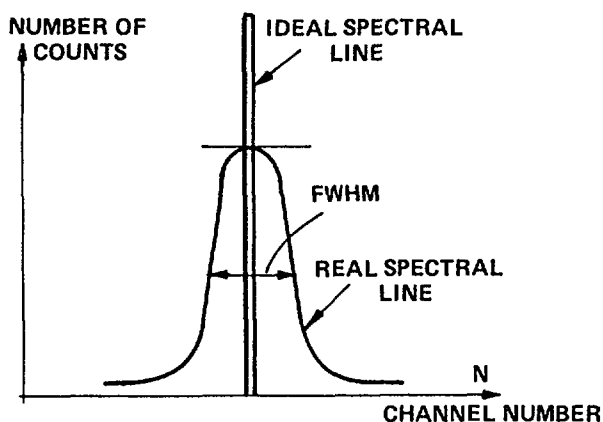Fig. 3.2: Equivalent circuit of detector and charge integrator

Fig. 3.3: MCA display

expect according to the strictly monoenergetic nature of radiation. A line broadening will occur.

The real spectral line has a Gaussian shape and its Full Width at Half Maximum (FWHM) is the result of a combined action of the varius noise contributions that affect the subsequent steps of the incoming energy-to-measured voltage conversion process.

Although the radiation is strictly monoenergetic, Q is a statistical variable, as the energy-to-charge conversion is affected by fluctuations (noise of the first group).

Moreover, the voltage signal obtained by integrating the detector current pulse $i(t)$ is added to the stochastic and interference noise of the preamplifier and amplifier as well as to the pick-up disturbances in the connecting cables.

The FWHM of the amplitude distribution defines the resolution of a spectrometry system. To determine it, two radiations of known energies E_I and E_2 have to be employed. The spectral lines of both radiations are stored and displayed on the multichannel analyser. Let N_I and N_2 be the centroid positions of the two spectral lines. The energy calibration of the multichannel scale is given by (see Fig. 3.4):
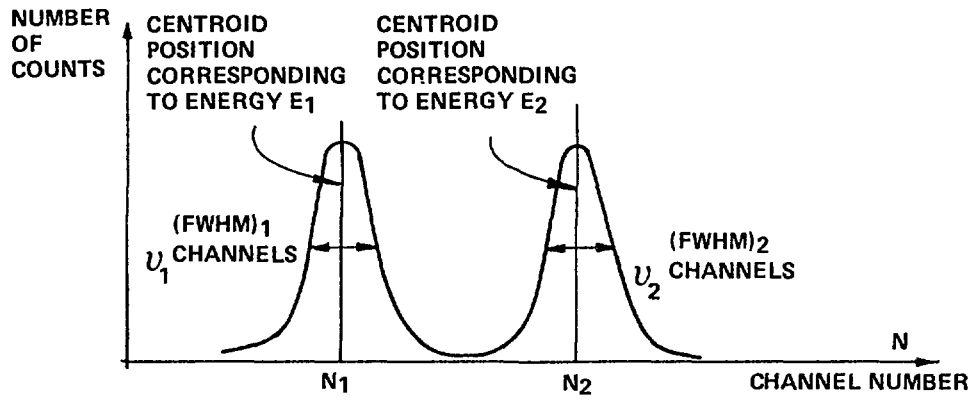
Fig. 3.4: Measurement of the FWHM

$$\frac{E_2 - E_1}{N_2 - N_1} \quad [eV/CHANNEL]$$

and the FWHM of the spectral lines is:

$$FWHM = \frac{E_2 - E_I}{N_2 - N_I} \cdot v \; [eV]$$

A spectrometry system with poor resolution, that is with a large FWHM, is unable to distinguish closely spaced spectral lines: two energies that differ by an amount smaller than the FWHM cannot be resolved by the system.
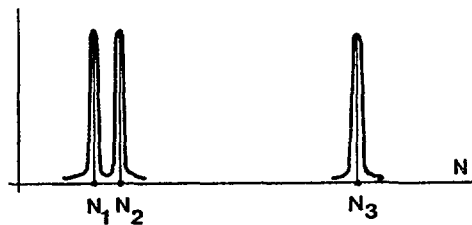
Fig. 3.5 shows how three spectrometers of highly different resolution affect a spectrum of three energies.
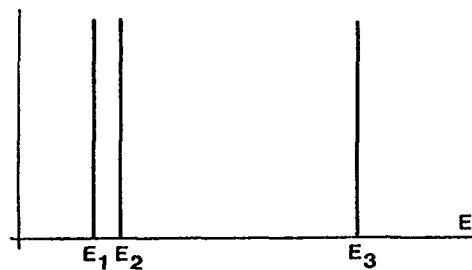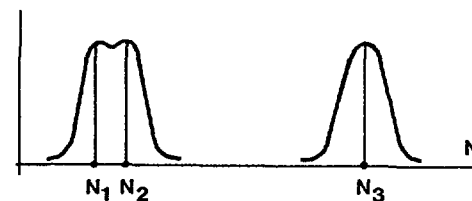

Number of Events — ENERGY DISTRIBUTION AS DETECTED BY A SPECTROMETER WITH INFINITE RESOLUTION

Number of Counts — THIS IS WHAT THE THREE SPECTRAL LINES LOOK LIKE WITH A SPECTROMETRY SYSTEM OF HIGH RESOLUTION

Number of Counts — THIS IS WHAT THE THREE SPECTRAL LINES LOOK LIKE WITH A SPECTROMETRY SYSTEM OF LOW RESOLUTON. THE ENERGIES E_I, E_2 ARE HARDLY DISTINGUISHABLE

Fig. 3.5: Spectrum obtained by instruments with different resolution

57

It is therefore very important to introduce all the measures available in order to keep the resolution adequate to the requirements of the experiment.

In what follows we shall discuss some simple rules aiming at optimising the behaviour of low-noise instrumentation. We shall neglect here noise belonging to the first group which depends on the detection process and, to a lesser extent, on the material and technology employed in the detector development. We shall instead concentrate our attention on the electronic noise of groups 2 and 3.

3.3 DESIGN RECOMMENDATIONS

There are a few simple considerations to bear in mind in order to understand and, when possible, to improve low noise equipment.

1. The extent to which noise degrades an amplitude measurement is expressed by the SIGNAL-TO-NOISE RATIO. In other words, the accuracy of the amplitude measurement is not specified separately by the value of the signal or by the value of the noise, but IT IS THEIR RATIO THAT COUNTS.

2. Consider a linear amplifying system with several blocks in cascade. Assume that each block has the gain higher than one and comparable noise characteristics. The first block affects the noise behaviour of the cascade to the largest extent, the second to a lesser extent than the first one and so on. This conclusion can be made evident by evaluating the total noise at the output of the cascade (Fig. 3.6).

Noise Gain e_1 A_1 Noise Gain e_2 A_2 Noise Gain e_3 A_3

Output signal

$$A_1 \cdot A_2 \cdot A_3 \cdot e$$

Output noise

$$A_1 \cdot A_2 \cdot A_3 \cdot e_1 +$$
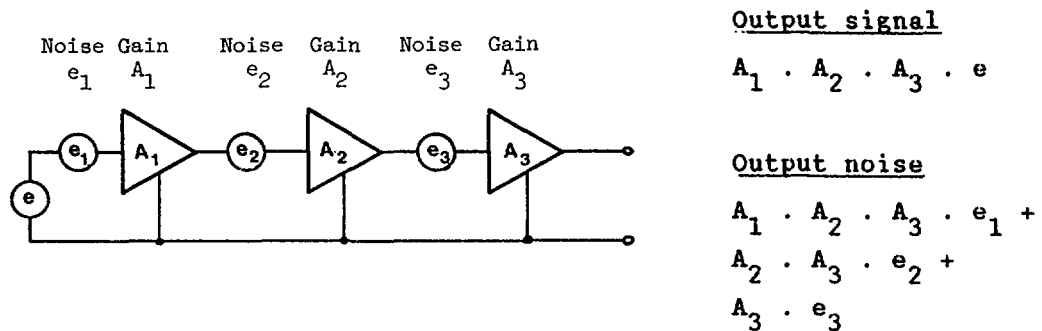$$A_2 \cdot A_3 \cdot e_2 +$$
$$A_3 \cdot e_3$$

Fig. 3.6 Noise contributions from cascaded stages

THE GAIN OF THE FIRST BLOCK OF AMPLIFICATION MUST BE KEPT AS HIGH AS POSSIBLE, in order to reduce the importance of the noise contributions coming from the following blocks.

3. As a consequence of the 2nd rule, avoid in low-noise amplifiers input stages with unit gain. They would remarkable worsen the noise behaviour of the system.

4. It must be remembered that the lower the sensitivity of the radiation detector, the lower the charge delivered when a given energy is released in its sensitive volume, then THE NOISE OF THE PREAMPLIFIER IS MORE IMPORTANT.

5. As a consequence of the 4th point, detectors that do not have an internal charge multiplication, like IONIZATION CHAMBERS and SEMICONDUCTOR DETECTORS need low-noise preamplifiers. For detectors

that have an internal charge multiplication, like PROPORTIONAL and SCINTILLATION COUNTERS, the noise of the preamplifier has little or no importance at all.

6. Get used to checking the noise performances of your preamplifiers and amplifiers. The output noise can be measured in one of the following ways (see Fig. 3.7).

   a) with a wideband rms volt/meter. A suitable instrument is a 3400 A Hewlett-Packard.

   b) sending a signal to your preamplifier or amplifier and storing the output amplitude distribuiton into a multichannel analyzer.

For a quick evaluation of the noise characteristics of your preamplifier or amplifier, you can simply measure with a scope the peak-to-peak noise band:
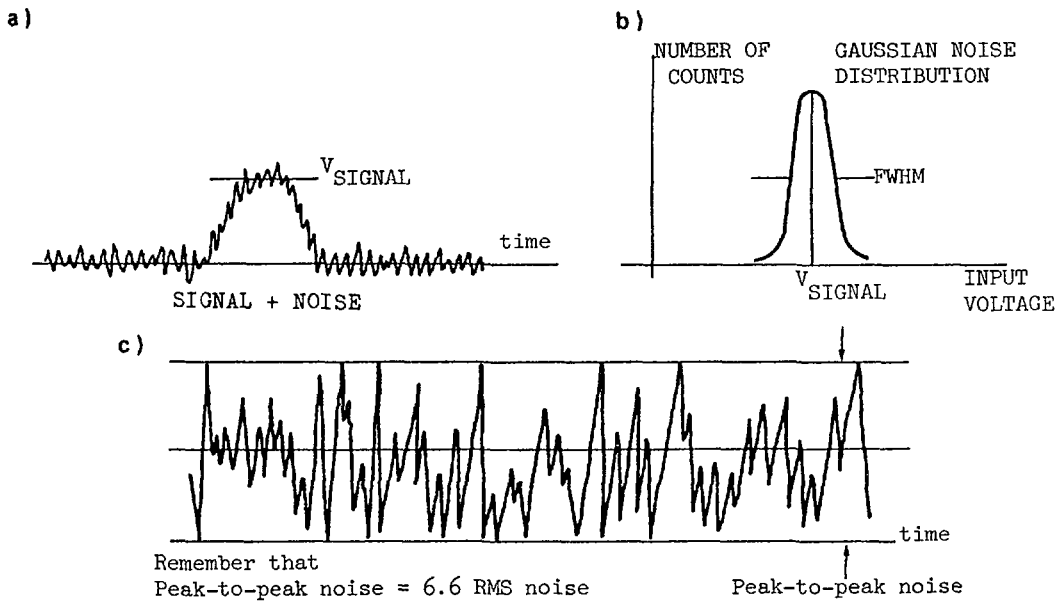
a)

V_{SIGNAL}

time

SIGNAL + NOISE

b)

NUMBER OF COUNTS   GAUSSIAN NOISE DISTRIBUTION

FWHM

V_{SIGNAL} INPUT VOLTAGE

c)

time

Remember that
Peak-to-peak noise = 6.6 RMS noise

Peak-to-peak noise

Fig. 3.7: Noise measurements. a) Remember that FWHM = 2.355RMS, b) MCA display, c) Peak-to-peak noise = 6.6 RMS.

To reduce the effects of deterministic noise pay attention to the following points.

7. Use well filtered power supplies for the preamplifier.

8. If necessary, introduce on the supply lines of the preamplifier the filter, as shown in Fig. 3.8.
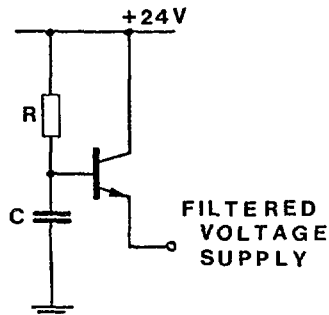
+24V

R

C

FILTERED
VOLTAGE
SUPPLY

Fig. 3.8: Power supply filtering

9. If the distance between preamplifier and main amplifier exceeds one or two meters, use a differential signal transmission between them, as shown in Fig. 3.9.
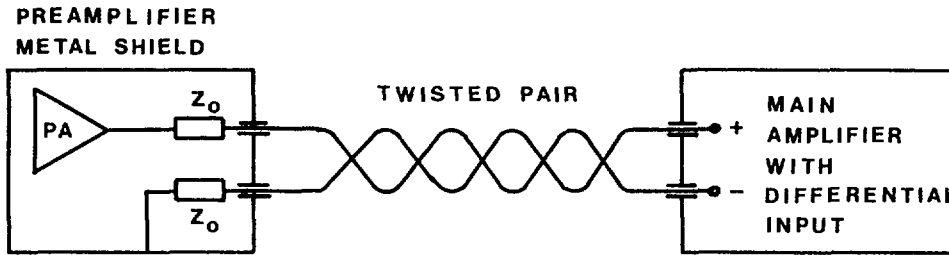

Fig. 3.9: Connection by twisted pair cable

3.4 REPRESENTATION OF NOISE SOURCES

Purely stochastic noise is usually represented in the frequency domain by specifying the power spectrum of the equivalent voltage or current source:

Fig. 3.10: Noise represented by power spectra. a) Noise voltage source, b) Noise current source

The power spectrum of either generator has the following meaning (see Fig. 3.11): e_N^2 is the average power delivered to a 1 ohm – load in the frequency interval Δf; i_N^2 is the average power delivered to a 1 ohm – load in the frequency interval Δf:
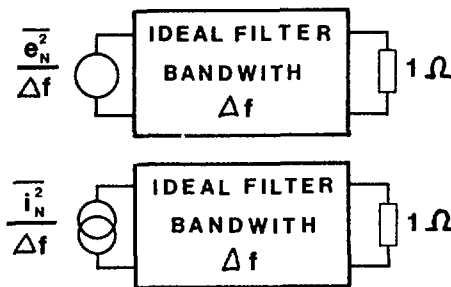

Fig. 3.11: Noise filtering

3.5 NOISE IN ELECTRON DEVICES

The sources of purely stochastic noise in electron devices belong to the following basic models:

1. Thermal noise in a resistor. It is due to the thermal agitation of free charge carriers in the resistor.

2. Shot noise. First analysed on a vacuum diode in the current-saturated region, it is related to the discreet nature of electricity.

60

3. Partition nise. First observed in the division of the cathode current of a tethrode or penthode between screen grid and anode.

4. $\frac{I}{fn}$ low frequency noise. To this type belong the long-term drifts of electronic components, the noise associated with lossy dielectric and lossy magnetic cores, the noise due to charge trapping in the oxide of a MOS transistor.

We shall now review very briefly the noise sources associated with electron devices.

PASSIVE COMPONENTS

Thermal noise in a resistor is represented by an emf source in series or by a current source in parallel to the real resistor. The RMS voltage or the RMS current in the frequency interval Δf are:
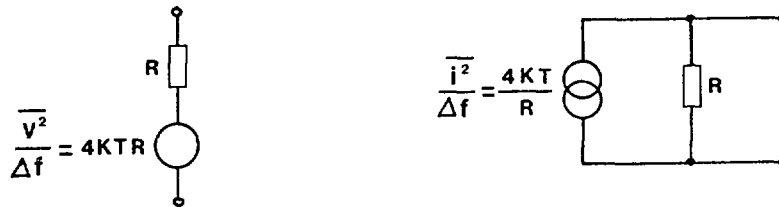

Fig. 3.12: Representations of thermal noise

An ideal capacitor as well as an ideal inductor are NOISELESS.

A real capacitor has noise associated with losses in the dielectric.

In low-noise circuits use preferably METAL-FILM RESISTORS: non-metallic resistors might have additional noise when standing current flows through.

ACTIVE COMPONENTS

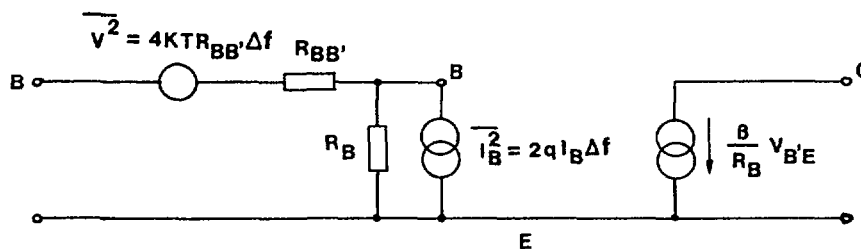In bipolar transistors the dominant noise sources are (see Fig. 3.13):

Fig. 3.13: Equivalent circuit of the bipolar transistor showing the related noise sources.

thermal noise associated with the base spreading resistance,

$$\overline{v^2} = 4kTR_{BB'}\Delta f$$

shot noise associated with the base current

$$\overline{i_B^2} = 2qI_B\Delta f, \quad (q \text{ is electron charge, } I_B \text{ is the standing base current}).$$

61

In field effect transistors the dominant noise sources are:
thermal noise of the conducting channel which can be represented with a
voltage source in series with the gate,

$$\overline{v^2} = 4kT \cdot \frac{0.7}{g_m} \Delta f \quad (g_m \text{ is the transconductance}),$$

shot noise associated with the gate current, $\overline{i^2} = 2 q I_G \Delta f$, I_G is
the standing gate current, i.e. current of the reverse-biased
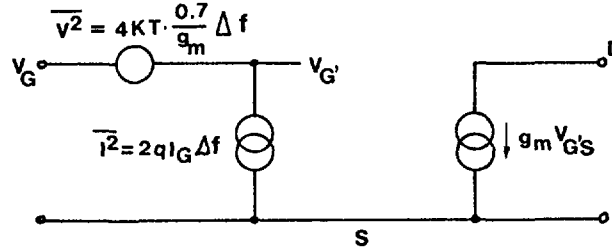gate-to-channel junction.

Fig. 3.14:   Equivalent circuit of a junction field-effect transistor
              and related noise sources.

3.6   SIGNAL, NOISE AND FILTERING

In this chapter, for the sake of simplicity, we shall represent signal
and noise as being delivered by voltage sources. However, all the
conclusions that we will reach can be easily extended to the case in which
either or both variables are currents. Fig. 3.15 is an example of the
head-part of a data-acquisition system, with signal source and preamplifier
in a noisy environment.

The environment noise is generated by a power cable nearby, carrying a
current I(w) at line frequency. The magnetic flux created by I(w) is linked
with the loop 1 connecting the signal source to the preamplifier and with
the loop 2 connecting preamplifier and main amplifier.

The noise induced in these loops is represented by two deterministic
generators e_{NI}, e_{N2}. The purely stochastic noise of the preamplifier is
accounted for by e_N. The signal-to-noise ratio at the preamplifier input
port (a,a') is given by:

$$\frac{S}{N}\bigg|_{aa'} = \frac{e_s}{e_{NI} + e_N}$$

The noiseless preamplifier, which is also supposed to have a
frequency-independent transfer function does not alter the signal-to-noise
ratio. So we have:

$$\frac{S}{N}\bigg|_{bb'} = \frac{A_p e_s}{A_p e_{NI} + A_p e_N} = \frac{S}{N}\bigg|_{aa'}$$

The signal-to-noise ratio at the input of the main amplifier is
affected by the additional source e_{N2}:

$$\left(\frac{S}{N}\right)_{cc'} = \frac{A_p e_s}{A_p e_{NI} + A_p e_N + e_{N2}} \ll \left(\frac{S}{N}\right)_{aa'}$$

As the last equation shows, the higher is A_p, the lower is the effect
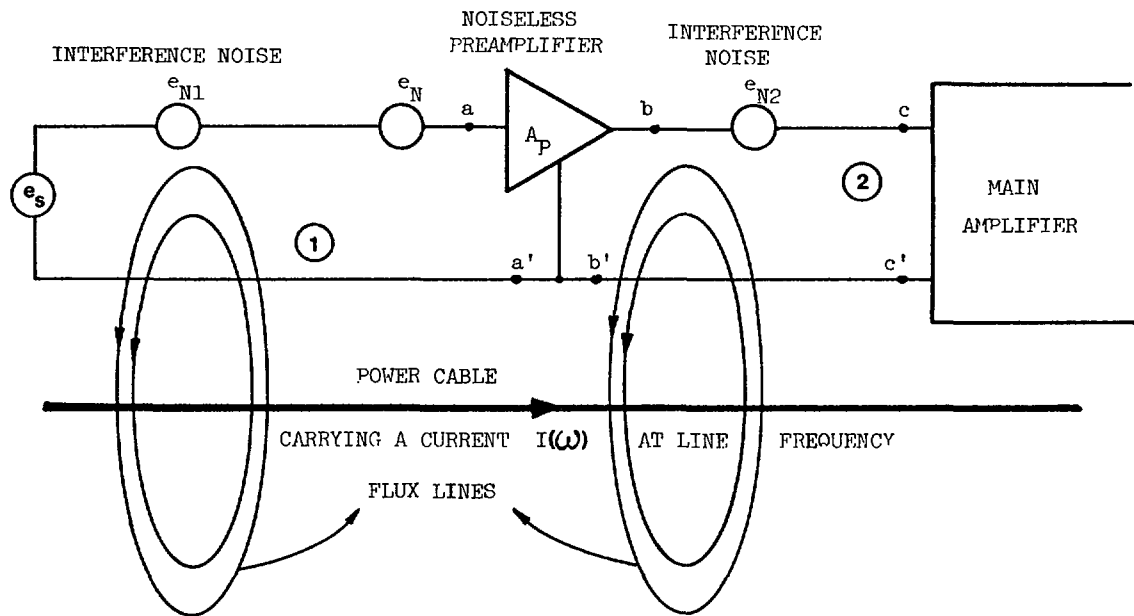of e_{N2} on signal-to-noise ratio.
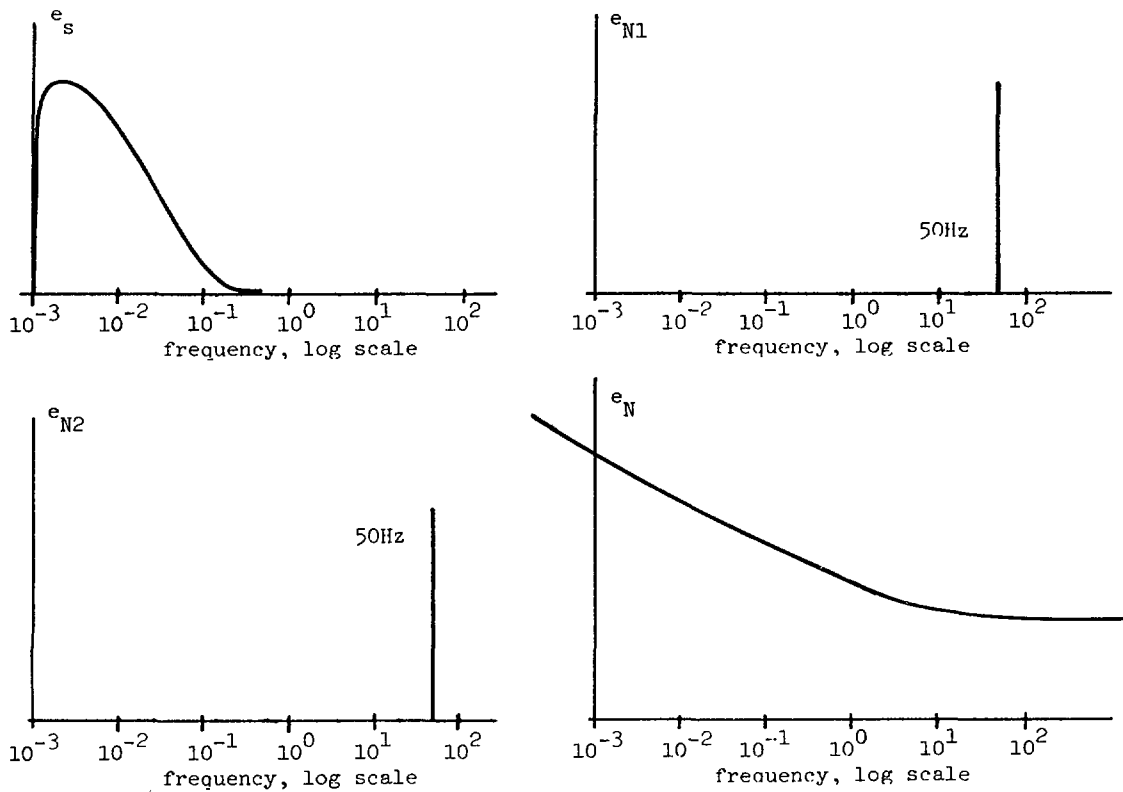
62

Fig. 3.15   Example of noise pick-up.

Fig. 3.16:   Examples of different noise contributions and related spectra.

The discussion above is oriented to clarify the basic mechanism of noise induction and the relative importance of the various noise sources. In order to make the discussion simple, we have implicitly assumed that e_S, e_{NI} and e_{N2}, e_N are instantaneous values of signal and noise. We must not forget, however, that while e_S, e_{NI} and e_{N2} are deterministic variables, e_N is a stochastic process which cannot be represented by an a priori predictable time-dependence.

Going back to the circuit diagram of Fig. 3.15 we can introduce the
following remarks. For a fixed distance between power cable and the
connecting loops for a given value of I(w), the induced noise emf's e_{NI},
e_{N2} are proportional to the loop area. Therefore, it is very important to
keep e_{N2} low when the loop 2 cannot be reduced in length; the signal from
the preamplifier should be connected to the main amplifier by a twisted
cable, as already shown in one of the previous paragraphs.

If SIGNAL AND NOISE HAVE DIFFERENT FREQUENCY COMPOSITIONS then a filter
following the main amplifier can be employed to increase the signal-to-noise
ratio. A filter operates a frequency-selective processing of signal and
noise. It should pass the Fourier frequencies of the signal with the least
possible attenuation, while it should depress as much as possible the
Fourier frequencies of the noise.

The design of a filter requires the knowledge of the frequency spectra
of signal and noise. Signal and deterministic noise are usually described
by their Fourier amplitude spectrum, while stochastic noise is represented
by its power spectrum. Fig. 3.16 shows what the spectra of the four
generators e_S, e_{N1}, e_{N2}, e_N look like in the case in which e_S is
the signal delivered by a thermocouple.

3.7 SIGNAL, NOISE AND RESOLUTION IN NUCLEAR SPECTROMETRY

We are now going to analyse the specific case of low-noise measurements
on nuclear detector pulses. The first example that we shall discuss is
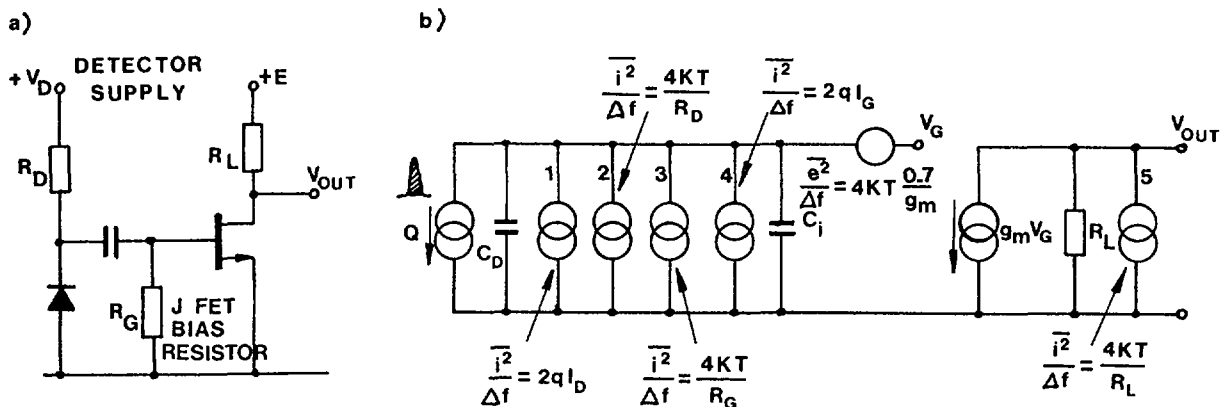based upon a voltage-sensitive preamplifier.

a)   b)

Fig. 3.17:   a) Simplified preamplifier configuration
             b) Noise sources associated with detector and preamplifer

Part a) of Fig. 3.17 shows the actual circuit with the detector AC
coupled to the preamplifier and b) is the equivalent circuit with all the
noise sources.

The resistors R_D and R_G are usually so large that their effect on
the signal shape is negligible. For this reason they have been omitted in
part b) of Fig. 3.17, but the noise they generate is accounted for. The
detector is represented in b) by a current source which delivers a short
pulse of area Q.

C_D is the detector capacitance and C_i is the input capacitance of
the field-effect transistor.

The four parallel noise generators represent:

1. Shot noise associated with the reverse current I_D in the detector.
2. Thermal noise in the detector biasing resistor R_D.
3. Thermal noise in the JFET biasing resistor R_G.
4. Shot noise associated with the gate leakage current I_G.

The channel thermal noise in the field-effect transistor is represented by the series voltage source.

The voltage-controlled source $g_m v_G$ describes the actual effect of the field-effect transistor.

Thermal noise of load resistor R_L is taken into account by current source 5.

The detector coupling capacitor C is very large and has been omitted in b), Fig. 3.17.

We are now going to evaluate signal and noise and the preamplifier output. The short current pulse delivered by th detector is integrated by the parallel combination of C_D and C_i. A voltage step appears therefore at V_G and its amplitude is $Q/(C_D + C_i)$

The output signal, consequently, is a step of amplitude $g_m R_L$ $Q/(C_D + C_i)$

The noise power spectra at the preamplifier output consists of the following terms:

1. $\quad 2qI_D \quad \cdot \quad \dfrac{1}{\omega^2 \cdot (C_D + C_i)^2} \quad \cdot \quad g_m^2 \cdot R_L^2$

2. $\quad \dfrac{4kT}{R_D} \quad \cdot \quad \dfrac{I}{\omega^2 (C_D + C_i)^2} \quad \cdot \quad g_m^2 \cdot R_L^2$

where $\omega = 2\pi f$ is the angular frequency

3. $\quad 2qI_G \quad \cdot \quad \dfrac{I}{\omega^2 \cdot (C_D + C_i)^2} \quad \cdot \quad g_m^2 \cdot R_L^2$

4. $\quad \dfrac{4kT}{R_G} \quad \cdot \quad \dfrac{I}{\omega^2 (C_D + C_i)^2} \quad \cdot \quad g_m^2 \cdot R_L^2$

5. $\quad 4kT \ R_L$

6. $\quad 4kT \quad \cdot \quad \dfrac{0.7}{g_m} \quad \cdot \quad g_m^2 \cdot R_L^2$

The noise spectra written above have been calculated observing that the noise currents delivered by the four parallel sources are integrated by $C_D + C_i$ and the resulting voltages are amplified by the JFET. The series noise source is transferred to the output multiplied by the voltage gain $g_m R_L$ of the JFET, while the thermal noise in the load resistor R_L appears directly at the output.

Note that the SIGNAL HAS THE SAME NATURE AND THE SAME FREQUENCY SPECTRUM OF THE PARALLEL NOISE SOURCES, 1, 2, 3, 4.

Therefore, if the series noise and that due to R_L were not present, filtering at the output of the preamplifier would be meaningless and the best SIGNAL-TO-NOISE ratio would be directly obtained at (a, a') port.

Owing to the presence of series and R_L-noise, filtering can improve the signal-to-noise ratio.

Here we want to evaluate the signal-to-noise ratio at the output of the simple RC - CR filter shown in Fig. 3.18.
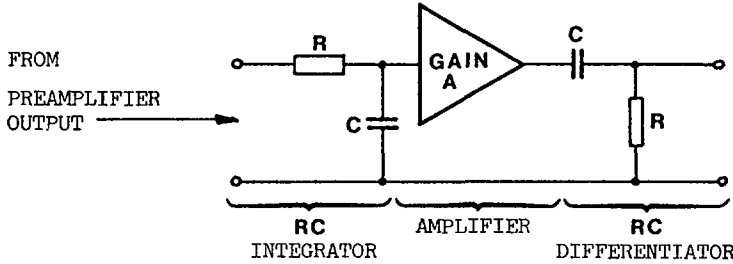

Fig. 3.18: RC-CR shaper

The transfer function of this filter for sinusoidal wave is:

$$T(j\omega) = A \cdot \frac{1}{1 + j\omega RC} \cdot \frac{j\omega RC}{1 + j\omega RC}$$

transfer function of the RC integrator   transfer function of the RC differentiator

To evaluate the noise at the filter output we must remember that the noise sources considered here are UNCORRELATED. The resulting power spectrum at the preamplifier output is simply obtained by adding the power spectra of the single sources:

$$g_m^2 \, R_L^2 \, 4kT \cdot \frac{0.7}{g_m} + \frac{I}{\omega^2 (C_D + C_i)^2} \cdot \left[2qI_D + 2qI_G + \frac{4kT}{R_D} + \frac{4kT}{R_G} \right] + 4kTR_L$$

Remember moreover, that the power spectrum at the output of a linear filter with transfer function $T(j\omega)$ is given by the following rule:

$$\begin{bmatrix} \text{POWER SPECTRUM} \\ \text{AT THE OUTPUT} \end{bmatrix} = \begin{bmatrix} \text{POWER SPECTRUM} \\ \text{AT THE INPUT} \end{bmatrix} \cdot | T (j\omega) |^2$$

In our case, the power spectrum at the filter output is:

$$\left[g_m^2 \cdot R_L^2 \cdot 4kT \cdot \frac{0.7}{g_m} + \frac{I}{\omega^2 (C_D + C_i)^2} \left(2qI_D + 2qI_G + \frac{4kT}{R_D} + \frac{4kT}{R_G} \right) + 4kTR_L \right] A^2 \cdot \frac{\omega^2 R^2 C^2}{(1+\omega^2 R^2 C^2)^2}$$

The total RMS noise at the filter output is obtained integrating the above power spectrum on the $0 \div \infty$ frequency interval. We obtain:

$$[\text{RMS NOISE}]^2 = A^2 \cdot \left[g_m^2 R_L^2 \cdot 4kT \cdot \frac{0.7}{g_m} + 4kTR_L \right] \cdot \int_0^\infty \frac{\omega^2 R^2 C^2}{(1+\omega^2 R^2 C^2)^2} \, df +$$

$$\left[2qI_D + 2qI_G + \frac{4kT}{R_D} + \frac{4kT}{R_G} \right] A^2 \cdot g_m^2 \cdot R_L^2 \cdot \frac{R^2 C^2}{(C_D + C_i)^2} \cdot \int_0^\infty \frac{I}{(I + \omega^2 R^2 C^2)^2} \, df$$

Evaluating the two integrals and putting $\tau = RC$ we finally obtain

$$[\text{RMS NOISE}]^2 = A^2 \left[4kT \cdot \frac{0.7}{g_m} \, g_m^2 \, R_L^2 + 4kTR_L \right] \cdot \frac{I}{8\tau} +$$

$$+ A^2 g_m^2 \, R_L^2 \left[2qI_D + 2qI_G + \frac{4kT}{R_D} + \frac{4kT}{R_G} \right] \cdot \frac{\tau}{8} \cdot \frac{I}{(C_D + C_i)^2}$$

The steplike signal delivered by the preamplifier appears at the output of the filter shaped as shown in Fig. 3.19.
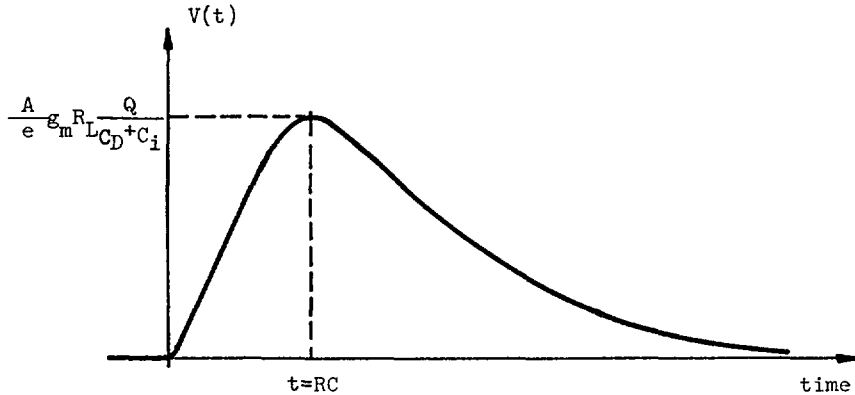

Fig. 3.19:   Signal at the Output of an RC-CR filter when a step signal is applied at the input

Its peak amplitude is $Qg_m \cdot R_L \cdot A/e \, (C_D + C_i)$, e being the base of natural logarithms.

We can now evaluate the resolution of our spectrometry system in the following way. Let ϵ be the average energy required to create an electron-hole pair in our detector.

$\epsilon = 3.67$ eV/pair in silicon
$\epsilon = 2.97$ eV/pair in germanium
$\epsilon = 13$ eV/pair in diamond
$\epsilon = 25$ ev/pair in gaseous argon

The amplitude $Qg_m \, R_L \, A/e \, (C_D + C_i)$ at the output of the filter corresponds to a charge of Q coulombs injected at the detector port of our spectrometry system. Now, the equivalent energy required in order to make such a charge available is:

$$E = \frac{Q}{1.6 \cdot 10^{-19}} \cdot \epsilon \quad \text{eV or}$$

$$Q = 1.6 \cdot 10^{-19} \cdot \frac{E}{\epsilon}$$

The peak amplitude of the pulse shaped by the filter can therefore be expressed as:

$$1.6 \cdot 10^{-19} \, \frac{E}{\epsilon} \cdot \frac{1}{C_D + C_i} \cdot g_m R_L \cdot \frac{A}{e}$$

We define the resolution of the spectrometry system as the energy E which gives rise to a pulse at the output of the filter equal in amplitude to the FWHM of the noise linewidth. In other words, we write:

$$1.6 \cdot 10^{-19} \frac{E_{RES}}{\varepsilon} \cdot \frac{1}{C_D + C_i} \, g_m R_L \cdot \frac{A}{e} = 2.355 \, [RMS \, NOISE]$$

Introducing the expression of RMS NOISE, solving with respect to E_{RES} and rearranging the coefficients we obtain:

$$E_{RES} = 4.09 \cdot 10^{+19} \varepsilon \left\{ \left[kT \cdot \frac{0.7}{g_m} + kT \frac{I}{g_m^2 R_L} \right] \frac{(C_D + C_i)^2}{2\tau} + \left[qI_D + qI_G + \frac{2kT}{R_D} + \frac{2kT}{R_G} \right] \frac{\tau}{4} \right\}^{1/2}$$

E_{RES} is in eV.

As stated in the last equation, the shorter τ is, the more important is the series noise (thermal noise in the channel of the JFET), the less important is the parallel noise. Conversely, the longer τ is, the more important is the parallel noise, and the less important the series noise.

Let us make a numerical example. We want to evaluate the resolution at the output of a spectrometry system based upon:

a)   a Si surface-barrier detector with
     $C_D = 100$ pF, $I_D = I\mu A$, $R_D = IM\Omega$

b)   A JFET voltage preamplifier with
     gm = 10 mA/v, $C_i = 10pF$, $R_L = 1k\Omega$, $I_G = 10^{-11}A$,

$$R_G = 10^9 \Omega$$

c)   a RC-CR filter with $\tau = 1\mu s$

Assume T = 300°K and remember that

k = $1.38 \cdot 10^{-23}$ Joule/°K

q = $1.6 \cdot 10^{-19}$ Coul.

Then:

$$E_{RES} = 4.09 \cdot 10^{19} \varepsilon \left\{ 1.38 \cdot 10^{-23} \cdot 300 \left[\frac{0.7}{10^{-2}} + \frac{I}{10^{-4} \cdot 10^3} \right] \cdot \frac{1.21}{2} \cdot 10^{-14} + \right.$$

$$\left. + \left[1.6 \cdot 10^{-19} (10^{-6} + 10^{-11}) + 2.1 \cdot 38 \cdot 10^{-23} \cdot 300 (10^{-6} + 10^{-9}) \right] \frac{10^{-6}}{4} \right\}^{1/2}$$

$$= 4.09 \cdot 3.67 \cdot 10^{-19} \left[2 \cdot 10^{-33} + 4 \cdot 10^{-32} \right]^{1/2} = 30.6 \, KeV$$

In nuclear spectrometry with semiconductor detectors, the voltage-sensitive preamplifier is not employed because the energy-to-voltage relationship would depend on the detector capacitance C_D. C_D is a function of the applied voltage as well as of the temperature and therefore the coefficient of the energy-to-voltage conversion would not be adequately

reliable. It is for this reason that with semiconductor detectors charge-sensitive preamplifiers are employed. A complete spectrometry system based upon a charge-sensitive preamplifier is shown in Fig. 3.20.
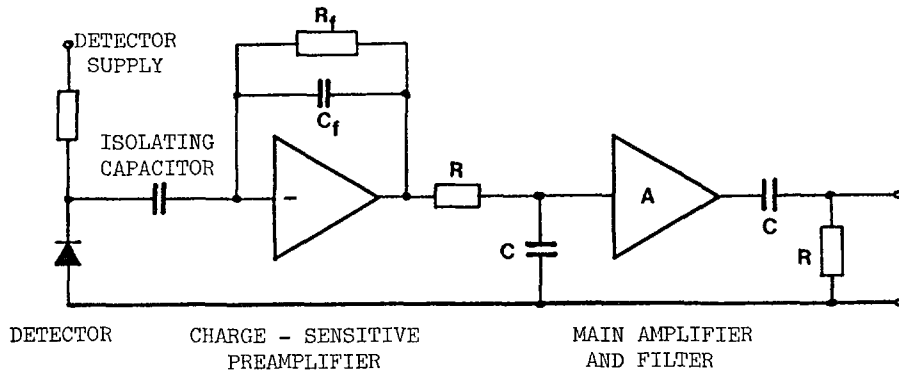

Fig. 3.20: Spectrometry channel employing charge-sensitive preamplification and RC-CR shaping
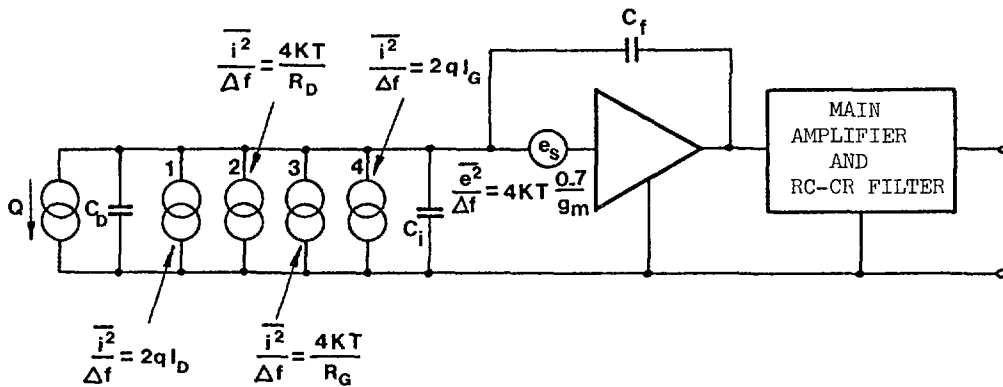

Fig. 3.21: Charge-sensitive preamplifier and related noise sources.

The analysis of a charge-sensitive preamplifier can be made with the help of Fig. 3.21. As in the case of the voltage-sensitive preamplifier, part a) represents the actual circuit and part b) the equivalent noise sources. The resistance R_f serves the purpose of providing a dc feedback around the charge-sensitive preamplifier and replaces the JFET BIASING RESISTOR R_G of the previous case. R_D and R_f have been omitted in part b), but their noise contribution is accounted for.

The signal current generator as well as the four parallel noise sources are integrated by the feedback capacitor C_f. At the output of the preamplifier we have a steplike signal with an amplitude $\dfrac{Q}{C_f}$ and the following noise spectrum due to sources 1, 2, 3, 4:

$$\left[2q \ (I_D + I_G) + 4kT \ (\frac{I}{R_D} + \frac{I}{R_f}) \right] \cdot \frac{I}{\omega^2 C_f^2}$$

In order to evaluate the contribution to the power spectrum at the preamplifier output coming from the series source, we remember that the inverting input of the charge-sensitive preamplifier is a virtual ground. Then the voltage across C_i is equal to e_s. The total charge on the parallel-connected capacitors C_i and C_D will be $(C_i + C_D) \ e_s$. The same charge will be present on C_f.

We therefore have:

$$(C_i + C_D)e_s = C_f(V_u - e_s),$$

that is:

$$V_u = \frac{C_i + C_D + C_f}{C_f} \cdot e_s$$

It can be concluded that the series generator contributes to the spectrum at the preamplifier output with the term:

$$\frac{(C_i + C_D + C_f)^2}{C_f^2} \cdot 4kT \cdot \frac{0.7}{g_m}$$

From this point on, the evaluation of the RMS noise, of the signal amplitude, of the energy resolution proceeds along the same path of the previous example.

We obtain:

SIGNAL, PEAK AMPLITUDE
AT THE FILTER OUTPUT: $\qquad\qquad\qquad\qquad \dfrac{A\,Q}{e\,C_f}$

RMS NOISE
AT THE FILTER OUTPUT: $A^2 \left[kT \cdot \dfrac{0.7}{g_m} \cdot \dfrac{(C_i + C_D + C_f)^2}{C_f^2} \cdot \dfrac{1}{2\tau} + \right.$

$$\left. + \left(q(I_D + I_G) + 2kT\,(\tfrac{1}{R_D} + \tfrac{1}{R_f}) \right) \cdot \tfrac{\tau}{4} \cdot \tfrac{1}{C_f^2} \right]$$

RESOLUTION: $E_{RES} = 4.09 \times 10^{19} \times \varepsilon \left\{ kT \cdot \dfrac{0.7}{g_m} \cdot (C_D + C_i + C_f)^2 \cdot \dfrac{1}{2\tau} \right.$

$$\left. + \left[q\,(I_D + I_G) + 2kT\,(\tfrac{1}{R_D} + \tfrac{1}{R_f}) \right] \cdot \tfrac{\tau}{4} \right\}^{1/2} \quad \text{[ev]} \qquad\qquad \text{(Eq. 3.1)}$$

We shall now compare the resolution obtainable from the charge-sensitive arrangement with that of the voltage-sensitive one. In the charge-sensitive setup we did not include the noise contribution coming from the load resistor of the JFET, R_L; this contribution is usually negligible if $g_m R_L$ is made large enough.

The resolution is practically the same for the two types of preamplifier. In the charge-sensitive one there is a small increase in the total capacitance, $C_D + C_i + C_f$ against $C_D + C_i$ of the voltage-sensitive case, but usually C_f can be neglected in comparison to C_D.

70

3.8 COMMENTS ABOUT THE RESOLUTION

The expression of the energy resolution, as shown by the relationship in Eq. 3.1 can be written in a form much more useful for numerical calculation if the most convenient units are adopted for the system parameters:

g_m measured in mA/V, τ in μs, I_D and I_G in nA, R_D and R_f in M ohm, C_D, C_i, C_f in pF.

The energy resolution becomes:

$$E_{RES} = 40.9 \cdot \varepsilon \cdot \left\{ 1.45 \cdot \frac{(C_D + C_i + C_f)^2}{g_m} \cdot \frac{1}{\tau} + \left[40 (I_D + I_G) + 2.07 . 10^3 (\frac{1}{R_D} + \frac{1}{R_f}) \right] \tau \right\}^{1/2} \quad \text{[FWHM, eV]}$$

The energy resolution depends, through ε, on the detector. To arrive at an expression which is independent on the detector and is therefore more suitable to describe the intrinsic noise performances of the preamplifier, the concept of equivalent noise charge ENC is adopted. The expression of ENC is obtained by dividing E_{RES} with ε and then by 2.355, on account of the fact that the equivalent noise charge is usually expressed in rms electrons. Then:

$$ENC = 17.36 \left\{ 1.45 \frac{(C_D + C_i + C_f)^2}{g_m} \cdot \frac{1}{\tau} + \left[40(I_D + I_G) + 2.07 \ 10^3 (\frac{1}{R_D} + \frac{1}{R_f}) \right] \tau \right\}^{1/2}$$

The above equation deserves some comments. In a well-designed spectroscopy system, the noise due to the parallel sources I_D, I_G, R_D, R_f, is usually negligible. For instance, in an γ-ray spectrometer, where detector and first amplifying stage are at cryogenic temperature and R_D, R_f are of the order of G ohm (10^9 ohm), the parallel noise is negligible. Then, the ENC becomes:

$$ENC \approx 17.36 \left\{ 1.45 \frac{(C_D + C_i + C_f)^2}{g_m} \cdot \frac{1}{\tau} \right\}^{1/2} =$$

$$= 21 \left\{ \frac{(C_D + C_i + C_f)^2}{g_m} \cdot \frac{1}{\tau} \right\}^{1/2} \quad \text{[rms electrons]}$$

The equivalent noise charge can be rewritten, by neglecting C_f, which is usually very small compared to C_D and C_i in the following form:

$$ENC = 21 \cdot (C_D C_i)^{1/2} \cdot \left\{ \left[\left(\frac{C_D}{C_i}\right)^{1/2} + \left(\frac{C_i}{C_D}\right)^{1/2} \right]^2 \cdot \frac{1}{g_m \tau} \right\}^{1/2} \quad \text{[rms electrons]}$$

71

The ratio $(C_D/C_i)^{1/2} = m$ is called mismatch factor. Introducing m into ENC, the following expression for ENC is obtained:

$$ENC = 21 \cdot (C_D \, C_i)^{1/2} \cdot \left[\left(m + \frac{1}{m} \right)^2 \cdot \frac{1}{g_m \tau} \right]^{1/2} \quad \text{[rms electrons]}$$

ENC is at minimum for $m=1$:

$$ENC_{OPT} = 42 \left(C_D \right)^{1/2} \cdot \left[\frac{C_i}{g_m} \cdot \frac{1}{\tau} \right]^{1/2} \quad \text{[rms electrons]}$$

The optimum condition, i.e. $m=1$, means $C_D = C_i$; the detector is capacitively matched to the preamplifier. As a conclusion, for a given detector capacitance C_D and a given shaping time constant τ the minimum ENC occurs when the input capacitance of the amplifying device is equal to C_D. If an amplifying device is available, with a fixed gain bandwidth product $\omega_T = g_{mo}/C_{io}$, the condition of capacitive matching is achieved by paralleling as many devices as required by the condition $C_D = m \, C_{io}$, with n = number of devices required. The optimum ENC can be written as

$$ENC_{OPT} = 42 \cdot C_D^{1/2} \cdot \left[\frac{n \, C_{io}}{n \, g_m} \cdot \frac{1}{\tau} \right]^{1/2} = 42 \left[\frac{C_o}{\omega_T \tau} \right]^{1/2} \quad \text{[rms electrons]}$$

The ENC_{OPT} is proportional to the square root of the detector capacitance and inversely proportional to the square root of the product of the ω_T of the device employed and the shaping time constant τ. Remembering that the units adopted for g_m and C_i are respectively mA/V and pF, the gain bandwidth product ω_T turns out to be automatically expressed in Grad/s. So, for instance, in condition of capacitive matching, with C_D = 10 pF, τ = 1µs, ω_T = 1 Grad/s, the ENC_{OPT} would be about 132 rms electrons.

3.9 CONFIGURATION OF CHARGE-SENSITIVE PREAMPLIFIER

The charge-sensitive preamplifier is an operational integrator designed around a low noise amplifier. One important feature of an amplifier suitable for this kind of application is that it must have a single-ended, inverting-only configuration. The balanced input structure which is so common among current operational amplifier would be unsuitable in the actual application, for it would increase the noise referred to the input by $\sqrt{2}$, this being the ratio between the noise referred to the input of a long-tailed pair and the noise referred to the input of a grounded-source stage employing the same kind of device.

