

BNL--30214

BNL 30214  
OG 599

DE82 010252

CONF - 810787 --1  
MASTER

Detectors and Signal Processing for  
High-Energy Physics

P. Rehak  
Brookhaven National Laboratory, Upton, NY 11973

Talk given at 3rd Course of

International School of Physics, Enrico Fermi  
Varenna, Lake Como, Italy 7/28 - 8/1/81

DISCLAIMER

This manuscript has been authored under contract DE-AC02-76CH00016 with the U.S. Department of Energy. Accordingly, the U.S. Government retains a nonexclusive, royalty-free license to publish or reproduce the published form of this contribution, or allow others to do so, for U.S. Government purposes.

The submitted manuscript has been authored under contract DE-AC02-76CH00016 with the U.S. Department of Energy. Accordingly, the U.S. Government retains a nonexclusive, royalty-free license to publish or reproduce the published form of this contribution, or allow others to do so, for U.S. Government purposes.

DISTRIBUTION OF THIS IS UNLIMITED

429

OG599

Detection and Signal Processing in  
High Energy Physics \*

by

P. Rehak

Brookhaven National Laboratory, Upton, NY 11973

presented at

the Enrico Fermi International School of Physics

Varenna, Lake Como, Italy

July 28 to August 7, 1981 (3rd course)

\* This research supported by the U.S. Department of Energy under  
Contract No. DE-AC02-76CH00016.

### Summary

Basic principles of the particle detection and signal processing for high-energy physics experiments are presented. It is shown that the optimum performance of a properly designed detector system is not limited by incidental imperfections, but solely by more fundamental limitations imposed by the quantum nature and statistical behavior of matter

The noise sources connected with the detection and signal processing are studied. The concepts of optimal filtering and optimal detector/amplifying device matching are introduced. Signal processing for a liquid argon calorimeter is analyzed into some detail.

The position detection in gas counters is studied. Resolution in drift chambers for the drift coordinate measurement as well as the second coordinate measurement is discussed.

## Part I.

### 1. Introduction

While detectors and detection in high energy physics do not need an introduction here we have the feeling that the signal processing does.

Everybody is familiar with (or at least has seen) some detectors around an interaction region at a storage ring type machine or downstream from a target in a classical particle physics experiment. The title of this school is "Data aquisition in high energy physics" so everybody knows that the data from the detectors describing the event is stored in many bits in many registers in form of a digital information. The signal processing is a way of converting an obscure detector signal into useful information.

We will study the signal processing in a broad sense. Signal processing includes signal formation due to a particle passage within a detector, signal amplification, signal shaping (filtering) and read out. The basic goal of the signal processing is to extract the derived and pertinent information from the obscuring factors (or "noise") usually present. The relative strength of the signal and the noise is one way to measure the quality of the detection.

The two quantities of greatest importance to be extracted from detector signals are their amplitude and time of occurrence. The amplitude information is usually related to the energy or to the nature of the particle. Also some position sensing detectors give the localization of the particle via signal amplitude, however, the most popular method of particle localization in high energy physics experiment is based on time information.

## 2. Means of Detection

As will be shown later each detection method has to extract some energy from the particle to be detected. This is done via the particle interaction with the detector medium.

Nearly all detection methods (Čerenkov and transition radiation detector being exceptions) make use of ionization or excitation produced in a detection medium as a result of the particle interaction. This interaction changes some quantities characterizing the detected particles.

This very fact imposes the final limitation on the accuracy of some measurements as required by laws of quantum mechanics. In this lecture we will find that usually the other physical phenomena in particle detection impose larger limits on the precision of measurement.

Returning to the detection process; in the case of charged particles, ionization and excitation is produced directly by the interaction of the particle electromagnetic field with the electrons of the detection medium. The resultant ionization and excitation is distributed along the path of the particle. A typical particle energy in today's experiments is of the order of few GeV, while the energy loss can be below the MeV level. This is an example of a non-destructive method for detection of charged particles.

All neutral particles must first undergo some process which transfers all or part of their energy to charged particles. The detection method is destructive. The interaction cross section for the neutral particle can be quite small and therefore a large amount

of material may be needed for such a detector. A great variety of detectors for electromagnetic and hadron cascades have been recently developed.