A basic structure of a low noise, inverting-only amplifier is shown in Fig. 3.22.

It consists of a grounded source junction field-effect transistor acting as a transconductance amplifier, of a common base transistor T_1 acting as a gain 1 current follower and of a common base transistor T_2 which behaves as a current source, presenting a very large dynamic impedance to the collector of T_1. The idea, already employed in the operational amplifier of Fig. 2.9 is based upon achieving a large dc gain by injecting the drain signal current, delivered by J, into an ideal current follower,
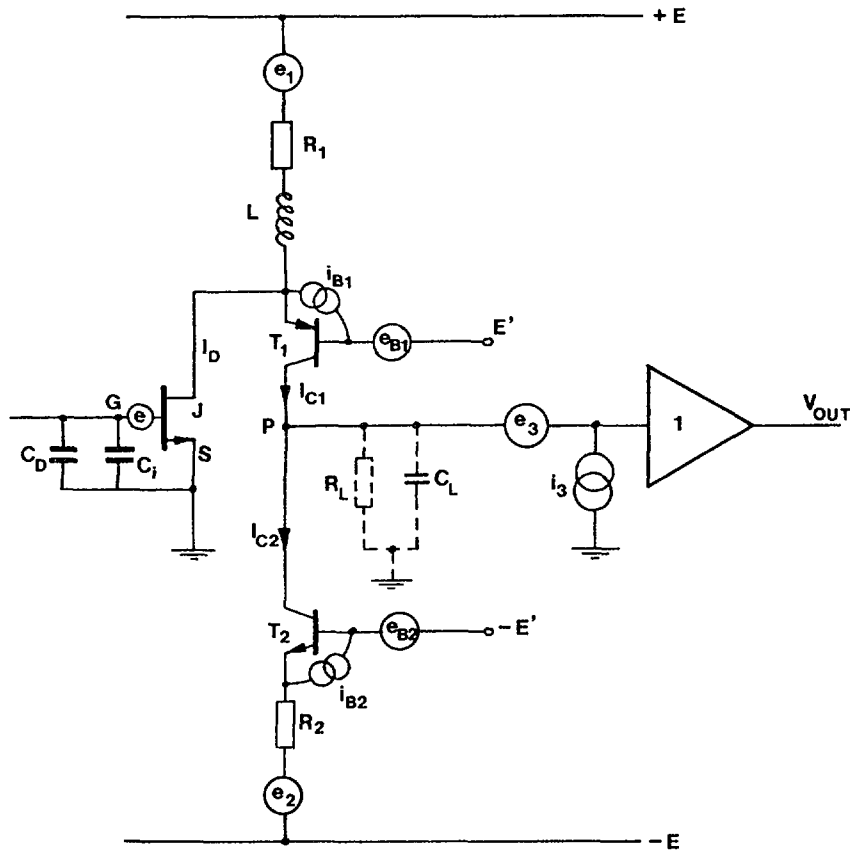
Fig. 3.22:   Basic circuit diagram of a low noise, inverting only amplifier showing the more important noise sources

T_1. This current flows across a very large impedance, the one appearing between node P and ground, owing to the presence of the two collectors in opposition.

The impedance appearing between P and ground is represented by the parallel combination of a resistor and a capacitance, components in dotted lines. The transfer function between the gate-to-source voltage V_{GS} of J and the voltage V_ρ is obtained in the following way. The current signal in the collector of T_1 is $-g_m \cdot V_{GS}$, with g_m transconductance of the FET. Therefore, in the Laplace transform domain we have:

$$\frac{V_p(S)}{V_{GS}(S)} = - \frac{g_m R_L}{1 + S R_L C_L}$$

The dc gain, $- g_m R_L$, by assuming $g_m = 10$ mA/$_V$, $R_L = 10$ M ohm, could be as large as 10^4.

The diagram of Fig. 3.22 can be analysed in the following way. In the standing condition, the current source T_2 absorbs a current given by:

$$I_{C2} = \frac{- E' - V_{BE} + E}{R_2} \qquad V_{BE} \sim 0.7V$$

Such a current has to be provided by T_1. Assume that the input bias current of the output buffer 1 is negligible. Then:

$$I_{C1} = I_{C2}.$$

Of course, this current balance will be met only when the amplifier is in the closed-loop situation, because in the open-loop condition indicated in Fig. 3.22, it would saturate. To determine the current balances analytically, the discussion can be carried on on the open-loop configuration, making the working hypothesis that it is dc stable. The current flowing across R_1 and L is given by $(E - E' - V_{EB})/R_1$ where $V_{EB} \sim 0.7V$ and this current provides I_{C1}, that is, I_{C2} and the drain current of the JFET. Therefore, once I_{CS} is fixed (a reasonable value for it is 1mA), the resistor R_1 must be chosen so as to meet the following current balance:

$$\frac{E - E' - V_{EB}}{R_1} = 1 \text{ mA} + I_D$$

Actually I_D depends on the type of FET employed and must never exceed $9I_{DSS}$, where I_{DSS} is the drain current flowing when $V_{GS} = 0$.

It is interesting now to calculate the contribution to V_{OUT} due to the various noise sources present in the circuit. The calculation will be done again in the open loop configuration. What matters is the relative importance of the various sources and it is unaffected by the feedback.

With reference to Fig. 3.22, we can define the following variables:
e the series noise in the field-effect transistor

e_1 the thermal noise associated with resistor R_1

i_{B1} the base current noise in T_1

e_{B1} the series noise of T_1

e_{B2} the series noise of T_2

i_{B2} the base current noise in T_2

e_2 the thermal noise source associated with R_2

e_3 the series of the output buffer noise source i_3 the parallel noise source of the output buffer.

The inductor L has the following function. If it is adequately large, so that over the whole frequency range of interest it can be considered an open circuit for the signal, both e_1 and e_{B1} are prevented from contributing to the noise current at P. The effect on e_{B1} is clear; it can be pointed out that it acts as a voltage driving on the base of an emitter follower which has an indefinitely large resistor on the emitter. The generator e_{B1} is accordingly unable to change the current in T_1 and therefore it does not contribute to the total noise current at P. The inductor L is a critical circuit component. To be really useful, it has to be large in value (0.1H), has to present a negligible stray capacitance, and be able to stand relatively large currents without saturating. If an inductor of these characteristics is present in the circuit, the effect of noise sources e_1 and e_{B1} becomes negligible. The spectrum of the output noise voltage can now be calculated introducing the values of the varous spectral densities, according to Fig. 3.12, 3.13 and 3.14.

$$\frac{dv_{OUT}^2}{df} = 4kT \cdot 0.7 \cdot g_m + 2qI_{B1} + 4KT\frac{R_{BB1}}{R_2^2} + \frac{4kT}{R_2} + 2qI_{B2} +$$

$$+ \frac{R_L^2}{1 + \omega^2 R_L^2 C_L^2}\frac{di_3^2}{df} + \frac{de_3^2}{df}$$

The term $4kTR_{BB1}/R_2^2$ (remember: $R_{BB} \ll R_2$) can be neglected in comparison to $\dfrac{4kT}{R_2}$.

As long as $g_m \gg 1/R_2$, which is true in most cases

(g_m 10 mA/V, R_2 1k ohm or more) this term is negligible with respect to $4kT \cdot 0.7 \cdot g_m$.

It is interesting, instead, to judge whether the three terms $2q\, I_{B1}$, $2qI_{B2}$, $2qI_{B3}$ added together may to some extent impair the noise performances of the preamplifier, which is expected to depend on the input device alone. Therefore

$$4kT \cdot 0.7 \cdot g_m \text{ has to be compared with } 2q\,(I_{B1} + I_{B2} + I_{B3}).$$

If $4kT \cdot 0.7 \cdot g_m$ has to be larger than $2q\,(I_{B1} + I_{B2} + I_{B3})$, the following condition has to be met:

$$g_m > (I_{B1} + I_{B2} + I_{B3}) \cdot \frac{q}{1.4\ kT}$$

Remembering that at room temperature kT/q 25m V we obtain

$$g_m > (I_{B1} + I_{B2} + I_{B3}) \cdot \frac{10^3}{1.4 \cdot 25} \simeq (I_{B1} + I_{B2} + I_{B3}) \cdot 30$$

Assume $I_{B1} = I_{B2} = I_{B3} = 10\mu A$, that means that these are base currents of bipolar transistors with $\beta = 100$ and $I_c = 1$ mA. In order to make the noise contribution from I_{B1}, I_{B2}, I_{B3} negligible, the transconductance g_m of the FET must be at least 10 times larger than 900 $\mu A/_V$, that is g_m must be at least 9 mA/V.

This condition is not always simple to fulfill. The noise contribution due to the base currents of T_1, T_2 and output stage in several cases is not negligible.

3.10 CONNECTION BETWEEN DETECTOR AND PREAMPLIFIER

As was pointed out in the previous sections, the bias network of the detector and the feedback resistor in the charge sensitive preamplifiers are sources of parallel noise of thermal nature. The detector leakage current and the input bias current of the preamplifier with their shot noise add to the parallel noise. As the parallel noise may seriously degrade the performance of the system at long time constants, it becomes important to reduce these contributions as much as possible.

Fig. 3.23 shows the open-loop circuit diagram of an inverting amplifier which represents a suitable solution to realise the low noise integrator
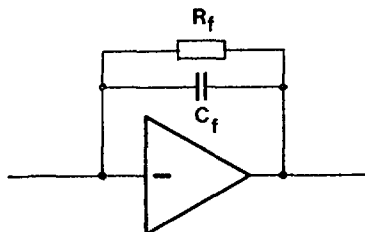

Fig. 3.23: Operational feedback closed around the low noise inverting amplifier.

called charge-sensitive amplifier. The feedback loop is closed around the circuit of Fig. 3.23 by connecting between output and input an impedance made of the parallel combination of the capacitor C_f and of the resistor R_f.

The capacitor C_f accounts for the integrating action of the circuit, while the resistor R_f serves a twofold purpose:

a)   it provides a dc path to the input bias current of the preamplifier, and

b)   it provides a leakage path for the charge deposited on C_f by the ionizing events, thereby preventing saturation in the output level of the preamplifier.

As pointed out in noise analysis, R_f should be made as large as possible in order to make its thermal noise contribution negligible. However, if the preamplifier is employed with a high intensity source of radiation, a large value of R_f may not be compatible with the need to prevent saturation. If I is the average current delivered by the detector, V_{SAT} is the saturation voltage at the preamplifier output, R_f must not exceed the value V_{sat}/I^- but must be close to this value.

To give a practical example, assume a germanium detector which receives γ-rays of fixed energy, 1 MeV each, at a rate of 10^5 events per second and suppose that the preamplifier has a saturation level of 8V. The largest usable value of R_f has to be determined.

Each ionizing event creates in the detector a charge of $5.3.10^{-14}$ Coulomb, whence I = $5.3.10^{-14}.10^5$ = 5.3 nA. The largest value for R_f is:

$$R_f = \frac{V_{SAT}}{I} \approx 2 \text{ G ohm}$$

which can be considered adequately large for most applications.

As to the coupling between detector and preamplifier, the following considerations apply.

If the detector has a large leakage, $I_D \approx 10^{-6}$A, as in the case of a big surface barrier silicon detector operated at room temperature, then the ac coupled connection between detector and preamplifier is advisable, Fig. 3.24.

Alternatively, the AC coupling can be realised as shown in Fig. 3.26, where the isolating capacitor C_{is} is inside the high frequency feedback loop.

The advantage of the circuit in Fig. 3.25 over that in Fig. 3.24 is that the charge delivered by the detector flows entirely through C_f and therefore no charge is lost, as would happen in Fig. 3.24 due to the division between C_D and the series connection of C_{is} and the feedback increased capacitance $C_f(1+A)$. The comparison between the two cases is shown in Fig. 3.26.

The biasing configuration of Fig. 3.26 has four terms of parallel noise, namely shot noise in detector leakage current, shot noise in the preamplifier input current, thermal noise in the detector bias resistor and thermal noise in the preamplifir feedback resistor. Besides, the isolating capacitor C_{is} which has to be introduced between detector and preamplifer is responsible for additional stray capacitance at the preamplifier input.
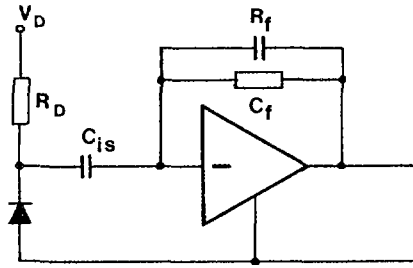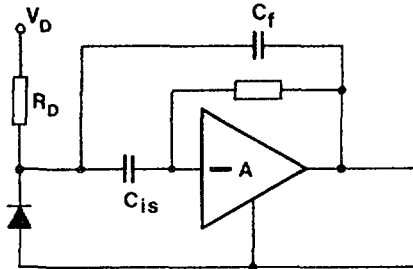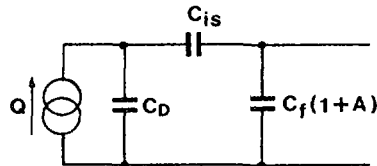
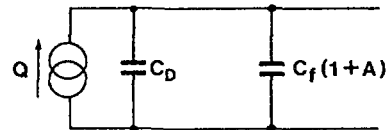Fig. 3.24:   AC coupling between detector and preamplifier

Fig. 3.25:   AC coupling between detector and preamplifier with C_{is}
inside the high frequency feedback loop.

equivalent circuit of the
connection Fig: 3.24

equivalent circuit of the
connection Fig: 3.25
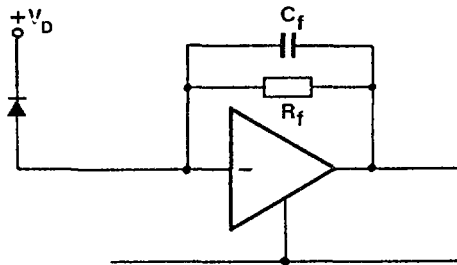
Fig. 3.26: Comparison between the two connections

Fig. 3.27: dc coupling between detector and preamplifier

If the detector leakage current is small enough, as in the case of a high purity germanium detector (I_D can be as small as 10^{-14}A), the dc coupling between detector and preamplifier is possible (see Fig. 3.27).

With respect to the ac coupling of Fig. 3.27, a source of thermal noise, the one associated with R_D is removed and C_{is} is no longer necessary.

If both detector and preamplifier input stage work at cryogenic temperature, then the only source of parallel noise is R_f.

A possible way of removing R_f is based upon the use of optoelectronic charge reset, Fig. 3.28.

77
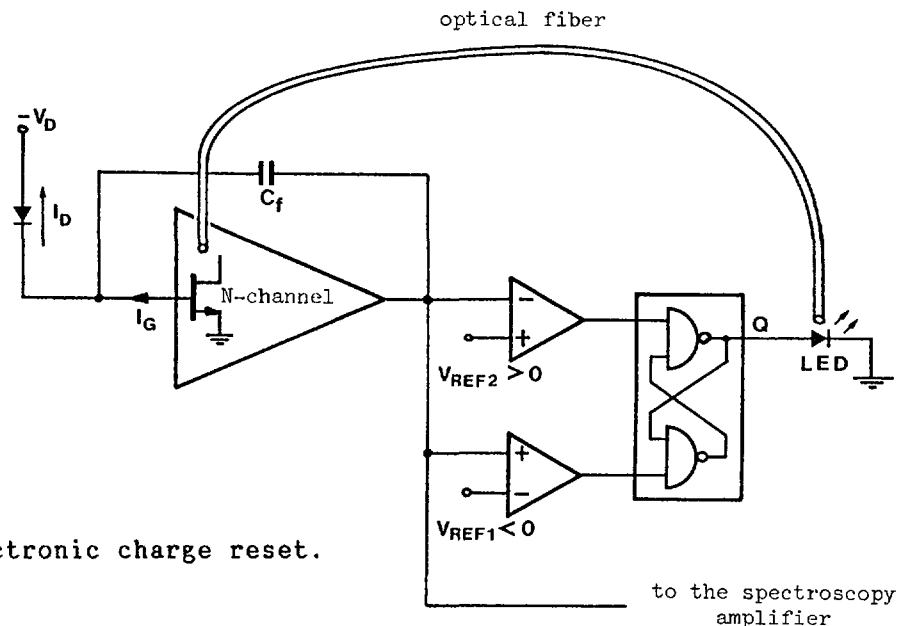
Fig. 3.28: Optoelectronic charge reset.

Both detector and input JFET are supposed to work at cryogenic temperature and I_D will be assumed to be larger than I_G. In the absence of radiation falling on the detector, C_f will slowly charge through the difference $I_D - I_G$, and the output of the preamplifier would tend to the positive saturation level if it were not for the comparator 2, which changes state as soon as the reference V_{REF2} is passed. The comparator sets the flip-flop with the Q output at logic HIGH, thereby turning ON the light emitting diode. The optical fibre couples the light emitted by the LED to the gate-to-drain junction of the FET. The light increases the rate of electron-hole generation in the FET well beyond the thermal level. There is an increase in the gate current which flows into C_f, quickly discharging it. The operation proceeds until the output of the charge sensitive preamplifier goes below V_{REF1}, thereby switching the comparator 1. The output transition of the comparator 1 resets Q to logic LOW, turning off the LED. From this instant, C_f starts charging slowly through I_D-I_G intil a new light pulse initiates a charge resetting operation similar to the previous one. The same procedure of charge reset by optoelectronic means occurs also when radiation falls on the detector. In this case the rise of the preamplifier output voltage towards V_{REF2} occurs in steps, rather than in a ramp-like fashion. The voltage at the preamplifier output, under the infuence of the ionizing events in the detector has the shape of a stochastic staircase, until V_{REF2} is reached and the reset action begins.

The optoelectronic charge reset is used in spectrometry systems (essentially those for x-ray energy dispersive analysis) avoiding a thermal noise source, like that associated with R_f, may sensibly improve the resolution.

3.11 CHOICE OF THE PREAMPLIFIER INPUT DEVICE

In the noise analysis developed so far, some simplifying assumptions have been made. The dominant series noise contribution in a bipolar transistor was assumed to be the one related with the thermal noise in R_{BB1}. The shot noise associated with the collection current, which referred to the input would be represented by a voltage source of spectral power density $0.5/g_{mB}$, with g_{mB} transconductance of the bipolar transistor, was neglected. Such an assumption is reasonable in conventional

78

bipolar transistors, featuring values of R_{BB1} of 100 ohm or more, when the standing collector current exceeds a few hundred μA and $0.5/g_{mB}$ turns out to be 30 ohm or even less.

The assumption, however, is no longer true for the modern microwave bipolar transistors, featuring values of R_{BB1} in the 10-15 ohm range. These transistors, that have values of gain-bandwidth product of some GHz at I_C around 5mA and still above 1GHz at I_C = 1 mA, can be competitive, in some applications, with the junction field-effect transistor. The curves of Fig. 3.29 are boundary curves in the $[C_D, \tau_M]$ plane that separate the regions where a single, state-of-the-art junction field-effect transistor exhibits smaller or larger ENC than a state-of-the art microwave bipolar transistor. The shaping adopted for this comparison is an almost Guassian one, with a peaking time τ_M. As shown in Fig. 3.29, the field-effect transistor at fixed τ_M behaves better, noisewise, than the microwave bipolar transistor at lower C_D values, while at larger C_D the bipolar transistor is better. At fixed C_D, the bipolar transistor is superior at short τ_M values, while the junction field-effect transistor becomes superior at larger values of τ_M.
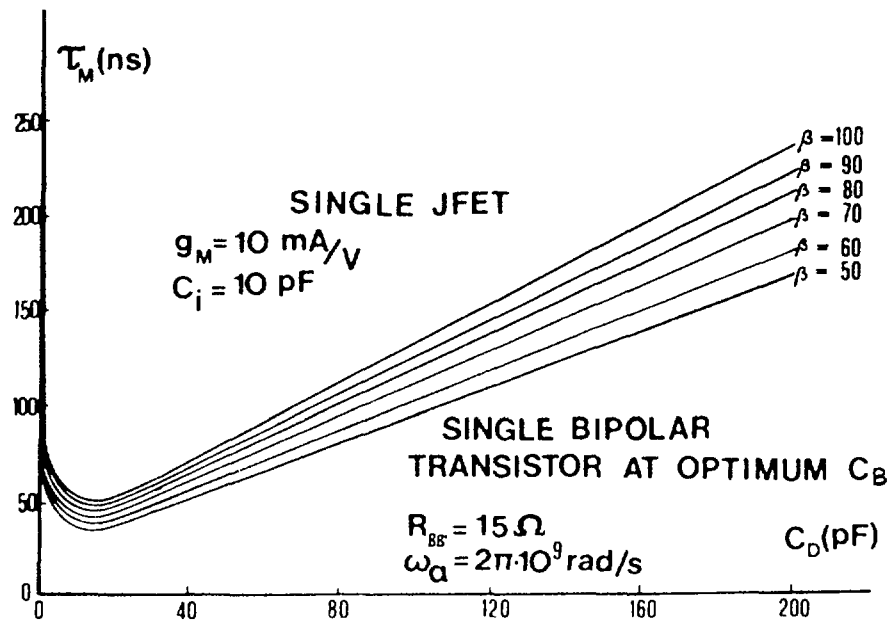
Fig. 3.29:   Boundary curves in the (C_D, τ_M) plane, comparing the behavior of JFET with microwave bipolar transistor. The β of the bipolar transistor is used as a parameter.

If the shot noise in the collector current of the bipolar transistor is accounted for, the series generator assumes the approximate expression

$$\frac{\overline{e^2}}{\Delta f} = 4kT \; R_{BB1} + \frac{0.5}{gm}$$

Another noise contribution which has been so far neglected, is a series noise term with a spectral power density inversely proportional to the frequency, the so-called $\frac{1}{f}$ -type of noise:

$$\frac{\overline{e^2}}{\Delta f} = \frac{A_2}{\omega}$$

79

This noise is negligible in most of JFET's and bipolar transistors, but it is not negligible in devices like the Ga As FET or the silicon MOSFET. It is therefore interesting to give a general expression of the equivalent noise charge, which describes all the active devices, JFET, bipolar transistors, Ga As FETs and MOSFETs. Such an expression will be given here for a rather general situation, that is, the one in which the shaping employed is a triangular one, with peaking time τ_M. The expression of ENC^2 in the case of the charge-sensitive preamplifier is:

$$ENC^2 = A_1 \cdot (C_D + C_i + C_f)^2 \cdot \frac{1}{\tau_M} + A_2 \cdot (C_D + C_i + C_f)^2 \cdot \frac{2 \cdot \ln 2}{\pi} + \frac{1}{3} \cdot A_3 \cdot \tau_M$$

where for the bipolar transistor

$$A_1 = 4kT \left(R_{BB1} + \frac{0.5}{g_m} \right)$$

$$A_2 \simeq 10^{-15} \text{ V}^2$$

$$A_3 = 2q \, (I_D + I_B) + 4kT \cdot \left(\frac{1}{R_D} + \frac{1}{R_f} \right)$$

for the JFET

$$A_1 = 4kT \cdot \frac{0.7}{g_m}$$

$$A_2 \simeq 10^{-15} \text{ V}^2$$

$$A_3 = 2q \, (I_D + I_G) + 4kT \left(\frac{1}{R_D} + \frac{1}{R_f} \right)$$

for the Ga As Schottky Barrier FET

$$A_1 = 4kT \cdot \frac{0.7}{g_m}$$

$$A_2 \simeq 10^{-13} \div 10^{-12} \text{ V}^2$$

$$A_3 = 2q \, (I_D + I_G) + 4kT \left(\frac{1}{R_D} + \frac{1}{R_f} \right)$$

for the silicon MOSFET

$$A_1 = 4kT \cdot \frac{0.7}{g_m}$$

$$A_2 \simeq 10^{-11} - 10^{-10} \text{ V}^2$$

$$A_3 = 2q \cdot I_D + 4kT \left(\frac{1}{R_D} + \frac{1}{R_f} \right)$$

In the general expression of ENC^2, all the parallel noise contributions, due to either thermal noise or shot noise components in the detector in the preamplifier and in the relevant bias networks, have been

considered. Therefore, the expression of ENC^2 accounts for the most general case of ac coupling between detector and preamplifier. For the silicon MOSFET, $I_G = 0$.

It is interesting to note that in the expression of ENC^2, there are three terms with the following τ_M dependence:

TERM ARISING FROM SERIES WHITE NOISE: it is inversely proportional to τ_M, which means, this term becomes important at short shaping times, in spectroscopy amplifiers intended for high counting rates.

TERM ARISING FROM SERIES $\frac{1}{f}$ -type NOISE: it is independent of τ_M

TERM ARISING FROM PARALLEL NOISE: it is proportional to τ_M and it therefore becomes important at long shaping times.

The effect of the different noise sources on ENC and the behaviour of the active devices considered so far are summarized in Fig. 3.30. The two figures represent in a log-log plot the equivalent noise charge as a function of the peaking time τ_M at two different values of the detector capacitance C_D ($C_D = 10$ pF and $C_D = 30$ pF, respectively).
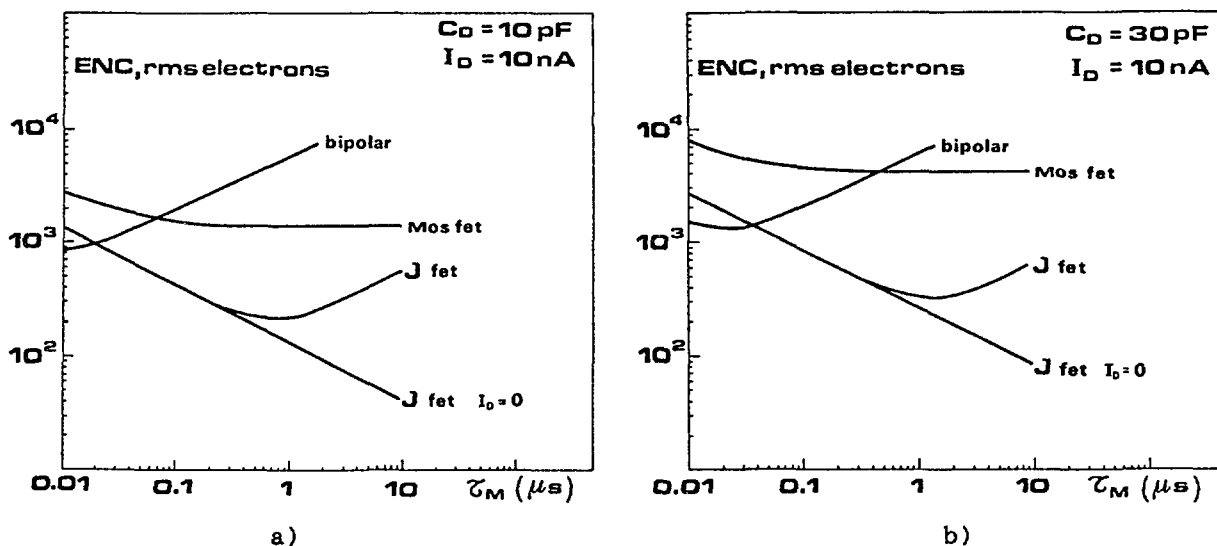
Fig. 3.30:   Noise charge dependence on peaking time: a) $C_D = 10pF$,
b) $C_D = 30pF$

As shown by both plots, the microwave bipolar transistor is the best device at τ_M in the 10 ns range. As soon as τ_M is increased, the ENC also increases as a consequence of the large parallel noise in the bipolar transistor.

The junction field-effect transistor presents, on a broad τ_M range, a decreasing ENC as τ_M is increased. In absence of thermal noise in the bias network ($R_D \to \infty$, $R_F \to \infty$) and neglecting the detector leakage current, ENC is still decreasing at $\tau_M = 10\mu s$ or longer. A 10 nA detector leakage current would make ENC for the JFET increase, after a minimum occurring around $\propto 1\mu s$. For the MOSFET, after an initial decrease of ENC as a function of τ_M, it settles to a constant value, determined by the series $\frac{1}{f}$ - type of noise. So, while the bipolar and the junction field-effect transistors have their characteristic domain of application as a function of τ_M, the MOSFET has a very annoying uniform

behaviour of ENC versus τ_M. For this reason, the MOSFET has not been considered so far a useful low noise device. However, should the $\frac{1}{f}$ - noise be improved, as expected, by future technological advancement, the MOSFET might be considered with interest, because a MOS monolithic chip might include, besides low-noise amplification, several other analog functions, among which analog switching, sampled-data filtering and analog-to-digital conversion. So, a single chip could include an entire nuclear pulse spectrometer.

CHAPTER 4

FURTHER ANALOG PROCESSING

4 FURTHER ANALOG SIGNAL PROCESSING BEFORE THE ANALOG-TO-DIGITAL CONVERSION

4.1 INTRODUCTORY REMARKS

The analog signals delivered by the linear amplifying and shaping section require further processing before analog-to-digital conversion. The following three operations are usually performed:

    a) baseline restoration
    b) linear gating
    c) analog storage.

Baseline restoration was introduced in Section 1 functional approach. Here attention will be concentrated on the circuits that implement it.

Baseline restoration has the purpose of reducing baseline fluctuations at high counting rates as well as that of removing low frequency noise and thermal drift. Baseline fluctuations are introduced by the long tails, generally of opposite polarity with respect to the signal. Although the AMPLIFYING SYSTEM IS SUPPOSED TO USE ONLY POLE-ZERO CANCELLATION NETWORKS and not differentiations to clip the signals, improperly adjusted pole-zero pairs may still give rise to unwanted tails. The baseline restorer aims at further reducing these tails. Moreover, some low frequency noise, 50 Hz pickup and microphonic noise may be preesent at the amplifier output, despite the action of high frequency filters implemented by the pole-zero cancellation networks. Baseline restoration can be very effective in removing noise of this nature. The baseline-restoring circuits in most cases are inside the amplifier module. All contemporary, commercial spectroscopy and research amplifiers have a built-in baseline restorer.

The other two operations, that is, LINEAR GATING and ANALOG STORAGE of the peak amplitude are performed by the analog section of the analog-to-digital converter.

Linear gating has the purpose of selecting on the pulses to be transmitted to the analog-to-digital converter. It may happen, in some measurements, that out of many pulses delivered by the amplifier, only those that meet some particular requirements, have to be analyzed. Or, it may be required that the pulses coming from the amplifier have to be analyzed with the exception of those to which some specific rejection rules are applied. In either case, the amplifier signals are selected by a circuit called LINEAR GATE. The signals transmitted by the linear gate are ready for pulse amplitude analysis. As already pointed out, the information of interest, that is, the information related with the energy released in the detector is carried by the peak amplitude of the signal at the amplifier output. Whatever principle is employed for analog-to-digital conversion, (with the exception of the FLASH-type) a temporary storage of the peak amplitude of the signal is required, for the operation of analog-to-digital conversion is not an instantaneous one. The temporary storage of nuclear pulses is often referred to as PEAK STRETCHING and the circuit which performs it is called a PEAK STRETCHER.

The sequence of three analog operations described so far is summarized in Fig. 4.1.
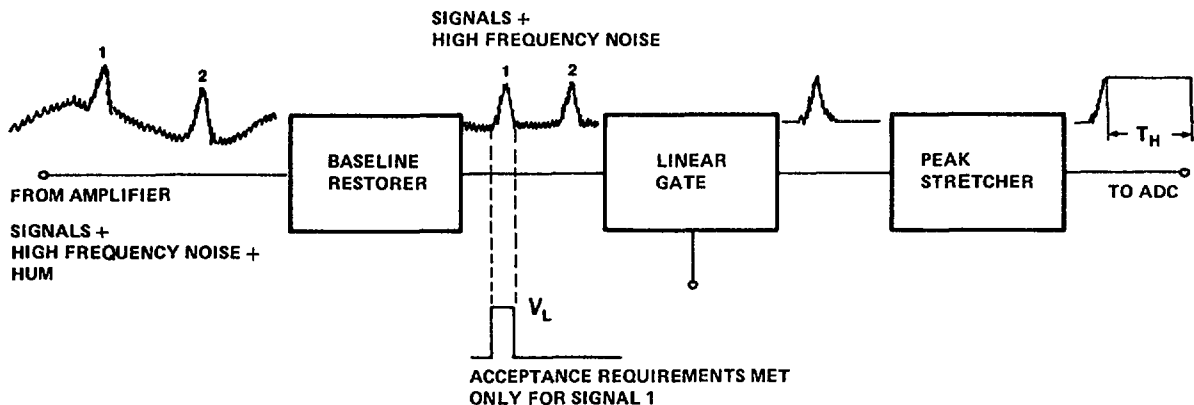
Fig. 4.1:   Sequence of operations performed prior to analog-to-digital
            conversion on the signals at the output of the amplifier

As shown in Fig. 4.1, the signals at the output of the amplifier are
assumed to be superimposed to high frequency statistic noise, such as the
noise due to the preamplifier and to a low frequency deterministic noise,
for instance a 50Hz pick-up. The baseline restorer keeps the baseline free
from low frequency noise. It has also been assumed in Fig. 4.1 that only
those nuclear signals are accepted, for which the positive logic signal V_L
is present. The presence of V_L states that the experiment acceptance
requirements are met. The linear gate, transmits signal 1 and stops signal
2, as for the latter the acceptance signal is not present. Finally, the
peak stretcher holds the peak of the signal passed by the linear gate, for
the time T_H needed by the analog-to-digital converter. The basic circuits
that implement the three functions will now be described.

4.2 BASELINE RESTORERS

A baseline restorer can be thought of as a nonlinear differentiator
which presents a long time constant for the signal and a very short time
constant for the baseline in the absence of signal. The principle of a
baseline restorer is described by the circuit of Fig. 4.2.
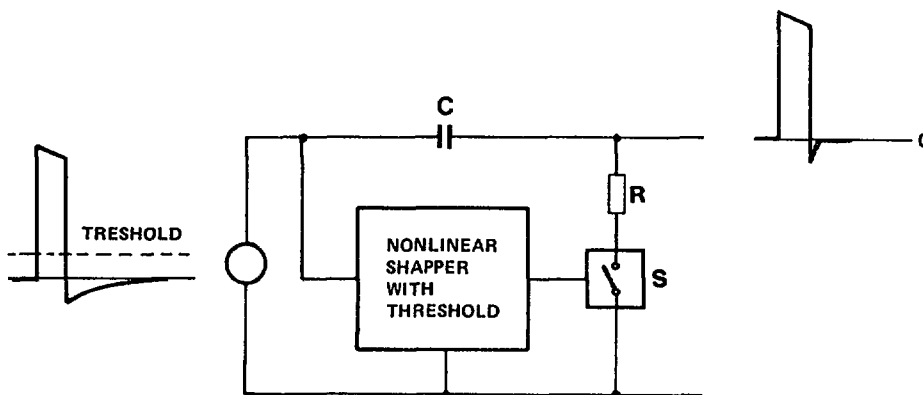
Fig. 4.2: Basic diagram of a baseline restorer

In the absence of a signal the switch S is closed and the circuit
behaves for noise and drift like a low pass filter with a cut-off frequency

$$f_c = \frac{1}{2\pi RC}$$

The time constant RC is made small enough $(C = 10^4 \div 10^5 pF,$
$R \sim 10\Omega)$ to give values of f_c in the 100 KHz - 1MHz range.

Therefore, low frequency variations in the baseline are strongly attenuated. The incoming signal, if its amplitude exceeds the threshold of the nonlinear shaper which controls the state of the switch, makes S go into the OPEN condition. The time constant of the differentiator becomes very long and the signal is transmitted unchanged to the output. As soon as the positive part of the signal is over, the nonlinear shaper triggers back to the previous state. The switch S closes and the time constant of the differentiator reaches again the small value. The negative tail after the signal, therefore, recovers quickly to zero. As a result, the original negative tail following the signal, which may give rise (owing to tail-on-tail superposition) to a negative baseline component, is transformed by the baseline restorer into a short tail. The effect of tail-on-tail superposition is strongly reduced. The choice of the time constant RC is a very important aspect in the optimization of the performances of a baseline restorer. According to the discussion developed so far, it might seem that the smaller the RC, the better, for the baseline restorer becomes a more effective filter for the low frequency baseline fluctuations and it also gives a faster tail recovery after the pulse.This is, however, not true if the effect of the value of RC on the high frequency noise is also taken into consideration. It is easy to show that, like the baseline subtractor of Fig. 1.13 in Section 1, a baseline restorer with a very small value of RC can introduce a $\sqrt{2}$-factor increase in the high frequency noise. The explanation is given in Fig. 4.3, where the time constant RC has been made equal to zero for the sake of simplicity.
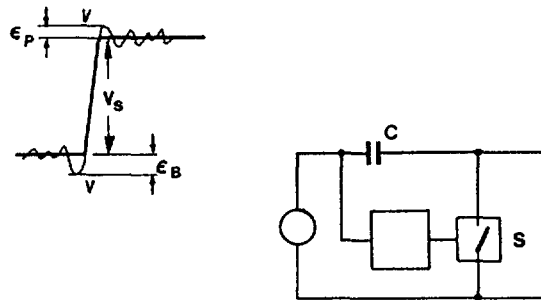
Fig. 4.3:   Effect of a very small RC on the high frequency noise.

Owing to the negligible small value of RC, when S is closed the voltage across C reproduces the input variations, no matter how fast they are. The leading edge of the input signal C is charged to the negative voltage sample $-E_B$. This negative voltage sample is nothing but the instantaneous noise value before the pulse. For high frequency noise, noise values taken at short time distance are uncorrelated. So, for instance, the noise sample immediately after the leading edge of the signal might be positive. The signal amplitude which appears at the output of the baseline restorer once S is opened is the full excursion, from $-v$ to v, that is,

$$V_s + | E_B | + | E_p |$$

which accounts for the noise increase due to the baseline restorer when the time constant at S closed is too small. It has to be emphasized once again, that the described effect is similar to the one discussed in Section 1 talking about the pulse amplitude measurement performed as a difference between the pulse peak amplitude and the instantaneous baseline value before the pulse. The baseline restorer with a very small value of the time constant rC performs exactly the same kind of operation: it, takes the difference between the peak amplitude of the signal, which is affected by the noise and the value of the baseline before the signal. All those noise

and drift components that do not change appreciably from an instant
preceding the pulse to the peaking time of the pulse are effectively
cancelled out. Those that may change considerably in this time interval
(like high frequency noise) appear with enhanced importance. The worsening
effect of baseline restorer, in the limiting case may be as large as the
$\sqrt{2}$ factor in the noise linewidth, that is, a 40% reduction in resolution.

The diagram shown in Fig. 4.2 corresponds to a rather general
configuration of baseline restorer, from which the more modern circuits
have been developed. Other types are simplified versions of the general
principle, from which they differ because the switch S is realized simply
by a diode, which is turned ON and OFF by the analog signal itself.

To this type belong the two baseline restorers described below.

The first, is a symmetric restorer, based upon two diodes and an
operational amplifier, (see Fig. 4.4.)
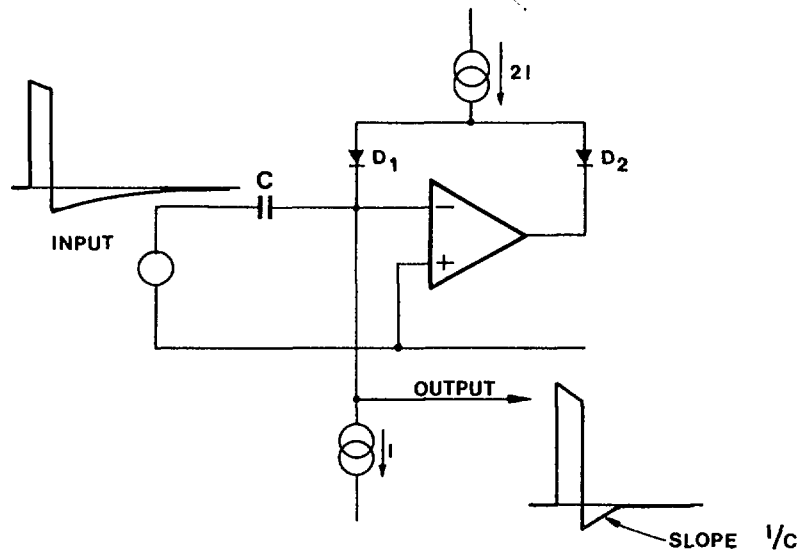
Fig. 4.4: Symmetric baseline restorer

Prior to the arrival of the pulse, with the circuit in the waiting
state, no current flows into the capacitor C and therefore the diodes D_1
and D_2 work at equal currents I. If they are supposed to be equal and
the operational amplifier has zero offset and very high gain, the voltage
at the inverting input as well as the amplifier output voltage are zero.
Assume now a positive signal with a slow negative tail at the input. If
the capacitor C is adequately large, it can be considered as a shortcircuit
on the leading edge of the signal, even if the leading edge is not as fast
as that shown in the Fig. 3.2. On the positive-going edge, the diode D_1
is turned off. The current of the lower source, I, is now diverted into
the capacitor and charges it, with the effect of increasing the negative
slope on the top of the pulse.

The negative-going edge of the signal pulls the cathode of diode D_1
negative, thus turning off D_2. It has to be pointed out that as soon as
the cathode of D_1, which is connected to the inverting input of the
operational amplifier, becomes slightly negative, an amplified positive
voltage appears on the cathode of D_2, thereby quickly reverse biasing
it. The current delivered by the upper source 2I flows through D_1. A
part I of this current is absorbed by the lower generator. The difference,
2I-I=I flows through the capacitor. The output signal rises linearly with

88

a slope I/C towards the zero line. The long tail of the input signal is
transformed into a short, linearly recovering tail.

The described restorer may work equally well for positive or negative
input pulses.

A baseline restorer of non symmetric nature, designed to cope with
very high counting rates is shown in Fig. 4.5. Its main feature is a very
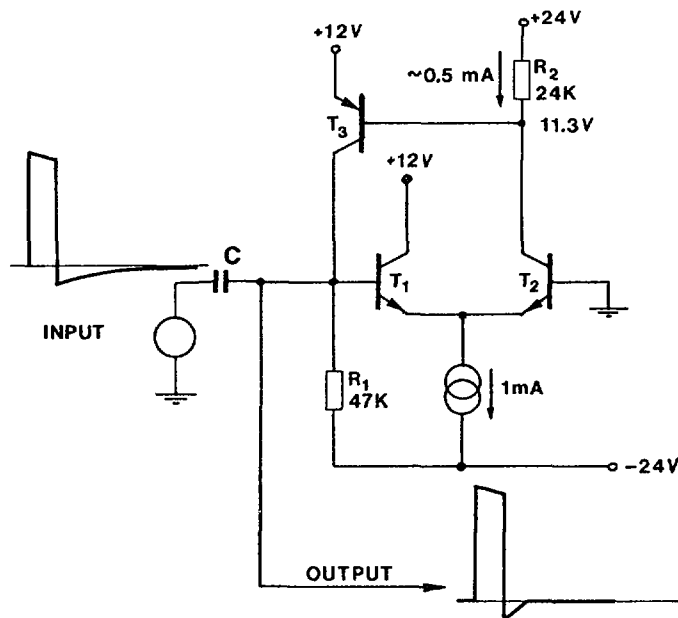fast recovery of the tail.

Fig. 4.5:   Highly asymmetrical baseline restorer

The dc analysis of the circuit of Fig. 4.5 shows that the transistors
T_1, T_2 work in a balanced condition, with about 0.5mA each. The base of
T_1 is near 0V if T_1 and T_2 are identical. The collector current of
T_3 is determined by the 47 K resistor which has a voltage drop across it
of 24V. Therefore, the collector current of T_3 is about 0.5mA.

The positive input signal increases the current in T_1 and reduces
that of T_2. The reduction in the collector current of T_2 takes place at
the expense of the base current of T_3. T_3, therefore, goes off as soon
as the input signal exceeds a few mV. Once T_3 is off, the 0.5mA current
flowing through R_1 is diverted into the capacitor, thus increasing the
slope on the top of the signal. It is worth pointing out that the current
which flows into the capacitor on the positive signal is limited to about
0.5mA.

When the input signal recovers to the negative value of the tail, there
is a reduction in the current through T_1 and an increase in the current
through T_2. Such an increase flows entirely into the base of T_3 and
multiplied by its ñ (a factor 100) flows into C, thus forcing a quick
recovery of the negative tail towards the zero baseline. With the values
shown in Fig. 4.5, a 10 mV negative tail would force out of the collector of
T_3 into the capacitor C a current of about 10 mA. There is, therefore, a
ratio 20 between the current available to restore the baseline on the
negative tail and the current available to charge the capacitor during the
signal. The asymmetrical baseline restorer ensures a quick recovery of the
tail. However, its worsening effect on the high frequency noise is more
pronounced than that of a symmetrical restorer. Therefore, the use of

highly asymmetrical baseline restorers should be restricted to the case in which it is strictly necessary. For instance, where the counting rates exceed 10^5 pulses per second. At lower counting rates it is advisable to employ a symmetrical restorer.

4.3 LINEAR GATES

In this section the concept of analog switch will frequently be employed. The analog switch is represented by the symbol shown in Fig. 4.6. The terminal V_L is the control input. The state, either OPEN or CLOSED of the analog switch is controlled by the logic state, either LOW or HIGH of the voltage at the control input. It will be assumed, according to Fig. 4.6 that the logic state V_L = HIGH corresponds to the switch CLOSED and that the logic state V_L = LOW corresponds to the switch OPEN.
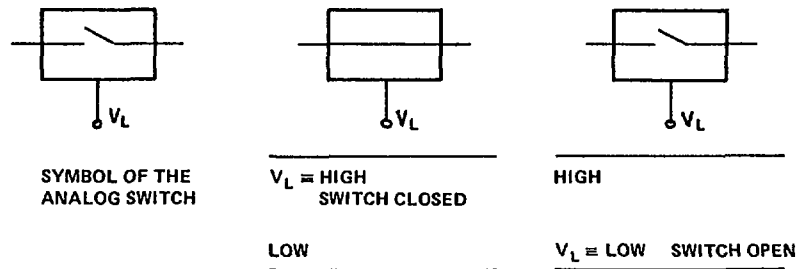
Fig. 4.6: Analog switch

The two selection functions operated on a sequence of nuclear events are shown in Fig. 4.7.

a)   The analog nuclear events are accepted provided that an acceptance logic signal (V_L = HIGH) is present. Therefore, according to the time diagrams, nuclear pulses 1 and 4 are rejected, 2 and 3 are accepted.
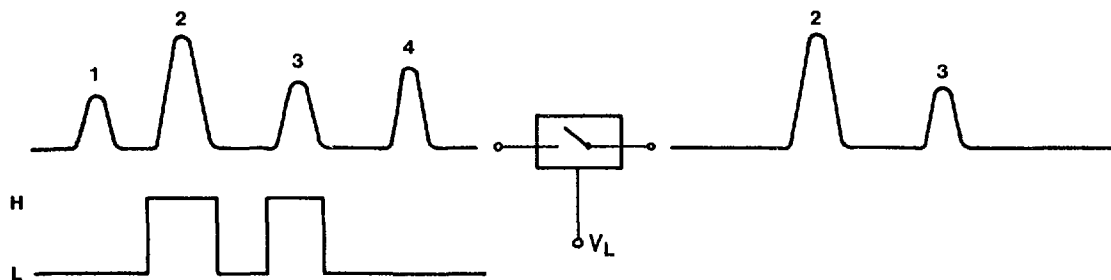
b)   The analog nuclear events are accepted unless a rejection logic signal (V_L = LOW) is present. According to the time diagrams, the nuclear pulses 1,2,4,6 are accepted, while 3 and 5 are rejected.
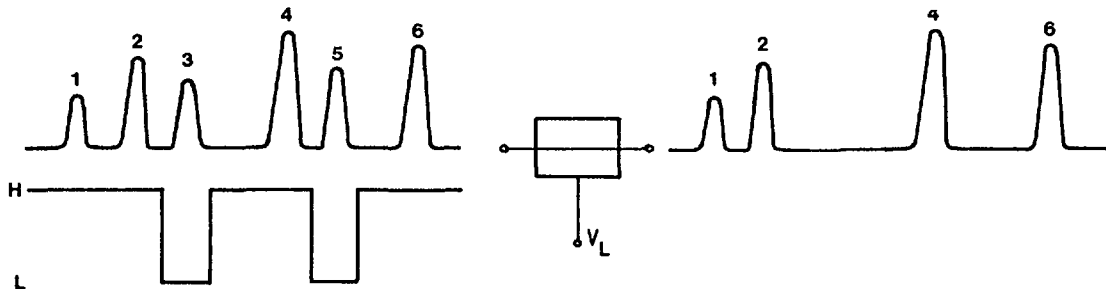
Fig. 4.7:   Acceptance conditions of nuclear pulses and their realization with an analog switch.

90

In practical circumstances, the transmission and rejection functions shown in Fig. 4.7 are better realized by circuits called LINEAR GATES. A linear gate is a composite structure employing one or more analog switches. Linear gates belong to three different types, as shown in Fig. 4.8.
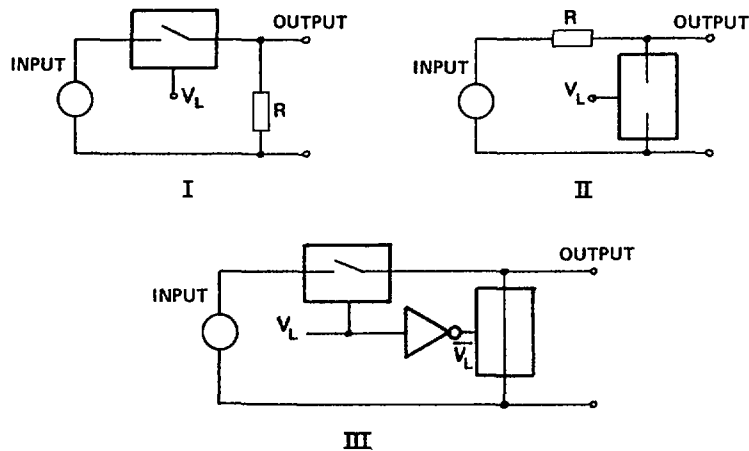
Fig. 4.8: Basic configurations of linear gates

Linear gate types are described as follows:

(I) series-type:
- the linear gate is OPEN when the switch is CLOSED
- the linear gate is CLOSED when the switch is OPEN

(II) parallel-type:
- the linear gate is OPEN when the switch is OPEN
- the linear gate is CLOSED when the switch is CLOSED

(III) series-parallel type:
- the linear gate is OPEN when the series switch is CLOSED and the parallel switch is OPEN
- the linear gate is CLOSED when the series switch is OPEN and the parallel switch is CLOSED

Observe the logic inverter between the control inputs of series and parallel switch.

Now, the problem arises of how to realize the analog switches. There are several possibilities:

- saturated transistors,
- junction field-effect transistors,
- complementary MOS transistors, and
- diode bridges.

Junction field-effect transistors in series-parallel configurations offer very good performances, but they require rather sophisticated control circuitry.

CMOS switches are simple, are available in integrated structures with multiple switches and for this reason are becoming very popular in industrial and consumer applications. In the field of nuclear pulse processing they present the disadvantage of fairly large switching transients.

For the above explained reasons, attention will be restricted here to two linear gates: one based upon saturated transistors and suiting the field of applications with relatively slow pulses, some microseconds base-width and one employing the diode bridge principle for fast applications, down to a few tens of nanoseconds base-width.

The analog switch employing a saturated transistor is shown in Fig. 4.9.
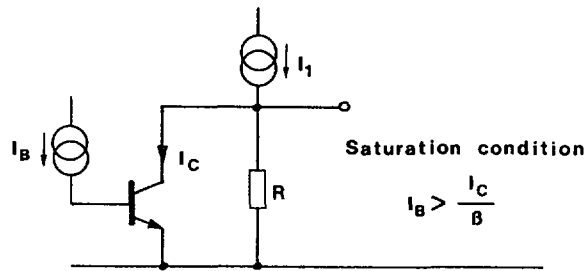
Fig. 4.9: Saturated transistor as analog switch

Assume that the currents I_1 and I_B are fixed by two independent generators - and assume that the fraction of I_1 which flows through the resistor R is negligible with respect to the transistor collector current I_C. If the control current I_B is chosen in such a way that: $I_B > I_C/\beta$ the transistor becomes saturated. It presents a low collector-to-emitter voltage and a low collector-to-emitter dynamic resistance. It is possible to show that, if $I_B \gg I_C/\beta$ the saturation V_{CE} approaches the value:

$$V_{CE,SAT} = \frac{kT}{q} \cdot \ln (1/\alpha_{INV})$$

In the equation above, $kT/q = 25$ mV at T = 300 K and α_{INV} is the reverse current gain, that is, the current gain between the COLLECTOR WORKING AS EMITTER and the EMITTER WORKING AS COLLECTOR. It has to be pointed out that the emitter and collector regions of a bipolar transistor, having the same type of doping, can be exchanged. Unless purposely designed to be used indifferently in the NORMAL and REVERSE condition, a bipolar transistor has a much lower current gain in the REVERSE than in the NORMAL connection. So, while αNORM, that is, the current gain between the EMITTER working as EMITTER and the COLLECTOR WORKING AS COLLECTOR approaches unity, α_{INV} may be as low as 0.1. The value of the saturation voltage, accordingly, may be of the order of

$$\frac{kT}{q} \cdot \ln 10 \gg 60 \text{ mV}$$

When it is important to reduce the saturation voltage, it is advisable to employ the REVERSE CONNECTION, as shown in Fig. 4.10.
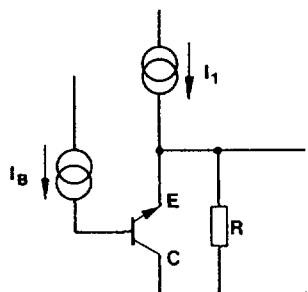
Fig. 4.10: Saturated transistor in REVERSE CONNECTION as analog switch.

In this case $V_{CE,SAT}$, approaches, at large values of I_B the value $\frac{kT}{q}$. ln $(1/\alpha_N)$ that is, for $\alpha_N = 0.99$, less than 1 mV.

In both cases of Fig. 4.9 and 4.10, when I_B is present the saturated transistor acts as a closed switch and the voltage across R equals the transistor saturation voltage.

Once I_B is turned OFF, the transistor goes OFF, thus reproducing the condition of OPEN SWITCH and the current I, flows across the resistor R, developing through it the output signal.

The complete circuit diagram of a linear gate employing two cells of parallel-type is shown in Fig. 4.11.
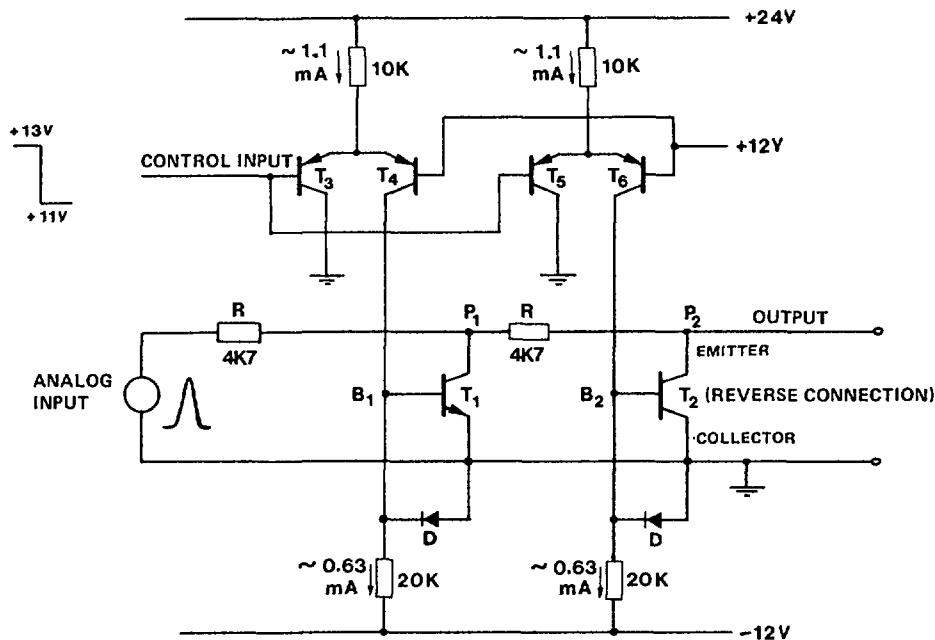
Fig. 4.11:   Details circuit diagram of linear gate employing
             saturated transistors.

The two cells, resistor R and transistor T_1, resistor R and transistor T_2 constitute the parallel-type linear gate. The first cell, which performs the first, rough attenuation of the input signal when the gate is closed, employs a transistor in NORMAL CONNECTION, while the second cell, which defines a precise output level when the gate is closed, employs a transistor in the REVERSE CONNECTION. The long tailed pairs (T_3, T_4 and T_5, T_6) represent the control circuits. The linear gate is intended for unipolar, positive input pulses with amplitude ranging between 0 and 10V. The control levels, + 11V, + 13V are applied to the bases of T_3 and T_5. When the control input is at + 13V, T_3 and T_5 are OFF, T_4 and T_6 are ON. The collector currents of T_4 and T_6 are given by 11.3V/10K = 1.1mA.

These currents are injected into the nodes B_1 and B_2. The currents leaving the nodes through the 20 k resistors are of the order of 0.6mA each. There is, therefore, a difference current of about 0.5 mA flowing into the base of both T_1 and T_2. The transistors T_1 and T_2 are therefore saturated and the linear gate is accordingly closed.

To open the gate, the control input must be decreased down to 11V. In this condition T_3 and T_5 are turned ON, while T_4 and T_6 are turned OFF. The collector currents of T_4 and T_6 stop flowing. At the nodes B_1 and B_2 only the two currents fixed by the 20 k resistors keep flowing. These currents are in the opposite direction to the one required to maintain T_1 and T_2 saturated. It means that these currents quickly drain out the charge stored in the base of the transistors T_1, T_2 during the previous saturation phase. As a matter of fact, the two 20 K resistors connected between B_1 and B_2 and the negative supply voltage serve the purpose of PROVIDING THE REVERSE CURRENTS THAT ENSURE A QUICK REMOVAL OF THE STORED CHARGE AND, THEREFORE, A QUICK SWITCHING TO THE OFF STATE. Once the stored charge is removed, B_1 and B_2 would move towards the negative supply and might cause reverse breakdown in the base-to-emitter junctions of T_1 and T_2, if they were not clamped to about -0.7V by the diodes D_1 and D_2. The described circuit is simple and accurate. It is not very fast in the transmitting state, because of the bandwidth limitations arising from the two RC integrators determined by the 4.7K resistors and by the stray capacitances shunting P_1 and P_2 to ground.

The design of a linear gate capable of operating on short pulses can be based upon a diode-bridge analog switch. Sets of matched diodes, sometimes already connected in a bridge configuration, are made available by some manufacturers. Some of these bridges are based upon fast diodes of the Schottky-barrier type. If a matched set of four diodes is not available, it can be realized by selecting from a batch of diodes four units with similar characteristics. A simple criterion is that of choosing the four diodes that present the closest forward voltages when a 5 ÷ 10 mA current flows through them.

It will be assumed in the following analysis that the four diodes employed in the bridge are perfectly identical.

The circuit diagram of a linear gate employing a diode bridge as a series switch is shown in Fig. 4.12.

The diode bridge is ON when the two currents 2I injected and absorbed respectively by upper and lower generator are present. Under the hypothesis of perfectly identical diodes, with the input at zero volt, the output will also be at zero volt and the four diodes work at equal currents I. If the resistor R, which represents the load is large enough in comparison to the incremental resistance of the diodes, the same conditions occur even if a signal is present at the input. To meet the requirements of $R \gg r_d$, for instance $R = 10k$, $r_d = 5\Omega$, an isolation buffer may be required between diode bridge and load. If the above mentioned condition for R is respected, the output signal tracks the input with a high degree of accuracy.
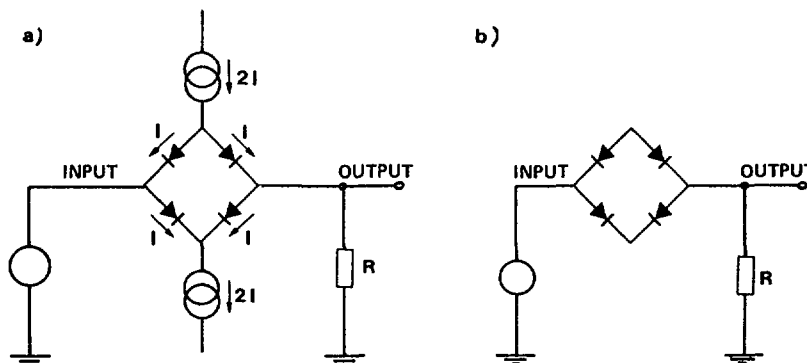

Fig. 4.12: Linear gate. A series switch is using a diode bridge:
a) gate open, b) gate closed

To open the transmission between input and output it is enough to remove the current sources 2I, Fig. 4.12.b. In this condition, the input signal does not reach the output as along both upper and lower paths, it finds two BACK-TO-BACK DIODES. Then, the output voltage is equal zero regardless of the input amplitude.

In the design of a linear gate employing a diode bridge, due attention must be paid to the design of the two switchable current sources 2I. To avoid spurious transients at the output at the opening and closing instants, these current sources must be turned OFF and ON simultaneously. Two long-tailed pairs, one made of fast PNP transistors, the other made of fast NPN transistors may suit the purpose, as will be shown soon. To improve the behaviour of the linear gates based upon diode bridges, a configuration of the series-parallel type can be employed, according to the detailed diagram of Fig. 4.13.
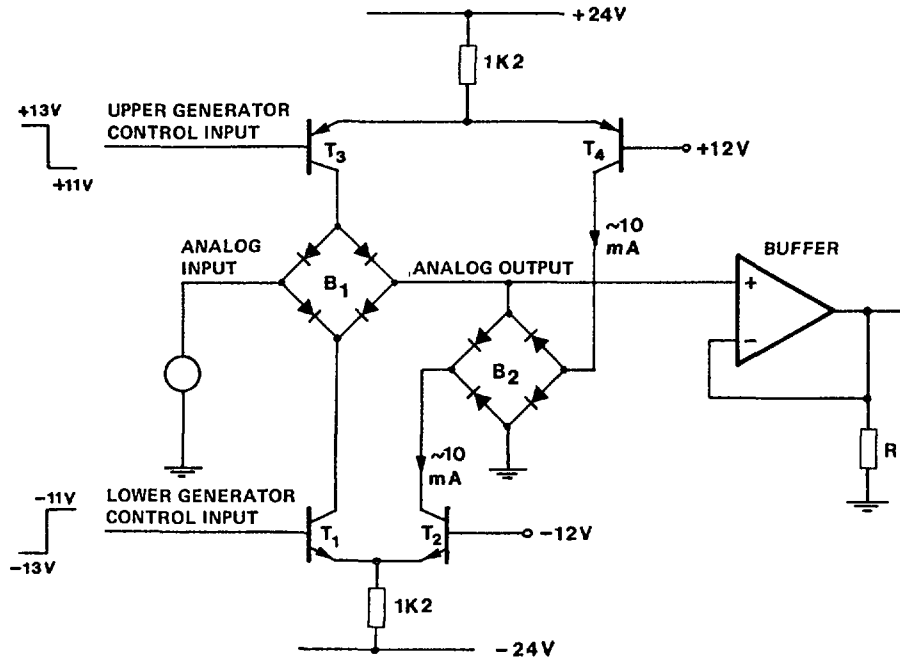
Fig. 4.13: Series-parallel diode bridge linear gate

The circuit shown in Fig. 4.13 employs two diode bridges, B_1 and B_2 and two control long-tailed pairs: (T_3, T_4) made of PNP transistors and (T_1, T_2) made of NPN transistors. The control inputs are at the base of T_1 and at that of T_3. If the voltage on the base of T_1 is -13V and the voltage on the base of T_3 is + 13V, T_1 and T_3 are OFF and T_2, T_4 are ON. In this condition no current flows into the bridge B_1, which is open. The collector currents of T_2, T_4, flow through the diodes of B_2, thus keeping B_2 closed. The current injected by T_4 and absorbed by T_2 is nearly 10 mA, therefore each diode in the bridge B_2 works at about 5 mA.

To reverse the condition of the two bridges it is necessary to switch simultaneously the base of T_3 to + 11V and that of T_1 to -11V. Doing so, T_2 and T_4 are turned off, thereby opening the bridge B_2, while T_1 and T_3 are turned ON, thereby closing the bridge B_1.

As shown in Fig. 4.13, an output buffer is provided to avoid that current be drained by the load R from the series bridge when the gate is OPEN. This improves the linearity and the stability in the gate transmission.

95

Provided that B_1 and B_2 are realized with adequately selected diodes, the linear gate just described, probably represents the best circuit for nuclear applications when the requirements of high linearity, good stability and ability to operate on short pulses are simultaneously present.

It has to be pointed out that the diode bridge gates of Fig. 4.12, and Fig. 4.13 work equally well on positive and on negative pulses.

4.4 ANALOG STORAGE

A widely employed method for analog storage is the peak stretcher.

The principle of the peak stretching operation can be described with reference to the simplified circuit of Fig. 4.14.
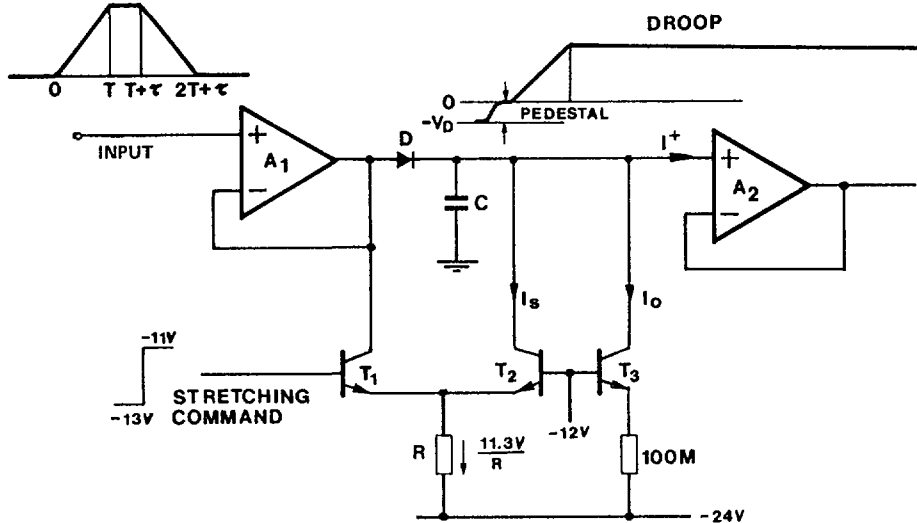
Fig. 4.14: Basic diagram of a peak stretcher.

Prior to the arrival of the input pulse, the base of T_1 is at -13V and accordingly T_1 is OFF and T_2 is ON. A relatively large current, some mA, flows through the diode D as fixed by T_2. The value of this current is chosen with the criterium that as long as T_2 is ON, the voltage across the capacitor C must be able to follow the input voltage when this either rises or falls. As long as T_2 is ON, no signal stretching takes place. With the current in T_2 switchable, it is possible to avoid stretching of unwanted signals and to control the stretching operation by sending a STRETCHING COMMAND to the base of T_1 prior to the arrival of the pulse.

Assume that the input signal, has a trapezoidal shape and its maximum amplitude is $V_{i,MAX}$. It is possible to allow the circuit to follow the input pulse on its trailing edge in the absence of STRETCHING COMMAND provided that

$$I_S > C \cdot \frac{V_{i,\ MAX}}{T}$$

this being the condition under which the diode D does not turn off on the trailing edge of the input signal. If, for instance, $V_{i,MAX}$ = 10V, T = 1 µs, C = 1000 pF, the current in T_2 must be at least 10 mA to enable the circuit to follow the input signal during its trailing edge.

96

During the leading edge of the positive input signal, the diode would always be forward biased and the voltage across C would follow the input signal even if the current in T_2 were zero. It can be concluded, therefore, that turning off T_2 makes the circuit incapable of following the trailing edge of the input signal, though it is still able to follow the leading edge of the positive signal. So, turning off T_2 disables the TRACKING OPERATION and enables the PEAK STRETCHING OPERATION. This transition is realized by the STRETCHING COMMAND which raises the base of T_1, thus turning OFF T_2 shortly before the signal to be stretched comes in. Once the current of T_2 is removed, a pedestal appears at the output because the current in the diode has changed from some mA to about 10 nA, this being the current determined by T_3. This current has the purpose of keeping D forward conducting during the time, usually a few hundreds of ns which elapses between the stretching command and the appearance of the input signal. The voltage across C jumps from about -0.7V, determined by the current in T_2 to about -0.4V corresponding to I_0.

When the input signal arrives, the diode increases its forward conduction during the leading edge of the pulse, thereby charging the capacitor C to the peak of the input pulse.

As soon as the flat top of the trapezoidal signal is over and the input signal starts falling, the diode gets reverse biased and the peak is stored. During the storage operation the voltage across C presents a "drop" which is determined by:

- the current I_0
- the bias current of the output buffer
- the reverse current of the diode
- the leakage current of T_2.

The output buffer must be realized with a JFET input operational amplifier.

If the stretcher is intended for long HOLDING TIMES it is possible to make I_0 switchable and to turn it off shortly after the peak of the input pulse. Besides, the diode D must be a low leakage specimen.

The stretching operation ends when the STRETCHING COMMAND stops. At that instant the base of T_1 drops again to -13V, the current in T_2 is restored and this discharges the storage capacitor by resetting the voltage across it to the value $-V_D$ determined by T_2's current flowing through the diode. The voltage across C_D recovers linearly from the stretched level to $-V_D$ at a rate given by $- I_s/C$, where I_s is the standing current across T_2.

The described simple configuration is rarely used in highly accurate instrumentation because it would lead to large charging errors on short signals of small amplitude.

This depends on the fact that the diode is a nonlinear device and therefore the change of a capacitor across a diode is a nonlinear process. Actually, the circuit of Fig. 4.14 becomes faster for signals of larger amplitude, slower for signals of lower amplitude.

To reduce the charging inaccuracy, it is necessary to employ a FEEDBACK STRETCHER, that is, to introduce diode and storage capacitor into a feedback loop of the type shown in Fig. 4.15, where the current sources I_s and I_0 are represented in a simplified way.
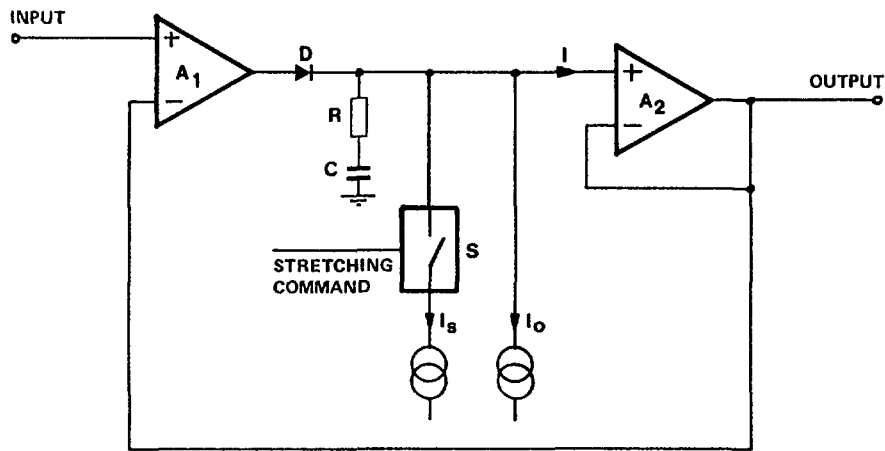
Fig. 4.15: Feedback stretcher

The loop including A_1, D, C, A_2 can be thought of as a unity gain buffer which follows the input signal and its RISING AND FALLING PORTIONS when I_S is present and which instead follows only the LEADING EDGE of the positive signals and HOLDS THE PEAKS when I_S is absent.

Besides reducing the charging errors, the feedback stretcher has the advantage, that the pedestal appearing at the output when I_S is turned OFF is strongly reduced. To understand this point, assume that no signal is present at the input, that is, the noninverting input of A_1 is exactly at 0V and S is closed. When S is turned off and the current in the diode drops from some mA to a few nA, the feedback loop maintains the inverting input of A_1 close to the noninverting input. The output is close to zero. The voltage variation across C appears at the output of A_1.

The resistor R in series with the storage capacitor C has the purpose of avoiding nonlinear overshoots in the loop that might lead to a wrong value in the stored voltage. This compensating resistor is usually in the 10 ÷ 100 ohm range and its value is determined experimentally, looking at both input and stretched signals and changing R until the output signal has the same amplitude of the input.

It has to be pointed out that, when R is present in the circuit position shown, the charging time is not determined by the risetime of the output pulse, but it is rather determined by the time constant RC. It is evident from the circuit that the stored voltage, that is, the voltage across C is obtained from the output voltage through an approximate integration of time constant RC.

A second type of analog memory, the so called track-and-hold amplifier, is used to store the instantaneous value of a continuous waveform. The function required, and a very basic circuit able to implement it are shown in Fig. 4.16.

In the tracking condition, $t < t_0$, the logic level of the variable V_L which controls the switch S is HIGH. S is accordingly closed and the voltage across C reproduces the input voltage V(t) as long as the time variations of V(t) are slow enough not to be smoothed by the finite gain - bandwidth product of the A amplifier. Switching from track-to-hold condition is achieved by setting V_L to logic LOW. The capacitor C becomes isolated from the amplifier, and it holds the voltage $V(t_0)$ existing on it at $t = t_0$. As long as S is open, C is insensitive to any variation which may occur on V(t).

98
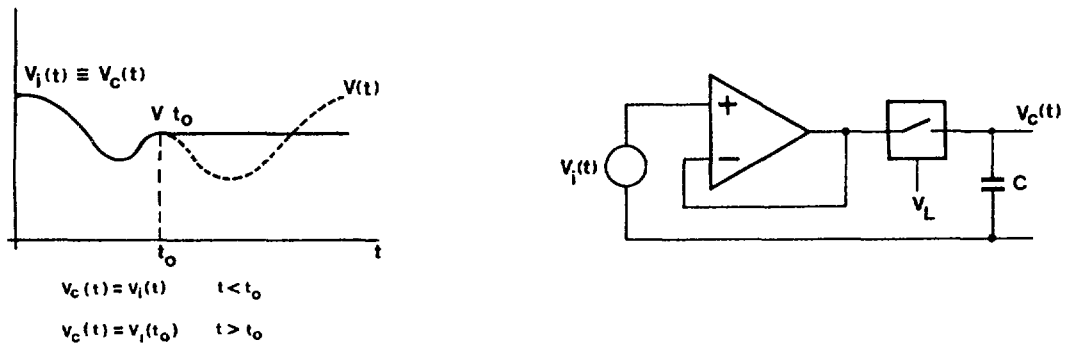
Fig. 4.16: Track and hold function and circuit which implements it.

The holding mode lasts until V_L is restored to logic HIGH with the results of closing S and forcing the storage capacitor to start tracking V(t) again. The tracking-to-holding-to tracking sequence is illustrated in Fig. 4.16.

In the actual operation, several undesired effects make the transitions from track condition to hold condition, and the inverse one, by far less neat than one may think by looking at the simplified diagrams of Fig. 4.16. Charge pumping effects through the stray capacitances of the switch S introduce an inaccuracy on the held voltage $V(t_0)$, while the finite slew rate and settling time of A slow down the process by which, at the end of the hold period, $V_C(t)$ starts tracking V(t) again.
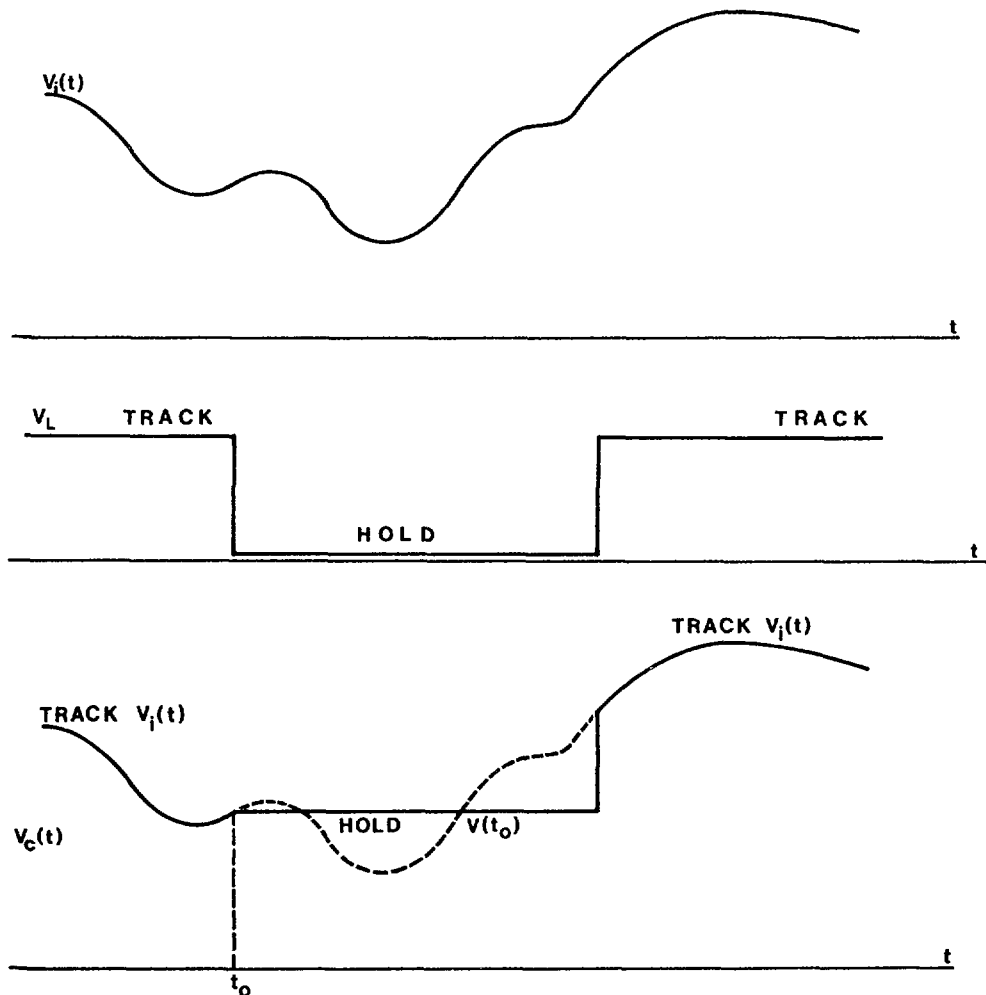
Fig. 4.17: Operational sequence in an ideal track-hold amplifier

A track-and-hold amplifier can be based upon a circuit diagram similar to the one shown in Fig. 4.15 for the feedback stretcher. The diode D has simply to be replaced with a switch, Fig. 4.18. The switch S in Fig. 4.18 is realised with a junction field-effect transistor and its state is controlled by the long tailed pair T_1, T_2. In the standing state the base of transistor T_1 is at about $-17.3V$ ($-24V + 6V + 0.7$) and the base of T_2 is one diode voltage drop more negative, $-18V$. T_1 which is ON, drives a current of about 6mA, as determined by the 1kΩ resistor connected between its emitter and ground, while T_2 is off.

No current, accordingly flows through the resistor R_1, which means that gate and source of the field effect transistors are at the same potential and it behaves, therefore, as a closed switch. In this condition the (C, R) impedance is connected to the output of A_1 through the small ON resistance of S. The feedback loop has the structure shown in Fig. 4.19.
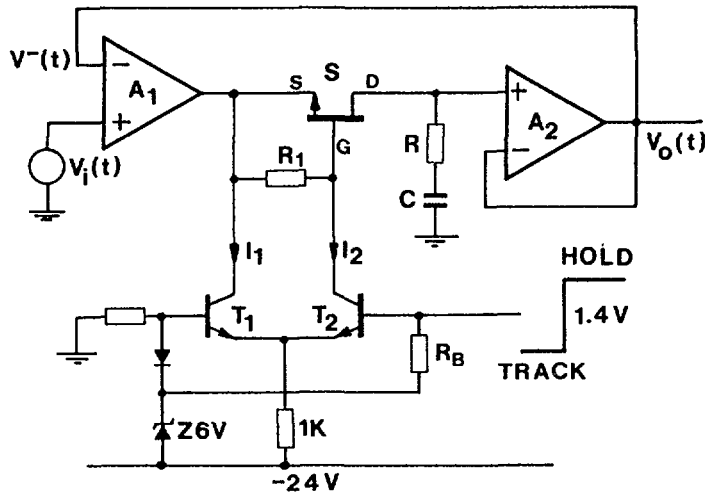
Fig. 4.18:   Track-and-Hold Amplifier
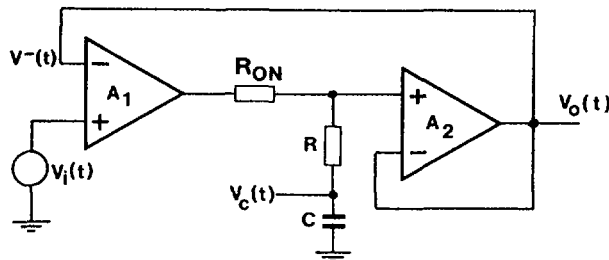
Fig. 4.19:   Feedback loop in the "track" condition

If τ_{ON} is sufficiently small and A_1 is large enough, the voltage $V^-(t)$ is identical to $V_i(t)$, while the voltage across C is related to $V^-(t)$ by the equation in the Laplace-transform domain:

$$V_c(s) = V^-(s) \cdot \frac{1}{1 + sRC}$$

In other words, the relationship between the input voltage and stored voltage has the nature of an approximate integration, with time constant RC. The resistance R is required, as it would be easily demonstrated by an elementary feedback theory approach, to avoid complex poles and therfore a ringing-response to the input step. It can be said,

therefore, that the time constant RC limits the speed at which fast variations in the input signal v(t) are tracked by C. The transition to the hold condition takes place as soon as the base of T_2 is made less negative of 1.4V, that is, as soon as it passes from -18V to -16.6V.

In this case T_1 goes off and the current determined by the 1K resistor flows entirely through T_2, thereby developing a voltage drop across R_1, which keeps the gate G of the field-effect transistor more negative than its source. The connection of Fig. 4.18 makes sure that in either situation, that is, (T_1 ON T_2 OFF) or (T_1OFF, T_2 ON), the GATE-to-SOURCE voltage difference is independent of V(t). Such a difference is 0 when T_1 is ON AND T_2 is OFF, and is $- R_1 I_2$ when T_1 is OFF and T_2 is ON. In the latter case R_1 has to be chosen so as to guarantee that the field effect transistor is OFF. The circuit in this situation is represented in Fig. 4.19, with the capacitor C floating and the voltage on it transferred to the output by the buffer A_2, Fig. 4.20.

Fig. 4.20: Holding condition
in the track-and-hold amplifier
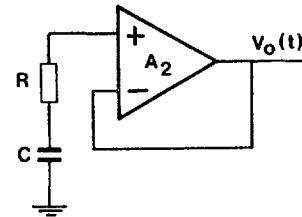
It has to be borne in mind that the buffer A_2 must be implemented with an operational amplifier featuring a low input bias current, to reduce the charge drained from C during the "hold" phase to the smallest possible amount. For this reason, the buffer A_2 is usually realized with an amplifier with a junction field-effect transistor at the input.

Some types of nuclear pulse spectrometers employ as an analog memory a circuit called gated integrator, which besides implementing the store function can also perform a shaping function. The basic diagram of a gated integrator is shown in Fig. 4.21.

Fig. 4.21: Gated integrator
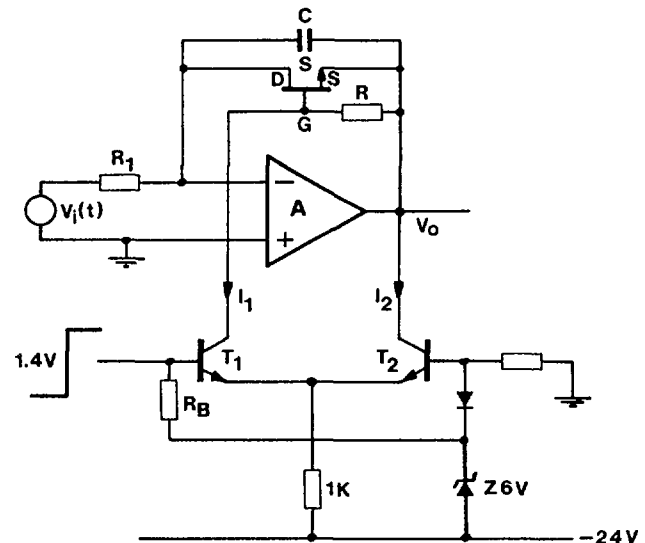
The resistor R_1, the capacitor C and the operational amplifier A implement a well known operational amplifier configuration. The switch S connected across C and controlled in the same way as the switch S of Fig. 4.20 adds to the integrator the gating facility. Again, in the standing state, T_1 is ON and T_2 is off, so that no current flows

across R and the switch S is closed. In this situation, any signal v(t) present at the input appears at the output strongly attenuated, for the switch S, connected as feedback impedance of the operational amplifier has a suitably small ON resistance. Integration of the incoming signal is enabled by a logic command which makes the base of T_2 less negative of at least 1.4V. In this situation T_2 becomes conducting, while T_1 is turned off. The current driven by T_2 flows through R and turns S off. The input signal is integrated on C and held on it as long as S remains open. In some applications, it may be necessary to make the holding time longer than the integration time. This can be achieved with the circuit of Fig. 4.22.
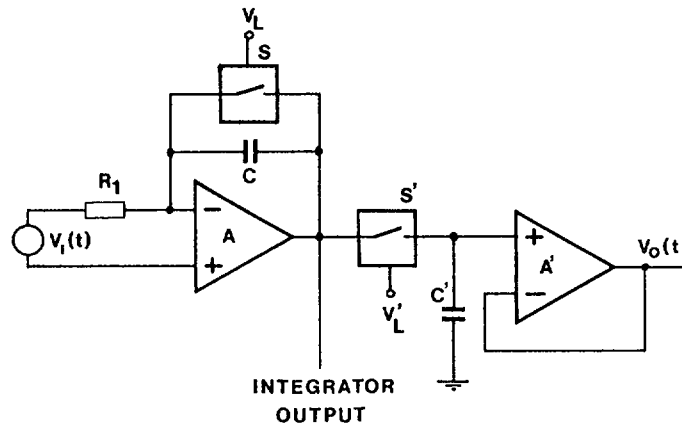
Fig. 4.22: Gated integrator with holding time longer than the integration time.

According to Fig. 4.22, the group S', C', and the buffer A' realise a track-and-hold amplifier which receives at the input the voltage present at the output of the integrator.

When S is turned off, so that the integration starts, S' is turned on. In this way, the voltage which appears across C is instantaneously transferred to C'. When the integration ends, S' is turned off, thereby isolating C' from the integrator. The final voltage of the integrator is held on C'.

After a short delay, which has the function of protecting the charge stored on C', S is turned on and quickly resets the charge on C, while the hold function is implemented by C'.

4.5 ORTEC 572 SPECTROSCOPY AMPLIFIER

4.5.1 Panel Specifications

As an example of complete spectrosocpy amplifier, ORTEC 572 will be described here. In order to introduce the functions it is able to implement, its front panel will be described at first. According to the panel settings, its gain is variable through six coarse setps and a fine adjustment through a helipot potentiometer between 10 and 1500. the spectroscopy amplifier accepts at the input slowly decaying exponeutial signals like those provided by a charge-sensitive preamplifier and provides at the output both unipolar signals: BNC connector UNI on the front panel or bipolar signals, BNC connector BI on the front panel. The unipolar signal has a quasi gaussiau shape with a time constant variable from 0.5 µs to 10 µs. The bipolar one is obtained from the nipolar signal

102
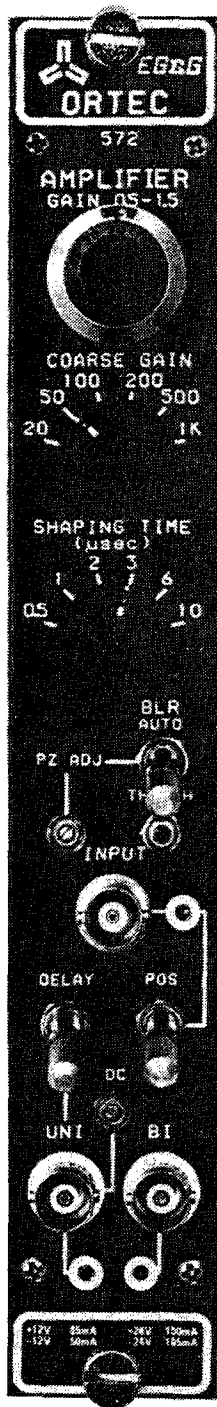
Fig. 4.23: Front panel
configuration
of 572 ORTEC
spectroscopy
amplifier
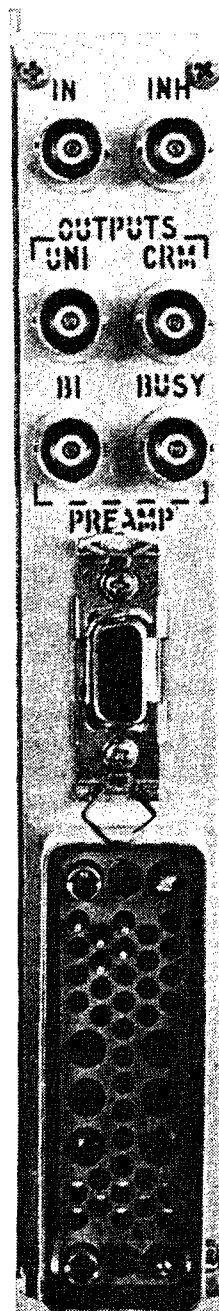
Fig. 4.24: Rear panel
configuration
of 572 ORTEC
spectroscopy
amplifier

through differentiation. The 572 spectroscopy amplifier has a crewdriver
adjustable P-Z cancellation. It has a built-in baseline restorer of a type
which will be explained later and which is called "time variant", with
adjustable threshold. The dc level at the unipolar output is also variable
through a front-panel potentiometer. The possibility of introducing a
μs delay on the unipolar output is foreseen.

As shown in Fig. 4.24, also on the rear panel, one BNC unipolar
output connector and one bipolar output connector are available. The

INPUT V_1

| CLIPPING BY P-Z CANCELLATION AND TIME-CONSTANT SELECTION S_1 | FIXED GAIN FIRST AMPLIFYING STAGE | SECOND AMPLIFYING STAGE | THIRD AMPLIFYING STAGE AND FINE GAIN CONTROL |

V_2

COARSE GAIN CONTROL

| 1 ST SECOND ORDER INTEGRATOR TIME-CONSTANT SELECTION COUPLED TO S_1 | 2 ND SECOND ORDER INTEGRATOR TIME-CONSTANT SELECTION COUPLED TO S_1 | DIFFERENTIATOR TIME-CONSTANT SELECTION COUPLED TO S_1 |

P

BIPOLAR OUTPUT

2µs DELAY

— — DELAY OUT — — —

UNIPOLAR OUTPUT

BLR

+E

FINAL LINEAR STAGE

-V   +V

C

GENERATION OF LOGIC SIGNALS FOR BLR CONTROL PILE-UP MONITORING AND BUSY TIME GENERATION

-E

Fig. 4.25: Block diagram of the 572 ORTEC amplifier

corresponding outputs on the front panel have low output impedance, those on the rear panel have a 93 Ω output impedance.

From the rear panel some logic commands are also available.

4.5.2 Block Diagram Description

The block diagram of 572 spectrscopy amplifier is shown in Fig. 4.25.

A slowly decaying expotential signal arriving from the preamplifier to the amplifier input is clipped by the pole-zero cancellation network, to a time constant which can be selected with the front panel "shaping time" to one of the following values: τ_s = 0.5, 1, 2, 4, 6, 10, according to the operation shown in Fig. 4.26.
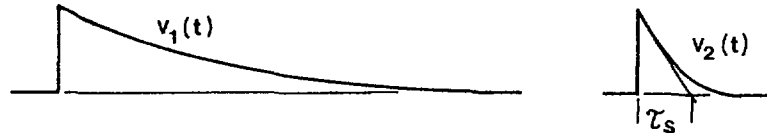
Fig. 4.26: Signal clipping at the input of 572 spectroscopy amplifier

The signal $V_2(t)$ is then amplified by the first, sixed gain stage which through the switch POS, NEG on the front panel can be set in either inverting or non-inverting configuration.

The second and third amplifying stages are wideband units implementing the variable gain function. The fine variation relies upon the third stage, while the coarse variation involves both. Following the signal path, the next two blocks realize the shaping integrations. These blocks are both second order integrators, based upon a circuit configuration which is frequently employed in the realization of semigaussian shapers. Due to their importance, a diversion will be made here to explain their behavior. Their circuit diagram is shown in Fig. 4.27.
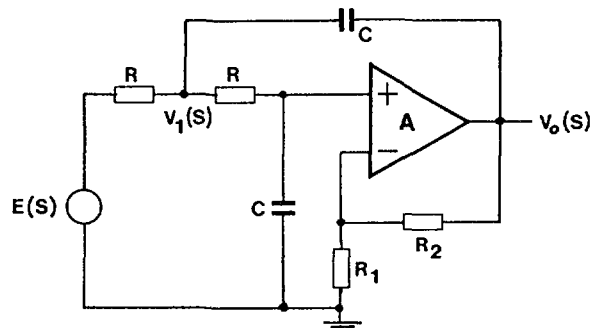
Fig. 4.27: Second order integrator

By applying the Laplace transform rules to the circuit of Fig. 4.27, the following relationships can be written.

$$\frac{E(S) - V_1(S)}{R} = SC\ [V_1(S) - V_o(S)] + V_1(S)\ \frac{1}{R + \frac{1}{SC}}$$

whence

$$E(S) = V_1(S)\ [1 + SCR - \frac{SCR}{1 + SCR}] - SCRV_o \qquad\qquad \text{(Eq. 4.1)}$$

If the operational amplifier is assumed to be ideal and have infinite gain, then

$$V_o(S) = \frac{R_1 + R_2}{R_1} \cdot V_1(S) \cdot \frac{1}{1 + SRC} \qquad \text{(Eq. 4.2)}$$

Introducing Eq. 4.2 into Eq. 4.1, the following result is obtained:

$$E(S) = \left\{ \frac{R_1}{R_1 + R_2} \cdot (1 + 3SCR + S^2 R^2 C^2) - SCR \right\} \cdot V_o(S)$$

$$E(S) = \left\{ 1 + \left(3 - \frac{R_1 + R_2}{R_1}\right) SCR + S^2 R^2 C^2 \right\} \frac{R_1}{R_1 + R_2} V_o(S)$$

The transfer function $\dfrac{V_o(S)}{E(S)}$ becomes accordingly:

$$\frac{V_o(S)}{E(S)} = \frac{\dfrac{R_1 + R_2}{R_1}}{1 + \left(3 - \dfrac{R_1 + R_2}{R_1}\right) SCR + S^2 R^2 C^2} \qquad \text{(Eq. 4.3)}$$

In the two integrators of the block diagram of Fig. 4.27, the value of $\dfrac{R_1 + R_2}{R_1}$ is kept low enough that the impulse response, is quickly damped, thereby simulating a Gaussian response with a shorter tail than the one achievable with integrators having real poles. The resistors R_1 R are front-panel selectable through the same switch S_1 which controls the time constants.

The output of the second integrator of Eq. 4.3 provides the basic unipolar signal, point P. From P the following further operations can take place. The unipolar signal is sent to a differentiator whose time constant is selected through S with the result that a bipolar signal becomes available at the relevant output.

Along the other way the unipolar signal goes through a path which may include a delay line (switch "delay" on the front panel) to the final output stage where baseline restoration occurs.

The baseline restoration employed in the 572 amplifier is a time-variant one, somewhat different in conception from the circuit employed in the laboratory spectroscopy amplifier, and therefore it is worth describing. The description will be referred to as Fig. 4.28. The loop of three amplifiers A_1, A_2, A_3 is designed in such a way that the dc output level of A_1 is fixed by the voltage at the inverting input of A_2 and can be adjusted through the potentiometer O. The amplifier A_3 is a transconductance amplifier, whose output can be represented as consisting of two current sources in opposition. A turn-off command leads these two generators in the open condition, by virtue of which the capacitor C remains isolated and holds the charge stored on it. To understand the behavior of the time-variant baseline restorer, suppose at first that no signal is present at the input I of A_1. In this case the feedback loop around A_1, A_2, A_3, J is closed and the voltage at A_1 output is equal to the voltage fixed by potentiometer O. A voltage is stored on C and this voltage is transmitted
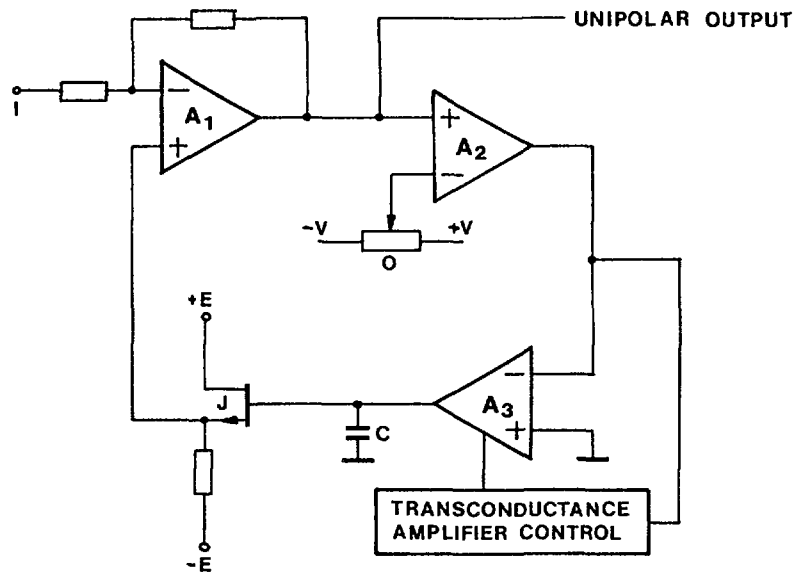
Fig. 4.28: Time-variant baseline restorer

through the field effect transistor J to the non-inverting input of A_1. In this linear situation the baseline restorer behaves as a differentiator, as it contains an integrator in return path of the feedback loop. So, low frequency noise of small amplitude does not appear at the output of A_1. When a signal is applied at the input I, such a signal through A_2 and the circuit which controls the transconductance amplifier turns off A_3, leaving C isolated. The voltage stored on C, which acts as an analog memory holds the output of A_1 stable as long as A_3 is off. As soon as the baseline at I recovers to its quiescent value, the feedback loop gets closed and the stabilization action starts again.

The time-variant baseline restorer can be thought of as a linear differentiator which fixes the output voltage of A_1 as long as no signal is present. When the signal comes in, the loop is interrupted, but the standing condition is held by virtue of the voltage stored on C. As during the signal presence A_3 is off and C is isolated, the signal cannot alter the dc standing condition of the loop.

Actually the control of A_3, which determines the behaviour of the baseline restorer is somewhat more involved than it may appear from the simplified description given here. The upper and lower baseline limits within which the restorer loop works in a linear fashion can be fixed either automatically or manually.

The ORTEC 572 amplifier has also an effective system to detect the pile-up of events. The logic sequence is summarized in Fig. 4.29. Suppose that two events come from the preamplifier too close to each other (Fig. 4.29.a), and make the hypothesis that they still overlap after the input clipping (Fig. 4.29.b). The overlapped signals, after gaussian shaping, have the shape shown in Fig. 4.29.c. The problem is then to recognize when pile-up occurs. To achieve this, the signal Fig. 4.29.b is strongly differentiated (Fig. 4.29.d) and then shaped by a trigger (Fig. 4.29.e). The rectangular signals of Fig. 4.29.e trigger an updating monostable multivibrator which generates an output pulse whose width is sufficient to cover the linear pulse (Fig. 4.29.f). The rectangular pulse (Fig. 4.29.f) is used as a busy output to show the total time duration during which the amplifier baseline is occupied by linear pulses and to enable the experimenter to carry on the dead time correction. The signal

107

a)

b)

c)

d)
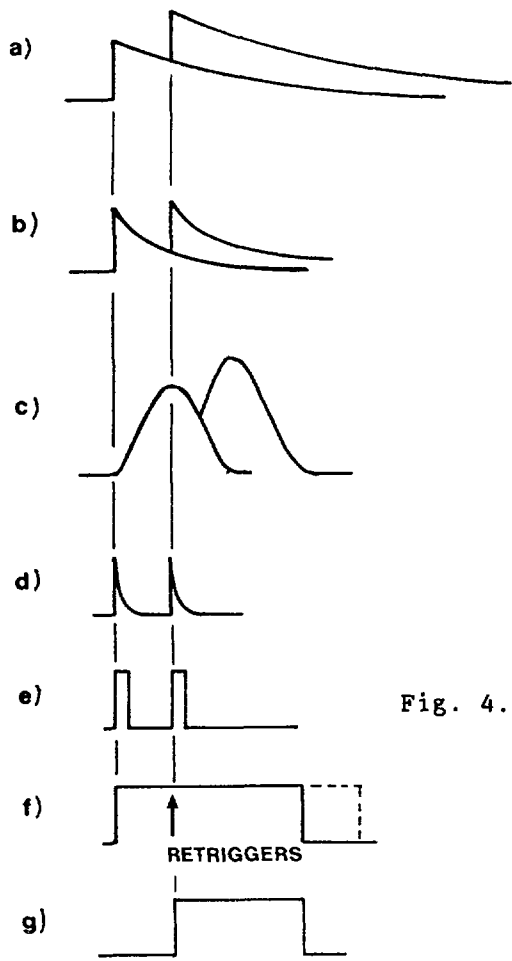
e)　　　　　　Fig. 4.29:　Logic sequence for pile-up detection

f)
RETRIGGERS

g)

(Fig. 4.29.g), which is generated only if the Figs. 4.29.e and 4.29.d come at a mutual distance shorter than the basewidth of the linear pulse, has the purpose of rejecting overlapped pulses.

CHAPTER 5

AMPLITUDE ANALYSIS

5   AMPLITUDE ANALYSIS

For most of the detectors, the output signals are proportional to the energy loss in the sensitive region of the detector.

After amplification and filtering as discussed in Sections 2 and 3, the amplitude of the analog signals is therefore an indication of the energy of the radiation particle. So the pulse height of these signals is the most interesting parameter in nuclear radiation measurement. As also discussed in previous sections, the pulses are of random occurence in amplitude as well as in time.

A typical sequence of pulses as seen on the output of an amplifier is shown in Fig. 5.1.
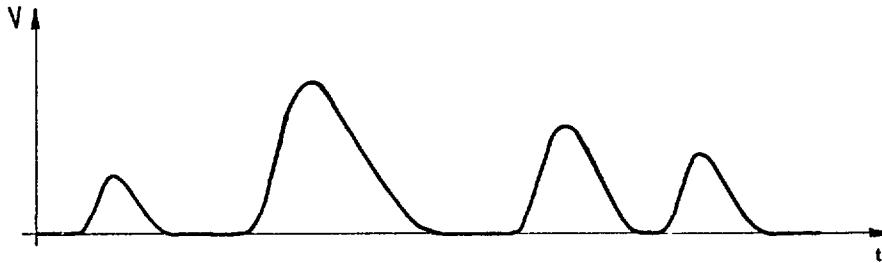
Fig. 5.1:   Typical pulses at the amplifier output

The amplitude V is a function of the energy E of the detected particle. We are interested in the distribution N(E) of all the pulses and call this an ENERGY SPECTRUM.

Typical energy spectra are shown in Fig. 5.2. Two different detectors were used, NaI scintillation detector and hyperpure Ge. Spectra are accumulated taking the analog pulses and sorting them according to their amplitude into energy channels. During the measurement all pulses of the same amplitude are summed up (accumulated) so that at the end of the measurement time the distribution of the energy can be shown, as in Fig. 5.2. Observe that the shape of the spectra depends on the detector used showing sharp (better resolved) peaks the solid state detector compared with the NaI scintillation detectors.

Fig. 5.2b shows a typical spectrum obtained with a Si(Li) detector used for x-rays.