### 3. Detection of Ionization and Excitation

In most ionization detectors the total ionization is collected using an externally applied electrical field. Sometimes an amplification process by avalanche formation in a high electrical field is used.

Examples of detectors are

- a) Proportional chamber (drift chamber)
- b) Liquid argon chamber
- c) Semiconductor detector

Detection of excitation is a multistep process. The excitation is first converted into luminescent light by emission, and then the light into an electrical signal. The most used light to electrical signal converters are photomultipliers, however, the use of micro-channel plates and vacuum and solid state photodiodes in high energy physics is steadily increasing.

As can be seen from the short list of detectors, all of them provide a certain amount of charge onto an output electrode. The electrode represents a certain capacitance so from subsequent signal processing point of view these detectors are capacitive sources, i.e. their output impedance is dominated by the capacitance.

This common feature of all detectors for particle physics allow rather unified approach to the signal processing.

The science of signal processing for semiconductor detectors was basically developed during the last 20 years.<sup>2</sup> Most of the rules are directly applicable to our detectors. However, in spite of common

features among various detectors used in high energy physics we should not underestimate the differences among them. The typical total charge at the detector outputs can differ by six orders of magnitude for two different kinds of detectors and the output capacitances can differ by the same factor. What is important for one kind of detector can be irrelevant for another and vice versa.

In a majority of detectors the signal produced by the detected particles is relatively small and is contaminated by noise generated by the detector itself and by the amplifier. We can ask a few "fundamental" questions. What is minimal signal which we can detect? (That means to measure its amplitude and the time of its occurrence). How does it depend on the shape of the waveform? What is the amplifying element which we have to use to obtain this optimum condition? Is the optimum solution also practical for a big detector system so typical in a high energy physics experiment?

We will try to answer such questions. We will see that the optimum performance of a properly designed detector system is not limited by incidental imperfections but solely by more fundamental limitations imposed by the atomic structure and statistical behavior of matter. The presence of noise in the detection system is a manifestation of the phenomena governed by the laws of statistical mechanics which has to be taken into consideration at the design stage of the detector system to minimize its effects.

The meaning of the word "noise" is unfortunately different for different detectors. This is extremely confusing, but now it is already too late to change the tradition. For ionization detectors, noise has

the meaning of the electrical noise in the amplifier circuitry. Fluctuations due to other effects in particle detection (statistical fluctuation of collected charge for semiconductor detectors, sampling fluctuations in calorimeters, etc.) are usually treated as a separate factor limiting the detectability or detector resolution.

In photon detection the word noise usually includes all fluctuations including those due to the quantum nature of the light.

We would like to stress that noise in these notes always denotes the intrinsic noise related to the basic nature of the detection process. A broad class of man-made disturbances (radio and TV signals, switching, sparking, etc.) picked up by detectors is not a topic of these lectures. These interferences are sometimes called noise, but their presence can be reduced to the negligible level by proper shielding.

#### 4. Outline of the Paper

These notes are intended for the summer School of Physics and in principle their text should be understandable without any references. Some references are given for readers more deeply interested in a particular subject. Their list is not intended to be complete and we apologize to all of those who have contributed to the art of signal processing and are not quoted here.

In Part II we will develop a technique for treating the noise propagation in electronic systems. Shot and thermal noise sources will be defined and studied in the time as well as in the frequency domain.

Part III will treat the general problem of the charge measurement in the presence of noise. The concept of matched filters for the charge contained in the signal and for the time of the signal occurrence will be introduced.

In Part IV we will find the optimal filter for the charge measurement under additional constraints imposed by today's high-energy physics experiments. We will define the optimal matching between detectors and amplifying devices. The signal processing for a liquid argon calorimeter is treated in some detail.

Part V will give a short overview of methods of the position sensing in gas detectors. Basic phenomena of the gas counter detection will be presented. Limitations of the position measurement in drift chambers will be studied. The second coordinate read-out in drift chambers will be discussed.