In this chapter we discuss the various methods of measuring the pulse amplitudes.

5.1.   INTEGRAL DISCRIMINATOR

The easiest question we can ask is whether a pulse is larger or smaller than a given level. If we compare such a level E to the pulses of Fig. 5.1, we see, that some of them are below and others are above the level as shown in Fig. 5.3.

A circuit which delivers output pulses only for those input pulses which reach over a given level is called a pulse height DISCRIMINATOR. A Schmitt-trigger or an analog comparator can do the job. When we make the discriminator level adjustable, we can vary its value over the whole range of analog amplitudes. The output pulses are counted in a scaler for a given time. In Fig. 5.4 a block diagram of such a circuit is shown.
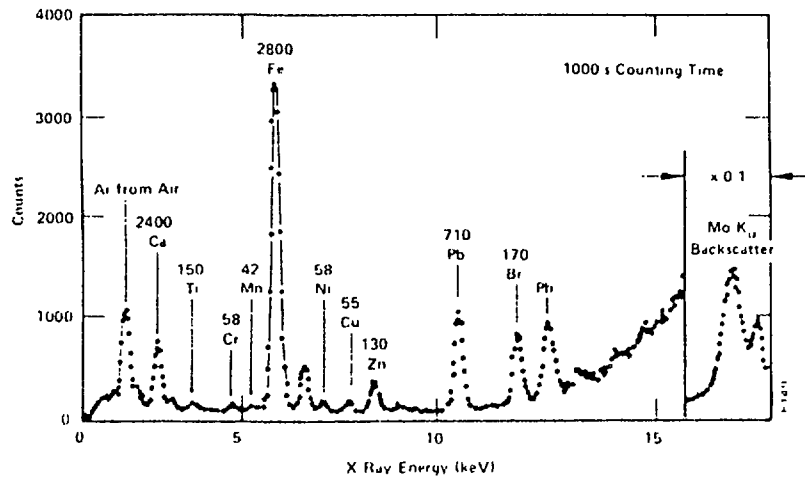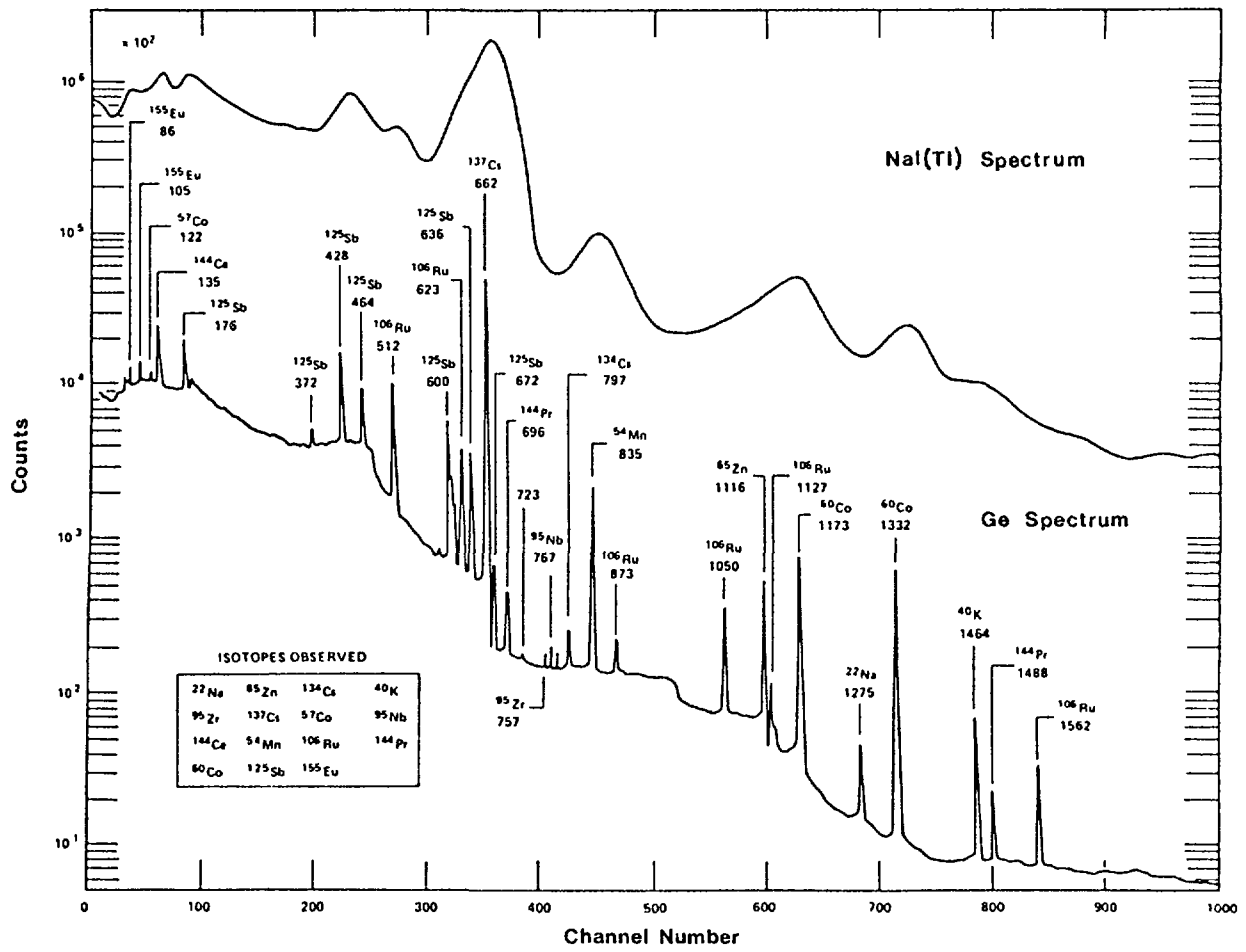
Fig. 5.2: Typical spectra, a) comparison of spectra measured by NaI and pure Ge detectors, b) spectrum obtained with a Si(Li) detector.
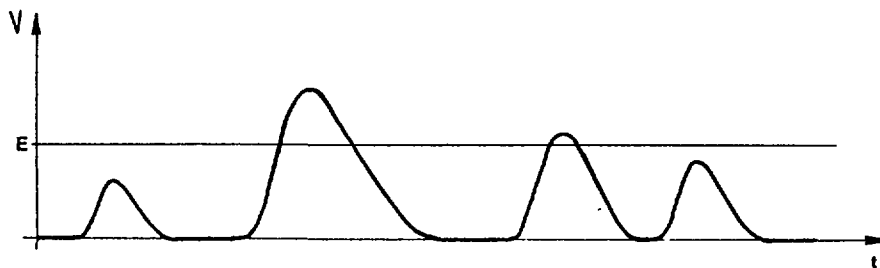
Fig. 5.3: Analog pulses with a discriminator level E

Fig. 5.4:   Block diagram
            of a discriminator
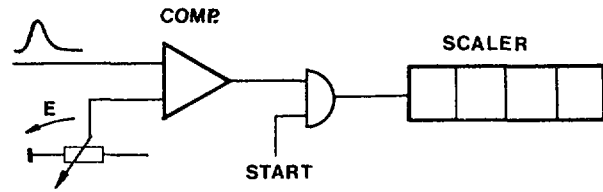
The measurement can be done by starting with the highest level and counting for a fixed time. Only few pulses, if any, will reach above the level. Now we step down with the level and count the output pulses for each level setting.

The number of the counted pulses will increase with every new setting i.e. we count the integral number of all the pulses above the level. That is why this circuit is called an INTEGRAL-DISCRIMINATOR.

When the result is plotted we get the integral or the spectrum. This means we can derivate the spectrum from the plot by differentation.

Fig. 5.5 gives an example.
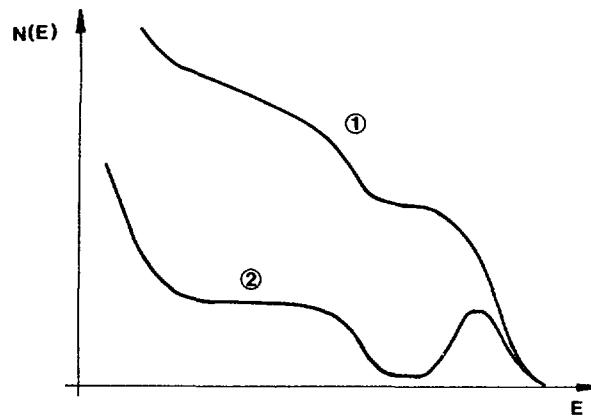
Fig. 5.5:   Integral result (1) and spectrum after differentiation (2)

The example makes clear that the evaluation of the energy spectrum with an integral discriminator is tedious. This kind of discriminators are therefore used, where two kind of pulses (e.g. real signals and noise pulses or gamma events and neutrons in a neutron counter) must be distinguished in amplitude.

5.2   SINGLE CHANNEL ANALYZER

From the previous chapter we see, that it would be more convenient to have a circuit with two adjustable levels and counting only these pulses which reach with their amplitude inbetween the two levels.
Introducing the second level leads to Fig. 5.6.
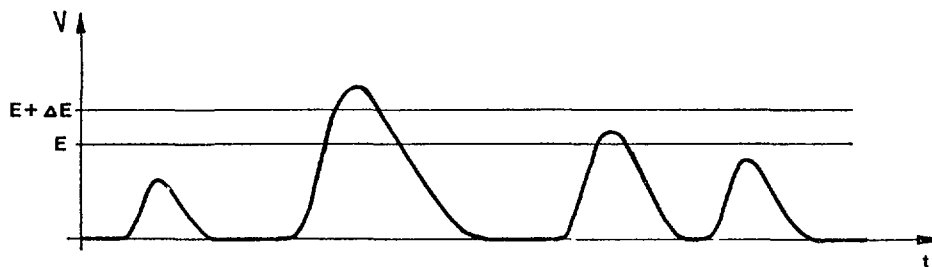
Fig. 5.6:   Pulse diagram with two discriminator levels

113

The two levels E and E +ΔE form a WINDOW of the width ΔE, which is called a CHANNEL.

To obtain the energy-spectrum, the channel is moved over the full range of pulse amplitudes. In each position all the pulses which fall in the channel in a given time are counted. The channel widthΔE is normally kept constant during the measurement.

An appropriate circuit is called a SINGLE CHANNEL ANALYZER (SCA) and its block diagram is shown in Fig. 5.7.
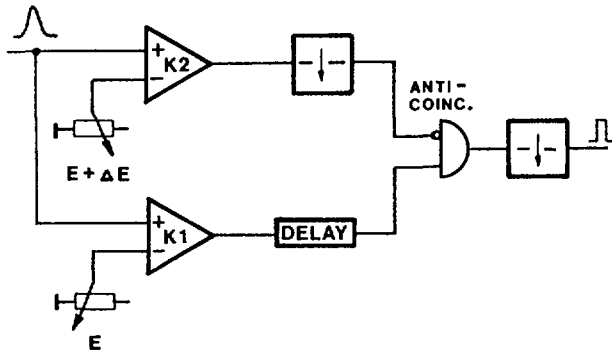
Fig. 5.7: Block diagram of a SCA

When the analog amplitude falls between E and E +ΔE (i.e. on the channel), only K1 is switched and the positive slope is fed to the input of the anti-coincidence gate through a delay.

When the analog pulse also switches K2 (i.e. the amplitude is higher than E + ΔE), the output swing of K2 prevents the pulse from K1 from going through the anti-coincidence gate.

The delay is inserted to ensure whether K2 was switched; the univibrator enlarges the pulse width from K2, so that it covers the K1-pulse at the anti-coincidence gate.

In the example of Fig. 5.7 the two levels are set independently. Obviously it would be advantageous to connect the upper level to the lower level in such a way, that the channel ΔE is kept constant.

This is done in the circuit-diagram of an SCA as shown in Fig. 5.8.

Transistor T1 forms a constant-current source. The current is adjusted with potentiometer P1 for 2,5 mA and flows through two ten-turn potentiometers of 2 kΩ each. With the lower one the baseline voltage E can be set, with the upper one the channel width ΔE can be set, which sits on top of E.

E is applied to comparator K1, E + ΔE is connected to K2. The output of K1 is connected to the clock-input of Flip-Flop 1 (FF1) which is triggered by a negative-going edge i.e. at the trailing edge of K1 output.

The output of FF1 is split in two branches: it is integrated at the base of T2 and fed back to the clear input (FF1 pin 13). So a short pulse at the Q output (FF1 pin 9) results. This short pulse is delayed by the two nand-gates (Z1a, Z1b) and fed to the anti-coincidence gate Z1c. As long as K2 is not triggered, the Q output of FF2 is high and a pulse occurs at the SCA-output.

114
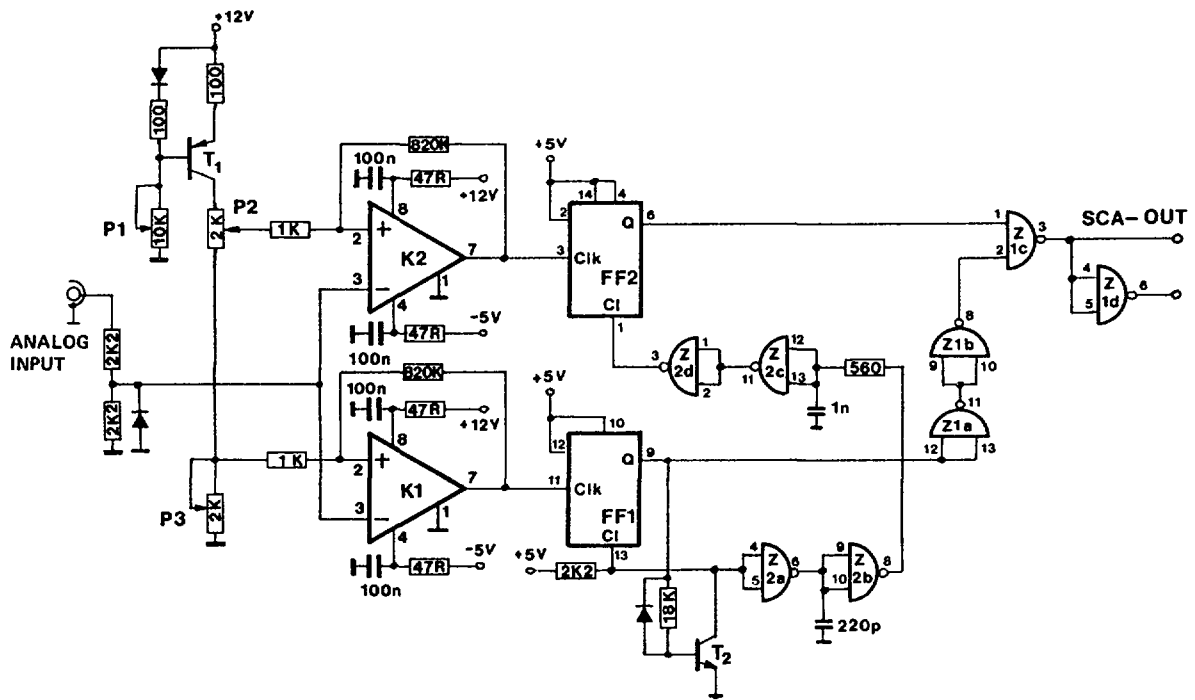
Fig. 5.8: Single channel analyzer

If K2 is triggered (i.e. the analog pulse is higher than
E +ΔE), flip-flop FF2 is set with the trailing edge of output from K2.
Therefore Q (FF2 pin 6) goes low and closes the anti-coincidence gate. FF2
is also reset by the output pulse of FF1 through a further delay (Z2c, Z2d),
so that it is ensured that a pulse from FF1 is fully overlapped by FF2, if
(E +ΔE) level is crossed.

An inverter at the output (Z1d) provides a positive output pulse.

In Fig. 5.9, a time diagram is given for three input pulses, where the
first is below E, the second in the channel and the third above the channel.
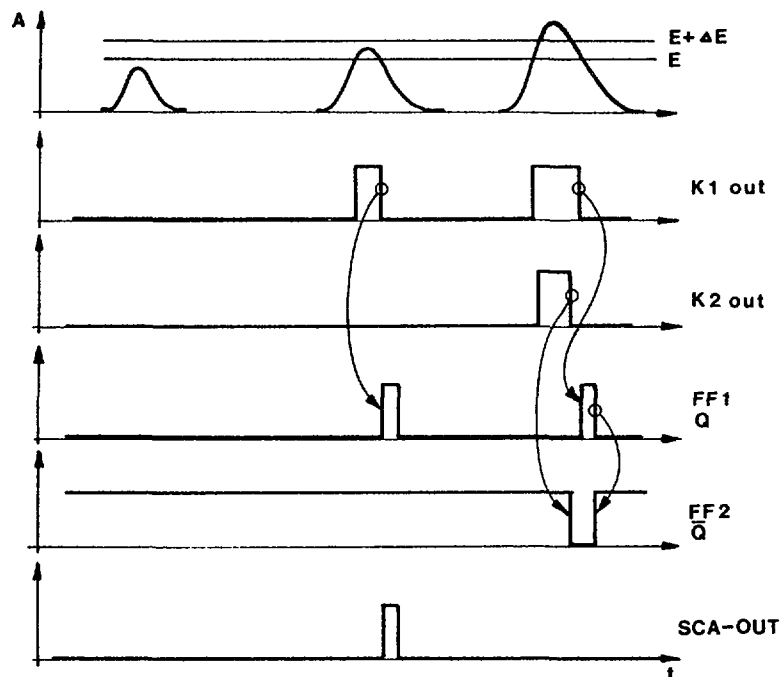
Fig. 5.9: Time diagram of the SCA

CHAPTER 6

ANALOG TO DIGITAL CONVERSION

6   ANALOG-TO-DIGITAL CONVERSION

6.1   INTRODUCTION

The analog-to-digital conversion is the link between the world of
analog electrical variables and the world of digital variables. Most
sensors and transducers of physical quantities deliver at their output
analog variables. The processing of the information carried by these
variables is implemented with digital techniques. It is easy, therefore,
to understand why the analog-to-digital conversion and the hardware which
does it, called analog-to-digital converter and labelled ADC in technical
literature, are becoming basic elements of a growing number of
instruments and systems. The increasing demand of ADCs, which has been
recorded over the past few years, has to be put in connection with the
more and more widespread use of digital processors, minicomputers in the
early seventies and later microprocessors and with the large diffusion of
digital displays in industrial and scientific instrumentation. For
instance, ADCs became basic modules of digital multimeters and digital
panel meters, of a large number of measuring instruments and of nearly
all data acquisition systems.

The discussion below will be restricted to the applications of
analog-to-digital converters in the nuclear field, where the ADCs made
their first appearance in the late forties, that is, much earlier than in
other domains of physical measurements. The reason for such an early
need of analog-to-digital conversion is explained by the remark that
nuclear physicists, with the aim of implementing systems able to record
and display probability density functions, among them the energy spectra
of ionising radiations, developed one of the first measuring instruments
based upon a partial digital processing of the information, the
multichannel analyzer (MCA). The problem therefore arose of making a
dialogue possible between a world of analog quantities (the charges
released by the radiation detectors) and the digital section of the MCA,
mainly the random access memory used to store the spectra.

6.2   ANALOG-TO DIGITAL CONVERSION FOR NUCLEAR APPLICATIONS

The analog-to-digital conversion can be performed either on the peak
amplitude of a pulse, like in the case of Fig. 6.1 or on the sample taken
at the instant t_o from a continuous waveform, as in Fig. 5.1. The case
presented in Fig. 6.1.a occurs whenever the pulses at the output of a
spectroscopy amplifier have to be converted. Their peak amplitude is
proportional to the charge released by the ionizing radiation in the
sensitive region of the radiation detector working in the "single event" mode.

Situation in Fig. 6.1.b refers to the pulse at the output of a
time-to-amplitude converter. The amplitude of interest is V_f, which is
proportional to the time-interval $t_{STOP} - t_{START}$.

The case of Fig. 6.2 arises whenever instantaneous samples taken from a
continuous waveform have to be converted. This kind of application, which is
not specific to the nuclear field, is common in association with radiation
detectors employed in the charge-integrating mode, for instance in nuclear
reactor monitoring and health physics instrumentation and in the measurements
of accelerator beam intensity.

While in the cases of Fig. 6.1 the analog-to-digital conversion is
intended for spectrometry applications, either energy spectrometry in the
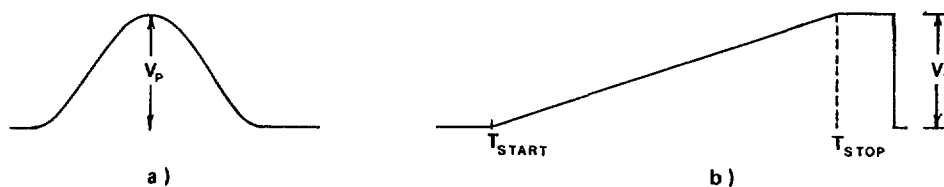
a)   b)

Fig. 6.1: Pulses to be converted in nuclear electronics applications.
a) Output pulse from a Gaussian spectroscopy amplifier. The
amplitude to be converted is V_p. b) Output pulse from a
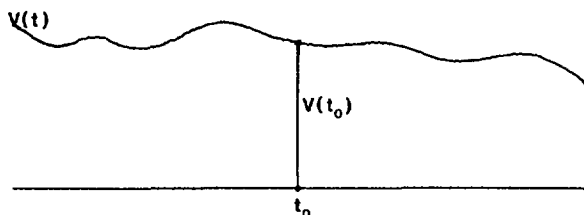time-to-amplitude converter. The amplitude to be converted
is V_f.

Fig. 6.2: Continuous waveform. The amplitude to be converted is the
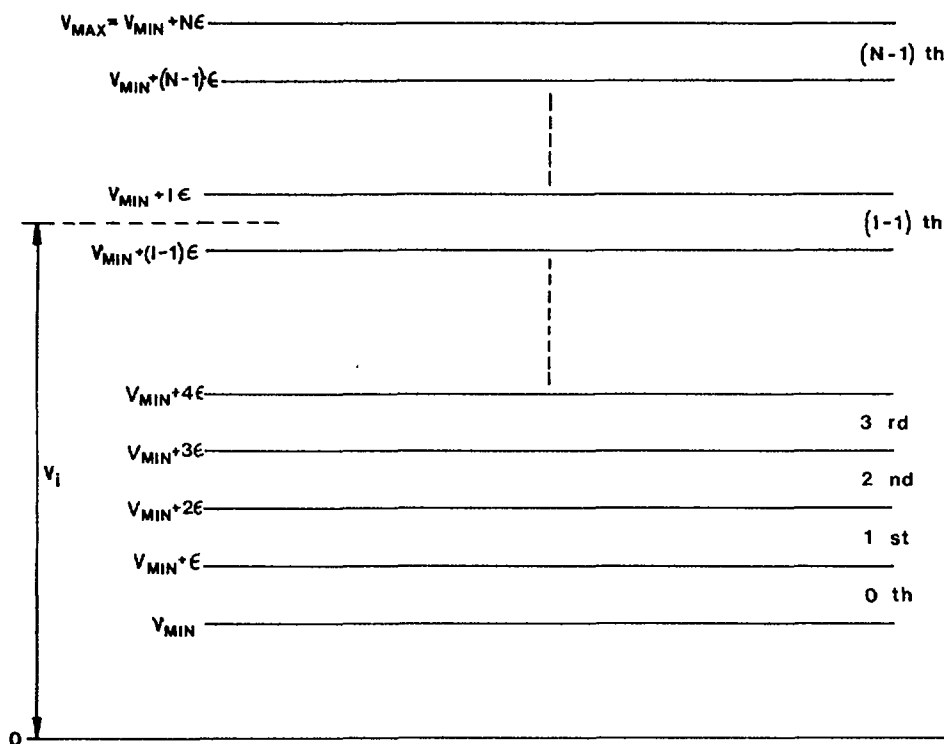sample $V(t_o)$, taken from $V(t)$ at $t = t_o$.

Fig. 6.3: Channel boundaries and sorting of V_i.

case a) or time-spectrometry in case b), the applications to which Fig. 6.2
refers are generally non-spectrometric. There is, however, an interesting
example in which a spectrometric application is connected with a continuous
waveform of the type shown in Fig. 6.2. This is the case in which V(t) is
the noise waveform at the output of a linear spectroscopy channel, for
instance, and the probability density function of the noise samples has to be
determined in order to investigate the properties of the noise.

The analog-to-digital conversion will be introduced in the following
way. Let V_i be the amplitude to be converted, either V_p or V_f or
$V(t_o)$. Let V_{MIN}, V_{MAX} be the lower and upper limits of the ADC input

120

range. The analog-to-digital conversion consists in subdividing the input range V_{MIN}, V_{MAX} into N equal parts, called channels, and in sorting the input amplitude V_i, provided that it falls within the ($V_{IN} \div V_{MAX}$) range into the corresponding channel. It will be assumed, as a starting hypothesis that the operation is ideal, that is, all the channels have equal width,

$$\in = \frac{V_{MAX} - V_{MIN}}{N}$$

The number of channels N is usually an integer power of 2, $N = 2^K$, where K is the number of bits. For the sake of simplicity, the input range $V_{MAX} - V_{MIN}$ will be made equal, to 10V. So an ADC with 10 bits, has $2^{10} = 1024$ channels and the channel width is slightly less than 10 mV. Similarly, an ADC with 12 bits has $2^{12} = 4096$ channels and the channel width $\in$ is about 2.5 mV. At present, the largest number of channels in ADCs for nuclear spectroscopy is $2^{14} = 16.394$, but such ADCs are somewhat rare.

The operation of the ADC can now be figured out by looking at the diagram in Fig. 6.3, which shows the input range (V_{MIN}, V_{MAX}) split into channels of equal width by equally spaced boundaries V_{MIN}, $V_{MIN} + \in$, ...,

The channels are numbered from the lower edge to the upper edge of the input range. The channel defined by the boundaries V_{MIN} and $V_{MIN} + \in$ is conventionally labelled channel of order 0, the one defined by the boundaries $V_{MIN} + \in$ and $V_{MIN} + 2\in$ is the first channel, the one defined by $V_{MIN} + 2\in$ and $V_{MIN} + 3\in$ the second one, the channel defined by $V_{MIN} + (N-1)\in$ and $V_{MIN} + N\in$ is the last one.

Analog-to-digital conversion consists of assigning the input amplitude V_i to the channel within the boundaries of which V_i lies. It is equivalent to say, in the example of Fig. 6.3, that V_i falls into the (ℓ-1)th channel or that V_i is put in correspondence with the number ℓ-1, which is the channel order. Figuring out the input amplitude which falls into a certain channel is more usual among designers and users of ADCs for nuclear applications, while designers and users of ADCs for general purposes speak more commonly about correspondence between input amplitude and output number. There is nothing substantial in the difference between the two ways of describing the same operation.

As to the electrical symbol of the ADC, the one of Fig. 6.4 is widely employed, with one analog line on the input side, where V_i is applied and K output lines that give in binary parallel form the number of the channel into which V_i is sorted.

Fig. 6.4: Electrical symbol for an ADC.

The K output lines can assume two logic states, 0 and 1 and therefore they can represent, in parallel binary form, all the integers from 0 to 2^K-1, that is 2^K-1 + 1 = N numbers, as many as the channels.

6.3 CHARACTERIZATION OF AMPLITUDE-TO-DIGITAL CONVERTERS

Two input amplitudes $V_{i,1}$ and $V_{i,2}$ that differ by at least one channel width will be assigned TO TWO DIFFERENT CHANNELS in the ADC, or, in other words, they will be RESOLVED. The channel width, therefore can be assumed to specify the resolving capability of the ADC. The resolving capability is sometimes stated in relative terms, referred to the full input

range $V_{MAX}-V_{MIN}$. It will be said, for instance, that an ADC with N channels is able to resolve two amplitudes as long as they differ by at least 1/Nth of the full input range.

The need of ADCs with high resolution capabilites came along with the introduction of high resolution detectors, like Ge detectors for γ-rays, spanning a broad energy range and capable of separating closely spaced spectral peaks. So, the requirement of ADCs with 8000 channels is rather usual in γ-radiation spectrometry with solid state detectors, while the so call γX-detectors, able to cope with photon energies ranging from some KeVs up to a few MeVs may need ADCs with 16,000 channels.

ADCs for spectrometric applications, especially if they are intended to operate with detectors featuring a broad linear range, should have an adequately high degree of INTEGRAL LINEARITY.

A simple method of checking the integral linearity of an ADC is outlined in Fig. 6.5.
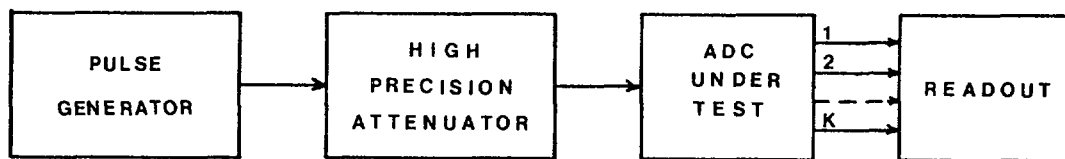
Fig. 6.5: Integral linearity test on an ADC.

The set-up is based upon a pulse generator delivering a 2μs wide rectangular signal with an amplitude exceeding the upper boundary V_{MAX} of the range of the ADC under test, and having a high precision attenuator. At the beginning of the test the attenuator is set in the transmission-1 position (0dB attenuation), and the amplitude of the generator pulse is adjusted by acting on the fine amplitude knob until the last ADC channel is read out: all the K output lines declare a logic $\underline{1}$ at the end of each conversion. This is equivalent to state that at that particular setting of the generator fine adjustment, with no change in amplitude due to the attenuator, the generator amplitude V_{GS} is sorted into the last channel of the ADC. That is, V_{GS} corresponds to the channel N--1.

The measurement then goes on by keeping the generator setting fixed and by introducing at each step, a gradually increasing attenuation.

At every step, the attenuator setting and the relevant digital readout at the ADC output must be recorded. A plot is then drawn, which presents on the horizontal axis the attenuator settings and on the vertical axis the numbers readout at the ADC output. For an ideal ADC, such a plot should be a straight line. For a real ADC, instead, deviations from linearity will be observed (Fig. 6.6).

In Fig. 6.6 a reference straight line is shown connecting the (0dB, N-1) point with the point at the lower boundary of the ADC range $(-\alpha_K dB, K)$. Note that an attenuator setting able to bring the attenuated signal exactly into the 0-th channel of the ADC may not even be available. The measurement will be concluded by setting a value of the attenuation such that the corresponding channel be of sufficiently low order, less than 10, for instance in a 1024 channel ADC.

Once the plot of Fig. 6.6 is obtained, the maximum deviation σ_{MAX} (number of channels) between actual and ideally linear behaviour can be
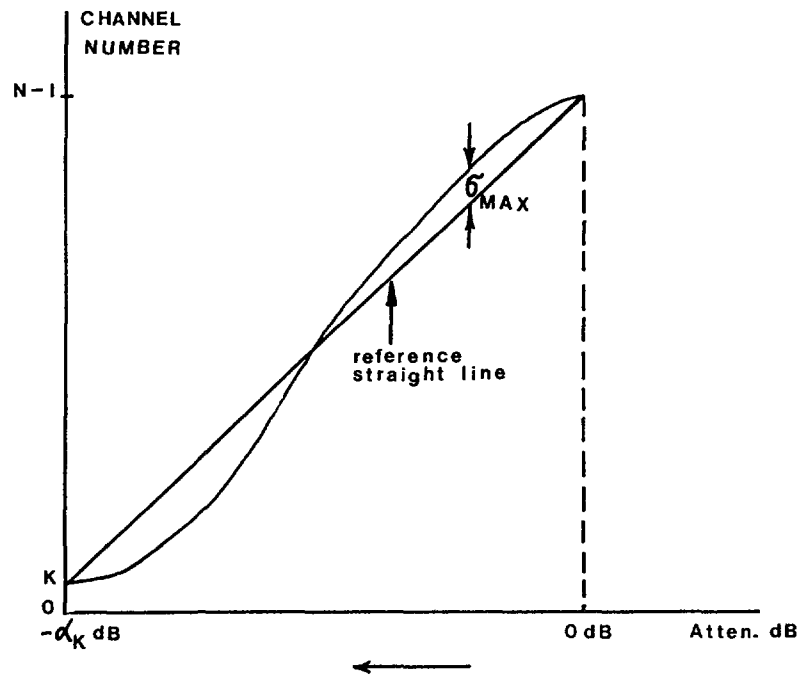
122

Fig. 6.6: Plot of channel number versus attenuation, showing deviations from linearity.

defined. The integral nonlinearity is given as $(G_{MAX}/N) \cdot 100$; it is usually expressed in percent of the total channel number.

Remember that the integral nonlinearity is just a number. An ADC featuring 0.1% integral nonlinearity means that a 4096 channel ADC has a real behaviour which deviates from the reference straight line by four channels.

A small integral nonlinearity is a feature of great concern to designers and users of ADCs in a number of application fields, not only in the nuclear one. The integral nonlinearity impairs the accuracy with which an electrical variable is measured. For instance, the problem of achieving a highly linear behaviour is common to measuring instrumentation of whatever nature, from the accurate multimeter, to the low current meter, from the MCA to the instruments for quantitative chemical analysis.

A peculiar requirement for ADCs intended for nuclear spectrometry applications is the channel width uniformity. It was assumed so far in the discussion that the N channels all had equal width $\in$. This may be untrue in real ADCs. The effect of a nonuniform channel width in spectrometry applications can be understood by comparing the behaviour of two ADCs when both operate on a distribution of pulse amplitudes which has, as a feature, equal number of pulses per equal amplitude intervals. Such a distribution is often referred to as a white amplitude distribution.

Of the two ADCs being compared, one has, like in the ideal case, channels of equal width, while the other one has large channel width irregularities.

The comparison between the two ADCs is carried out in Fig. 6.7. The ADC with equal channel width does not alter the feature of the input amplitude distribution, that is, equal numbers of pulses channels are contained wihin all amplitude channels. According to Fig. 6.7, in all the channels of the first ADC, equal number of pulses are counted, 6. For the ADC with large irregularities in channel width the pattern of the number of pulses counted
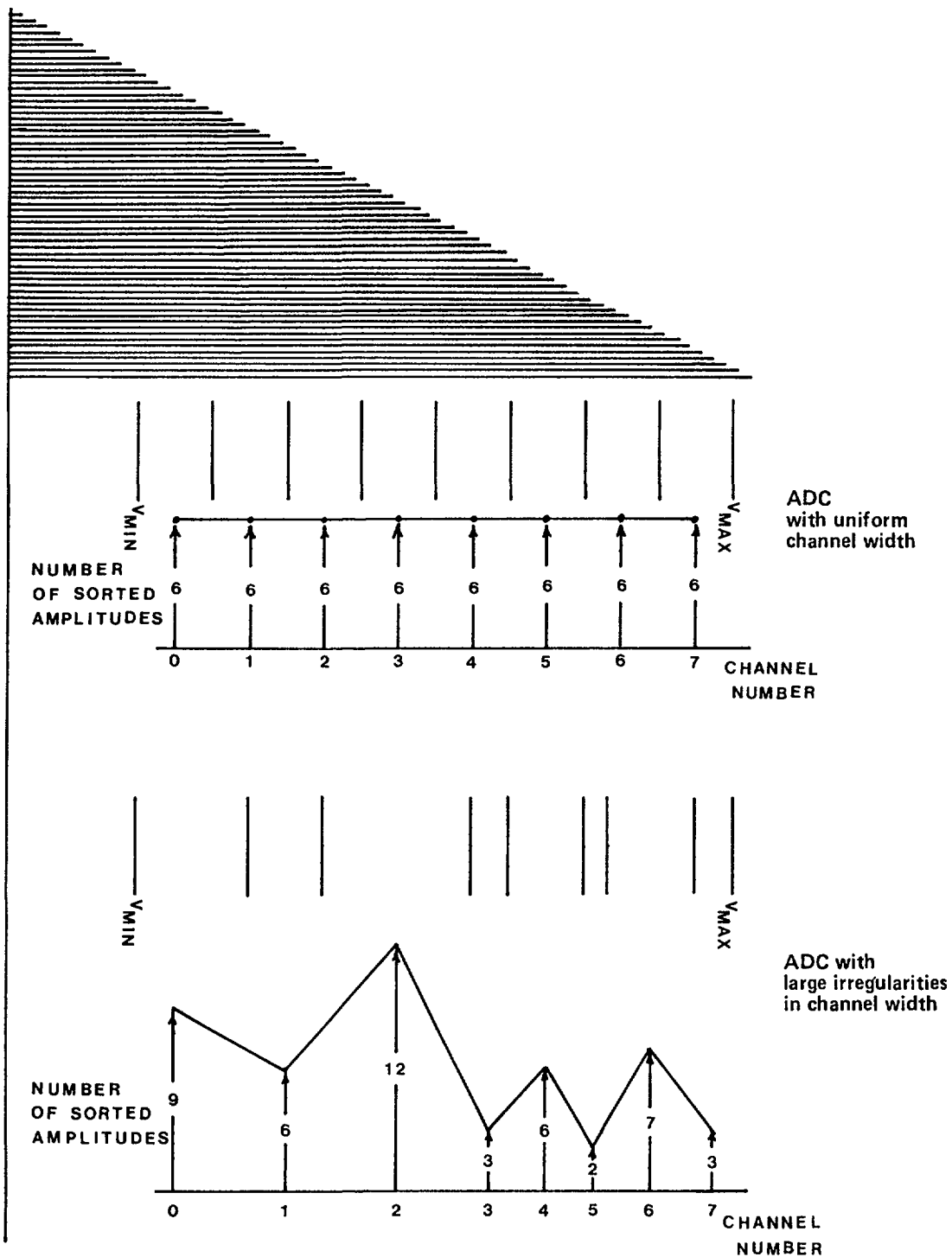
Fig. 6.7: Comparison of two ADCs having uniform and nonuniform channel
widths.

channel by channel is far from being constant, for it presents peaks and
valleys depending on the local behaviour of channel width. As a result, the
spectrum of the input amplitude distribution is badly distorted.

Strictly speaking, whenever variations in channel width occur, also
nonlinearity of an integral nature has to be present. In some cases,
however, irregularities in channel width of the type shown in Fig. 6.7,
though large, may be restricted to a very small number of channels, so that
the integral nonlinearity of the ADC is affected to a negligible extent,
while the portion of spectrum stored in those channels gets completely

124

distorted. A typical example of very annoying effect related with channel
width irregularities is the so-called even-odd effect, according to which the
odd channels are systematically wider or narrower than the even ones. The
situation is shown in Fig. 6.8.

V_{MIN}
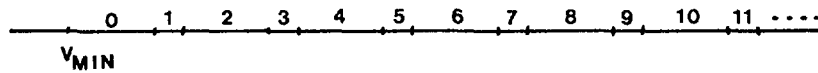
Fig. 6.8: Even-odd effect.

The even channels 0, 2, 4, 6, 8, 10 are systematically three times wider
than the odd channels, 1, 3, 5, 7, 9, 11. As the average of an even channel
and of the next odd one gives the correct channel width, the integral
linearity of the ADC is practically unaffected. However, the white amplitude
distribution of Fig. 6.6 would be split into two distributions, one
corresponding to the even channels, with 9 pulses sorted into every channel
and one corresponding to the odd channels with 3 pulses sorted into every
channel (see Fig. 6.9).
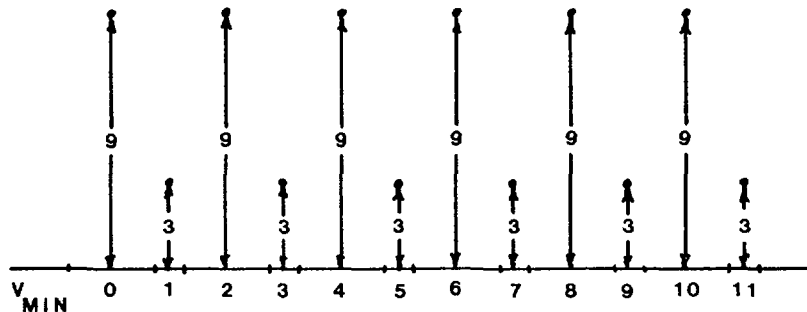
Fig. 6.9: Periodic irregularities in the channel content arising from
even-off effect.

The even-odd effect is frequently observed in multichannel pulse height
analyses with a defective ADC, when they are used in the analysis of a gamma
ray spectrum. The Compton edge of the gamma radiation appears to be split in
two distributions, like in Fig. 6.10.

Fig. 6.10: ^{60}Co spectrum distorted by an ADC with even-odd effect.

A test of the channel width pattern is a highly qualifying one for ADCs
intended for nuclear applications. For this kind of test, the ADC has to be
connected to a multichannel pulse analyser, which provides the RAM where the
content of the various channels of the ADC can be accumulated, displayed and
printed. Alternatively, the ADC can be associated with a microprocessor
which can do the same job as the MCA. The test can be performed using as
input signal source the so-called "sweeping pulser" which generates the
"white" amplitude distribution of Fig. 6.7. A sweeping pulser is made of a
long term integrator which creates a slowly rising and decaying triangular
waveform and a linear gate which samples the triangular voltage at a fixed
rate and for a relatively short time duration. Sweeping pulsers are
commercially available. If it does not exist among the testing facilities of

the laboratory where the test on channel width uniformity has to be performed, it can be easily implemented by employing existing instrumentation. Two possible realisations of sweeping pulser are shown in Fig. 6.11.
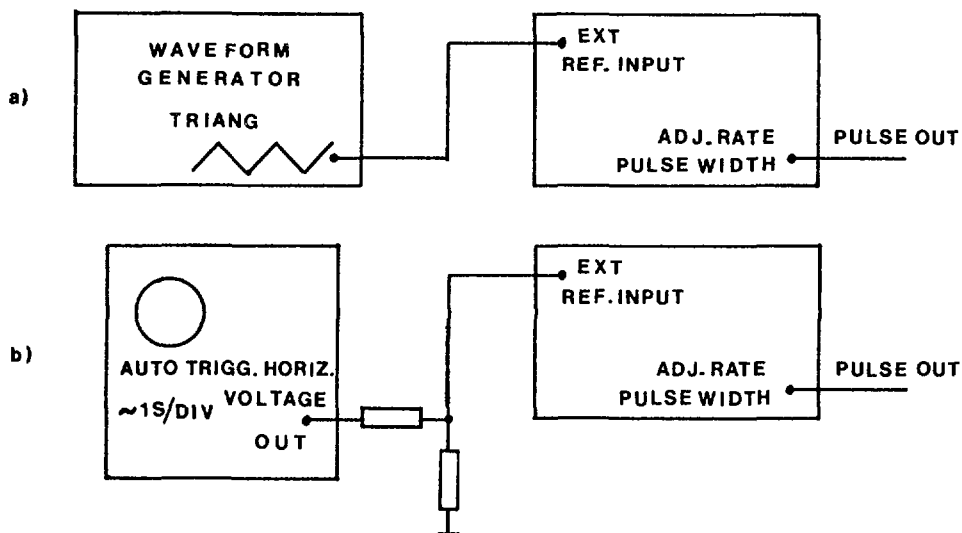
Fig. 6.11:   Sweeping pulser implemented with existing instruments.
a)   Generator of triangular waveform and pulse generator
b)   Oscilloscope and pulse generator

The first set-up makes use of a waveform generator operating in the triangular output mode. The triangular waveform is applied to the external reference input of a pulse generator which rate and width can be adjusted.

The second solution employs as linearly rising slow waveform the sawtooth generated in an oscilloscope to realise the horizontal deflection. The sawtooth, after attenuation, is applied again to the external reference input of the pulse generator.

For the channel width tests, the triangular waveform or the sawtooth should be adjusted to a slope of 1V/s or less. The sampling rate of the pulse generator should be the highest possible compatibly with the ADC conversion time. The pulse width should be about $2\mu s$.

The linearly changing amplitude distribution should be accumulated in the MCA memory for an integer number of triangles or sawtooth periods. Let υ_K be the number of pulses stored in the K-th channel. Let:

$$\bar{\upsilon} = \frac{\sum_\upsilon \upsilon_K}{N}$$

be the average number of pulses stored in the channels. The channel width error in the K-th channel is defined as $(\upsilon_k - \upsilon)/\upsilon$.

It is useful to point out that the sweeping pulses allows also a straightforward integral linearity test. For this purpose, the content of the various channels at the end of the sweeping pulser operation has to be processed in the following way.

The content υ_0 of the 0-th channel is stored in channel 0. The sum $\upsilon_0 + \upsilon_1$ of the contents of channels 0.1 is stored in channel 1. The sum $\upsilon_0 + \upsilon_1 + \upsilon_2$ of the contents of channels 0, 1, 2 is stored in channel 2 and so on.

126

A plot is accordingly drawn, which has on the horizontal axis the channel orders 0, 1, 2 ... and on the vertical axis the sums v_0, $v_0 +$ v_1, $v_0 + v_1 + v_2$ ...., as in Fig. 6.12.
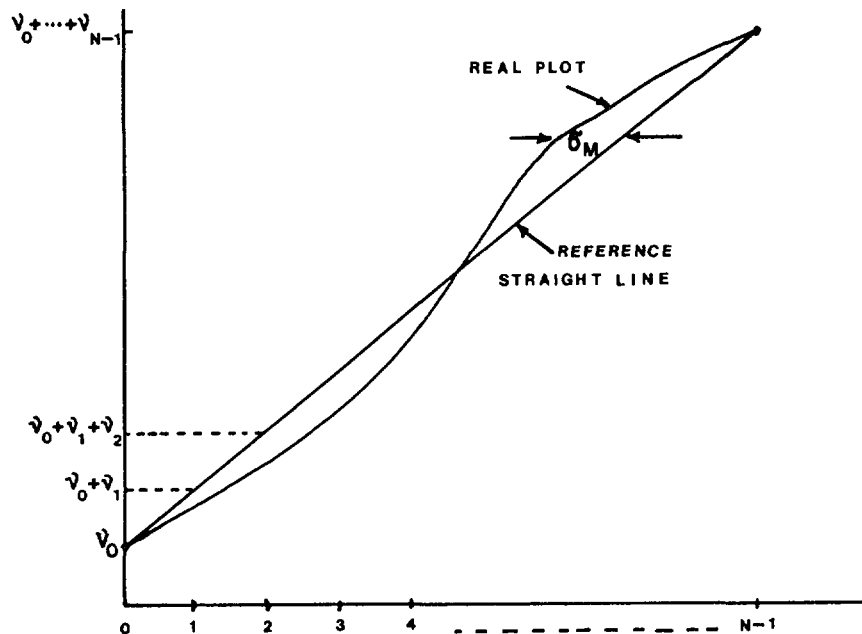
Fig. 6.12: Integral nonlinearity measured with the sweeping pulser

The deviations of the real plot (h, $v_0 + v_h + v_2 ... +$ v_K) from the reference straight line connecting the points (0, v_0) and (N-1, $v_0 + v_1 + v_2 + ... + v_{N-1}$) provide a way of evaluating the integral nonlinearity.

A further information of interest in ADCs conceived for nuclear applications is related to the so-called "channel profile". The concept of channel profile emerges quite spontaneously in answering to the the following questions: what is the probability that the input amplitude V_i be stored in the corresponding channel, say K? The question can be answered, to begin with, in the case of ideal channels of equal width $\in$. The lower and upper boundaries of channels K-1, K, K+ 1 are shown in Fig. 6.13.
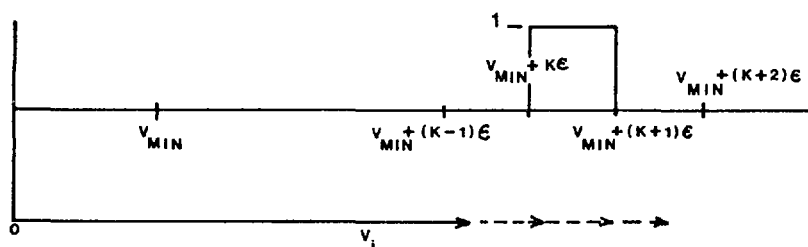
Fig. 6.13: Channel profile in the ideal case of equal channels.

If the input amplitude V_i is continuously increased, the signal is stored with probability 1 in the K-th channel, as soon as V_i exceeds the lower boundary of the K-th channel and keeps being stored in that channel with probability 1 until V_i reaches the upper channel boundary $V_{MIN} + (K+1) \in$.

The probability that the input amplitude V_i be stored in the K-th channel is called PROFILE OF THE CHANNEL K. In the present case, all the channel profiles are equal rectangles of unity height.

127

In the real cases, the channel profiles are not rectangular. Even if in a well designed ADC the channel profiles are equal, some degree of overlapping always exists, as shown in Fig. 6.14.
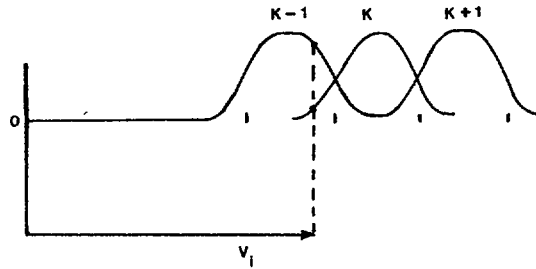
Fig. 6.14: Overlapping, nonrectangular channel profiles.

This means that the input amplitude V_i shown in Fig. 6.14 can be assigned with finite probabilities to channel K and channel K-1 as well. With the value V_i assumed in Fig. 6.14, assignment to (K-1)th channel is actually more likely than to K-th channel.

The channel profile of an ADC can be investigated with the arrangement shown in Fig. 6.15. Assume the following example, where the 1000-th channel of a 2.5mV channel width ADC has to be tested. Let V_{MIN} be 100 mV. According to these data, the ideal value of the lower boundary of the 1000-th channel would be:

$$V_{MIN} + 1000 \, \in \, = 2600 \text{ mV}$$

and the upper boundary of the same channel would be

$$2600 \text{ mV} + 2.5\text{mV} = 2602.5 \text{ mV}.$$

In order to explore the channel profile of the 1000-th channel, an adequately broad amplitude range, including at least two channel widths (5 mV) below the lower boundary and at least two channels (5 more mV) above the upper boundary of the 1000-th channel should be covered.
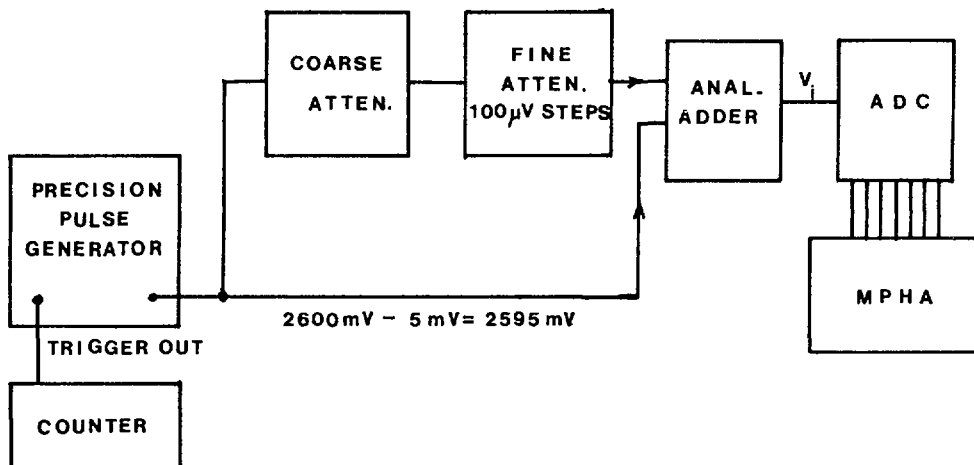
Fig. 6.15: Channel profile measurement.

As shown in Fig. 6.15, the set-up for the channel profile investigation includes a precision pulse generator, whose output amplitude is adjusted to be about 5 mV below the lower boundary of the channel to be tested, a system of two attenuators to generate pulses with amplitude variable in steps of 100 μV each from 0 to 12.5 mV and an analog adder. The signal at the adder

128

output is variable between 2595 mV and 2607.5 mV in steps of 100 μV, which means that the range of interest, 5 mV below the lower boundary of the 1000-th channel to 5 mV above the upper boundary of the same channel can be explored with an amplitude resolution of 100μV.

The different amplitudes in this range, 100 μV apart from each other are set and the generator is kept running at that attenuator setting for a given time, for instance 1 minute. The ratio between the number of pulses stored in the 1000-th channel at every amplitude setting in the selected range and the number of pulses delivered by the generator and counted in an external counter is plotted as a function of the ADC input amplitude. The resulting curve is the desired channel profile (see Fig. 6.16).
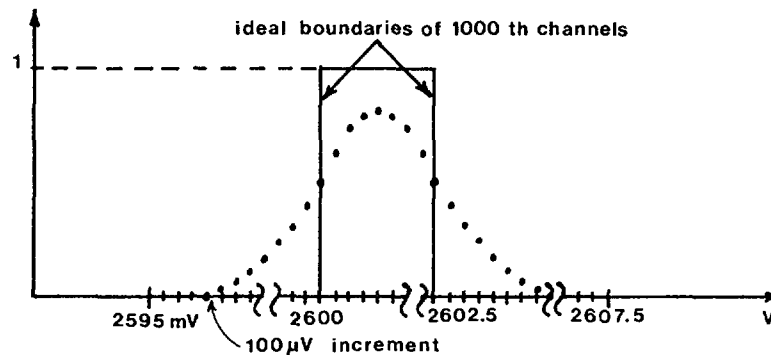
Fig. 6.16: Channel profile obtained with the set-up of Fig. 6.15.

In the example of Fig. 6.16, the channel profile is so broad that it extends well outside the nominal channel width of 2.5 mV. The channel overlapping is so large that nowhere the input amplitude has 100% probability of being stored into the 1000-th channel.

Obviously the quality of an ADC is expressed by the extent to which the channel profiles of all the channels approach the ideal rectangular shape with a width close to the nominal value ∈ .

It is true that measurement of the channel profile is not a simple neither a short one. However, in several cases, it is sufficient to perform it on a limited number of channels, less than 10, say, chosen at different locations in the ADC range to detect irregularities that may strongly affect the behaviour of the ADC.

6.4 DIGITAL TO ANALOG CONVERSION

The digital-to-analog conversion makes a dialogue possible between the world of digitally represented variables, as they appear at the output of digital processors, like mini- and microcomputers and the world of analog variables. In general industrial applications, it is employed to interface the output of digital processors to elements, like the actuators, that require analog control at their input. The hardware which implements the digital-to-analog conversion, the digital-to-analog converter, or DAC, has in the more common version, in which the input variable is in parallel binary form, the symbol shown in Fig. 6.17.

The input digital word consists of K bits, and each bit is applied to the relevant input line. The B variables can either be at logic level HIGH, in which case, they are assumed to represent logic 1 or al level LOW In which case they represent the logic 0.
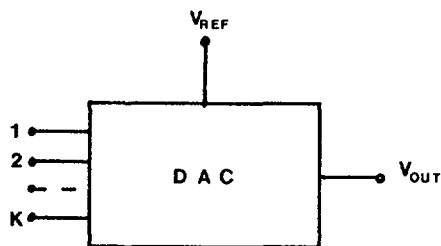
Fig. 6.17: Digital-to-analog converter.

The DAC establishes a correspondence between the input digital word and the output analog voltage V_{OUT}: such a correspondence is expressed by the relationship:

$$V_{OUT} = V_{REF} \left\{ \frac{B_1}{2^1} + \frac{B_2}{2^2} + \frac{B_3}{2^3} + \ldots + \frac{B_K}{B^K} \right\}$$

in which the B's can either be 1 or 0. So, for instance, if
$B_1 = B_2 = B_3 = \ldots = B_K = 0$, then $V_{OUT} = 0$.
If $B_1 = 1$, $B_2 = B_3 = \ldots = B_K = 0$, then $V_{OUT} = V_{REF}/2$.
Here B_1 is the MOST SIGNIFICANT BIT.
If all the B's from B_1 to B_{K-1} are put equal to zero and $B_K = 1$, then B_K, LEAST SIGNIFICAN BIT, produces at the output the voltage
$V_{OUT} = V_{REF}/2^K$.
When all B's are set at 1, then the DAC gives the largest analog value

$$V_{OUT} = V_{REF} \left[\frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \ldots + \frac{1}{2^K} \right] = V_{REF} \left[1 - \frac{1}{2^K} \right]$$

The input-output characteristic of a DAC can be dynamically explored by using the setup of Fig. 6.18.
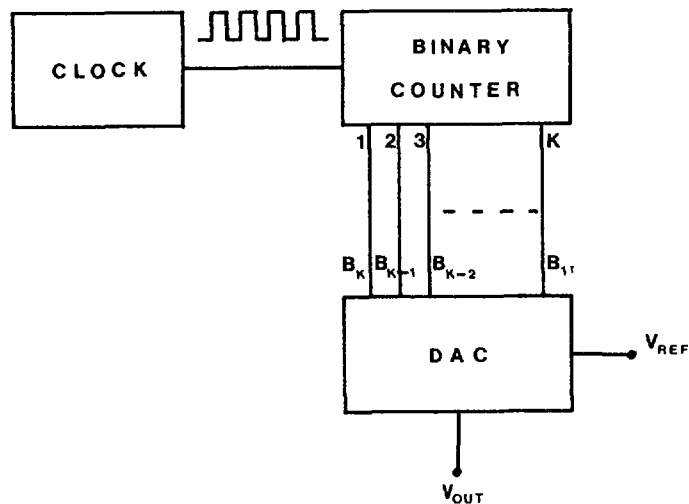
Fig. 6.18: Setup to measure the DAC input-output characteristic

In the connection of Fig. 6.18, the digital word applied at the DAC input is incremented by one bit at every clock pulse received by the binary counter. If the binary counter was reset before the first clock pulse V_{OUT} starts from zero, is incremented by an elementary step equal to $V_{REF}/2^K$ whenever a clock pulse is counted, reaches the maximum value
$V_{REF} (1 - 1/2^K)$ when all the digital outputs of the counter are at 1 and then is suddenly restored to zero on the next clock pulse (Fig. 6.19).

130
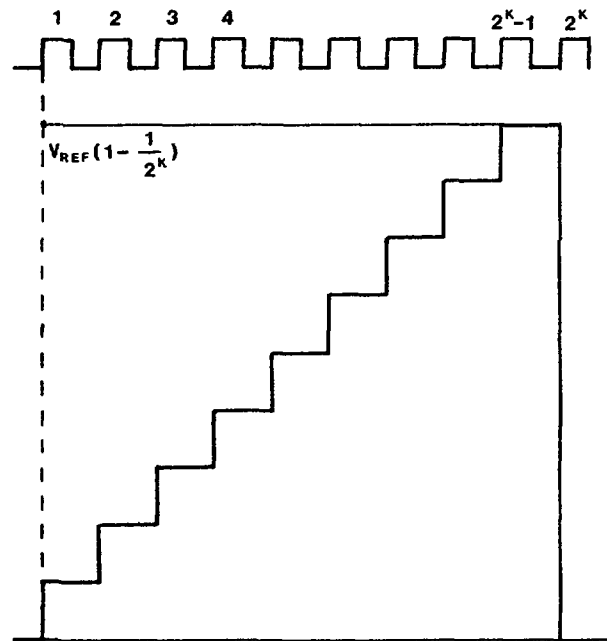
Fig. 6.19:   Input-output characteristic of a DAC

According to the diagram of Fig. 6.19, the range of the output analog variable, $0 \div V_{REF}$ $(1 - 1/2^K)$ is subdivided into $2^K - 1$ elementary steps, each equal in amplitude to the LEAST SIGNIFICANT BIT. The least significant bit defines the resolution of the DAC. Among the commercial DAC's, units in which the staircase of figure is resolved into up to 256,000 steps, that is, units featuring a resolution of 1 part over 256,000 are available. Assuming a value of V_{REF} of 10V, this means that the addition of one decimal unit to the input digital word changes the output analog voltage by about 40 μV.

In most of the commercially available DACs the output variable is a current delivered by an adequately good current source rather than a voltage. The voltage-to-current transformation is usually obtained by injecting this current into the virtual ground of a transresistance-type operational configuration, Fig. 6.20.
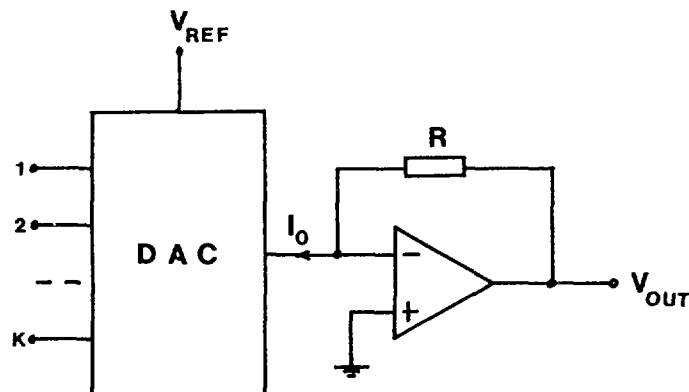
Fig. 6.20:   Current-to-voltage transformation at the DAC output.

In choosing the operational amplifier A, attention must be concentrated on its dynamic characteristics, specifically slewing rate, gain-bandwidth product, settling-time. If not carefully chosen, the output transresistance amplifier may become a potential source of degradation of the intrinsic speed characteristics of the DAC.

131

6.5 ADC PRINCIPLES

ADCs for nuclear applications are complete instruments that accept at the input signal of assigned shapes, either that of Fig. 6.1.a or that of Fig. 6.1.b or the one of Fig. 6.2 and provide at the output a digital information. This means that they have built-in analog processing circuits able to perform on the incoming pulses or waveforms the required pre-conditioning. The most important part of such a pre-conditioning is the analog storage of the amplitude to be converted. Analog storage circuits have already been described in Section 4.4. They are of the peak-sensing type for pulses like those of Fig. 6.1 of the sample & hold type for continuous waveforms of the nature shown in Fig. 6.2. In either case, it can be assumed that the voltage to be converted into a number stored on a capacitor and this capacitor is disconnected from the signal source and connected to the input of the digitizing section of the ADC. The digitizing section becomes therefore the distinguishing feature of the ADC. The three more popular types of digitizing principles employed in nuclear physics will be described here.

6.5.1   Flash-type or multicomparator digitizers

The block diagram of a flash-type or multicomparator digitizer is shown in Fig. 6.21.

The signal stored on C is transmitted through a buffer made of a JFET input operational amplifier to keep the charge draining from the capacitor as small as possible, to a set of threshold comparators. The thresholds start with V_{MIN} for the lowest comparator, which sets the lower boundary of the ADC range and then increase by voltage steps equal to the channel width ϵ from one comparator to the next one up to the value $V_{MIN} + N\epsilon = V_{MAX}$, which fixes the upper boundary of the ADC. According to the diagram of Fig. 6.21, N+1 comparators are required to define N channels. The top comparator has an overrange function, that is, whenever it triggers, it states that the input amplitude has exceeded the upper limit of the linear range.
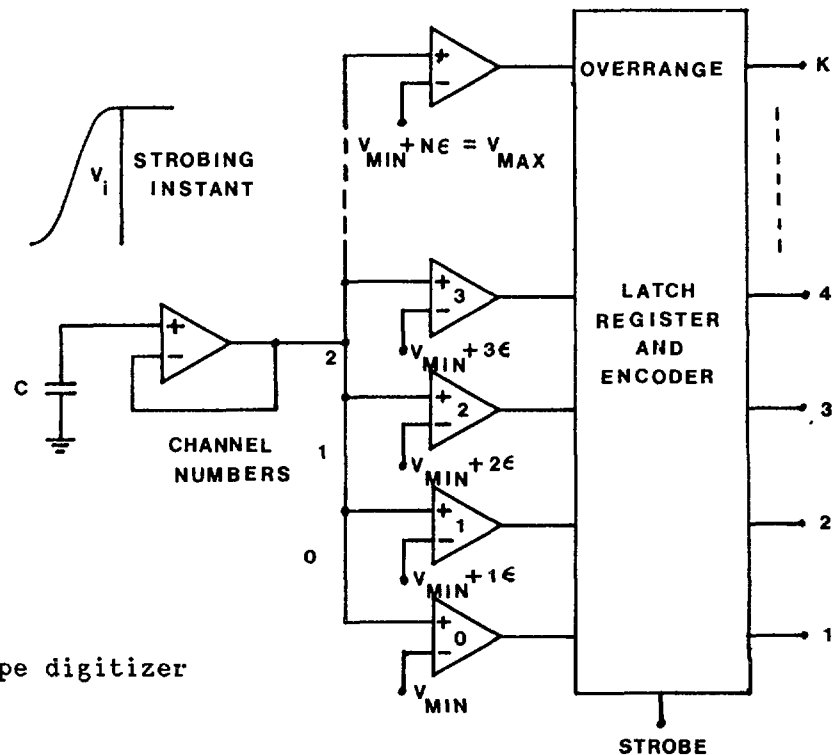
Fig. 6.21: Flash-type digitizer

The digitizing operation proceeds as follows. The stretched pulse is presented simultaneously to the inputs of all the comparators. Once the stretched pulse has settled to its final value, so that no more variations are expected, the configuration of the comparators that have changed state under the influence of the input signal applied to their noninverting terminals has to be considered stable. Such a configuration is frozen by sending a STROBE COMMAND to the LATCH REGISTER AND ENCODER MODULE. The output state, either LOW or HIGH of every comparator is transferred into the LATCH REGISTER.

To clarify the operation with an example, suppose that the amplitude V_i at the input of the digitizing system of comparators settles to a value higher than $V_{MIN} + 2 \in$ and lower than $V_{MIN} + 3 \in$. The comparators will be in the following configuration: outputs of comparators 0,1, 2 at logic level HIGH, while the outputs of all the comparators from 3 on will be at logic level LOW. Upon receipt of the STROBING COMMAND, this configuration of the comparator outputs is transferred into the latches. The assignment of the incomig amplitude to the relevant amplitude channel is done on the base of the last comparator changing state. So, in the example before, the comparator having changed state are 0, 1, 2, that is, the input amplitude is sorted in channel 2.

The decimal representation based upon the status of the latches is then converted into a parallel binary one by the encoder and passed on to the K output lines, that can represent any number between 0 and 2^K-1.

The flash-type digitizer is probably the oldest one in the history of nuclear electronics; in the old days it was realised with a set of vacuum tube Schmitt triggers. It was subsequently abandoned for simpler and more precise digitizing circuits and this can be understood on account of the fact that the flash-type digitizer for a total number of channels N requires N + 1 comparators. So, the principle does not lend itself to the realisation of high resolution ADCs.

Forgotten about for several years, it was rescued in recent times thanks to the improvement in monolithic technologies that made the integration of large numbers of comparators on a single chip possible. Besides, the quality of monolithic comparators has remarkably improved, leading to smaller and smaller offset voltages and shorter and shorter switching times.

Thanks to this technological improvement, the flash-type digitizer is now being looked at with interest by users in several fields, including nuclear physics. Owing to the growing demand, several companies are now producing flash-type ADCs of very good quality and with a comparatively large number of channels (512 and more). It is therefore useful to outline advantages and disadvantages that this digitizing principle has in the case of nuclear applications.

The flash-type digitizer has a parallel structure; the amplitude to be digitized is presented simultaneously to all the comparators and therefore the digitizing can be extremely fast. Actually the flash-type ADC is the fastest commercially available one. Conversion times of a few tens of ns for 8 bit are possible.

The flash digitizing may even operate without a stretcher. If the input signals have all the same shape, as in the case of Fig. 6.22 and the variable to be digitized is their peak amplitude, the input stretcher can be avoided provided that a strobing signal is generated, precisely coincident in time with the peak of the input pulses. In nuclear applications interest in this

kind of operation may be connected with analog-to-digital conversion in spectrometry applications at very high counting rates, with very short pulses. Another case in which the flash type digitizer is well suited occurs when the time-evolution of the current signal from a detector has to be investigated. This is done by taking repeated samples of the current waveform and by digitizing them with a flash-type digitizer.
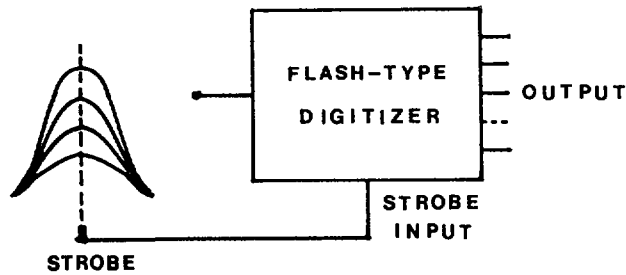

Fig. 6.22: Flash-type digitizer not preceded by a stretcher.

Conversely, the flash-type digitizer has a limitation which may be important in spectrometric application. As it is evident from the diagram of Fig. 6.21, the lower and upper boundaries of every channel are defined by independent physical elements. As a consequence, the uniformity in channel width cannot be very good. To give an example, assume that a 128 channel digitizer has to be realized on the 0-10V input range and that the comparators available have an offset voltage which may be positive or negative and reach a maximum value of 5 mV. The case of Fig. 6.22 can be considered, where the offset voltages e_K, e_{K+1} of the two comparators defining the K-th channel are now accounted for.
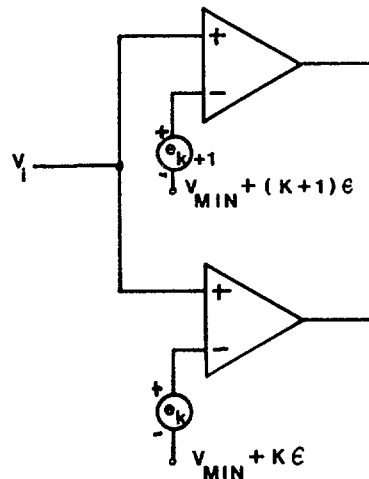
Fig. 6.23: Effect of offset voltage on channel width.

According to Fig. 6.23, V_i is attributed to the K-th channel if:

$$V_i > V_{MIN} + K\epsilon + e_K$$

and

$$V_i < V_{MIN} + (K + 1)\epsilon + e_{K+1}$$

The channel width is accordingly $\epsilon + e_{K+1} - e_K$. Using the values of our example, $\epsilon \approx 80mV$. In the worst case, when e_{K+1} is positive, e_K is negative (or vice versa) and both are at the maximum value of 5mV, the channel width inaccuracy due to the offset in the comparators might be as large as 10mV, thus leading to a 12,5% error in channel width. Such an error

is not acceptable in any spectroscopic measurement. The situation, of course, would become more serious if by using the same kind of the comparator, the number of channels should be increased.

6.5.2   Wilkinson-Type ADC

The principle introduced by D.H. Wilkinson in 1950 is based upon a linear discharge of the storage capacitor. A time-interval T is generated in this way, which is linearly related to the initial capacitor charge. This time interval is compared with the period of a reference clock; the number of clock pulses contained in the time-interval provides the digital value of the input amplitude. The block diagram of a Wilkinson-type ADC is shown in Fig. 6.24, which also provides an explanation of the ADC operation.
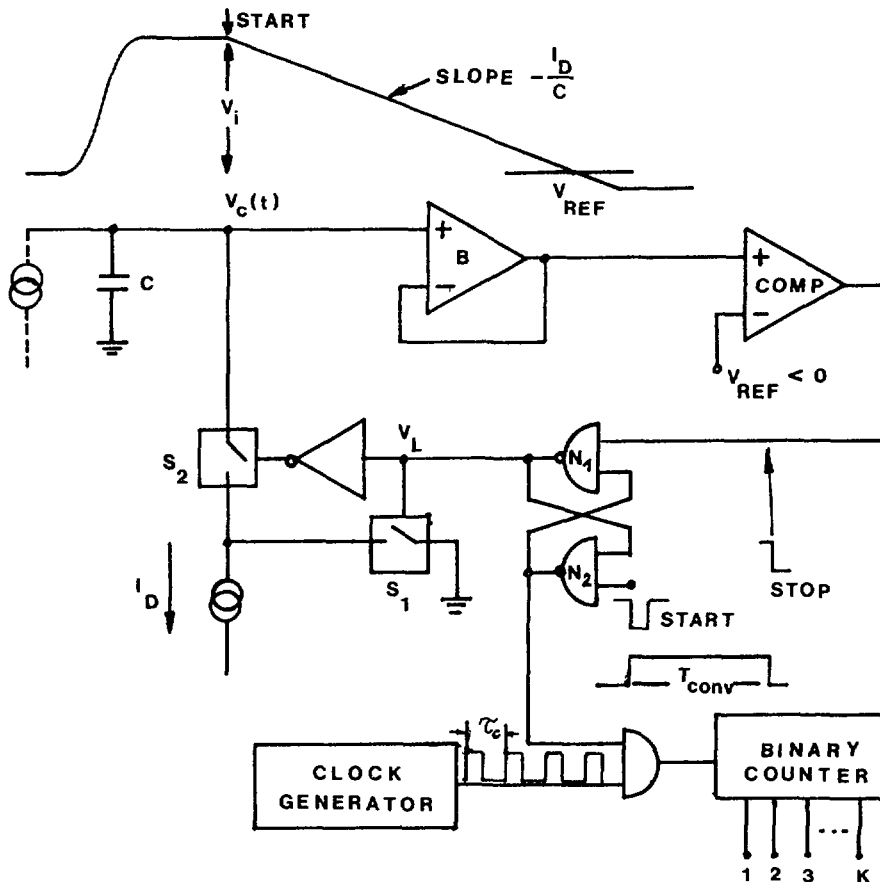
Fig. 6.24:   Block diagram of a Wilkinson type ADC.

The diagram shows the storage capacitor C disconnected from the body of the analog memory, either peak sensing or sample & hold, employed to charge it to the value V_i. In the charging phase of C, the flip-flop (N_1, N_2) is preset in such a way that the logic level V_L is HIGH. The switch S_1 is accordingly closed on the base of the convention, assumed so far for the control variable of the switches and the corresponding states, while S_2 is open. The storage capacitor C is practically isolated, the only two paths of charge leakage being represented by the input bias current of the buffer B, which, as previously, is supposed to have a JFET input and by the dashed source which accounts for the reverse currents of the charging circuit. As soon as the voltage accross C settles to the final value, a short negative triggering pulse, START, at the lower input of the flip-flop initiates the conversion operation. The flip-flop changes its state; V_L drops to logic LOW and accordingly S_1 opens and S_2 closes. The constant current I_D is

now applied to the storage capacitor. The voltage accross C follows the linear relationship:

$$V_c(t) = V_i - \frac{I_D}{C} \cdot t$$

and V_C is transmitted, through the buffer B, to the noninverting input of the comparator, whose other input is fixed to a slightly negative voltage V_{REF}. Since the (N_1, N_2) flip flop changed state upon receipt of the start command, the upper input of the AND gate A is at logic level HIGH and therefore the oscillation train from the clock generator is passed on to the binary counter.

The linear discharge of C and the clock counting go on until $V_C(t)$ reaches the reference voltage on the comparator, that is, until

$$V_c(t) = V_{REF} \quad \text{or} \quad V_i - \frac{I_D}{C} \cdot T_{CONV} = V_{REF} \; .$$

As soon as $V_C(t)$ crosses the V_{REF} level, the comparator changes state. The negative going edge at the comparator output resets the flip-flop and stops the counting. The conversion time T_{CONV} is equal to the width of the rectangular pulse on the upper input of the AND gate A:

$$T_{CONV} = (V_i - V_{REF}) \cdot \frac{C}{I_D} = V_i \cdot \frac{C}{I_D} - V_{REF} \frac{C}{I_D} \qquad \text{(Eq. 6.1)}$$

The relationship of Eq. 6.1 shows that as long as C and I_D can be considered independent of the voltage $V_C(t)$, T_{CONV} is linearly related to the amplitude to be converted V_i. The small negative V_{REF} has the effect of making T_{CONV} different from zero even at $V_i = 0$, it acts equivalently as a small pedestal added to V_i. If, instead V_{REF} were assumed to be positive, that is, of the same sign as V_i, then the conversion time would be

$$T_{CONV} = (V_i - V_{REF}) \cdot \frac{C_D}{I_D} \qquad \text{(Eq. 6.2)}$$

This shows that the effect of a positive V_{REF} is equivalent to a back bias applied on the input amplitude. The number counted by the binary counter during T_{CONV} is the channel number. It is given by:

$$n = \frac{T_{CONV}}{T_{CLOCK}} \qquad \text{(Eq. 6.3)}$$

where τ_{CLOCK} is the period of the clock oscillator.

Combining Eq. 6.2 and Eq. 6.3 the dependence between V_i and n is easily found:

$$n = \frac{1}{\tau_{CLOCK}} \cdot (V_i - V_{REF}) \frac{C}{I_D} \qquad \text{(Eq. 6.4)}$$

Eq. 6.4 states that thermal and long term drifts in the reference level V_{REF}, in the clock period τ_{CLOCK} as well as in the discharge current I_D affect the stability of the V_i - n relationship and so does a possible temperature dependence of C. Consequently, in high resolution, high precision Wilkinson-type ADCs, I_D must be provided by a highly stable generator, the digitising clock should be quartz-controlled and the capacitor C should have the smallest possible temperature coefficient.

136

The current source should be realised according to the diagram in Fig. 6.21. The current I_D is delivered by the drain of a field effect transistor around which a feedback loop based on an operational amplifier is designed. The noninverting input of the operational amplifier is connected to a stable voltage provided by a reference element. Let V_O be the reference voltage and e_O the amplifier offset voltage. The voltage drop across the source resistor R of the field-effect transistor is $V_O \pm e_O$ where the double sign in front of e_O accounts for the fact that the offset may be positive or negative.
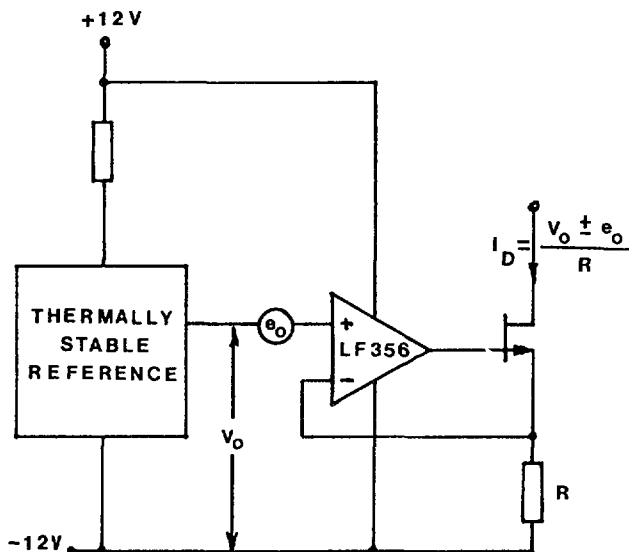
Fig. 6.25: High precision current source.

The source current and therefore the drain current are given by:

$$I_D = \frac{V_O \pm e_O}{R}$$

As e_O has a thermal coefficient of a few $\mu V/°C$ and R may be of the order of 10 Kohm, the thermal drift on I_D due to e_O may turn out to be as small as $10^{-10} A/°K$ for an indefinitely stable resistor. The major source of instability is still the reference element which might introduce on I_D a thermal drift of about 10 ppm/$°K$.

As pointed out by Eq. 6.4, the major source of integral nonlinearity are the voltage dependences of C and I_D. As long as I_D is generated by a feedback circuit of the type of Fig. 6.25, which is an almost ideal current source, I_D can be assumed to be strictly constant. The storage capacitor C, if properly chosen, may be of no concern. The diagram of Fig. 6.20 may feature integral nonlinearity smaller than 0.01%. It has to be made sure that such a high linearity of the digitizing section is not spoiled by a less linear behaviour in the analog memory, stretcher or sample & hold.

The main point in favour of the Wilkinson type ADCs is their inherently high uniformity in channel width. Such a high uniformity is related to the fact that a single physical parameter, that is, the period of the digitizing clock determines the width of all the channels.

Conversely, the Wilkinson principle is comparatively slow, for its operation is of a serial nature. According to Eq. 6.3, written in the form

$$T_{CONV} = T_{CLOCK} \cdot n \qquad (Eq. 6.5)$$

it can be observed that the higher the channel order, the longer the conversion time. Let n_{MAX} be the order of the last channel and $T_{CONV, MAX}$ the time required to digitize an amplitude which falls into the channel of order n_{MAX}. If, for instance, $n_{MAX} = 8,000$ and $T_{CLOCK} = 2.5$ ns, this being the shortest digitizing period of commercial Wilkinson type ADCs presently available, Eq. 6.4 would yield

$$T_{CONV, MAX} \simeq 20\mu s.$$

More conventionally, digitizing at $t_{CLOCK} = 10$ ns (clock frequency = 100 MHz), $T_{CONV, MAX}$ would be $100\mu s$. This might turn out to be too long in several spectrometric applications.

It might be argued that new Ga As logic chips are now becoming available; they are expected to reduce soon the value of τ_{CLOCK} down to 1 ns and therefore, the value of $T_{CONV, MAX}$ for an 8,000 channel ADC to $8\mu s$.

However, the general consideration is still valid, namely a digitizing principle of serial nature has, as an intrinsic limitation, a conversion time which increases linearly with the total number of channel.

6.5.3   ADCs employing successive approximation with binary steps

To introduce the principle of successive approximation with binary steps with an example, it will be assumed that the input range is 10V and the desired number of channels is 16. The estimation of a signal amplitude V_i with an accuracy of:

$$\frac{10V}{16} = 0.625 \text{ V requires four binary scaled weights:}$$

5V	2.5V	1.25V	0.625V
MOST SIGNIFICANT BIT			LEAST SIGNIFICANT BIT
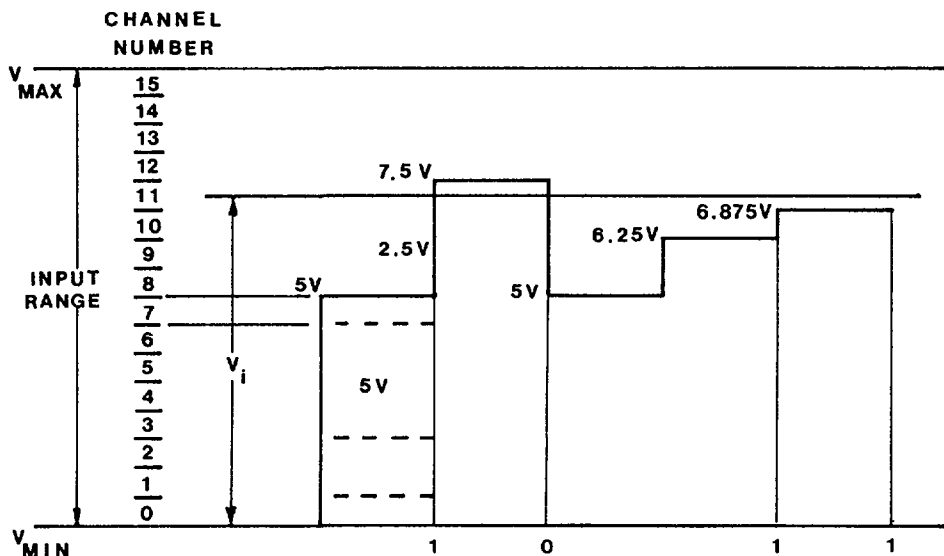
Fig. 6.26:   Successive approximation with binary scaled weights.

The attribution of V_i to the relevant channel takes place is illustrated in Fig. 6.26.

V_i is first compared to the MOST SIGNIFICANT BIT (MSB), 5V in the actual case. As V_i is larger than 5V, the 5V weight is left and the next

one is added to it. The resulting approximating voltage, 5V + 2.5V exceeds V_i, which means that the 2.5V weight has to be removed. The approximating voltage is back again at 5V. Then, the 1.25V weight is added and the resulting 6.25V level is compared with V_i. V_i exceeds the approximating voltage. The 1.25V weight is left and the next weight, LEAST SIGNIFICANT BIT (LSB) is added. The approximating voltage is now 5 + 1.25 + 0.625 = 6,875 and this represents the lower boundary of the channel where V_i is sorted. In units of LSB, such a lower bounary can be expressed as:

$$(1.2^3 + 0.2^2 + 1.2^1 + 1.2^0) \; V_{LSB}$$

As in this case V_{LSB} represents the channel width, $V_{LSB} = \in$, it can be concluded that V_i is attributed to the channel whose lower boundary is $11 \in$, that is, to the 11th channel, as confirmed by Fig. 6.26.

Subdivision of the input range into 16 channels requires therefore 4 binary weights and the attribution of V_i to one of these channels requires just four comparisons. More general, subdivision of the input range into 2^K channels requires K binary weights and K comparisons. So, an ADC with 8,192 channels is realised with 13 binary scaled weights, and 13 comparisons are needed to attribute the input amplitude to one of the channels. Assuming again a 10V input range, the MSB is 5V, the LSB is

$$\frac{10V}{2^{13}} \approx 1.20 \text{ mV}.$$

As every comparison can be done in a reasonably short time, the conversion time of ADCs employing successive approximation with binary weights, besides being amplitude independent, can be considerably smaller than the conversion time of a Wilkinson-type ADC with the same number of channels.

Actually, the time necessary to make one comparison depends on the error band within which the switched binary weights have to be settled. Consider, for instance, the case of an ADC with 8192 channels and a 10V input range. According to the timing sequence of Fig. 6.26, the input amplitude has first to be compared with the 5V weight.

Switching a 5V step is not a simple operation as may be inferred from from Fig. 6.27, especially if the 5V step must be compared with V_i within the accuracy of less than 1 channel width, 1.25 mV in the actual case.

Fig. 6.27, shows the actual time-dependence of the 5V step where it is switched ON.
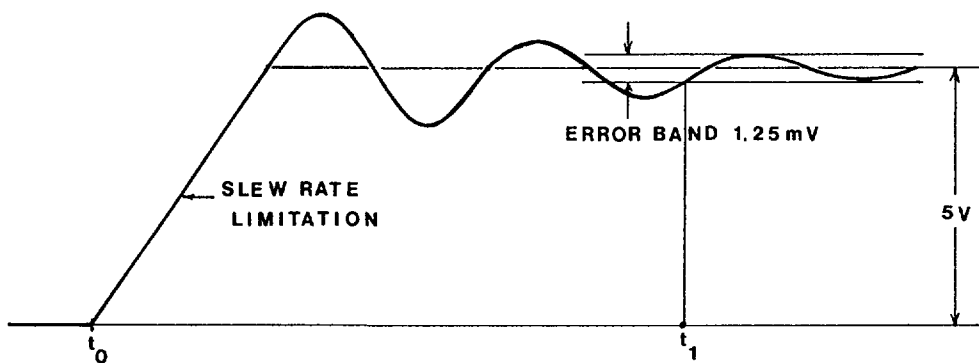
Fig. 6.27: Settling transient following the switching ON of a binary weight

Let t_0 be the instant at which the SWITCHING ON COMMAND is applied. There is an initial rise of almost linear behaviour, whose slope is fixed by the SLEW RATE limitations. Settling to the nominal value of 5V occurs with a damped oscillation in most cases. If an error band of 1.235 mV is drawn symetrically around the asymptotic line, the comparison with the input voltage cannot be performed until the damped oscillation enters the error band and remains within it forever. According to Fig. 6.27, therefore, the comparison cannot begin before the instant t_1.

As it is clear that the settling time, for given amplitudes of the binary weights increases as the channel width and therefore the error band is reduced; this explains why the time required for the single comparisons is longer for ADCs of smaller channel width, that is, for ADCs with a larger number of channels on a fixed input range. It is not surprising therefore, that a successive approximation ADC with 4096 channels can nowadays be realised with less than 2μs conversion time, while an ADC with 8192 channels can hardly be realised with a conversion time of less than 10μs.

Nonetheless, ADCs employing successive approximations with binary weights are surely faster than Wilkinson-type ADCs with the same number of channels.

On the other hand, successive approximation ADCs suffer from a limitation which restricts their extension to spectrometric applications. Such a limitation is related to the fact that the widths of the different channels are not determined by a single physical element and therefore they may exhibit large inaccuracies in channel width. Consider, for instance the case of the 7th channel, whose lower boundary is determined by the sum of three binary scaled weights, respectively

$$0.625V + 1.25V + 2.5V$$

and whose upper boundary is determined by the 5V weight. Assume that the 5V weight has a relative inaccuracy of 1%, that is, its value is 4.95V rather than 5.00. The width of the 7-th channel would be

$$4.95 - 4.375 = 0.575V$$

This means that the 7-th channel is affected by a channel width error of $\frac{0.625 - 0.575}{0.625}$ = 8%, which in spectrometric applications cannot be tolerated.

As a second example, it is interesting to evaluate the accuracy required from the 5V weight in order to make sure that, in an ADC with 8192 channels, the error in the width of channel 4095 is less than 1%, all the other weights being assumed indefinitely accurate.

The channel of order of 4095 has, as an upper boundary $(5 + \delta)V$, where δ is the error on the 5V weight. The channel width is, accordingly:

$$E = 1.25 \text{ mV} + \delta$$

and in order to keep the error on E lower than 1%, δ must be less than 12.5μV. To conclude, the required accuracy on the 5V weight must be of:

$$\frac{12.5 \text{ μV}}{5V} = 2.5 \text{ ppm.}$$

Resistors with such a degree of accuracy exist, but they turn out to be highly expensive.

It is for this reason that successive approximation ADCs, though fast in operation, cannot be extended in a straightforward way to spectrometric applications, where the request of a highly uniform channel width is essential.

A way of overcoming such a difficulty will be discussed later on. Now the detailed block diagram of a successive approximation ADC will be presented (see Fig. 6.28).
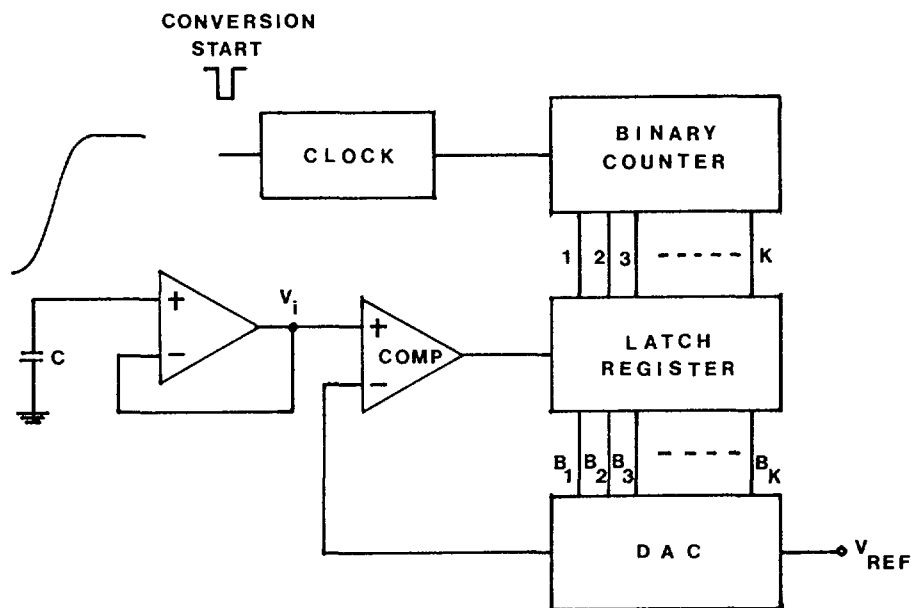
Fig. 6.28: Block diagram of a successive approximation ADC.

As shown in Fig. 6.28, once the charge storage transients have settled and the voltage across C can be considered steady, a CONVERSION START command initiates the operation, by activating the clock. The first clock pulse is counted and the relevant counter output sets in the LATCH REGISTER the flip flop which controls the DAC most significant bit, that is, B_1 goes to logic HIGH. At the DAC output the level

$\dfrac{V_{REF}}{2}$ is generated and comparison with V_i takes place.

If $V_i > \dfrac{V_{REF}}{2}$, no output appears from the comparator and the first latch remains at the logic state at which it was set before. If V_i, instead, is smaller than

$\dfrac{V_{REF}}{2}$, the transition which occurs at the comparator output resets the

LATCH and the $\dfrac{V_{REF}}{2}$ weight is accordingly removed. The second clock pulse sets the second latch in the register and therefore B_2 goes to logic HIGH. At the DAC output appear now the voltages

$$\dfrac{V_{REF}}{2} + \dfrac{V_{REF}}{4} \quad \text{or} \quad 0 + \dfrac{V_{REF}}{4}$$

depending on whether B_1 was in the previous comparison left at logic HIGH or reset to logic LOW. The operation proceeds until the attribution of the amplitude V_i is completed.

The parameters of the diagram of Fig. 6.23 must be chosen in the following way:

$$V_{REF} = V_{MAX} - V_{MIN} = 10V \quad if \quad V_{MAX} - V_{MIN} = 10V,$$

as is customary. The number K must be chosen in such a way that the total number of channels N = must be equal to 2^K.

6.5.4   Sliding-Scale Method to Correct Channel Width Irregularities

As pointed out in the previous paragraphs, converters like the one based upon a stack of comparators as well as the one employing successive approximations are limited in their applications to nuclear spectrometry by their intrinsic channel width irregularities. These irregularities come from the fact that the width of the channels is determined by different physical elements. Both types of converters, however, are attractive in some high-rate applications as they may be much faster than a Wilkinson-type ADC with the same number of channels. It is for this reason that attention was concentrated on a possible way of correcting the inaccuracies in channel width, without sacrificing the conversion speed of both multicomparator and successive approximation ADCs. The solution found in 1963 by Gatti and coworkers had the great advantage of making these converters suitable for nuclear spectrometry application. The method employed to correct the channel width inaccuracies was called "sliding scale".

The SLIDING SCALE method will be introduced with reference to Fig. 6.29, which shows 11 ADC channels with large width inaccuracies.
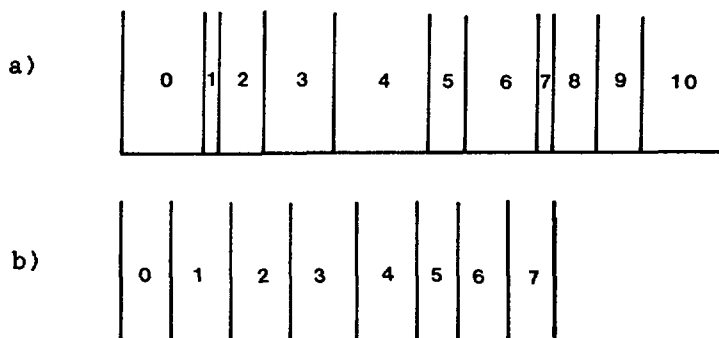
Fig. 6.29:   a) Channel width distribution of the original ADC
             b) Channel width distribution after averaging over 4
             channels of the original ADC.

Assume that, instead of 11 channels, only 8 are needed. then, the channel width irregularities of the original ADC can be smoothed by defining new 8 channels according to the following averaging procedure.

The width of the zero-th channel of the new ADC scale results from the average of the widths of channels 0, 1, 2, 3 of the old scale. The width of the first channel of the new ADC scale results from the average of the widths of channels 1, 2, 3, 4 in the old scale.

The width of the 7-th channel of the new ADC scale results from the average of the widths of channels 7, 8, 9, 10 in the old scale.

In other words, the width of any channel in the new ADC scale is obtained by averaging the width of the channel of the same order in the old

scale and the widths of the three channels that in the old scale follow the channel under consideration.

The effect of averaging is shown in Fig. 6.29.b. A considerable reduction in the channel width error has been introduced. The new scale has 8 channels, that is, three less than the original one, but this has no practical relevance.

The way in which the sliding scale is applied in a real case of a high resolution ADC is shown in Fig. 6.30.
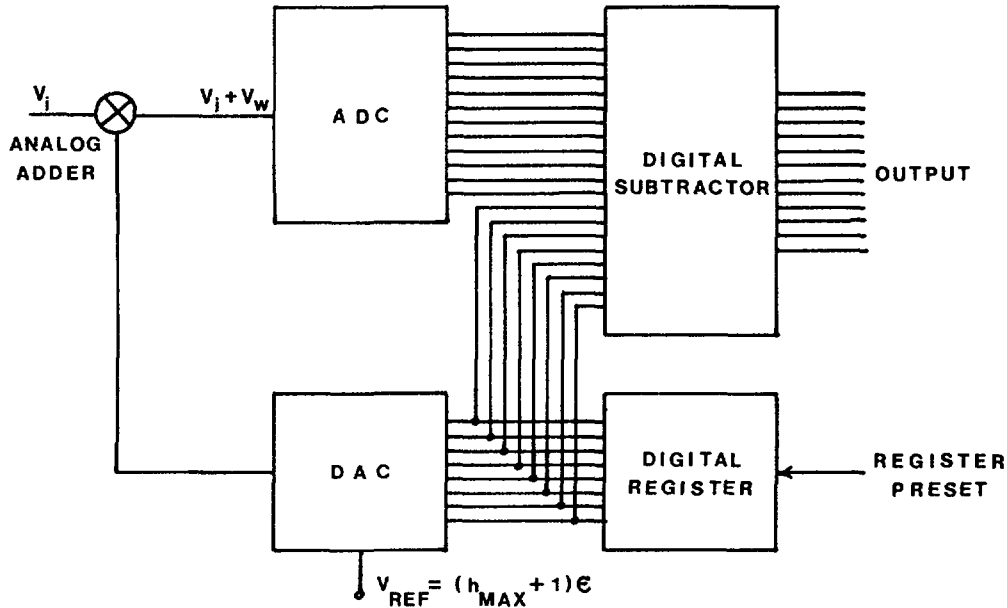
Fig. 6.30: Realization of sliding scale averaging

Starting from an ADC with a given number of channels, for instance, 4096, equivalent to 12 bits and wishing to perform the sliding-scale averaging over, say, 256 channels, the effective number of available channels in the final scale will be 4096-256 = 3840. Assuming a 10V input range for the original ADC, that is, a channel width of $\in$ = 2.5 mV, the sliding scale averaging will span 256 x 2.5 mV, or about 640 mV. The sliding-scale averaging is realised by generating analog voltages V_W that are integer multiples of the channel width $\in$ of the ADC $V_W = h.\in$, with h variable between 0 and 255, and adding them to the incoming pulses. The resulting voltage $V_i + V_W$ is encoded and from the obtained digital number, h is subtracted. In this way, to V_i is again attributed the correct digital value. Suppose now that the same amplitude V_i occurs repeatedly at the input and every time it occurs, a different value of V_W is added to it, until all the possible values of V_W are scanned. The digital number appearing at the output of the digital subtractor will always be the same, on account of the fact that if n_i is the digital number corresponding to V_i with a channel width $\in$, the digital number corresponding to $V_i + h \in$ is obviously $n_i + h$. The correct number n_i is found once h is subtracted. However, the analog amplitude $V_i + h \in$ at the ADC input, as h is varied, is sorted into different channels of the original ADC. If h spans the 0-255 range, for instance, the effective channel width results from averaging the widths of 256 consecutive channels in the original ADC.

The larger the number of channels upon which avering is performed, the lower the final inaccuracy in channel width.

143
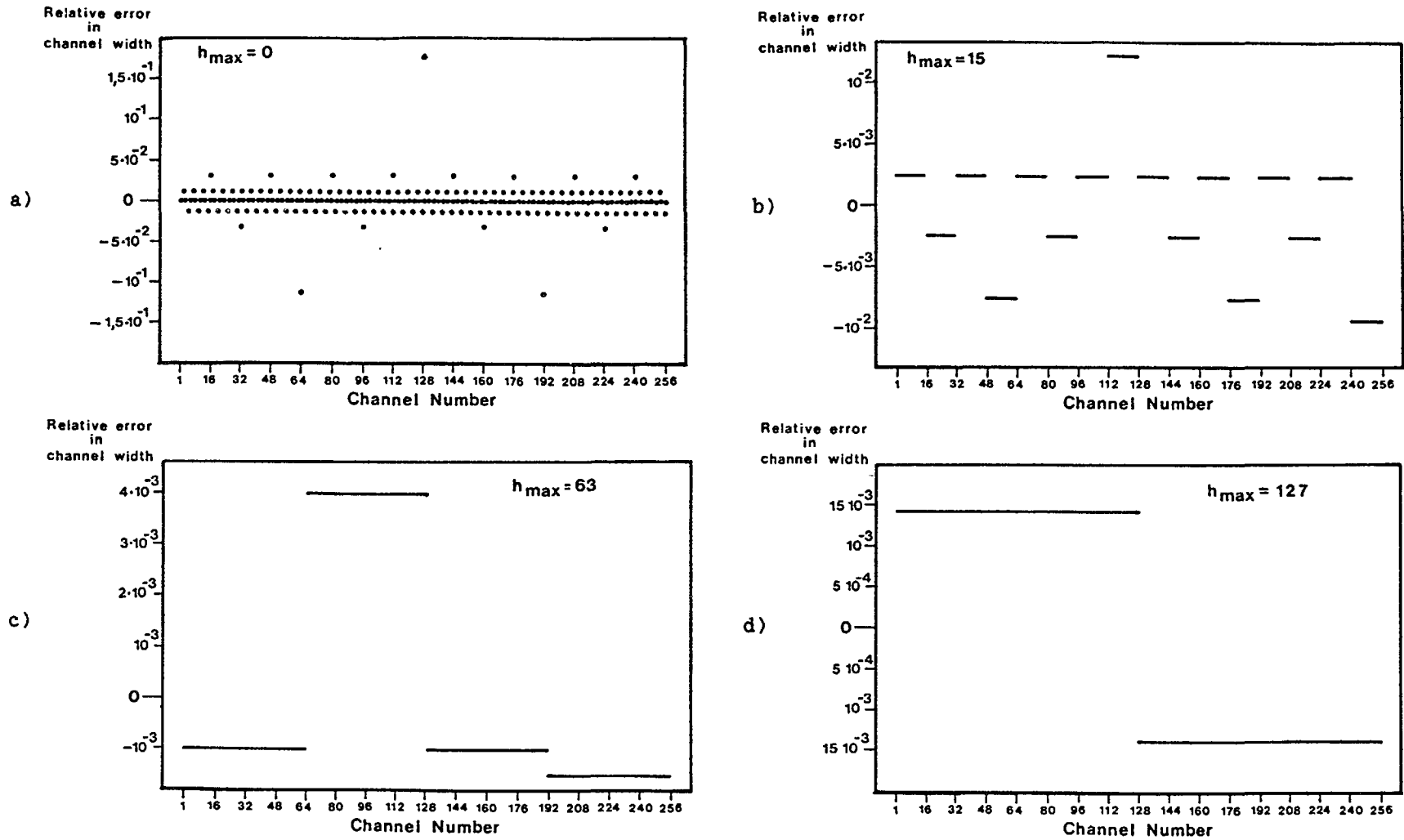
Fig. 6.31: Effect of the sliding scale averaging on the channel width inaccuracy of a successive approximation ADC.
a)   original inaccuracies
b)   results obtained by averaging over 16 channels
c)   results obtained by averaging over 64 channels
d)   results obtained by averaging over 127 channels (half range of the original ADC.

The curves a), b), c), and d) of Fig. 6.31 show the progressing
reduction in the relative channel width error ($\in$ - $\in$)/$\in$, and the
changes in the error pattern as the number of channels over which averaging
is performed is increased. The curves refer to a complete simulation of a
256 channel successive approximation ADC.

Going back to Fig. 6.30 it has to be pointed out that the operation of
the digital register can be sequential, that is, the register can be
replaced by a counter which is incremented by one unit after each conversion
or randomly by generating a random number after each conversion and writing
it into the register. The sequential preset is acceptable if the input
amplitudes are distributed randomly, as in the case of a nuclear spectrum.
If the input amplitudes occur with a deterministic sequence, like in the
case of a sweeping pulser, a random preset of the sliding scale register is
advisable.

The effect of sliding scale can be fully appreciated by comparing the
two spectra of Fig. 6.32. They are ^{60}Co scintillation detector spectra,
stored in a multichannel analyser which has at the input a 1024 channel
successive approximation ADC. In the spectrum of Fig. 6.23.a, the sliding
scale connection was excluded. The fluctuations in the content from channel
to channel are due to the intrinsically poor channel width accuracy of the
successive approximation ADC.

By adding the sliding scale correction, the spectrum has improved in
the way shown in Fig. 6.23.b.

a)

b)
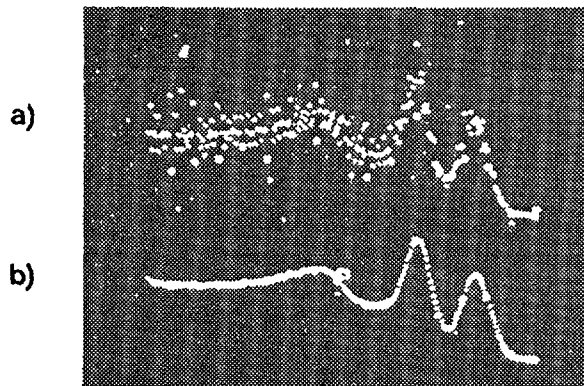
Fig. 6.32: Scintillation detector spectra obtained with a 1024 channel
            ADC.
            a) without sliding-scale correction
            b) with sliding-scale correction

6.6   DISCUSSION OF THE GENERAL DIAGRAM OF A WILKINSON-TYPE ADC

To conclude the section on analog-to-digital conversion, the detailed
analysis of the timing sequence which controls the operation of a
Wilkinson-type ADC will be carried on. The analysis will be refereed to
Fig. 6.33, where the input signal, supposed to be of Gaussian shape, through
a baseline restorer arrives to the linear gate. The gate can either be open
in the standing state or may open upon receipt of a control command,
correlated in time with the incoming signal. From the linear gate the
signal reaches the peak stretcher, which consists of the comparison
amplifier A_1, of the buffer A_2 of the charging diode D and storage
capacitor C. Three current generators, T_2, T_4, T_5 are connected to
the diode D cathode and they perform the functions described below. The

transistor T_7 is normally OFF and it will be employed for the fast
discharge of the storage capacitor at the end of the conversion.

The timing sequence of the ADC operation is shown in Fig. 6.34. The
inputer signal 1 reaches the peak stretcher. The stretcher has been made
ready to follow the input signal only in the positive-going direction, for
shortly before the arrival of signal 1 the current I_S, was truncated, (see
waveform 8). As I_S is usually of the order of some mA, as long as it is
present, it enables the stretcher to follow the positive-going and the
negative going input slopes as well. Removing I_S is the condition under
which the stretcher becomes a unidirectional amplifier. Once I_S is
removed, the forward conduction of the diode is guaranteed by a small
current I_L, which is of the order of a few nA. The forward current in the
diode substantially increases during the leading edge of the Gaussian
signal, until the peak is reached. As soon as the peak is detected, the
following sequence occurs. The output of the comparison amplifier drops to
the negative saturation level (waveform 10) thereby cutting off the diode
D. The sharp transition 10 is sensed by the PEAK DETECTING TRIGGER, which
generates a short rectangular signal (waveform 4). Such a signal is used to
generate the CLOSING COMMAND for the linear gate. At the output of the
linear gate the waveform has the behaviour shown at 2. Shortly after the
peak, the current I_2 is also removed and the storage capacitor remains
floating. The turning off command of I_2 is shown in diagram 7. The
circuit is now ready for the linear rundown. The command 9 activates the
linear discharge by turning off T_6 and consequently turning on T_5. The
accurately known, highly stable I_D through T_5 discharges C at constant
rate. As shown in Fig. 6.33, I_D can be varied in coarse steps by the
rotary switch S. This gives the possibility of changing, for a fixed analog
range and a fixed clockrate, the number of channels and the channel width.
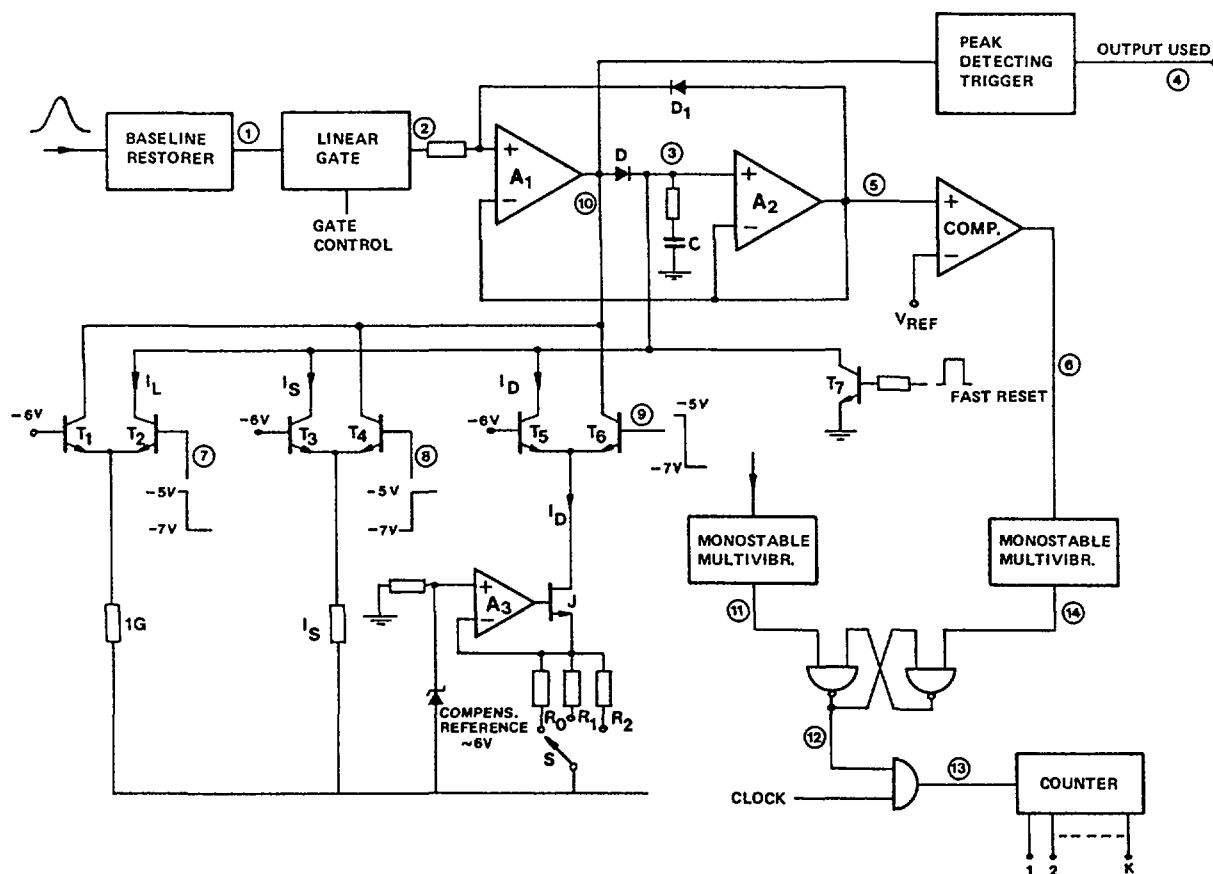
Fig. 6.33: Detailed diagram of a Wilkinson-type ADC
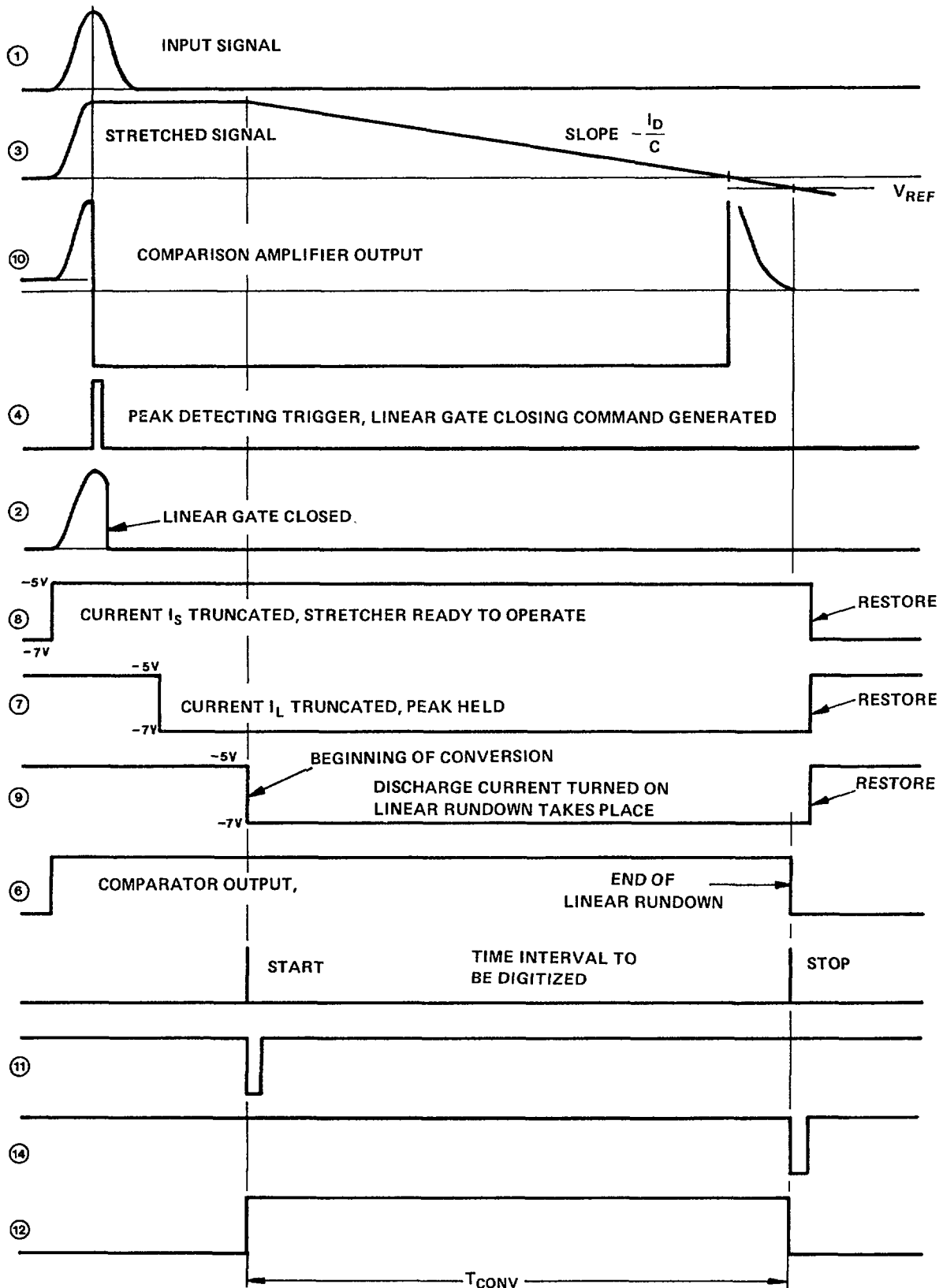
Fig. 6.34: Timing sequences in the Wilkinson-type ADC

Throughout the linear rundown, the input of A_1 is protected against possible damage which may arise from large differential voltages by the diode D_1, which keeps the noninverting input of A_1 one voltage drop below the voltage of the inverting one.

The time interval to be digitized is determined by two instants: a start instant defined by the HIGH-TO-LOW TRANSITION in waveform 9 and a stop instant defined by the switching of the comparator output from HIGH to LOW.

A rectangular signal can be easily generated by setting a flip-flop with the START transition and by resetting it with the STOP transition.

The number of clock pulses that were connected during the conversion interval remain stored in the counter whose K output lines represent in binary parallel form the order of the channel in which the input pulse was sorted.

Once the conversion is executed, after a short delay during which the status of the K output lines is transferred into a LATCH register, the currents I_L, I_S are restored, I_D is removed and a short rectangular pulse of positive polarity drives T_7 into saturation, thereby resetting any residual charge on the capacitor C.

The linear gate can now be opened or enabled to open on the next incoming event. The gate opening must be done under the control of a baseline inspector to avoid that the tail of a pulse present at the gate input during the conversion be transmitted to the stretcher.

CHAPTER 7

MULTICHANNEL ANALYZER

7   MULTICHANNEL ANALYZER

7.1   GENERAL DESCRIPTION

The ADC makes the conversion of the pulse amplitude into a digitial number. This number is expressed in binary form by setting the output lines in either binary state according to a certain code, normally pure binary. As previously mentioned, the multichannel pulse height analyzer (MCA) has to keep record of the number of pulses converted in each channel. A storage is needed for this purpose, and basically the MCA consists of a memory with as many addresses as channel numbers are provided. The content of each memory location is the number of times in which an input pulse corresponding to this channel number has been converted by the ADC.

In Fig. 7.1 a simplified block diagram of the memory organization of a MCA is given. Observe that the address of the memory location is selected by the external device, the ADC. This makes a difference with other systems in which the external device puts its data into a memory location whose address is selected by the system itself. In the MCA, the data converted channel determines the address of the memory location. The new content of the selected location will simply be the previous value plus one.
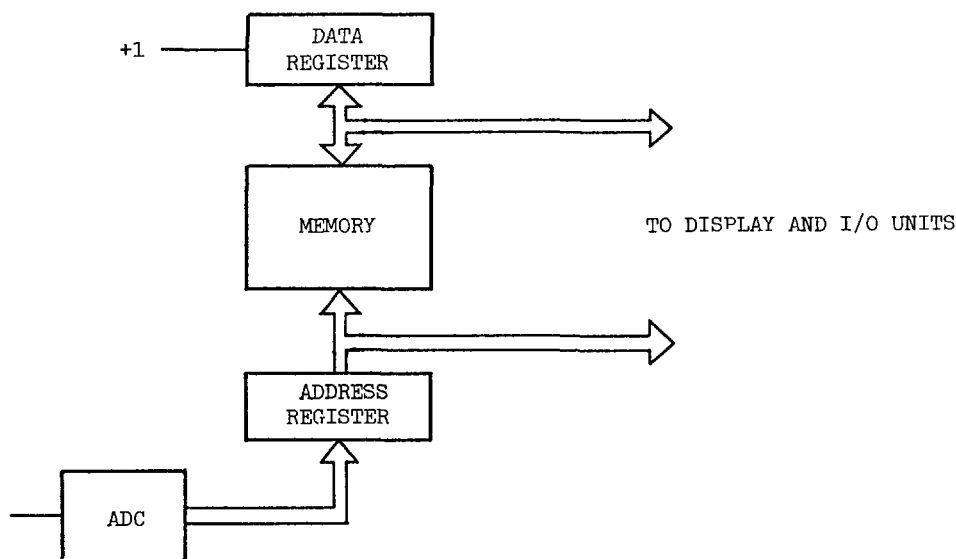
Fig. 7.1:   Simplified block diagram of a multichannel pulse height analyzer

A sequence of operations is performed each time a pulse is analyzed:

-    The ADC converts the pulse height into a digital number.

-    This number is stored in the address register of the MCA memory and one memory location is selected.

-    The content of the memory location is stored in the data register.

-    The data register's content is incremented by one unit.

-    The new content of the data register is stored back into the memory location.

151

This process is repeated a large number of times and at the end the memory contains information on the probability of a pulse to fall within each of the channel boundaries. If the pulse amplitude is proportional to the energy of the particle detected, the memory contains the energy spectrum.

There is now a need to visualize the spectrum. This can be done using the configuration described Fig. 7.2.

If the content of the address register is incremented very fast, each time by one unit, the data register will sequentially show the content of each memory location. Connecting a digital to analog converter to the data register, and another one to the address register, two analog signals are produced. One delivers an output voltage proportional to the memory content, while the other gives a voltage proportional to the address or channel. Connecting the output of the former DAC to the vertical deflection plates of a cathode-ray tube and the output of the latter to the horizontal ones, a display of the spectrum is obtained.
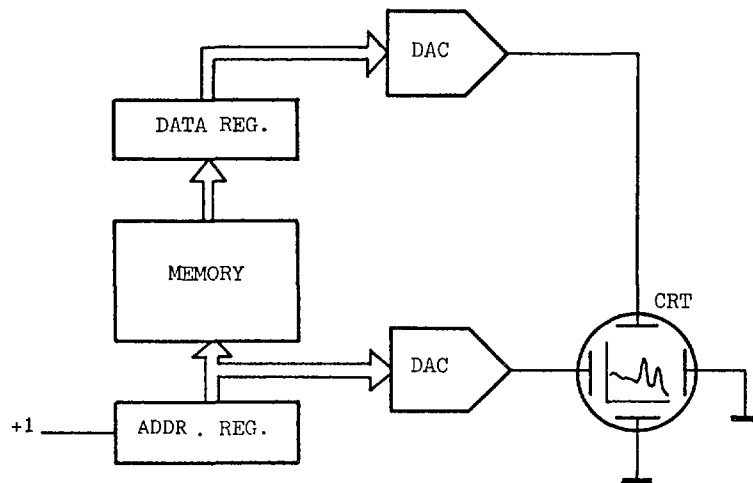
Fig. 7.2: Visualization of the memory content
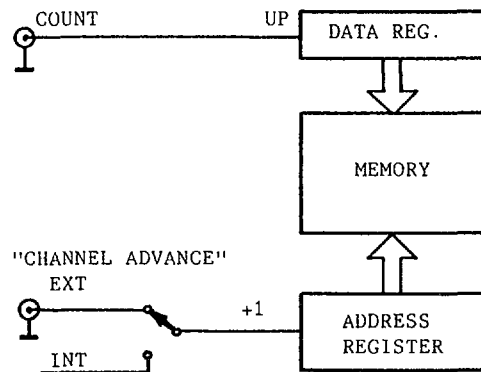
Fig. 7.3: Multichannel analyzer used in the multichannel scaling mode

Another mode of operation is possible in the MCA. The use of it as a multichannel scaler (MCS) is necessary when it is connected to a Mössbauer spectrometer or when measuring life times of radioisotopes.

In those cases each memory location is addressed sequentially for a fixed period of time called dwell time. The data register counts the external pulses and at the end of each period it stores its content into the

memory location. The next address is selected by an internal or external pulse generator and the process is repeated again.

In this way, each memory location contains a number which is a function of time as the address is sequentially selected.

7.2 MEMORY ORGANIZATION

The memory is implemented with RAM chips whose interconnection depends on the specific MCA. In all cases the number of addresses available corresponds to the required number of channels, but the number of bits used per channel depends on the code adopted. If the information is stored in pure binary, N bits per channel determines a 2^N counts per channel. Instead, if the code adopted is BCD then the number of bits per channel should be four times the number of decades.

To clarify, let us examine some examples of how a hypothetical MCA memory would be implemented.
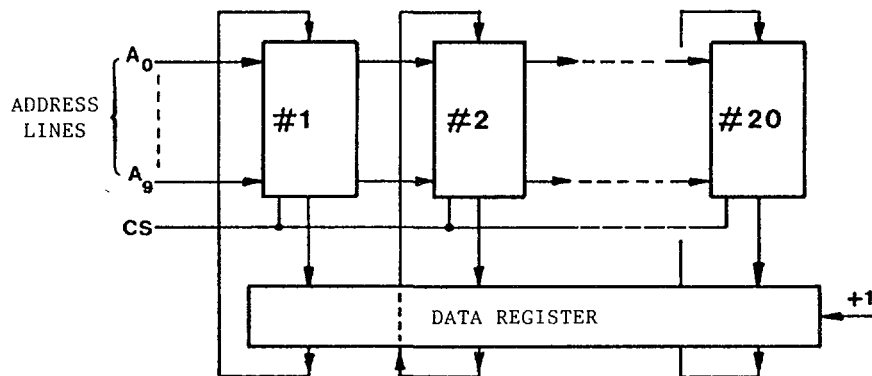
Fig. 7.4: 1K channels, 1 million counts per channel (20 bits) memory organization

In Fig. 7.4, a 1024 channel, 1 million counts per channel memory is shown. 1K x 1 memory chips such as 2114 were used. Note that twenty units were needed. All address lines are bussed, that is, equal weight address lines are connected together. In this way, every chip holds one bit of the total channel content. When selected each memory stores its contents in the data register, or loads data from the data register depending on the operation performed.

The memory content can also be coded in BCD format. In this case four bits per digit are needed and twenty-four chips are therefore required to hold six decades.

The BCD organization uses more chips compared to the pure binary one. This configuration is used by some manufacturers because the output operation is simplified as the use of binary to BCD converters is avoided.

Let us see how the memory is organized in the Canberra Series 30 MCA.

The memory has a capacity to store 1K channels of 6 BCD digits (maximum number of counts per channel 10^6). Also an additional bit is reserved for each channel. This special bit is used to identify whether the channel has been selected as ROI (Region Of Interest), as it will be described later on.

Not all the memory locations are used to store the number of counts per channel. Some are used for special purposes. For instance, the memory location which corresponds to channel zero is used to store the current live time, that is to say, the time elapsed since the accumulation period started, deducting the dead time.

Channels 1022 and 1023 are used to store the value of the integral, i.e. a sum of memory contents in a selected ROI. Observe that the integral has to have a greater capacity than the channel content. For that reason, the integral channels have a capacity of 10^8.

In Fig. 7.5, a block diagram of the memory is given.
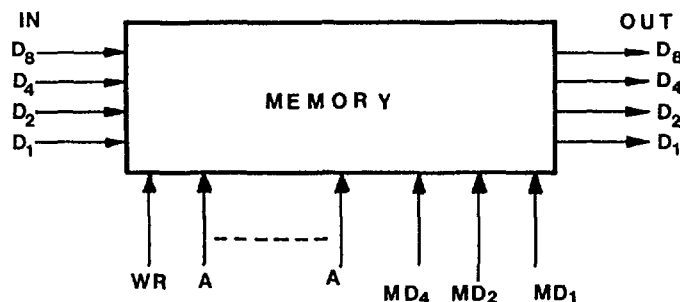
Fig. 7.5:   Block diagram of the memory organization used in the
            Canberra Series 30 MCA

Looking from outside, the memory presents the following characteristics.

One thousand channels can be accessed (more precisely 1024) and they are distributed as follows:

Channel #	Holds	Capacity (digits)
0	live time	6 BCD
1 to 1021	Normal counts	6 BCD + 1 ROI digit
1022, 1023	Integral	8 BCD

Address lines A0 to A9 determine the channel number.

Data enters through four lines of BCD codes. All digits are sequentially input. The digit selection is done with lines MD_4, MD_2 and MD_1 according to the following table.

MD_4	MD_2	MD_2	DATA CHANNEL	INTEGRAL CHANNEL
0	0	0	10 digit	10^0 digit
0	0	1	10^1 digit	10^1 digit
0	1	0	10^2 digit	10^2 digit
0	1	1	10^3 digit	10^3 digit
1	0	0	10^4 digit	10^4 digit
1	0	1	10^5 digit	10^5 digit
1	1	0	No function	10^6 digit
1	1	0	ROI tag bit (1 bit)	10^7 digit

Also data are transmitted to the output through four BCD lines and are selected according to the table above.

As an example, let us assume that the ADC converted a pulse in channel 324. One more count has to be added to the current content of channel 324. The ADC output lines will be connected, through a multiplexer, to the A_0 to A_9 address selection lines, according to the following code:

A9 ....A0
101000100

154

while the address lines are stable, the routing bits MD4, MD2, MD1 will be set to 000. The BCD content of channel 324 will output starting from the least significant decade, i.e. the units (10^0). Its value will go to a special circuit called ADD ONE LOGIC and the result will be written back into the memory location. If the previous content was a 9, then a carry is generated and a zero is written in the digit and the routing bits increments by one, selecting the next decade 10^1. The carry is added to the content of the selected decade. This scanning of the routing bits is done whenever a carry is detected. The control logic takes care of the timing of the whole process.

The memory is implemented using 4K x 1 CMOS RAM chips. Eight chips are used and the interconnection is shown in Fig. 7.6.
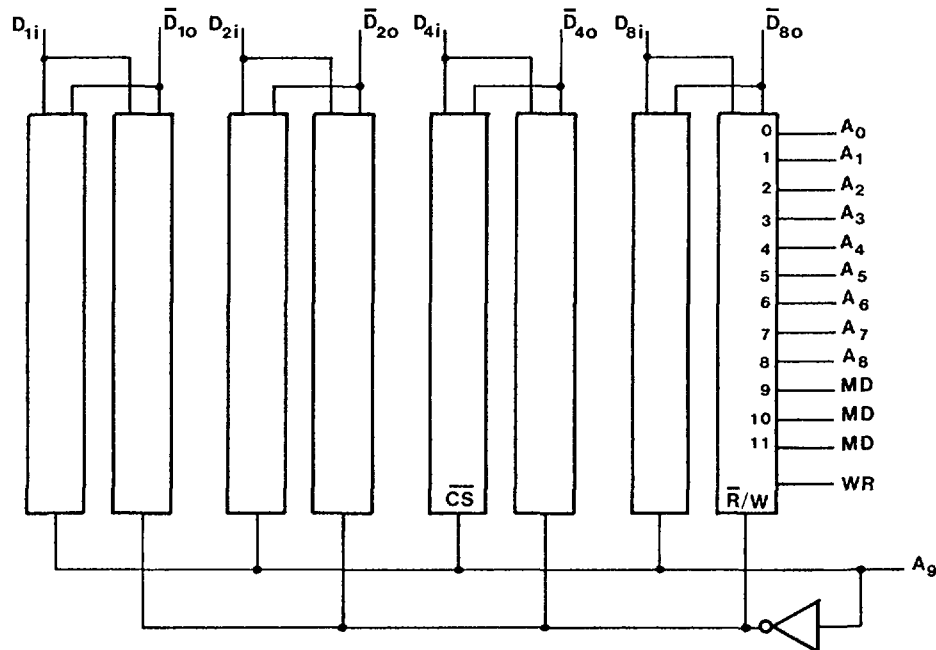
Fig. 7.6: Memory organization

Observe that the basic block consists of two chips sharing common input and output pins. Which chip will be active depends on the state of the chip select pin. This in turn depends on the logic state of A9.

Each memory chip has twelve address lines to select one out of 4K memory locations. The first nine lines are directly connected to the A0 to A8 address lines coming from the addressing device (the ADC for instance). The tenth input line (A9) is used to select one of the two chips of the basic block or memory bank.

The most significant address lines of the chip are connected to the Memory Routing bits MD1 to MD4.

In this way each memory bank holds one bit of every BCD digit. At the same time, each bank holds the corresponding BCD digit of all 1K channels, but A9 determines that channels 0 to 511 will be located in the left chip of the bank whereas channels 512 to 1023 in the right one.

It is important to understand the memory organization when troubleshooting. Let us give an example: when outputing the memory content into a teletype, we notice that although the memory is cleared, a number 4 is printed from channel 0 to 512. We can suspect that a failure is in the

left chip of the second bank starting from the right, as it is the one which holds the third BCD digits of the first 512 channels.

The memory map of one memory bank corresponding to bit D8 of a BCD digit is given in Fig. 7.7.
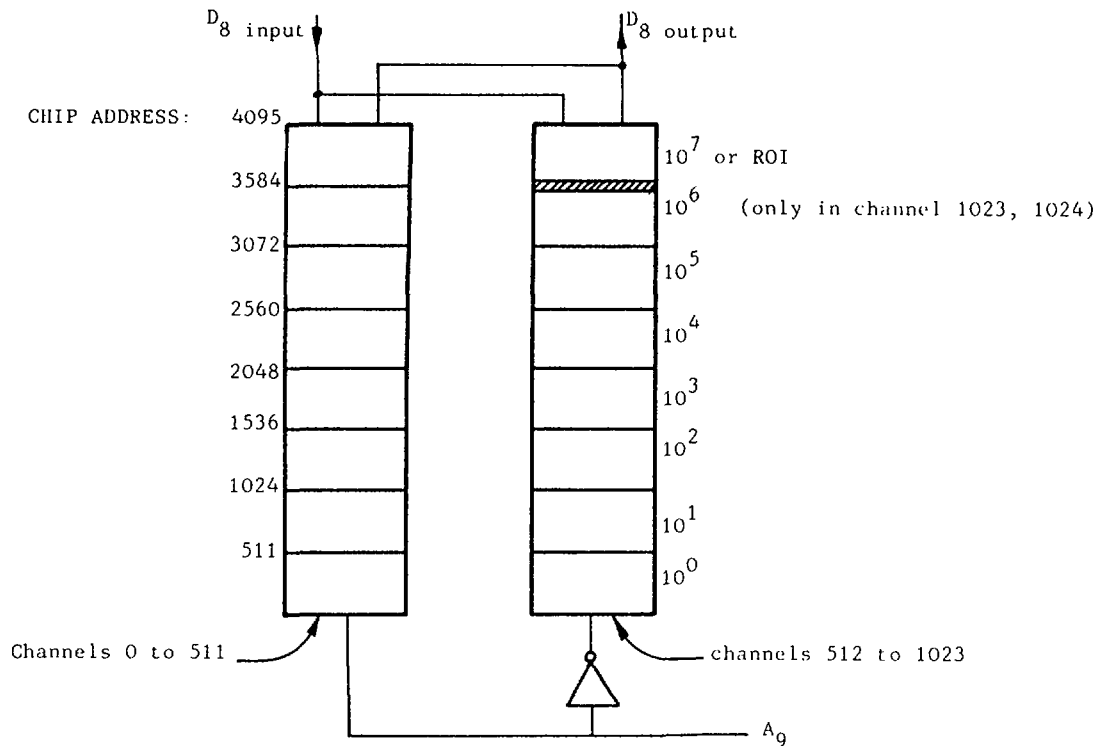
Fig. 7.7:   Memory map of digit D8

The addressing of the memory may come from different sources depending on the task being performed. A multiplexer selects the source which could either be the ADC or registers related to data display (ROI, cursor).

One important unit of any MCA is the display. Nowadays many MCAs use a TV tube for displaying data. For this application, the TV tube is operated in such a way that the raster is vertically oriented (Fig. 7.8).
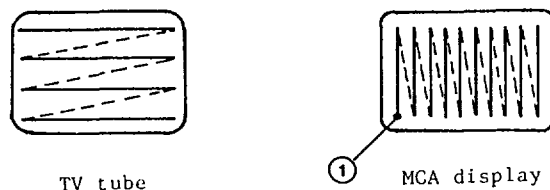
Fig. 7.8:   Display tube

When the field synchronisation pulse is present the beam is in position 1. When the field sweeps, 512 lines are displayed every 64 microseconds. If the memory address is scanned in sychronisation with the field, we can associate each vertical line with a given memory address.

The display unit consists of an address counter which addresses the memory in synchronisation with the display. The memory outputs BCD digits which are compared with a BCD counter fed by a 25 MHz clock. When both are equal a signal is sent to the video amplifier which produces a dot on the display.

156
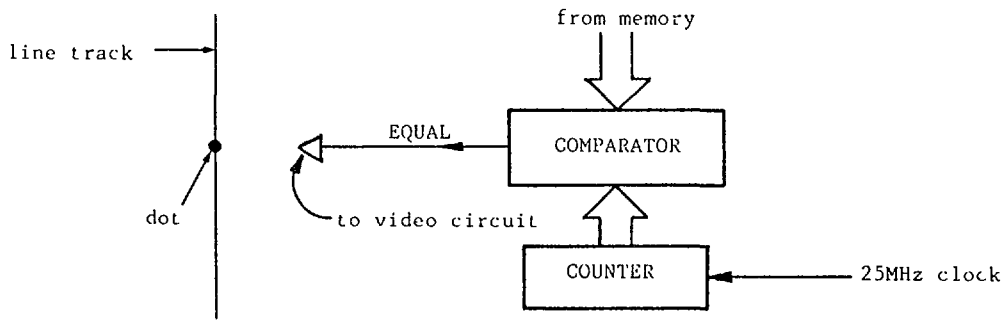
Fig. 7.9: Displaying memory content

The beam takes approximately 60 microseconds to go from bottom to top of the screen. The counter takes a time t_c to reach a value equal to the memory content. This time is proportional to it. Considering that while the counter is increasing its value, the beam sweeps from bottom to top, the displacement of the dot will be proportional to the current memory content.

Observe that if the clock is 25 MHz the counter reaches 999 counts in 40 microseconds. At this moment the beam is almost at the top (full scale position). The comparison is done only in the three most significant BCD digits, which gives a satisfactory visual resolution (1000 points).

If expansion of the vertical scale is desired, then the clock frequency has to be reduced, for instance, by a factor two. In this way, the counter takes twice the time to reach equality with the memory content and a dot can be produced at the top of the screen for only half of the previous count.

As the display described has only 512 lines, when a complete 1K channel spectrum is displayed, only the even channels are selected losing some information, but having a view of the complete spectrum.
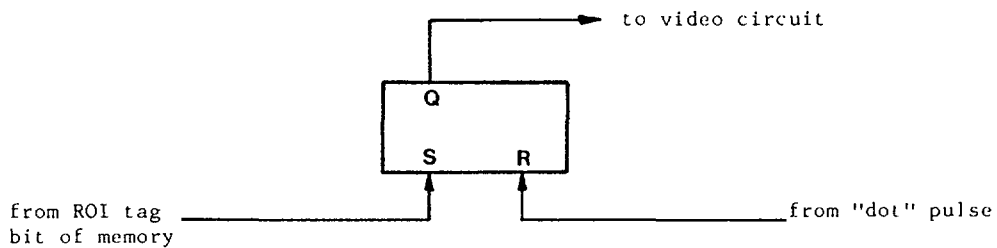
Fig. 7.10: Displaying ROI channels

NORMAL CHANNEL   ROI CHANNEL
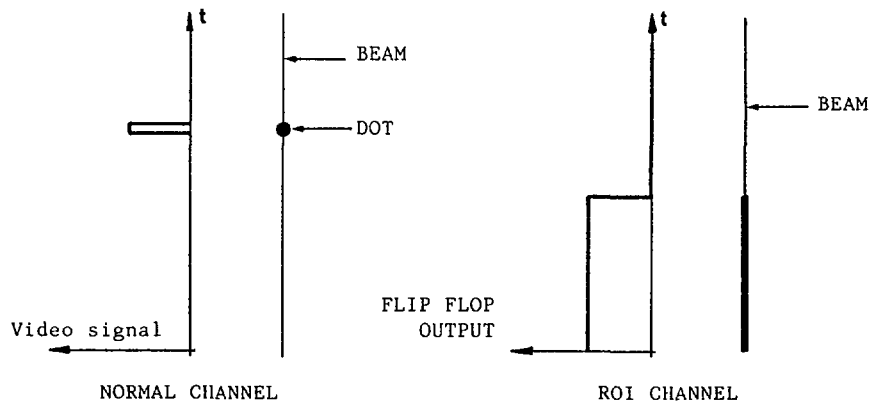
Fig. 7.11: Video signal in normal and ROI channels.

The Region of Interest (ROI): this feature is used to select groups of channels either to be displayed in a remarkable form, to be output selectivity or to perform any operation such as integration of the number of counts.

ROI selected channels are identified by setting to high logic level one bit of its content called ROI tag bit. When displaying, the ROI channels are intensified from bottom to top and this can be done with the simple circuit shown in Fig. 7.10. The RS flip-flop is set if the channel to be displayed has the ROI tag bit in high logic level. When the "dot" pulse comes, the flip flop is reset. The output is used to produce the intensified signal (see Fig. 7.11).

At this point a question arises and it is how can we select ROI channels. To answer it, we have to first see how the cursor function is implemented. Again, it must be remembered that there are several ways to do it depending on the manufacturer but all of them will approximate the proposed scheme.

When the display is being performed, the memory is addressed sequentially and there is a register or a counter which keeps record of the channel or address currently being displayed.

At the same time, an up-down binary counter (cursor register) can be loaded with the desired cursor position by counting clock pulses. A switch on the front panel can be actuated either way to determine up or down counting. Once loaded the situation is that the cursor register content remains stable while the address register continuously increases as the lines are being displayed.

A comparator signals out when both registers are equal and the pulse produced goes into the video circuit producing a dot (or a small bar) (see Fig. 7.12).
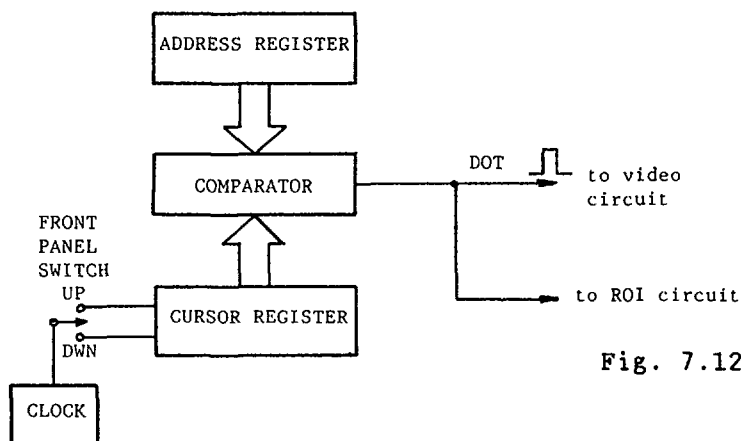
Fig. 7.12: Displaying the cursor

To select ROI channels a switch has to be operated indicating such a purpose. This activates a circuit which simply sets to high the ROI tag bit of the channel being addressed in coincidence with the comparator output.

Character generation: The characters displayed on the screen are generated with the use of a special circuit based on a Read Only Memory (ROM).

Characters are represented using a matrix of 7 rows by 5 columns. Each column is associated to a display line and seven dots can be displayed by driving the video circuit with seven short pulses.

158
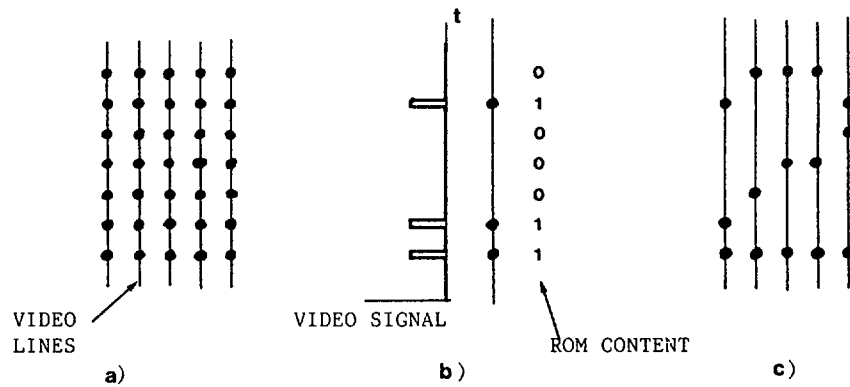
Fig 7.13: Matrix representation of characters: a) Basic matrix 7 rows by 5 columns, b) Video signal for the first column of character 2, c) Character 2

Any character can be represented by a 7 x 5 matrix with enough accuracy. In Fig. 7.13 the basic matrix is represented and a character "2" is formed by proper selection of the dots. Observe that a dot can be associated to a logic state "1" while a blank to a "0".

The character generator circuit consists basically of a ROM and a multiplexer. The ROM contains the pattern of ones and zeros to represent any character (see Fig. 7.16).
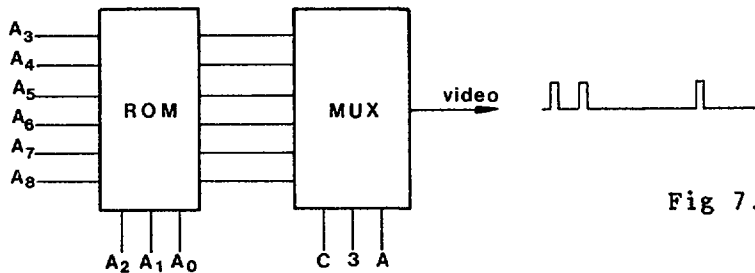
Fig 7.14: Character generator

Let us see the sequence of events which leads to the representation of the number 2 on the screen.

First, the binary code of the data is applied in the address lines A3, A4, A5, A6 of the ROM. Lines A7 and A8 are set by the control circuit which tells that a number is to be displayed. Let us suppose they are in states 0 and 1 respectively. So the address lines are in the states 100010XXX. The address lines select a field in the ROM which occupies eight locations. In these locations the bit pattern for the representation of the character "2" is stored.

279					272			ADDRESS
0	0	0	0	0	0	0	0	D_7
0	0	0	1	1	1	0	0	D_6
0	0	1	0	0	0	1	0	D_5
0	0	0	0	0	0	1	0	D_4
0	0	0	0	1	1	0	0	D_3
0	0	0	1	0	0	0	0	D_2
0	0	1	0	0	0	0	0	D_1
0	0	1	1	1	1	1	0	D_0
1	0	1	0	1	0	1	0	A_0
1	1	0	0	1	1	0	0	A_1
1	1	1	1	0	0	0	0	A_2

Fig. 7.15: ROM Content for the representation of character "2"

159

Observe in Fig. 7.15 that the "1"'s follows the shape of the character "2".

After the eight locations ROM field have been selected, the remaining addresses A0, A1, A2 are set to 000. The ROM output presents the first set of contents, in this case all zeros.

Next, the multiplexer selects all digits D0 to D7, so that its output is a train of pulses which drives the video amplifier.

The next step is a change in A0 so that the second column is selected. Again all bits are presented simultaneously to the multiplexer input which serialize them by fastly scanning the selection bits. A train of pulses are again sent to the video amplifier, in this case 10001100. The resulting video signal is a series of dots which form the most right column of the character "2". Let us calculate the speed at which the least significant addresses A0, A1, A2 and the selection bits of the multiplexer have to be scanned.

As every ROM column has to be in coincidence with the line sweep of the cathode ray tube, the addresses A0, A1, A2 may be driven by a counter receiving a clock input which could be the line synchronism pulses. In this way while a raster line is changed, the next address in the ROM is selected.

If we want to calculate the clock frequency for the multiplexer selection bits, we have first to decide the height of the characters. Let us suppose 0,5 cm. If we consider that the CRT beam travels about 20 cm in 50 microseconds (one line sweep), to travel 0,5, it will take 1,25 microseconds. During this time, eight dots have to be generated, so the clock period will be

$$T_{dot} = \frac{1250 \text{ ns}}{8} = 156 \text{ ns}$$

and the "dot frequency" 6,4 MHz. A complete circuit for the character generator will include also two counters as is shown in Fig. 7.16.
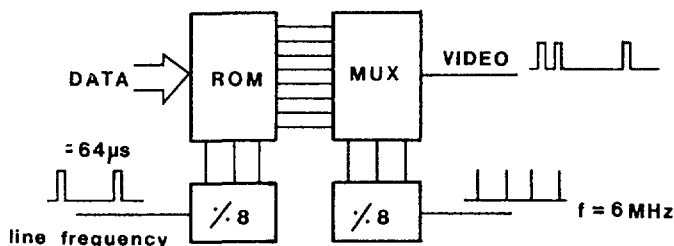
Fig. 7.16: Character generator.

Binary to BCD conversion: in some cases it is necessary to convert a binary number, such as the memory address, into a BCD number. One reason for doing that could be the need to printout the content of the memory and its corresponding addresses, and the peripheral has to receive data in BCD. The principle of operation of such a counter is to count fast clock pulses in a BCD counter and simultaneously in a binary counter. Comparing the content of the binary counter with the number to be converted, the count is stopped when both values are equal. At this moment, the BCD counter holds a BCD number which is equivalent to the binary one (see Fig. 7.17).

Timer circuits: all measurements of radiation activity are based upon the number of events counted per unit time. Measuring just the number of

160
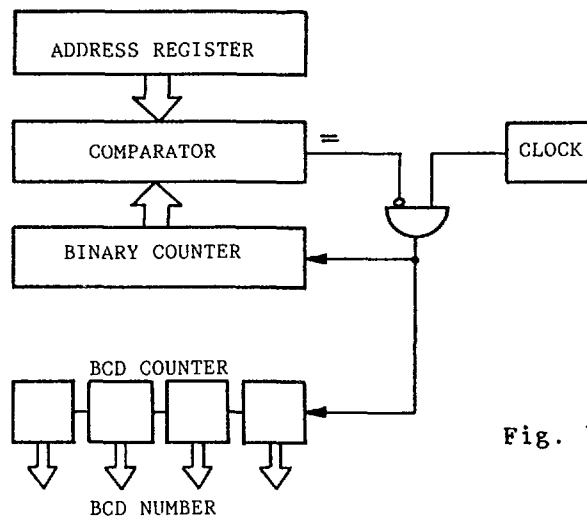
Fig. 7.17: Binary to BCD converter

pulses accumulated during a given time is a simplified form to measure activity as there are many other physically related factors to be considered.

Due to the time taken by the analog to digital converted to convert a pulse, it happens that not all the pulses ariving to the ADC are counted. Therefore, some sort of correction in the measuring time is required. The rising and the amplitude of the pulses reaching the ADC have a random nature. Therefore the conversion time is also random. A representation of a signal coming from the ADC which indicates when the ADC is busy has the following typical shape, Fig. 7.18.
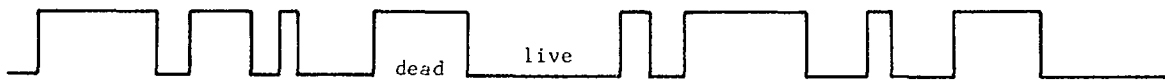
Fig. 7.18: "Busy" signal coming from the ADC.

We call the period of time when the ADC is busy "dead", whereas it is "live" the rest of the time. The total, real time is the "clock" time.

It is valid when making an experience and the dead time is considerable, to count only the live time and this can be realized with the circuit shown in Fig. 7.19.
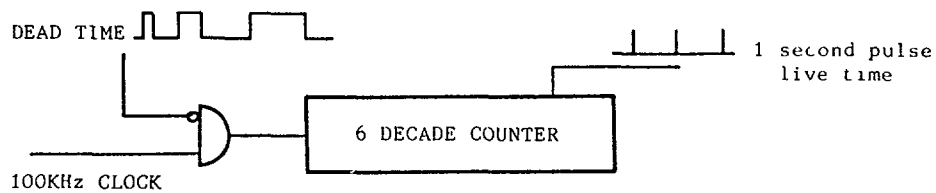
Fig. 7.19: Live time measuring circuit

A 100 KHz clock is counted only when the time is "live". Therefore dividing by 10^5 in the counter one pulse per "live" second is obtained.

The live time is counted in channel zero of the MCA.

Input/Output: the multichannel analyzer acquires data from the ADC and the memory content can be displayed on a screen. This is not enough, as we would also require to store the data in a magnetic tape or to print it on paper, or sent it to an external computer for further analysis of the collected data. Mechanism to input data from external devices are also required.

It is necessary to provide a way for MCA to communicate with the external world and this is done through the Input/Output Section.

Depending on the type of external device a different kind of interfacing is required.

If a high speed parallel printer is to be used, data will be outputed in BCD parallel. Instead, if a teletype or CRT terminal is connected, then a serial link will be adopted.

Parallel output: high speed printers with ten digit columns are frequently used for printing data of MCAs. They are very fast and permit printing out four digits for address and six for the content.
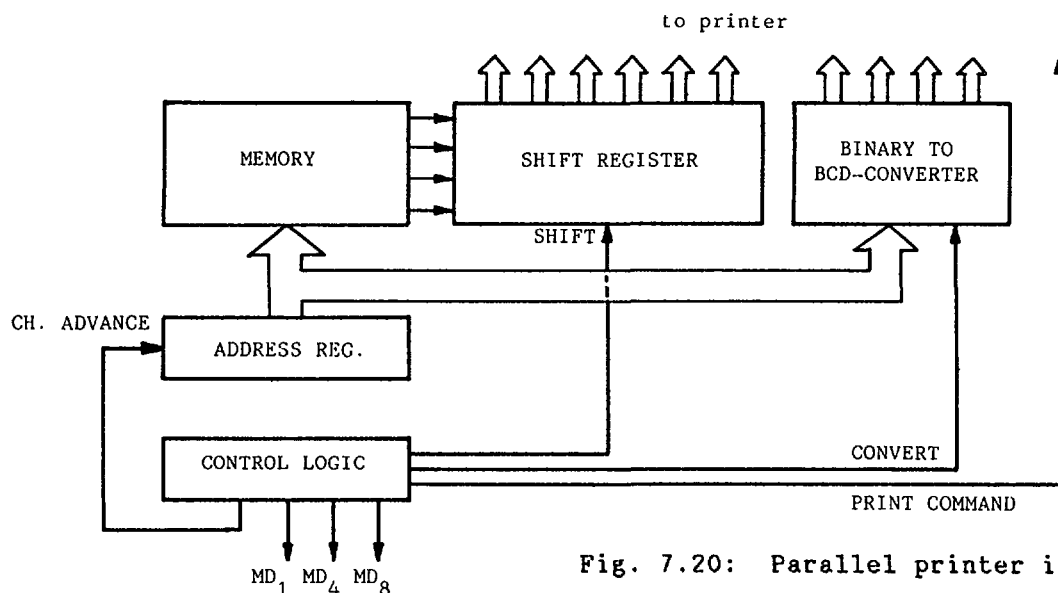
Fig. 7.20: Parallel printer interface

The printer interface (see Fig. 7.20) works in the following way: the first address is selected, by loading the address register. The first address could be channel zero or the beginning of a ROI.

The first BCD digit of the pointed memory location is loaded into the shift register. Next, memory routing MD bits are changed to select the second BCD digit. The process continues up to the moment the shift register is full. The address register content is converted into BCD code in the binary to BCD converter. When all BCD digits are present at the output lines, a point command pulse is sent to the parallel printer. The control logic advances one channel the address register and the process is repeated again up to the last channel is printed.

Observe that the interconnection with the printer uses more than 40 wires. Although the printing rate is fast, it cannot be used for long distances.

Plotter interface: by using analog to digital converters, it is possible to produce a graph of the memory content in an external plotter. The signals required by the plotter are the following:

- Analog voltage proportional to the channel content to move the pen to the proper vertical (Y) position.

- Analog voltage proportional to the current channel number to move the pen to the proper horizontal position. Observe that this X signal can

162

also be a digital pulse for one step advance as the X position is sequentially incremented. The type of signal, either analog or digital depends on the plotter used.

- Plot command: just a logic pulse to drive the pen down for a short time to print a dot.

One remark on the DAC's used is that the vertical one might take the three most significant BCD digits, giving a resolution of 1000 points. This DAC should be a 12 bit BCD digital to analog converter. The horizontal DAC might be a 10 bit DAC as the address is coded in pure binary and 1K channels are to be printed.

Serial Interface: this is the most popular way of interconnecting with the outside world as it is much more simple than the parallel interface. Only four wires are used for input and output of data. Of course date is present in the system in parallel form so it has to be serialized. For that purpose special chips are used called Universal Asynchronous Receiver Transmitters (UART, Fig. 7.21). It is worthwhile describing these devices as well as the asynchronous mode of data communication.

Before sending them each character has to be coded in ASC II. Converting a BCD digit into ASCIT code is very simple as the remaining three digits are 011 and they are added by the control logic. For instance the digit 6 in BCD is 0110 and in ASCII 0110110.
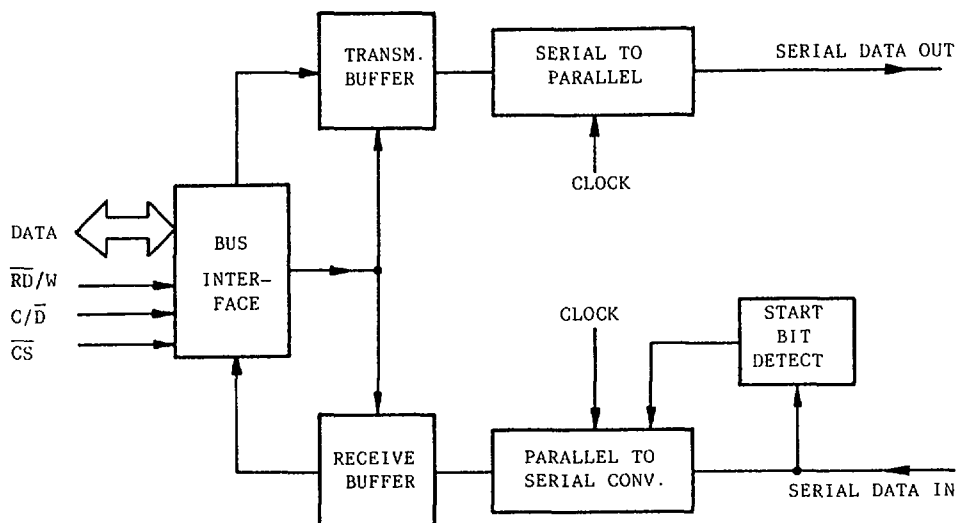
Fig. 7.21: Universal asynchronous receiver transmitter

Parallel data and control lines are connected to the bus interface. When a character is to be sent, a Write Command is sent to the chip. This action causes the data to be sent to the transmit buffer and then is shifted to the parallel to serial converter. This device adds start and stops bits at either end of the character code and then under control of the clock sends the bits out one at a time on the serial transmit line.

To receive the procedure is reversed. A start bit is detected (a high to low transition). The UART starts the assembling of a byte of data in the receiving buffer, having first undergone a serial to parallel conversion. It is held there until a Read command is received from the control unit of the MCA.

163

Both receiver and transmitter have to be set to the same speed of transmitter/reception or baud rate (bits per second).

The asynchronous communication is simple because the receiver just waits for the first high to low transition in interpreting it as a start. Then it samples the input at the clock rate and when all expected bits have been counted, it waits for the next high to low transition to start a new cycle.

Although both receiver and transmitter are clocked at the small baud rate, a large margin of error in the clock frequency is permitted.

Typical band rates are: 9600 for communication with a CRT terminal or 110 baud for communication with a teletype.

In Fig. 7.22, a waveshape corresponding to the transmission of a character is shown.

Fig. 7.22:   Asynchronous transmission of a character

The electrical connection with the external device can be done according to two different standards: EIARS-232C and the 20 mA current loop standard. The latter is, in fact, the standard that gained popularity due to the widespread use of teletype printers.

EIA RS-232C Standard: this standard defines the interfacing between data terminal equipment and the data communications equipment employing serial binary data interchange. The standard specifies 25 data lines but for most purposes, only a few are used. The typical lines used in the communication of a MCA and a peripheral are indicated in Fig. 7.23.
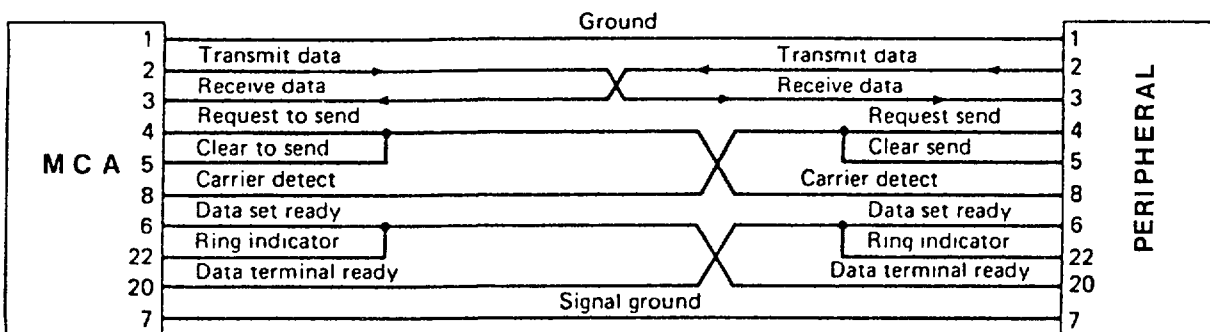
Fig. 7.23:   RS 232C Connection to a peripheral device.

The electrical levels are specified to be between -3V and -25V, and +3V to 25V. Systems typically uses +6V for "0" and -6V for "1".

Some integrated circuit like 74188/189 are used to interface TTL to EIA levels.

20mA current loop: this standard is much simpler than the RS-232C interface standard and consists only of four basic wires:

Transmit minus   Receive minus
Transmit plus   Receive plus

164

As Fig. 7.24 illustrates, four lines form two 20mA current loops. Logic 1s and 0s are sensed by opening and closing the current loop.
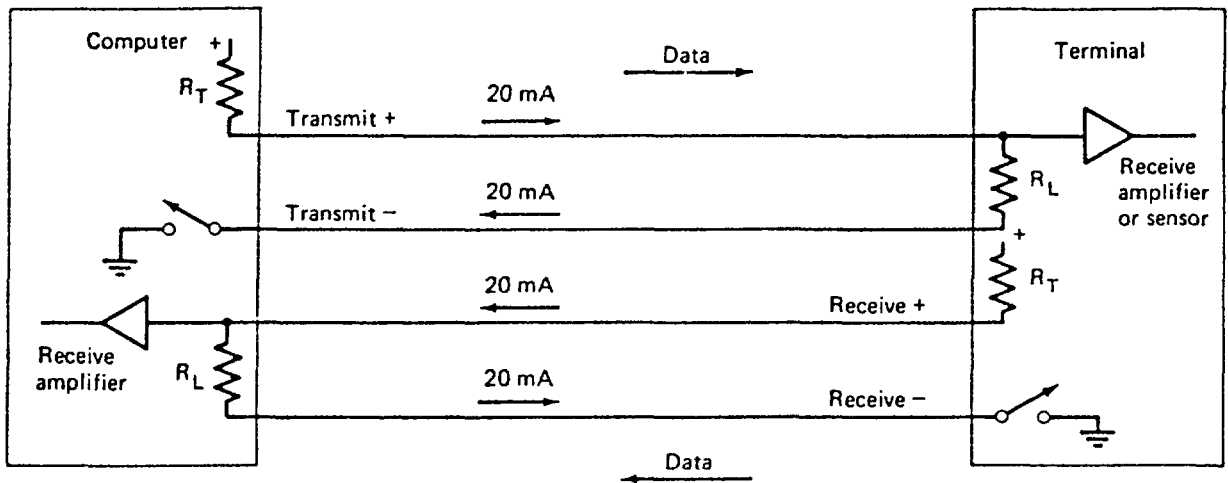
Fig. 7.24: 20 mA current loop interface
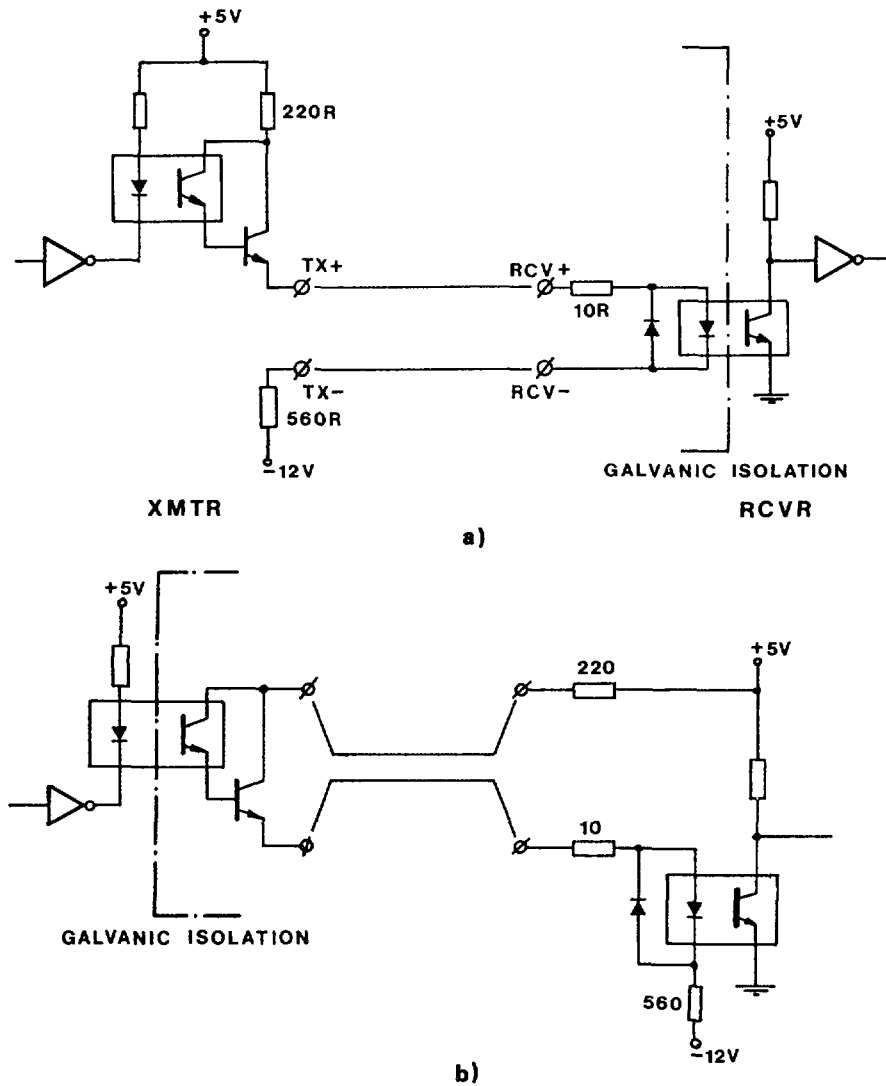
Fig. 7.25: 20ma - current loop interface
a) active transmitter - passive receiver
b) passive transmitter - active receiver

When the current loop was first used in teletypewriters, the loop was connected and broken by actual rotating switch contacts within the teletypewriter sending the data, and the 20mA signal drove a print magnet in the receiving teletypewriter. Today, most 20 mA current loops electronically simulate the opening switch and magnet arrangemnt.

A typical 20mA current loop inerface is shown in Fig. 7.24. Two modes of operation are possible; active transmitter-passive receiver and passive transmitter-active receiver depending on whether the tranmitter or the receiver provides for the current.

Optocouplers are used to assure galvanic isolation. Observe that only the device working in a "passive" condition effectively establishes the isolation.

Although it is quite normal to establish the loop current with two resistors of 220 and 560 ohm tied to +5V and -12V respectively, it is preferable to use current sources. In this way the loop current is made independent of receiver impedance and interconnecting cables resistance.

CHAPTER 8

TIME MEASUREMENT

8   TIME MEASUREMENT

Time is a continuous variable which can be digitized through
clocks. In nuclear electronics, time measurements are always related to
a time origin, assumed as a reference. Under most circumstances the time
interval between the origin, called START and a second instant called
STOP has to be measured.

In nuclear measurements time intervals to be measured can cover a
broad range: from thousands of seconds to less than a fraction of a
nanosecond. It becomes therefore evident that the accuracy of the timing
signals START and STOP must be related to the time interval to be
measured.

8.1   MACHINE TIME

The instants defined as START and STOP are not the real instants at
which the radiation interacts with the detectors; they are artificially
generated instants called machine times. The term "machine time"
emphasizes the fact that these are the instants at which some action is
taken by a circuit driven by the detector pulses and that these are
instants the experimenter has to refer to in order to define the time of
arrival of the event. To understand the meaning of machine time a
photomultiplier tube looking at a NaI(Tl) crystal, Fig. 8.1. The output
current of the phototube is integrated on its anode capacitance and the
resulting signal is presented to a threshold discriminator.

As we define machine time the instant at which the integrated
phototube output, $V_c(t)$, crosses the discriminator threshold V_T, the
time diagrams clearly show that between the instant at which the γ-ray
interacts with the scintillator and the instant at which the
discriminator threshold is crossed, several processes occur, like:

1.   conversion of the emitted light into photoelectrons.

2.   cascaded multiplication steps in the phototubes and generator of the
      output current pulse.

3.   integration of the current on C and charge accumulation until the
      threshold V_T is passed.

4.   appearance of the Signal and the comparator output.

As a result, there may be a considerable delay between the instant
at which interaction of the radiation with the detector occurs and the
instant at which the threshold discriminator senses the event, generating
the machine time t_u. This time shift would not affect the measurements
of time intervals if it were the same for both START and STOP pulses and
if the shift were stable in time.

Unfortunately this is not the case, and the machine time t_u defined
with respect to the true instant at which interaction between radiation and
detector takes place, is affected by inaccuracies.

These inaccuracies are basically of two types. One, referred to as
WALK is due to a systematic delay fluctuation in crossing a fixed threshold
level where the input signals have constant risetime, but different
amplitudes. The effect is explained in Fig. 8.2.
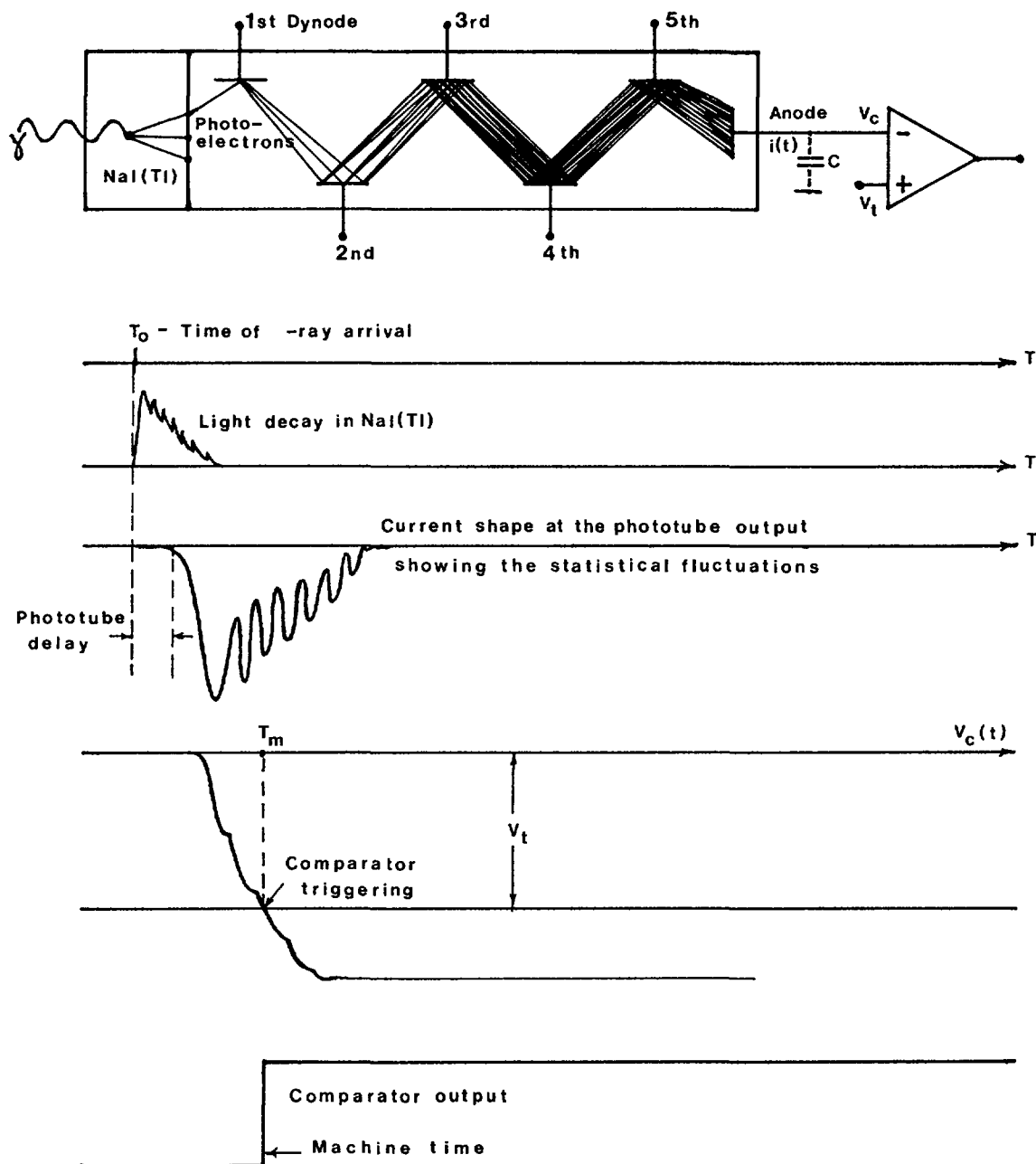
Fig. 8.1: Time sequence of electrical signals from a NaI(Te) detector

Here four signals with different amplitudes and equal risetimes are presented to an ideal discriminator with the threshold set at the value V_T. Signal 4 does not reach the threshold. Signals 1, 2, 3 trigger the timing discriminator but the instants at which the threshold is passed depend on the signal amplitude. The largest one, signal 1, gives the shortest triggering delay, while signal 3 gives the longest one. The best way of reducing the time WALK is based upon the zero crossing concept. Suppose that you shape the quasi-step signals of Fig. 8.2 into bipolar pulses, as can be done by using a double delay line amplifier. The resulting signals have the shape shown in Fig. 8.3.

THE ZERO CROSSING POINT IS INDEPENDENT OF THE PULSE AMPLITUDE. All the signals, provided that they have equal risetime cross at the same instant. The instant, as sensed by the so called "zero crossing trigger", provides a machine time. Such a machine time is delayed with respect to the beginning

170
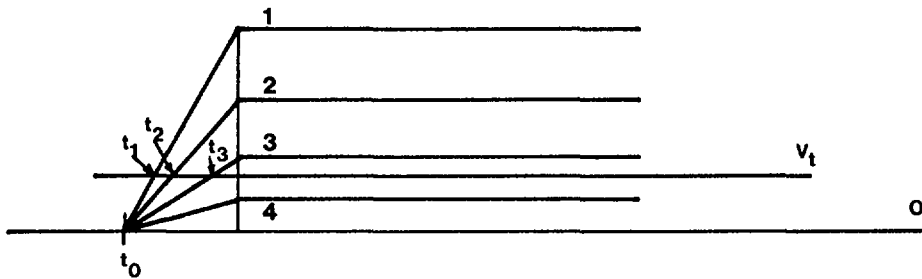
Fig. 8.2: Explanation of the WALK effect
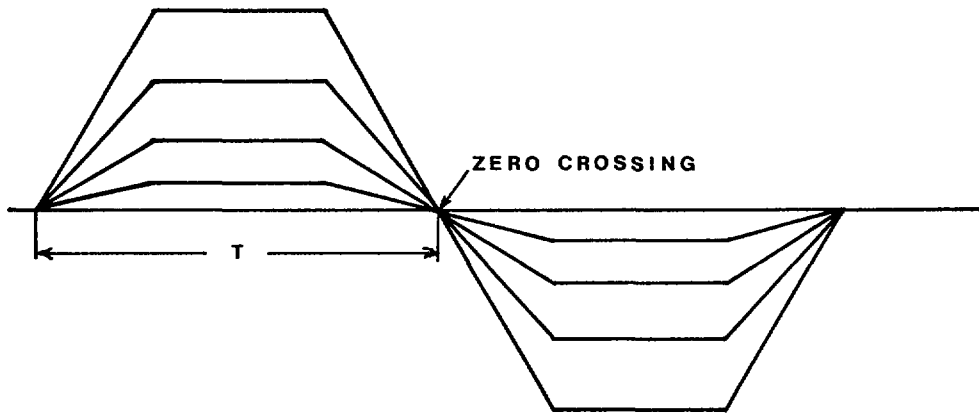
Fig. 8.3: Reducing the WALK effect by zero-crossing
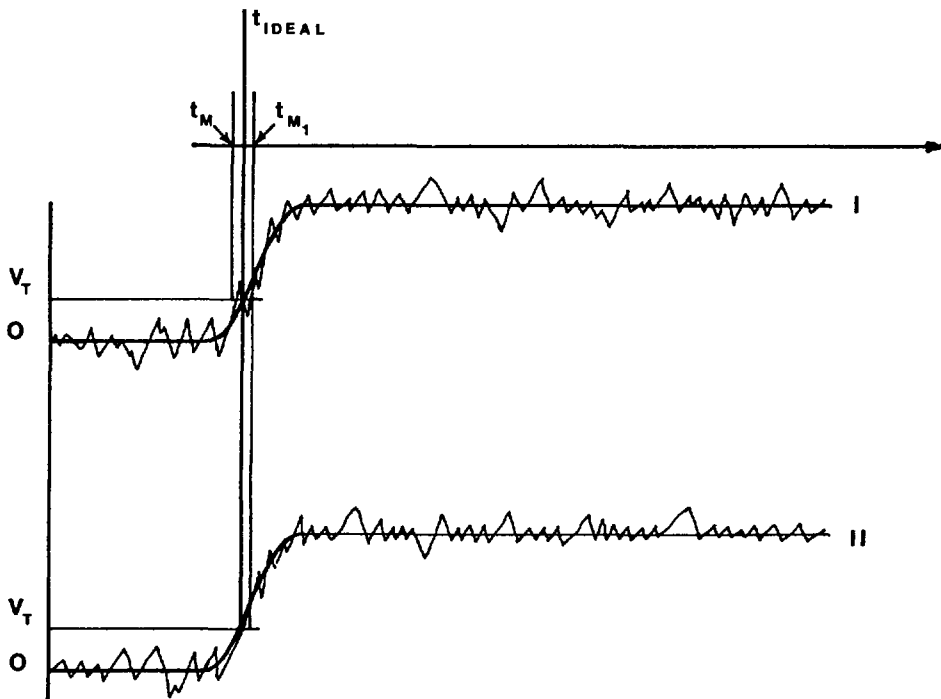
Fig. 8.4: Influence of noise on machine time signals

of the pulses at time T. A stable and reliable shift in the machine time with respect to the time instant of coert occurrence is unimportant, for it affects the start and stop events to be same extent, or at least to an extent which can be easily taken into account and corrected for.

A second source of inaccuracy in the machine time definition is called JITTER and is due to the noise superimposed on the signal. In order to

explain this kind of inaccuracy, which is of particular importance in timing problems with semiconductor detectors, reference will be made to Fig. 8.4. Two identical pulses represented by the heavier lines have been added to a noisy baseline. Such a situation may occur in two semiconductor detectors that deliver to the preamplifier two identical current signals; it will be assumed that these two signals are perfectly simultaneous. At the preamplifier output these two signals will not be identical any more, for the noise contributions of the two preamplifiers will be different. At the output of the preamplifiers the two clearn waveforms represented by heavy lines will not be observed any longer, for they are smeared by the noise generated in the preamplifiers. The described situation approaches that of semiconductor detectors, where, as a first approximation, the signal can be considered noiseless and the main contribution to the noise on the signal actually available is due to the noise sources inside the preamplifier.

The signals I and II of Fig. 8.4 are referred to their zero lines and equal thresholds V_T have been put on the diagrams. It can be seen that in the ideal noiseless case, the machine time would be represented by t_{IDEAL} and WOULD BE THE SAME FOR BOTH PULSES. This does not occur in the real case owing to the presence of noise. Fig. 8.4 shows that the noise in the proximity of the threshold for the pulse I can shift the crossing with respect to the ideal case. For the waveform II the noise around the threshold is such that crossing of the triggering threshold is delayed with respect to the ideal instant. Therefore, jitters affects the accuracy in the machine time in the way described by Fig. 8.4, introducing a scattering of values around the ideal one.

The above analysis of walk and jutter is oversimplified; no effect of real triggering discriminators is taken into account. Without going into details of circuit behaviour, it has to be borne in mind that for accurate timing, fast comparators, for instance MC 1650L, should be employed in the realization of timing discriminators of high performance.

One should remember that the STEEPER IS THE SLOPE ON THE LEADING EDGE OF THE SIGNAL AND THE SMALLER IS THE NOISE ASSOCIATED WITH THE SIGNAL, THE BETTER IS THE TIMING ACCURACY.

So far, two machine time definitions have been introduced, namely leading edge triggering and zero crossing triggering. They are compared again in Fig. 8.5.
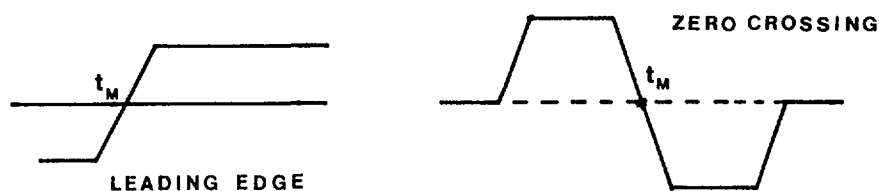
Fig. 8.5: Definition of machine time using different methods

The most widely employed trigger in modern experiments is called CONSTANT FRACTION. The idea which underlies the principle is to define the machine time as the instant at which A GIVEN FRACTION OF THE TOTAL CHARGE Q DELIVERED BY THE DETECTOR IS COLLECTED. This method, as can easily be seen, is walkfree. It also has some interesting features as far as reduction of jitter is concerned. To explain how the method can be realized, refer to the integrated pulses at the output of the detector and remember that the instantaneous value of the integrated signal is proportional to the charge collected up to the instant under consideration (see Fig. 8.6).

172
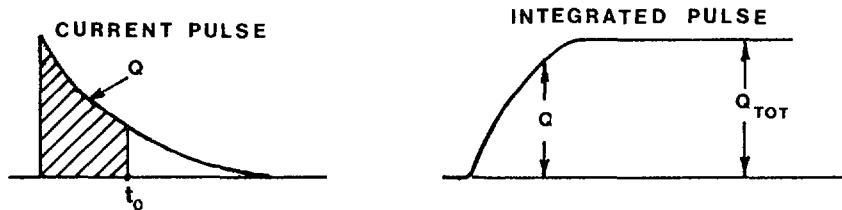
Fig. 8.6: Pulse at the detector output

In order to define the instant at which a given fraction of the total charge Q is collected, the amplitude Q_{TOT} has to be provided and then the actual pulse has to be compared with such an amplitude. The basic principle is shown in Fig. 8.7.
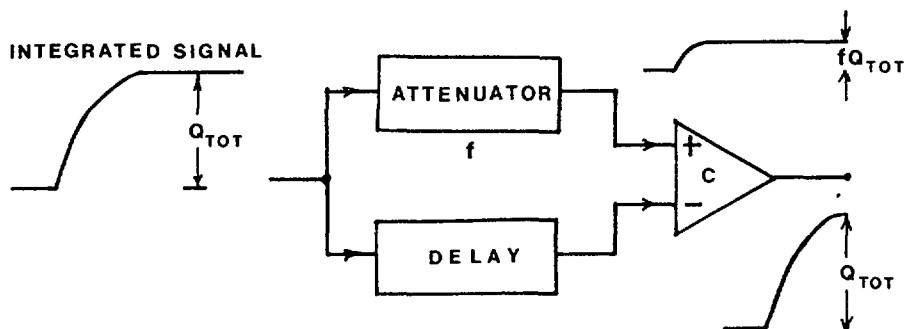
Fig. 8.7: Principle of constant fraction timing

The actual implementation of a constant fraction trigger is slightly different and is based upon the schematic shown in Fig. 8.8.

According to the diagram of Fig. 8.8, the zero crossing occurs when the unattenuated and delayed signal of final value Q_{TOT} equals the attenuated amplitude fQ_{TOT}.
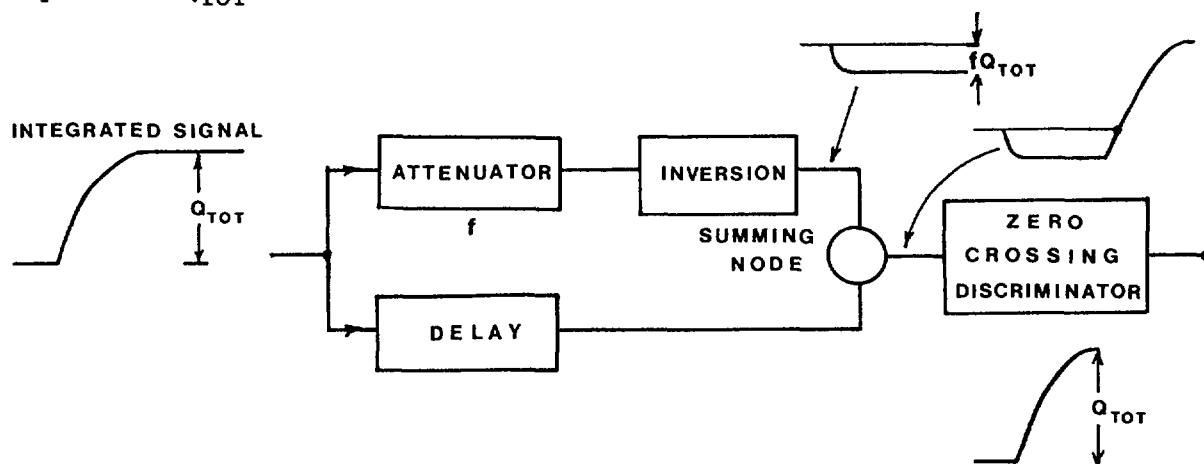
Fig. 8.8: Constant fraction trigger

8.2 CIRCUITS FOR TIME MEASUREMENTS

As already pointed out, the timing discriminator is a very important part in the time measurement problems.

A leading edge trigger is simply obtained from a comparator with the threshold set at the desired value V_T, as shown in Fig. 8.9.

173
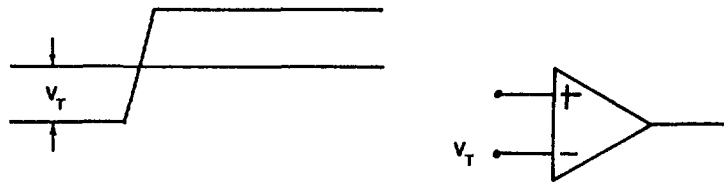
Fig. 8.9: Leading edge trigger

A zero crossing discriminator, of the type required to implement either true zero crossing operation or constant fraction triggering is obtained from a comparator with positive feedback, in order to achieve a hysteresis characteristics of the type shown in Fig. 8.10.
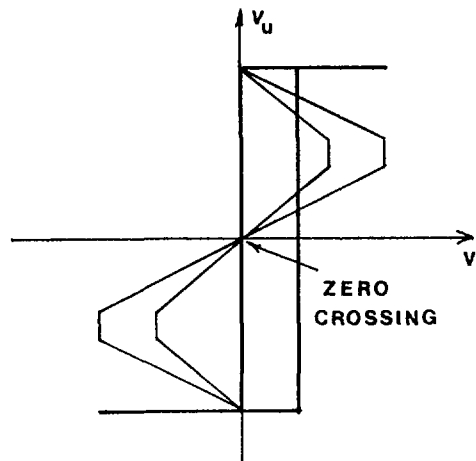
Fig. 8.10: Hysteresis obtained by a comperetor with positive feedbacks

So far, the problem of generating the machine time t_u has been considered. Now, it will be explained how this information can be utilized. We assume that detectors are used and that the trigger discriminators generate just two narrow spikes occurring at the machine times of the relevant detectors, Fig. 8.11.
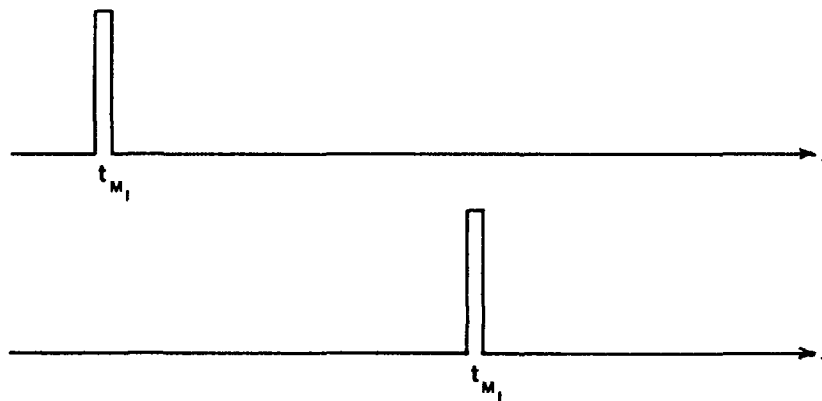
Fig. 8.11: Machine time signals

A first question which can be asked is whether the two pulses fall within a given time separation τ. Or, in other words, whether they are coincident in time within τ. Such a question can be answered by a coincidence experiment which is performed in the way shown in Fig. 8.12.

174
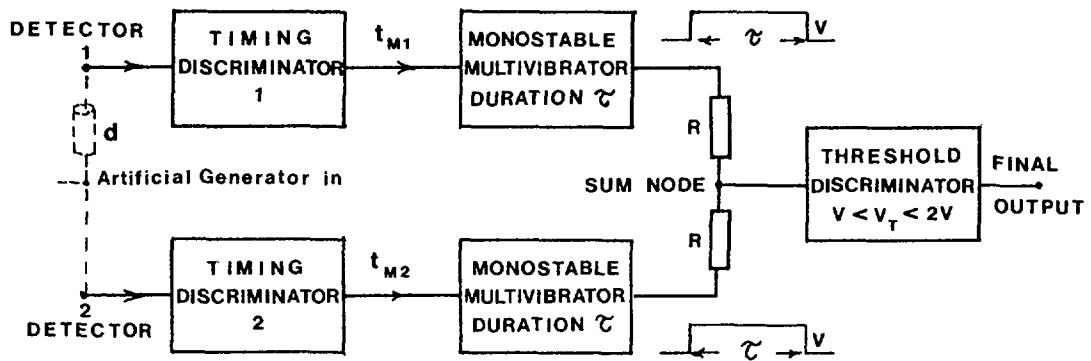
Fig. 8.12: Block diagram of a coincidence experiment

The meaning of the coincidence experiment is described by the diagrams in Fig. 8.13.

In case a) the time distance between t_{M1} and t_{M2} is shorter than τ. The outputs of the two monostables overlap and the threshold of the final discriminator, which is set between V and 2V is exceeded, with the result that a signal appears at the output. In case b) the time distance between t_{M2} and t_{M1} exceeds τ. The output pulses from the two monostables do not overlap and the threshold of the discriminator is not exceeded. No signal appears at the signal output. The experiment can be completed by disconnecting the detectors from the circuit and by sending to the inputs of the two timing discriminators the signal from an artifical pulser split along two lines, one with a delay of known value, one without delay. The signals at the final output have to be counted with a scaler for various values of the delay d. The delay has then to be put on the other input and the signal sent to the previous one without delay. The sign of the delay will be assumed to be positive, by arbitrary convention, when it is inserted on the input of the timing discriminator 1. The resulting curve will look like the one shown in Fig. 8.14.
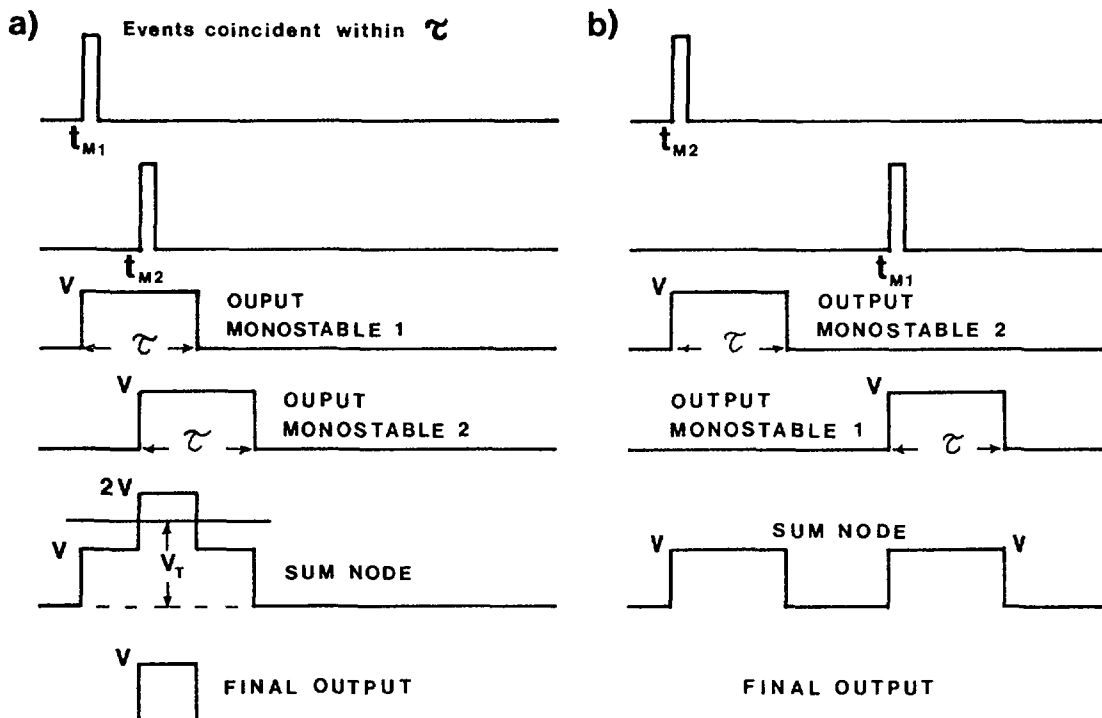
Fig. 8.13: Signals in a coincidence experiment

175
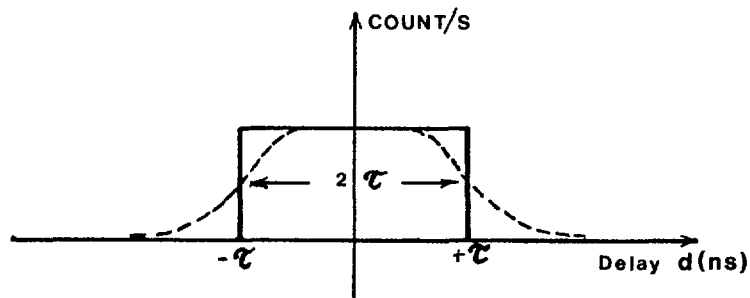
Fig. 8.14: Result of a coincidence experiment

Two really simultaneous events of largely different energies give rise to two signals like I and II in Fig. 8.15.

In the ideal case the counting be discontinued abruptly as soon as the delay is outside the $-\tau$, $+\tau$ interval. In the real case, the counting decrease more gently, according to the dotted curve. In either case, the full width at half maximum defines the resolving time of the coincidence which in the present case would be 2 τ.

Electronically, the resolving time of the coincidence is fixed by the width of the output pulses of monostable multivibrators. Attention must be paid to the values of walk and jitter on the related machine time. As an example of bad choice of resolving time, the case shown in Fig. 8.15 will be considered.
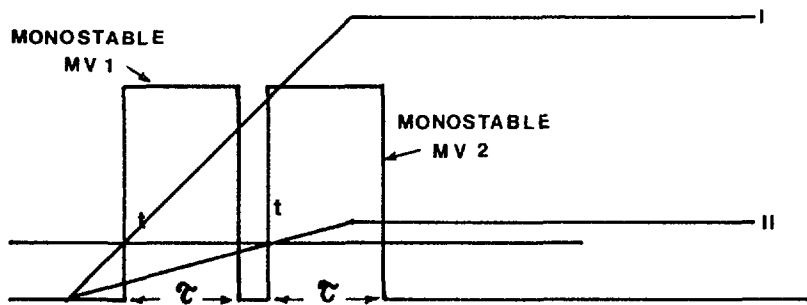
Fig. 8.15: Poor definition of resolving time
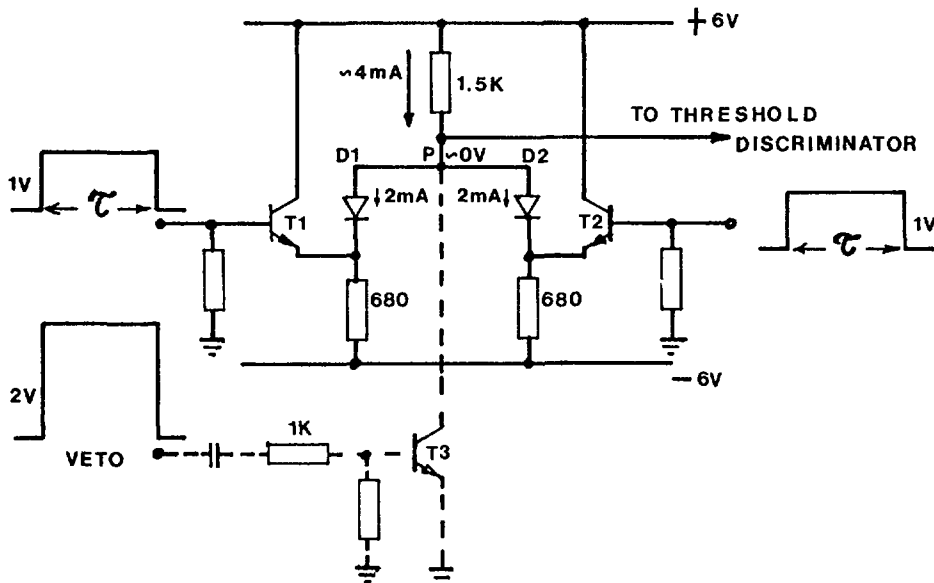
Fig. 8.16: Circuit with a high coincidence-to-single event ratio

Owing to the choice made of a machine time based upon leading edge triggering, large walk effects are present. The duration assumed for the two multivibrators is short compared to this walk. The multivibrator signals never overlap. Therefore, although the original events are really coincident, they will not be detected as coincident by the circuit employed. The coincidence circuit used in Fig. 8.12 is a simple resistive adder, which gives a 2:1 ratio between the coincidence and the signal event case. A circuit which has a higher coincident-to-single ratio is shown in Fig. 8.15. It will be assumed that transistors T_1 and T_2 are perfectly matched and equally matched are diodes D_1 and D_2; matching also exists between base-to-emitter voltage drops in the transistors and forward voltages in the diodes.

The anodes of D_1 and D_2 are approximately at OV and a current of about 4 mA flows across the 1.5 kohm resistor and then splits in equal parts between D_1 and D_2. This means that each diode works at a standing current of 2mA. Currents of about 7.8 mA flow through the 680 ohm resistor; two mA are provided by the diodes D_1 and D_2, and 5.8 mA are provided by the transistors T_1 and T_2. The multivibrator output signals are applied to the bases of T_1 and T_2. They must have a sufficient amplitude to be able, when separately applied, to turn off the corresponding diodes. It will be assumed that the multivibrator output signals have an amplitude of at least 1 V. If a signal of such an amplitude is applied to the base of T_1, while no signal is applied to T_2, diode D_1 will be turned off. There will be a current increase in D_2, but the voltage at point P will not appreciably change, for the current vs voltage characteristics of the diode is a logarithmic one. Consequently, a current of about 4 mA will flow through D_2 and the emitter current of T_2 will be accordingly reduced; owing to the relatively large current originally flowing through T_2, T_2 will keep conducting. A single signal applied to the coincidence circuit input will change voltage at P for a few mV as a result of current redistribution and current change in T_2. If two equal signals are simultaneously applied to the bases of T_1 and T_2, the cathodes of both diodes will rise to a voltage equal to the driving signals amplitude, and P will raise approximately to the same value.

Therefore, a large ratio is achieved at point P between the amplitude of the signal induced by a single base command and the signal induced by a simultaneous command on both bases. This feature makes the operation of the threshold discriminator, which has to distinguish between coincidence events and single events, less critical. Such a discriminator may be not even required and the coincidence-versus-single event separation can be made, for instance, by the input triggering circuit of the counter employed to measure the event rate at the output of the coincidence circuit.

The circuit shown in Fig. 8.15 has an additional facility which can be useful in some circumstances. Such a feature is made possible by the dotted transistor T_3. In the normal condition T_3 is off and the operation of the circuit proceeds as previously explained. In some cases, it may happen that a coincidence between two signals has to be suppressed if some unwanted condition takes place. For such reason, several practical coincidence circuits have the built-in VETO function, which in the case of the circuit of Fig. 8.16 is made possible by transistor T_3. If a coincidence takes place, but in the meantime the VETO signal is applied to the base of T_3, T_3 will go into saturation, D_1 and D_2 are turned off and point P will be clamped to ground. Obviously, the principle of the coincidence circuit can be easily extended to a number of inputs larger than 2 and, as a matter of fact, threefold, fourfold coincidences are also commercially available.

It has now to be explained how a time interval between a START pulse and a STOP pulse can be measured and how spectra of time intervals can be accumulated.

Time interval measurements can be done with the direct time digitizer shown in Fig. 8.17.
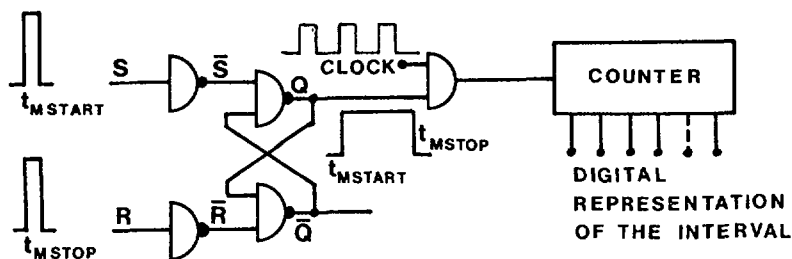

Fig. 8.17: Time digitizer

Assume that the flip-flop is reset before the measurement starts. As a result of this reset, the Q output of the flip-flop is at logic level HIGH, the Q output at logic level LOW and the clock transmission through the AND circuit is inhibited.

The START pulse sets the flip-flop with Q output at logic level HIGH. This enables the reference clock through the AND circuit to the counter. The digital number accumulated in the counter is proportional to the time measured from the START signal. When the STOP signal is received, the flip-flop is reset. The Q output returns to logic LOW, counting is disabled and the value of the $t_{mSTOP} - t_{mSTART}$ time interval is given by the number N stored in the counter multiplied by the period of the reference clock, τ_{OSC}.

$$t_{mSTOP} - t_{mSTART} = N \cdot \tau_{OSC}.$$

The oscillator period must be stable and accurately known. It determines the measurement accuracy. The digital number at the counter output can be directly used to address the RAM memory of a multichannel analyzer and therefore a time-interval spectrum will be obtained. It is worth pointing out that direct time interval digitizing is feasible only when the resolution required is not too high and the time interval to be measured is not too short. To clarify this pont it has to be borne in mind that one channel in the time measurement based on direct digitizing corresponds to one clock period. The highest clock frequency that can be employed in these applications is around 500 MHz. This means that the following inequality relates clock frequency, time interval and resolution or desired number of channels N_c:

$$\frac{t_{STOP} - t_{START}}{N_c} > \frac{1}{500 \text{ MHz}} = 2ns$$

So, for instance, measuring a 1 μs interval with a 200 channel resolution is feasible while measuring a 10 μs interval with a resolution of 10.000 channels is hardly possible

When the previous approach fails, the principle of analog time interval-to amplitude conversion can be successfully employed. It consists of switching into a capacitor the current supplied by an accurate reference source as soon as the START PULSE comes in (see Fig. 8.18). The voltage across the capacitor will then begin rising linearly. When the STOP PULSE arrives, the current source is turned off. The final voltage stored on the capacitor is proportional to the time interval under measurement.

178
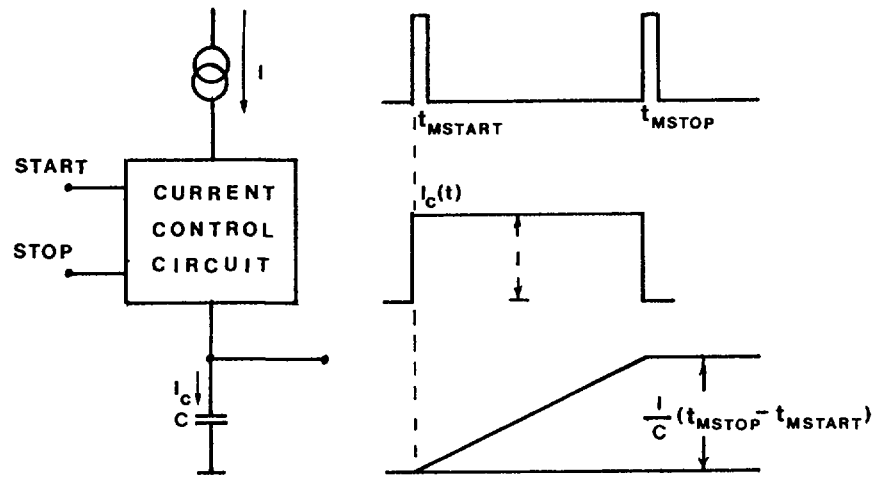
Fig. 8.18: Principle of analog time-to-amplitude conversion

Once the final voltage is reached on C, this capacitor can be discharged at a current much lower than I. Let t' be the instant at which the discharging current is applied and let t" be the instant at which the voltage across C reaches zero. From Fig. 8.19, it can be seem that time expansion is achieved and the expansion factor is determined by the ratio between the two currents I and I_R. The new time interval, t" - t' is proportional to the original one, $t_{M, STOP} - t_{M, START}$, but can be considerably longer and therefore suitable to be digitized with a clock of lower frequency.
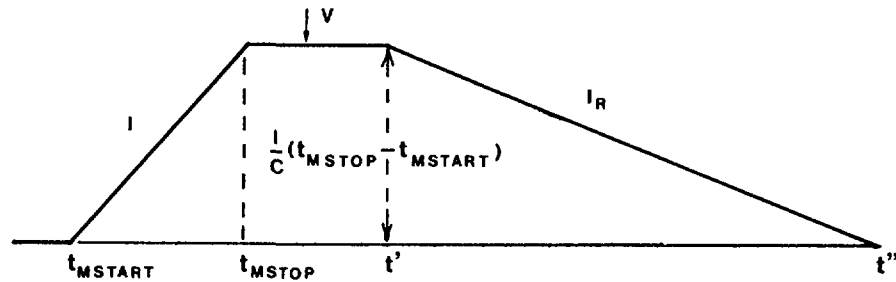
Fig. 8.19: Time expansion for time-to-digital conversion

$$V_{final} = \frac{I}{C} (t_{M, STOP} - t_{M, START}), \text{ but also}$$

$$V_{final} = \frac{I_R}{C} (t" - t') \quad \text{whence}$$

$$t" - t' = \frac{I}{I_R} \cdot (t_{M, STOP} - t_{START})$$

t"-t' is the expansion factor.

Direct time digitizers and time interval-to amplitude converters are basic building blocks in modern experiments requiring time measurements. They are made commercially available, and are produced by several qualified manufacturers.

CHAPTER 9

POWER SUPPLIES

9   POWER SUPPLIES

This chapter presents a discussion on the power supplies most
frequently used in nuclear electronics. It is a known observation that
most of the failures in nuclear instruments occur in the power supplies.
Therefore, a careful design of the power supply has a particular
importance, especially in developing countries where the electricity
power is neither reliable nor stable.

To be able to connect signals to the different (potential free)
nuclear equipments, they must be insulated from the mains. Nowadays,
this can only be achieved by batteries (cordless equipment) or by power
supplies with transformers. If the line separation is used, a
transformer must be used between the mains and the internal power supply.

9.1   POWER SUPPLIES USED IN NUCLEAR INSTRUMENTATION

Different instrumentation standards are used in nuclear
electronics. The most widely recognized is the Nuclear Instrumentation
Module (NIM) standard. The following voltages are used to supply the
modules with appropriate power:

    + 24 V/ 1 A
    - 24 V/ 1 A
    + 12 V/ 2 A
    - 12 V/ 2 A
    + 6 V/ 5 A
    - 6 V/ 5 A

For operation of a radiation detector, high voltage supplies are
required. The following values are usually selected:

For scintillation detectors (photomultipliers):
    0 - 3000 V/ 10 mA   (very stable)

For semiconductor detectors:
    0 - 5000 V/ 30 μA   (stable, low noise and very low ripple)

9.2   PERFORMANCE DEFINITIONS

A power supply must provide stable and ripple-free DC output voltage
independent of line and load variations. It has also to be stable with
time and with temperature variations. Let's remember some definitions
and typical values.

-    Line and load regulation: output voltage variation for
     combined line and load changes within the operating range.

-    Ripple: peak to peak AC component of output voltage related to
     DC voltage.

-    Temperature instability: percentage of output variation over
     the temperature range.

-    Long term drift: instability of output voltage for long periods
     of time.

Typical values

	Line and load regulation	Ripple	Temperature Instability	Long term Drift
NIM - power supply	± 0.5%	3 mV	0.01% per C	± 0.3% over 6 months
Photomultiplier power supply	± 0.001%	10 mV or 0.0003%	0.01% per C	± 0.015% over several hours
Semiconductor detector power supply	± 0.005%	10mV or 0.002%	± 0.02% per C	± 0.01% over several hours

9.3 RECTIFYING

To derive a DC supply from an AC, one of the following circuits has to be used:
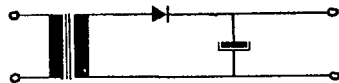
Fig. 9.1: Half-wave rectifier

The half-wave rectifier is illustrated in Fig. 9.1 and is used only for small load current. The advantages are: such a supply is very simple and low cost. The disadvantages are:

- high ripple, high current spikes during capacitor charging interval
- DC current flowing in the secondary winding of the transformer; consequently the core may saturate

This configuration may be found in low power, high voltage applications.

a)

b)
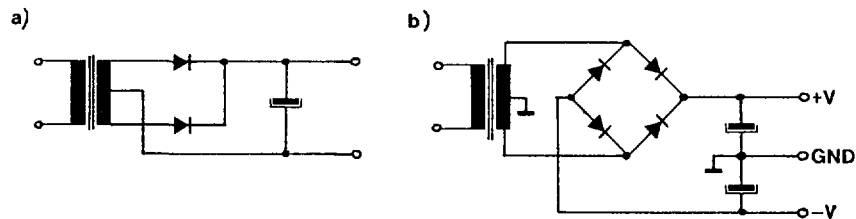
Fig. 9.2: Full-wave center tap

The type "a" of rectifier is used in supplies where a high current (10A and higher) is needed. The cost of a transformer with center tap is negligible compared to the higher price of such diodes. Nowadays, very often this type is used in small supplies also, as the same type of transformers can be used for both positive and negative supply voltages, thus allowing for a mass production.

The type "b" of rectifying is used in a double voltage supply (positive and negative voltage).
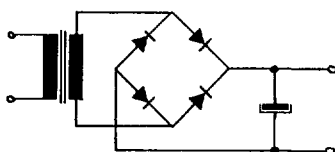
Fig. 9.3: Full-wave bridge

A full-wave bridge rectifier is used in medium power applications. Two extra diodes are the price for less windings on secondary transformer side compared with full-wave center tap configuration.

9.4 FILTER CAPACITOR

The rectified voltage will show waveform as is indicated in Fig. 9.4 and could be considered as a DC voltage plus a superimposed AC voltage shaped.

The ratio of the peak-to-peak voltage of the AC content divided by the peak voltage V_p will be the percentage ripple (see Fig. 9.4).
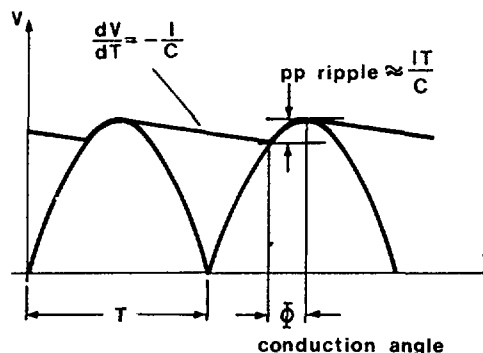
Fig. 9.4: Definition of ripple

From the drawing it is possible to derive:

$$C = I.T/V_{pp}$$

$$C = k.I \ (mA) \ / \ V_{pp} \ (V)$$

with k = 10 for 50 HZ and k = 8.3 for 60 HZ.

The requirement of such a capacitor is that every regulator needs a higher DC voltage component at the input compared to the output (minimum between 1.5V and 3V depending on the regulator circuit).

It can be seen that by using a large capacitor to improve the filtering for a given load, R_L decreases the conduction period for the diode. For a specified average load current, the diode current becomes more peaked. This may impose serious duty conditions on the rectifying diode since the average current may be well within the current rating of the diode and yet the peak current may be excessive. Similarly, a large peak current may saturate the transformer, which is well-rated for the average load currents. This demands higher ratings than the average load for both the diodes and the transformer.

This capacitor, as you can see in Fig. 9.4, has to handle an AC component (charge and discharge, which means losses and these losses are creating heat). Therefore, it is necessary to look into the data sheets concerning this maximum rating. Due to this capacitive load, the conductive angle (while current is flowing from the transformer via diodes to the capacitor) is very small (roughly 40 to 60 degrees) and all components like transformers and diodes must be enlarged (see Fig. 9.36). As a result, the transformer will become saturated and the diodes overloaded due to the recurrency (periodically) peak current (see data sheets of rectifier diodes).

The relation between transformer current, DC current and the different loads see Fig. 9.36.

9.5 TRANSFORMER

The transformer has to provide a minimum voltage to the regulator input despite the low input voltage, high ripple and heavy load conditions.

Although Vp, peak voltage, will apparently be the available voltage at the regulator input, there will be several reductions:

- 0.8 due to rectifier drop (1.6V for bridge circuit)
- $t.V_p$ due to transformer drop, typically t = 0.05
- $1.V_p$ due to low mains
- $r.V_p$ due to the ripple

From these considerations, the following formula is derived:

$$V_{rms} = 0.707 \; \frac{V_{DC} + V_{REG}}{1 - (t + 1 + r)} + 0.8 \; n$$

n = 1 in full wave rectifier
n = 2 in bridge rectifier

V_{DC} is the actual DC regulated voltage
V_{REG} is the minimum drop across the regulator

The RMS current for the transformer has to be calculated as follows:

Rectifier type	I_{RMS}
Full wave centre tap	1.2 x DC current
Bridge	1.8 x DC current

Once both V_{RMS} have been determined, the transformer can be calculated.

9.6 REGULATION

In regulated power supplies for small current the circuit, shown in Fig. 9.5, may be used.
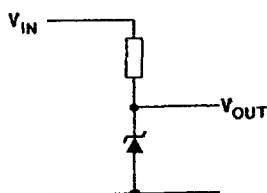
Fig. 9.5: Regulation by zener diodes

This circuit is only useful for small currents due to the maximum current I_{Dz}. If there is no load, the full current is taken by the zener I_{Dz} diode; which means: don't over-reach the maximum power dissipation. The maximum current can be calculated by $P = U_z \cdot I_{max}$ where U_z is the nominal voltage of the zener diode (reverse voltage) and P is the power dissipation of the zener diode.

At the maximum load there must still be a current flowing through the zener diode ~ 1 - 3mA (depends on the zener diode) to get a suitable regulation.
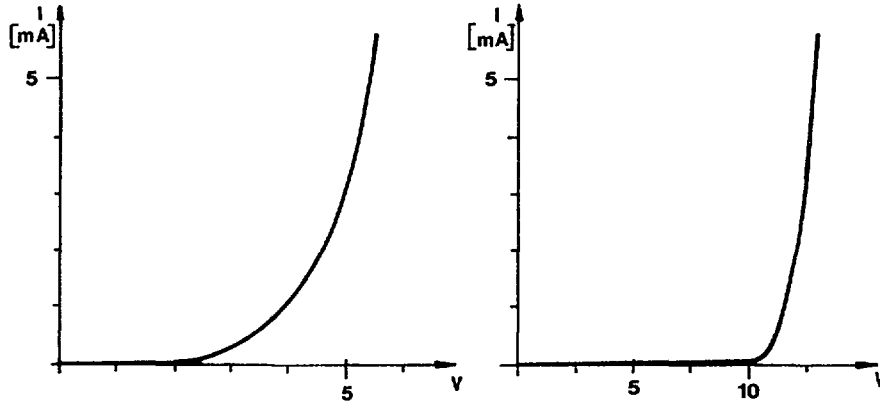
Fig. 9.6:   V-I characteristics of a zener diode
a) zener diodes below 5V
b) above 12V zener diodes

As you see in Fig. 9.6, it is necessary to always have a current flowing through the zener diode above the knee voltage.

For low current, also the base to emitter junction in the reverse mode can be used, as shown in Fig. 9.7.
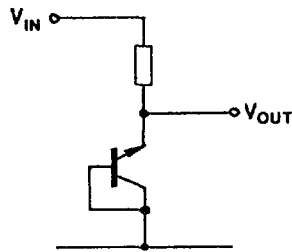
Fig. 9.7:   Stabilization by a reversed base-to-emitter junction transistor

This configuration has an advantage because of the good V-I characteristics of the device in the low current area.   The knee voltage is very sharp compared to a zener diode.
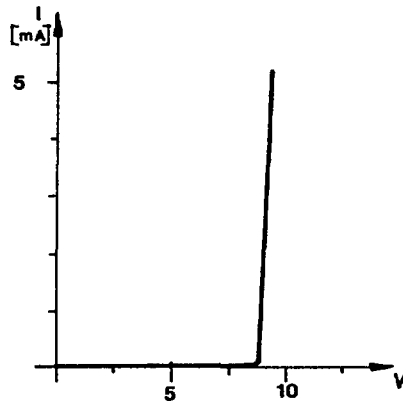
Fig. 9.8:   V-I characteristic of a base-emitter junction in reverse mode

Reference sources are now on the market. These devices demonstrate good properties in stabilization circuits.

Schematic Diagram
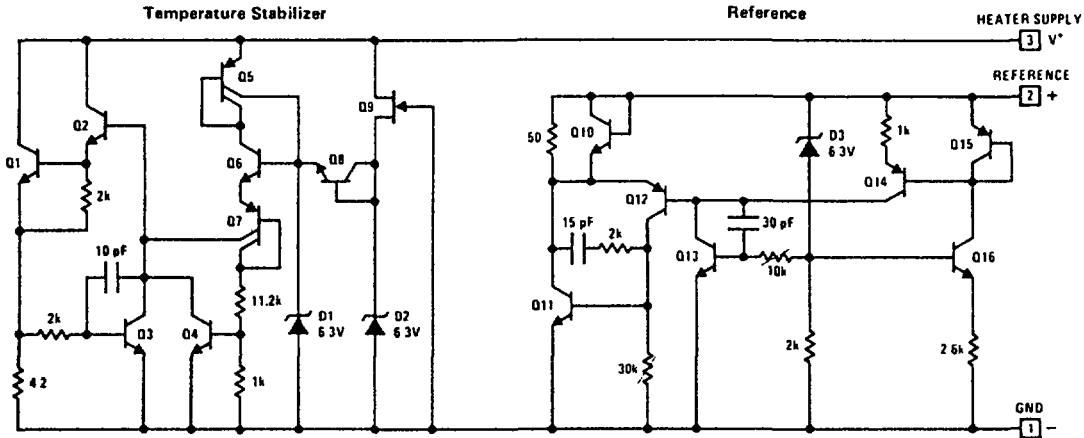
Functional Block Diagram
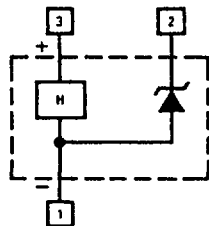
Typical Applications
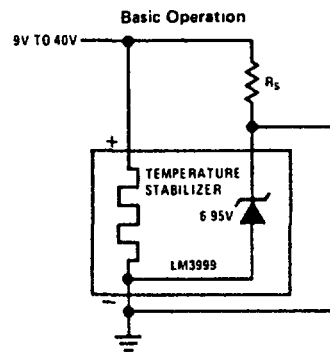
Fig. 9.9: Reference source with temperature stabilizer (LM3999)

As an example, consider LM3999 (Fig. 9.9). This device has the following characteristics:

- guaranteed 0'0005%/°C temperature coefficient
- low dynamic -0.5
- breakdown voltage about 5%
- sharp breakdown at 400μA
- wide operating range 500μA to 10mA
- wide supply range for temp. stabilizer (between 9 and 36V)
- low power for stablilization (~400mW at 25 °C)
- long term stability ~20$_{ppm}$

NOTE: Never use the stabilizer below 9V due to the high switch on current; the heater could be damaged.

For higher current the following circuit has to be used.
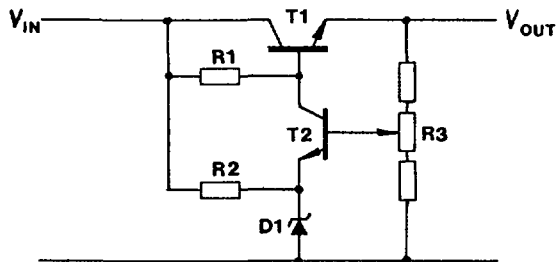
Fig. 9.10: A simple regulated power supply

Such a simple linear regulated power supply consists of several parts:

a) pass transistor T1 with bias resistor R1
b) error amplifier A1 (T2) and feedback resistor R3
c) reference source D1 with bias resistor R2

The pass transistor has to handle the output current. This means that the losses (if it is not a switching mode power supply) can be calculated by following formula $(V_{in} - V_{out}) \cdot I_{out}$. Therefore you need heat-sinks for the calculated power dissipation.

The error amplifier is necessary to compare the output voltage with the reference voltage and any change at the output must be regulated back. Therefore, a high open loop gain of such a system is needed to get good load and line regulation. Fig. 9.11 shows such a circuit diagram.
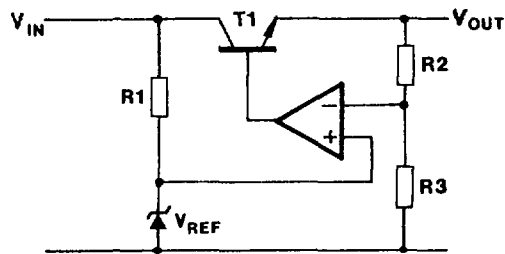
Fig. 9.11: Regulated power supply with operational amplifier and reference source

Fig. 9.11 shows a regulated power supply with an operational amplifier in the feedback loop. Such an operational amplifier has a high open loop gain and good temperature stability; therefore, it is able to get good load and line regulation.

This circuit has one big disadvantage: if there is at the output a short circuit, the pass transistor could be easily destroyed due to the high current. Therefore, a current limiter must be added. Such a circuit is shown in Fig. 9.12. It has two small disadvantages:

1. Due to the maximum power dissipation in the short circuit mode, the pass transistor could also be damaged.

2. An auxillary power supply is needed to supply the operational amplifier.
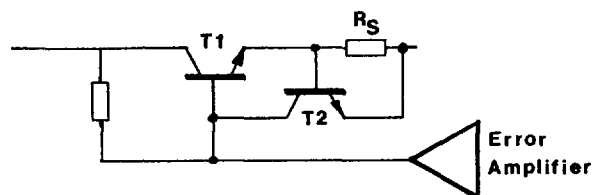
Fig. 9.12: Regulated power supply with current limitation

The current limitation can be calculated by $I_{max} = 0,6V/R_S$ and is shown in Fig. 9.13.

189
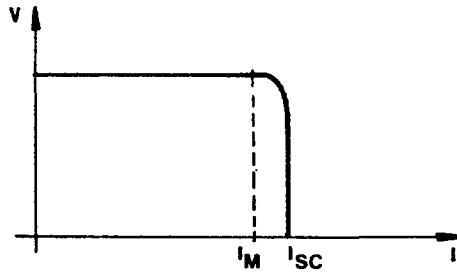
Fig. 9.13: V-I characteristics of current limiter

The current limitation in modern equipment is frequently made using operational amplifiers as shown in Fig. 9.14.
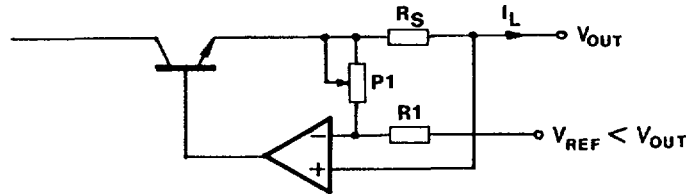
Fig. 9.14: Constant current limitation by operational 'amplifiers

The current limitation can be calculated by:

$$\frac{I_L \times R_S}{P_1} = I = \frac{V_{REF}}{R_1}$$

$$I_L = \frac{V_{REF} \cdot P_1}{R_1 \cdot R_S}$$

Again a linear variation of the regulated output is obtained, as P_1 is varied from zero to maximum. A combination of both operation modes is achieved using the configuration shown in Fig. 9.15.
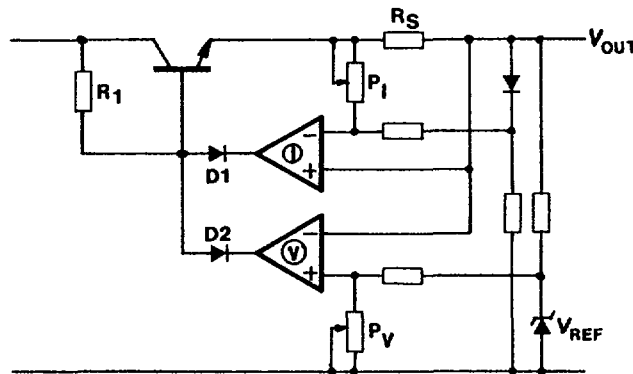
Fig. 9.15: Improved constant current constant voltage power supply

In this case two diodes, D1 and D2, and one resistor, R_1, are added forming an analog or gate. The operational amplifier whose output is less positive is controlling the loop.

The V-I characteristic of the power supply is shown in Fig. 9.16.

190
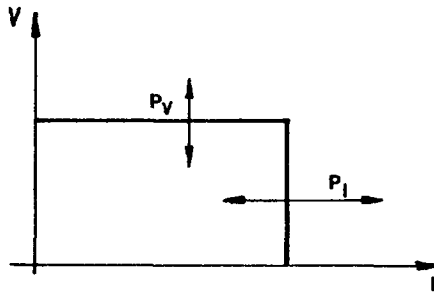
Fig. 9.16:   V-I characteristics of operational amplifier power supply

The first disadvantage can be eliminated by adding two resistors to get a so-called foldback current limiting, but the auxillary power supply for the error amplifier is still required.
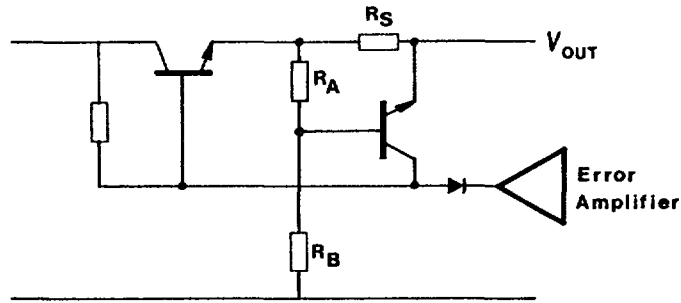
Fig. 9.17:   Regulated power supply with foldback current limiting
characteristics

The circuit starts limiting at the moment when T2 starts conducting.

$$V_{R_S} \quad kV_o \ + \ 0.5V$$

$$I_M \ \cdot \ R_S \quad kV_o \ + \ 0.5V \qquad \qquad (Eq. \ 9.1)$$

At short circuit

$$I_{SC} \ \cdot \ R_S \quad 0.7V \qquad \qquad (Eq. \ 9.2)$$

V_{RS} is voltage drop across R_S, I_M is maximum current,
I_{SC} is short circuit current, k is factor of voltage divider
R_2 and R_3, and V_o is output voltage.

Note that according to Eq. 9.1 and Eq. 9.2

$$I_M \ > \ I_{SC}$$
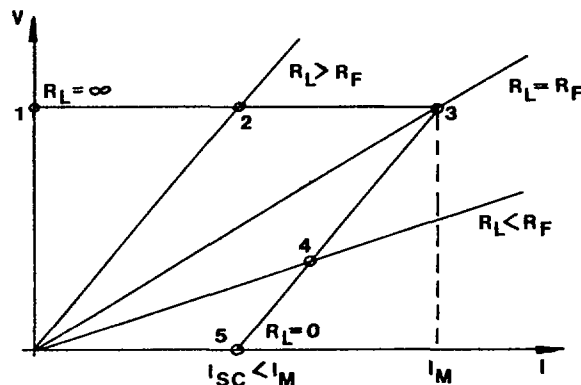
Fig. 9.18:   V-I characteristics of foldback current limiter

The criterion to establish I_{SC} is that the pass transistor power dissipation at short circuit be equal (or less) than the dissipation at maximum current I_M. If we define the dissipated power as:

$$P_{T1} = (Vin - Vo)\ I_M \quad \text{at maximum current, and}$$

$$P_{T1} \quad Vin \cdot I_{SC} \quad \text{at short circuit,}$$

we get the equation

$$\frac{I_M}{I_{SC}} = \frac{V_{in}}{V_{in} - V_o} > 1$$

Calculating I_{SC} and combining with Eq. 9.2 and Eq. 9.1, we obtain

$$k = \frac{I_M R_S - 0.5}{Vo} \qquad \text{(Eq. 9.3)}$$

In practice it is better to establish I_{SC} as small as possible with respect to I_M as P_{T1} will have a maximum between points (2) and (4) (see V-I characteristics, Fig. 9.18).

The limit to get low I_{SC} is given by the value of R_S and its voltage drop and power dissipation.

Advantages of a foldback current limiter are:
- unlimited short circuit operation
- low dissipation of T1
- automatic recovery from short circuit

Its disadvantages are:
- as R_S has a relatively large value, there is an increase in R_{OUT}
- relatively high power dissipation in R_S

Nowadays, for fixed voltages three-terminal regulators are available which have big advantages as well as disadvantages.

The advantages are:
- very small size,
- only few components required for the circuitry,
- they are fully short circuit proof (foldback characteristic with thermal foldback which means if, due to the current, the case temperature increases, the current goes down),
- low costs.

The disadvantages are:
- certain limitation in load and line regulation,
- input voltage (maximum 37V or 60V, depending on the regulator type),
- current limitation not adjustable.
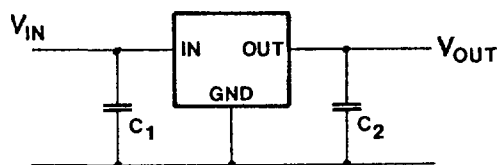
Fig. 9.19: Connection of a three-terminal voltage regulator

Fom application hints of such a regulator, very often C_1 and C_2 are requested to be situated nearby the regulator (Fig. 9.19).

Due to the high open loop gain of such a regulator it might happen without this capacitance near the terminals that it would oscillate. The values of C_1 and C_2 in the range of $0,1\mu F$ (ceramic capacitors) C_1 not always necessary if filter capacitator is not further than 10cm cable length away from regulator. These are linear regulators and therefore they must be mounted on heatsinks due to the maximum power dissipation.

9.7 HIGH VOLTAGE POWER SUPPLIES

There are mainly two types of HVPS in the nuclear field: those for very low current (up to 100 μA) used for biasing semiconductor detectors and gas-filled detectors, and those able to deliver higher currents needed when working with photomultiplier tubes.

In the first case the HVPS consists of an oscillator step-up transformer and rectifier circuit (see Fig. 9.20).
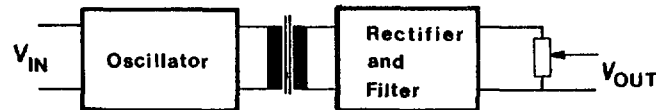
Fig. 9.20: Step-up tranformer and rectifier

In fact, this is a typical DC-DC converter.

The variable voltage is obtained by means of a potentiometric divider at the output. The use of high frequency leads to a smaller transformer and filter capacitors.

For higher current, a feedback loop is necessary to stabilize the output voltage (see Fig. 9.21).
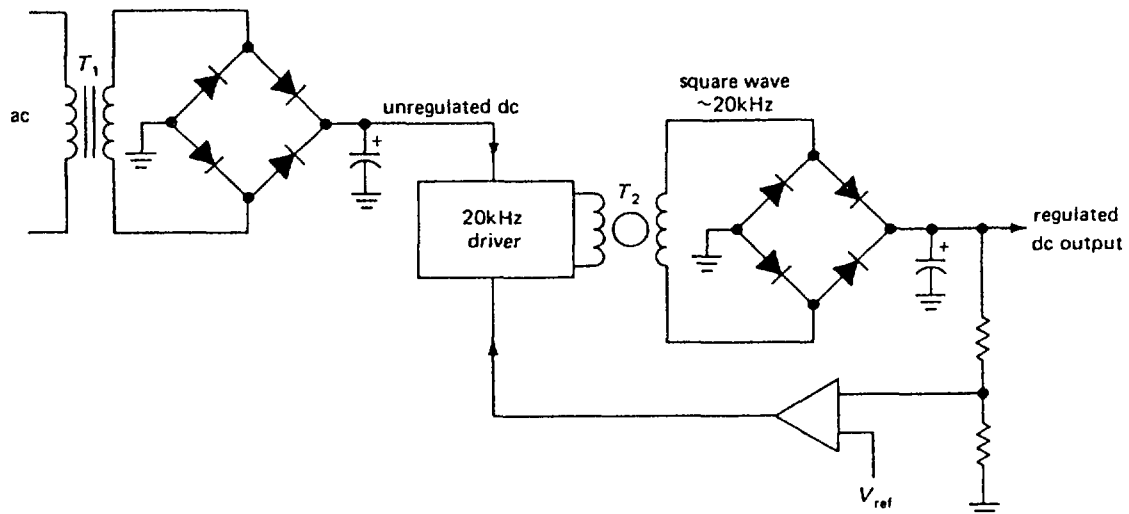
Fig. 9.21: High current high voltage power supply

The primary voltage at the step-up transformer is adjusted according to the required output voltage.

Also, voltage multipliers are very often used because of the insulation problem of such transformers. The wire insulation and the layer between the pot core (ferrite core) which should be grounded cannot handle more than 200V. Therefore, many insulation layers are needed for a 1000V transformer. More space is required for isolation than for windings, even for such thin wires.

These 1000V are more or less the standard values of such HV-transformers and afterwards the voltage multipliers are switched.

9.7.1 Voltage Multiplier

A typical configuration found in high voltage power supplies for detector biasing is shown in Fig. 9.22.

_____ doubler _____/   _____ tripler _____/   _____ quadrupler _____/
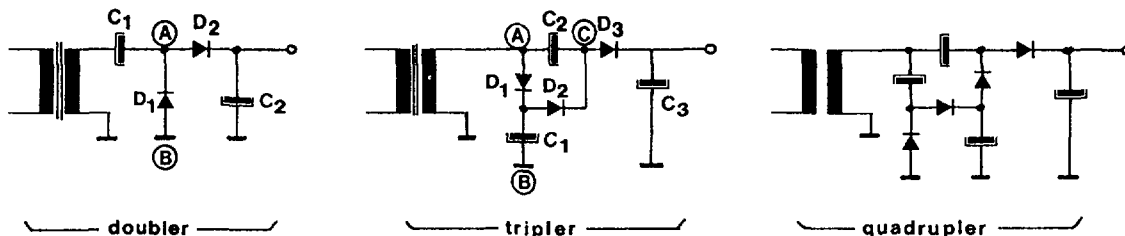
Fig. 9.22: Voltage multipliers

A voltage multiplier can be understood as a series of half-wave rectifiers connected in cascade. Observe that D_1 charges C_1 to the peak voltage V_p. The wave across D_1 is the result of adding up the AC between V_s plus the DC voltage across C_1 (V_p). As a result, between B and A there is a voltage presenting a peak value $2V_p$. This wave is rectified through D_2; and C_2 is then charged to $2V_p$. The same analysis determines that between B and C there is a voltage with a peak voltage $2V_p$, which rectified through D_3, charges C_3 to $3V_p$.

Of course several stages can be added up to increase the final DC voltage at the expense of increasing the ripple and self regulation.

In nuclear instrumentation manuals (5kV bias supply for solid state detectors), we frequently find a circuit diagram as shown in Fig. 9.23.
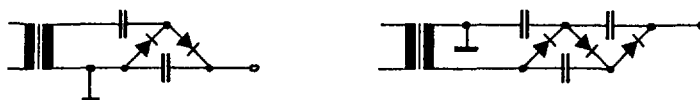

Fig. 9.23: Voltage multiplier as shown in many nuclear instrumentation service manuals

9.8 APPLICATION HINTS FOR POWER SUPPLIES

If a very high current is needed, this can be achieved by switching pass transistors in parallel (see Fig. 9.24). This is not so easy because of the slight difference in the base to emitter junction of the used pass transistors. All the bases see the same voltage. If one of these transistors would have a better h_{Fe} it would take over the whole current and it would burn. To avoid this problem, a resistor is to be inserted, in series with the emitter. If there is any assymetry in the characteristic of the transistor (assume an extremely high h_{Fe} of one

194

of the transistor), a high voltage drop on the resistor would appear; in this moment the base to emitter voltage decreases which means the collector current decreases and suddenly the voltage-drop across the resistor decrease. Then the reverse situation takes place until a balance is reached (this is so called negative feedback). The value of such resistors is in the range from 0·1 Ω to 2 Ω and depends on the current and the transistors used.
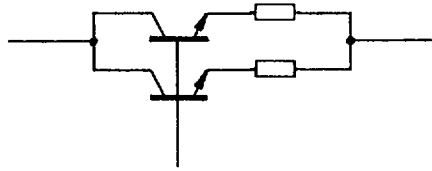

Fig. 9.24: Pass transistors in parallel mode

Note: if it is necessary to switch in parallel diodes, transistors or thyristors they have to have in series a resistor to balance the current in the different path (Fig. 9.24).

In some older equipment, cascaded pass transistors are used. The reason for this application was that old-generation transistors are not able to handle a voltage above 100V (V_{CE}). Therefore, it was necessary to cascade such transistors. To be sure that every transistor sees the same value of voltage, a resistor cascade with resistors bypassed by capacitors, was connected to the bases, as shown in Fig. 9.25.
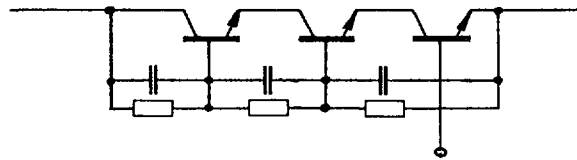
Fig. 9.25: Control of power dissipation of high voltage

If the current through the divider is greater than the base currents, the total voltage will divide equally across each transistor. This method is indicated when a relatively high voltage is being controlled.

The use of dissipating resistors is only necessary when the maximum power dissipation is higher than the P_{Tot} of the used pass transistor.
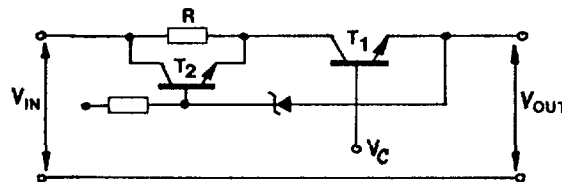

Fig. 9.26: Power dissipation on a resistor

If V_{OUT} is low (or zero in short circuit condition, Fig. 9.27) it is not necessary to have the whole input voltage applied across it. Only a small voltage is necessary to regulate T1. Therefore, a drop of voltage across R is permitted.

For R = Vin/I_{max} (unless the current is high), base emitter junction of T_2 is reverse biased and T_2 is disconnected from the circuit.

195

In this connection the maximum power dissipation of T_1 will occur when $V_{CE1} = V_R$ (see Fig. 9.27).
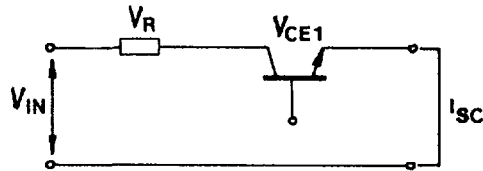

Fig. 9.27: Distribution of power dissipation

$$V_{CE1} = \frac{V_{IN}}{2} = V_{IN} - I_{SC} \cdot R = V_{IN} - I_{SC} \cdot \frac{V_{IN}}{I_M}$$

The power dissipation

$$P_{T1} = V_{CE1} \times I_{SC} = V_{IN} - I_{SC} \frac{V_{IN}}{I_M}$$

$$\Delta P_{T1} = I_{SC} \cdot V_{IN} - I_{SC}^2 \frac{V_{IN}}{I_M}$$

$$\Delta P_{I_{SC}} = V_{IN} - 2 I_{SC} \frac{V_{IN}}{I_M} = 0$$

$$I_{SC} = \frac{I_M}{2}$$

$$P_{T1max} = \frac{I_M}{2} V_{IN} - \frac{I_M^2}{4} \cdot \frac{V_{IN}}{I_M} = \frac{V_{IN} \cdot I_M}{4}$$

Half of the power dissipation is occurring on the pass transistor.

If more voltage is required at the output, T1 will conduct more reducing its V_{CE}. When this value is sufficiently low, T2 will start conducing, keeping V_{CE1} constant and deriving part of the load current through it.

The maximum power dissipation of T2 will occur when the current through it is equal to current through the resistor and both are equal to $I_M/2$ (see Fig. 9.26).

All linear regulated power supplies have a big disadvantage because of the power dissipation. Therefore, some manufacturer of nuclear equipment build in switching power supplies (Fig. 9.30).

Sensitive circuits do not tolerate increases in the supply voltage; they have to be protected by an overvoltage regulating system (Fig. 9.28).
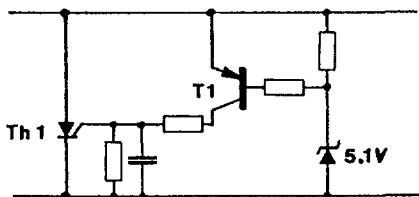

Fig. 9.28: Overvoltage protection

Due to an overvoltage in the supply of about 5.6V a base current of the transistor T1 is flowing through the 5.1V zener-diode. In the same moment the thyristor Th1 is fired by the transistor T1 and the supply voltage is shorted via Thyristor Th1.

In the case when the pass transistor of the voltage regulator or the current limitation gets damaged, a fuse has to be installed, otherwise the transformer and the rectifiers are endangered due to the short circuit.

The fuse will not blow if the pass transistor and the current limitation are working normally. There is, however, a voltage drop of about 1V on the supply line.

It is not easy to distinguish between overload and the triggered overvoltage protection. To determine the source of the troubles, the supply has to be fully separated from the load and switched on again because thyristor is only cut off when no current is flowing through it. If again the overvoltage protection is triggered the fault has to be sought in the voltage regulation system. Therefore, the overvoltage protection has to be disconnected to be able to find the failure in the regulating system.

The supply voltage for analog purpose $\pm12V$ and $\pm24V$ needs no overvoltage protection because of the wide range of the used transistors (refer to V_{CE}).

9.9 SWITCHING POWER SUPPLIES

To develop switching power supplies, extensive experience is necessary. The biggest advantage of such a system is the small weight. Due to the high frequency which is used, the transformer can be built very small. On the other hand, there is the spike problem. A lot of filtering, as well as a good design of the circuit board, is required to prevent the propogation of spikes.

In many applications of linear regulators, circuits for stepdown switching regulators are published. However, for these circuits one needs adequate fast recovery diodes and special chokes, which are very expensive. The advantage is that the pass transistor has nearly no power dissipation due to the fast switch on and off characteristic. Still a transformer with rectifying and filter capacitor is necessary to insulate the system from the mains (line). This is the price of the low power dissipation.

Fig. 9.29: Description of stepdown switching regulator

197
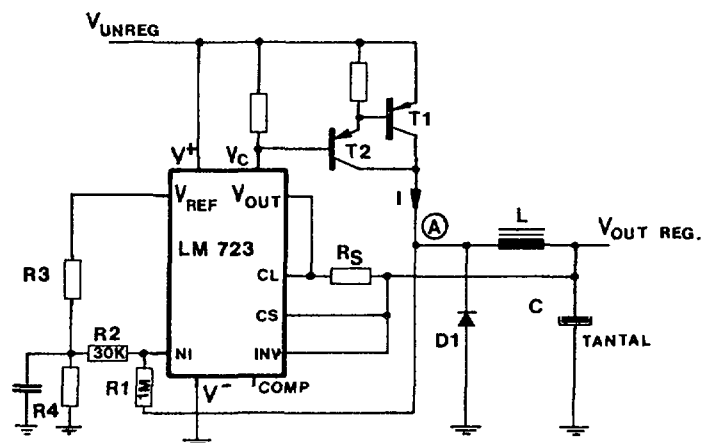
The operation of a stepdown regulator can be explained using Fig. 9.29.

At the beginning the voltage applied to the inverting input is smaller than to the non-inverting. The transistors T1 and T2 are fully conducting. When balance is reached, T1 will be cut off. The inductance will still keep the current I due to the stored energy. Therefore, the current I must still flow through the load or capacitor C and now through the diode D1 until the energy of the inductance is dissipated. This means that the node A is now negative compared to ground. Now the voltage at the NI input is slightly smaller because of the feedback resistor network R1 and R2 compared to the INV input so that the transistor becomes completely cut off. When the voltage across the capacitor C becomes smaller the switching cycle can start again.

To avoid the big mains tranformer, a switched mode regulated power supply has to be used as shown in Fig. 9.30.
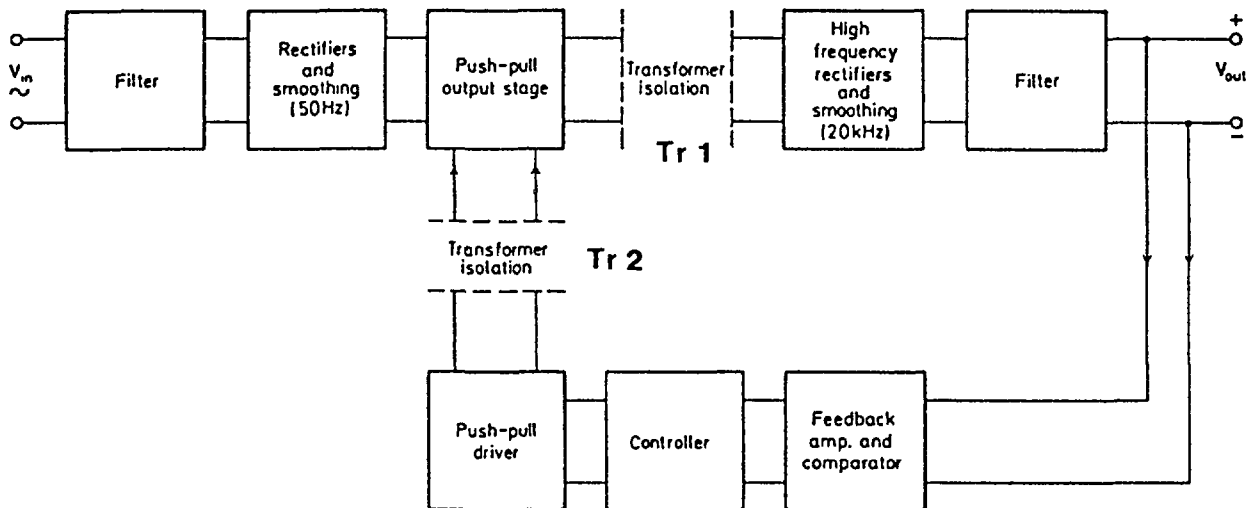
Fig. 9.30:   Block diagram of a switched mode regulated power supply to avoid high power dissipation as well as the big transformer

The function of such a supply is the following: on the primary side, there are rectifier diodes and a push-pull output stage which are controlled by the controller and push-pull driver via the transformer Tr2. If the DC voltage on the secondary side is not high enough triggering pulses are applied through Tr2. If the regulated DC voltage is high enough, no triggering pulses occur on the transformer Tr2. Nowadays, instead of the transformer Tr2, opto couplers are used. The isolation test voltage of such opto couplers between input and output are 1500 up to 4000V. In this case, the potential separation of such a system is achieved. The power is transferred by the transformer Tr1. The oscillator frequency is in the range 15 to 30k cycles. Such types of power supplies mostly sealed and in the several circuit diagrams as a box shown.

Such a device for a 5V supply with a maximum load of 1A and an input voltage of $200V_{ac}$ has a physical size of 80mm x 45mm x 25mm and is fully thermal and short circuit protected.

Due to the filtering problem, the switching power supplies are only rarely used in nuclear instrumentation as low voltage power supplies.

The most powerful supply which is used in nuclear instrumenation is the 5V supply of a multi-channel analyzer. Such a supply could easily reach 8 to 10A or more in the old fashion version for standard TTL logic. In the latest version of such MCA's, CMOS components are being used so that the current of the 5V supply is reduced to the values between 2 and 5Amps. Also an overvoltage protection has to be built in. In the case of a failure in the supply a switched short circuit has to protect the costly ICs. Fig. 9.28 shows such a overvoltage protection.

Nowadays, many manufacturers of MCAs use the operational amplifier regulated system for the low voltage power supplies. In such MCAs there is a need for +5V, ±12V and ±24V; so they are using the +24V for the MCA supply network as well as an auxillary supply for all other supplies. This means that the operational amplifiers are supplied by the +24V. The big advantage of such a system is that the pass transistor of the different supplies can be kept at a minimum voltage of about 1V. Therefore, the power dissipation is very small. Fig. 9.31 shows such a circuit diagram of a power supply.

Also, Darlington transistors are used as pass transistors. The sense resistor R_s can be very small because the current limitation is very often sensed by operational amplifier. The resistor in series to the emitter is to equilize the current through parallel switched pass transistor. Also, for current limitation, the voltage drop of Base-Emitter and the resistor in series to the emitter can be sensed. In this case, an extra sense resistor is not necessary. The supply for the base current is coming from a higher value supply voltage and all of these modern operational amplifier output stages have a built-in current limiter so that they can connect directly to the base. If this supply is not working, all other voltages are also not present. For troubleshooting it is necessary to find out which supply is the most important and from there the supply lines have to be followed step by step. Modern supplies indicate the presence of the voltages by LEDs so it is easy to find the fault.

To reduce the power dissipation of pass transistor, another circuit, as shown in Fig. 9.32, can be used.

This above circuit is also used in many applications. Also the power dissipation on pass transistor can be reduced., but a second transformer is needed or extra windings have to be made on the transformer. This circuit is useful when only one supply voltage is required.

Fig. 9.33 shows an option to reduce the power dissipation on the pass transistor. For this circuit, special voltage regulators are required; they are very costly and are not used anymore in modern equipment. To regulate the system a slight change of voltage at the base of T2 is enough to have a big change on the output voltage. If pass transistor T1 has a fault, as in most of the cases, through overload T2 is also immediately damaged.

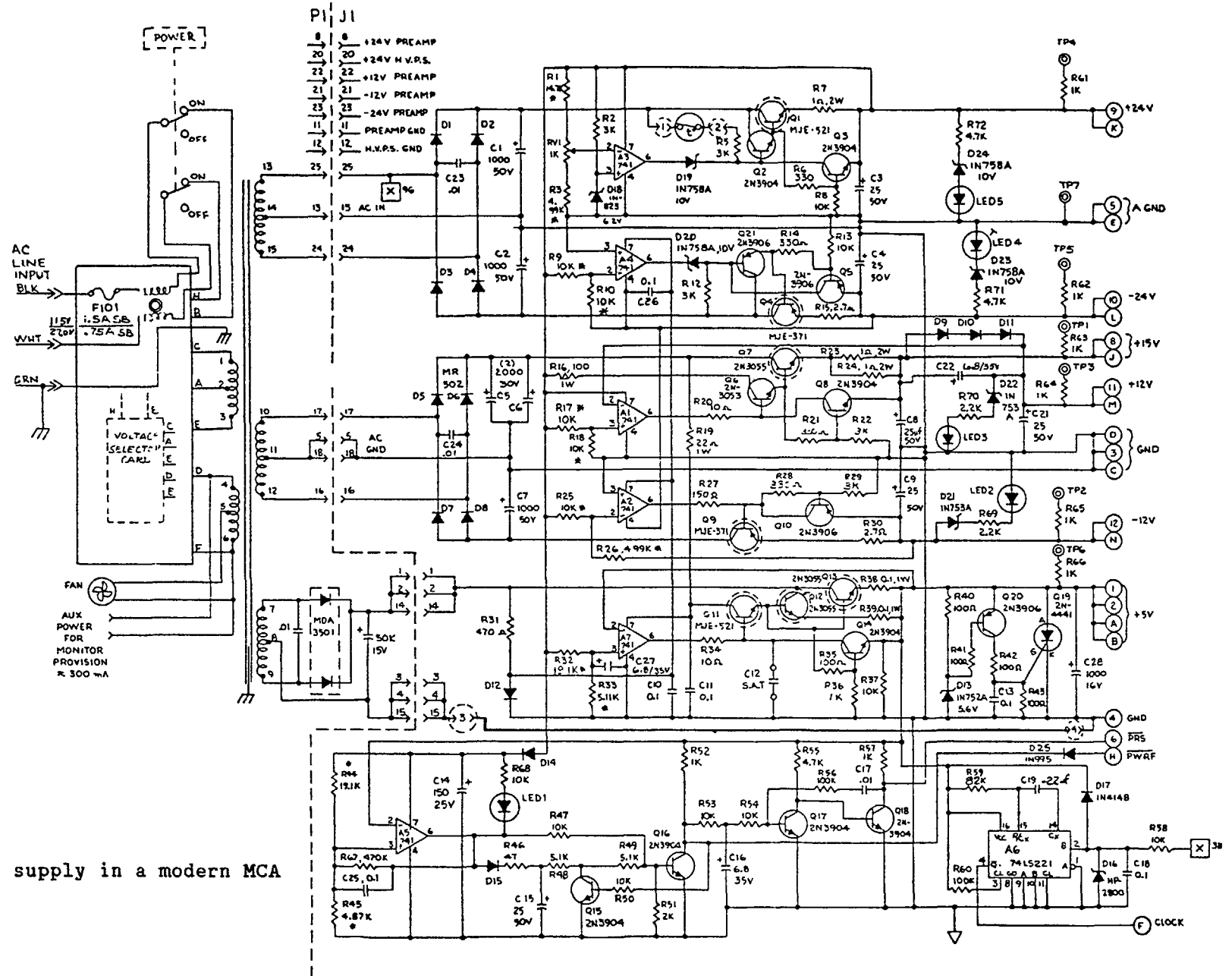Another circuit for current limitation is shown in Fig. 9.34.
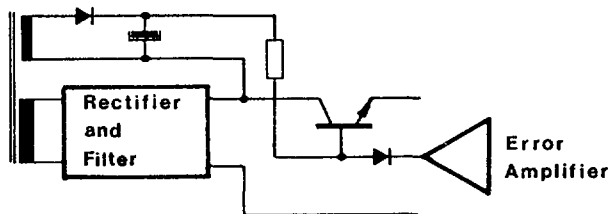
Fig. 9.31: Power supply in a modern MCA

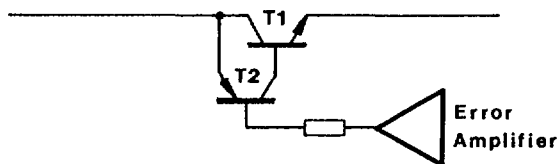Fig. 9.32: Auxillary supply for biasing pass transistor

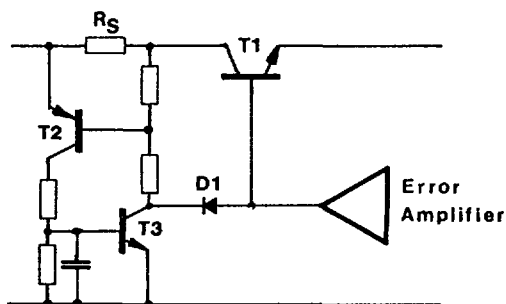Fig. 9.33: npn pass transistor biased by pnp transistor

Fig. 9.34: Current limitation with shut-down characteristic

If the voltage drop across R_S exceeds 0.7V, transistor T2 starts conducting. At the same moment transistor T3 is conducting and keeps transistor T2 in conducting condition and also takes over the base current via diode D1 from the pass transistor T1. Due to the current limitation features of the operational amplifier, it will not be destroyed. The V-I characteristic is shown is Fig. 9.35.
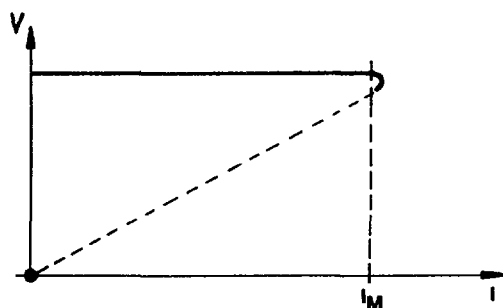
Fig. 9.35: Current limiting with shut-down characteristic

When current limitation is reached, there is almost no current flowing in the output. In case a short circuit occurs, there is only a small power dissipation in the pass transistor. The disadvantage of the circuit is the folowing: after overload or short circuit, the equipment has to be switched off and switched on again because of the memorizing of an overload by T2 and T3. Also the resistance of R_S can be kept to the minimum value, $R_S = 0.7V/I_{MAX}$, which gives the minimum voltage drop across the sense resistor.

Some manufacturers of nuclear equipment mount a thermo-coupled switch on the heatsinks of the pass transistors. This switch opens the contacts at a certain temperature. Such a contact can handle a very small DC current. Therefore, the switch must be connected into the control circuit. With this temperature switch it is possible to protect the pass transistor from overheating.

A similar thermo-coupled switch can be used to protect a transformer against overload. If a transformer has more than two independent windings, it is not so easy to protect the transformer against overload.

An overload on one of the secondary windings creates an overheating in the secondary winding but not a higher current on the primary side if only this winding is loaded. The maximum temperature of the windings is about 105°C, otherwise the isolating varnish which covers the wire will be damaged. If this happens a short circuit will result. Therefore, it is not possible to protect a transformer by fuses against overload if it has more than two secondary windings.

The power dissipation is calculated by the following formula:

$$P = I^2 \cdot R \qquad\qquad R \sim \text{windings}$$
resistance

A twenty percent higher current creates 45% more heat. This heat is present at the inside windings and there is no air stream; therefore, the transformer is easily destroyed. The fuses in front of a transformer should have slow blow characteristic, to protect the transformer at increased currents but not to react on short current spikes.

9.11 TRANSFORMER AND RECTIFIER RATINGS

When a transformer is used to supply other than a pure resistive load, it is necessary to derate its specified AC current rating to prevent overload. In Fig. 9.36, these current derating factors are shown below each circuit and relate the transformer AC current rating to the DC load current.

Note: where a transformer has a VA rating, this is the product of the secondary AC voltage and the secondary AC current.

For full-wave circuits, the average current per rectifier is $0.5 \cdot I_{dc}$. For half-wave circuits it equals I_{dc}. For all circuits, each reactifier should have a V_{RRM} rating in excess of $1.4 \times V_{ac}$ except the half-wave capacitor input circuit which requires a rating in excess of $2.8 \times V_{ac}$.

Relationships between DC voltage and AC voltage are shown in Fig. 9.36 for different rectifier circuits and do not include losses. The transformer AC output voltage is specified at the full load current - off load, the voltage will rise in accordance with the regulation specified for that particular transformer.

Relationships and circuits shown on the right are intended as a guide and assume no circuit losses. For low DC voltages, allowance should be made for the diode voltage drop.
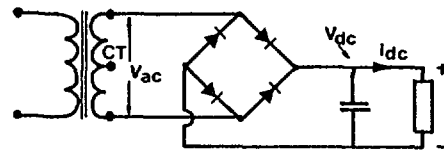
Half-Wave Capacitive Input Filter
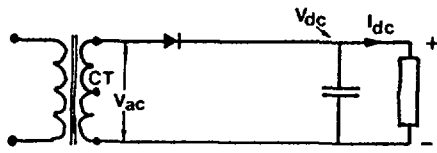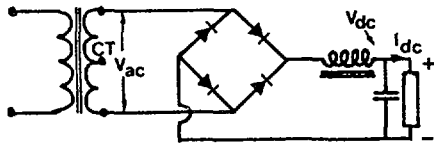

$V_{dc} = 1.41 \ V_{ac}$
$I_{dc} = 0.23 \ I_{ac}$

Full-Wave Bridge Capacitive Input Filter

$V_{dc} = 1.41 \ V_{ac}$
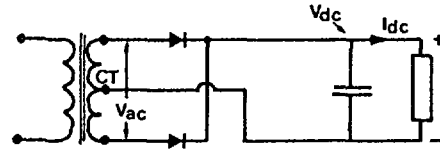$I_{dc} = 0.62 \ I_{ac}$

Full-Wave-Bridge Choke Inpute Filter

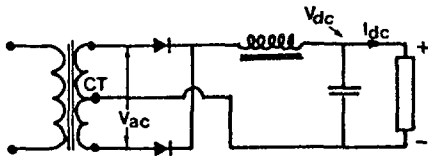$V_{dc} = 0.90 \ V_{ac}$
$I_{dc} = 0.94 \ I_{ac}$

Full-Wave-Capacitive Input Filter
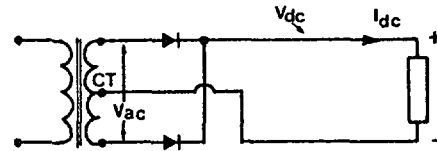

$V_{dc} = 0.71 \ V_{ac}$
$I_{dc} = 1.0 \ I_{ac}$

Full-Wave Choke Input Filter

$V_{dc} = 0.45 \ V_{ac}$
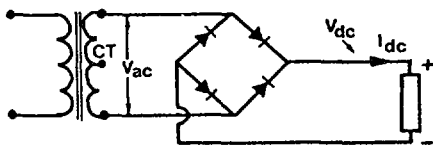$I_{dc} = 1.54 \ 1_{ac}$

Full Wave Resistive Load

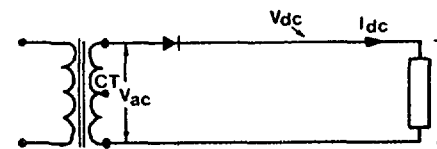$V_{ac} = 0.45 \ V_{ac}$
$I_{dc} = 1.27 \ I_{ac}$

Full-Wave Bridge-Resistive Load

$V_{dc} = 0.90 \ V_{ac}$
$I_{dc} = 0.90 \ I_{ac}$

Half-Wave-Resistive Load

$V_{dc} = 0.45 \ V_{ac}$
$I_{dc} = 0.64 \ I_{ac}$

Fig. 9.36: Relation between DC and AC voltages in different
configurations of power supplies