## Part II. Noise Analysis in Linear Systems<sup>3,4,5</sup>

### 1. Introduction

Noise analysis has a reputation for difficulty and obscurity. The reason may be that the study of noise is based on results from three different, rather disparate disciplines: statistical mechanics (thermodynamics), circuit theory and statistical analysis. The fundamental processes which give rise to noise in electronic components are in the realm of statistical mechanics. The noise signal enters and propagates through an electronics system requiring techniques of circuit theory. Finally, the results have to be interpreted in statistical terms.

In these lectures some fundamental knowledge of all three disciplines is assumed. Requirements on the knowledge of statistical mechanics and statistical analysis are well below the level of physics graduates. Circuit theory will be limited to the analysis of linear time invariant systems, and some knowledge of Fourier series, Fourier and Laplace transforms is assumed.

### 2. Random Processes

An example of a noise signal as could be seen on the oscilloscope is shown on Fig. 1. The random signal waveshape has an irregular, unpredictable appearance. We call this random time-varying function a random process  $v(t)$ . A sample of  $v(t)$  taken at an arbitrary time  $(t)$  is a random variable with some probability-density function  $f_t(v)$ . We

generally assume a gaussian noise distribution around a mean value  $E(v)$ .

(In these notes  $E(x)$  will denote the expectation value of  $x$  or ensemble average.) This assumption is based on the central limit theorem, but in any case, it is exceedingly unlikely that a single reading will depart from the mean value by more than a few times the r.m.s. (root mean square) for any "realistic" distribution. We, therefore, adopt r.m.s. to describe the probability-density function.

To determine  $f_t(v)$  we have to use the concept of an ensemble from statistical mechanics. Let us consider  $v$  as a voltage across a  $100\Omega$  resistor. A complete specification of the macroscopic state of the resistor would involve giving the coordinates or momenta of all the  $10^{23}$  particles constituting the resistor. This task is clearly beyond the bounds of practical possibility. Instead we have to accept a much cruder description of the resistor. We will assume it to be in thermodynamic equilibrium with the outside world.

In some of the possible configurations of the resistor, the potential difference across the resistor might be between  $v$  and  $v+\Delta v$ . We can also describe this by saying that if we had an ensemble of identical resistors specified macroscopically in an identical way, then for some fraction  $\theta_{MAC}(v) \cdot \Delta v$  of all the resistors the voltage would be between  $v$  and  $v+\Delta v$ .

In statistical mechanics we assume that a proper statistical ensemble contains one copy of the original system in every single possible macroscopic configuration compatible with the macroscopic description of the resistor. We can see immediately that  $\theta_{MAC}(v)\Delta v=f_t(v)\cdot\Delta v$ .

Another "experimental" approach to determine the probability density function  $f_t(v)$  would be to sample the value  $v(t)$  at intervals far

enough apart to ensure the statistical independence of the samples.

A random process  $v(t)$  is said to be an ergodic process if the probability density functions obtained by both the ensemble and time approaches are identical. We will encounter only ergodic processes in these lectures.

Before deriving the general relation between the noise description in the time and frequency domain, let us consider a special and important example - shot noise.

### 3. Shot Noise

Shot noise occurs whenever a noise phenomenon can be considered as a series of independent events occurring at random. It is important, for example, in emission of electrons by a photocathode or in minority carrier devices.

If we take a sufficiently short interval of time  $\Delta t$ , the probability  $\Delta p$  that an electron is emitted within this interval is proportional to  $\Delta t$

$$\Delta p = v \Delta t \quad (2-1)$$

where  $v$  is the mean rate of emission.

It can be easily shown that the mean number of electrons emitted during an interval  $\tau$

$$\begin{aligned} E(n_\tau) &= v \cdot \tau \\ \text{and} \quad E(n_\tau^2) &= v^2 \tau^2 + v\tau \\ \text{or} \quad E(\Delta n_\tau^2) &= v\tau \end{aligned} \quad (2-2)$$

(This is the well known result for the variance of Poisson distribution.)

Let us try to measure the current due to the electron emission during the time  $\tau$ . The mean current is

$$E(I) = \frac{q}{\tau} \cdot E(n_\tau) = qv \quad (2-3)$$

where  $q$  is the electron charge.

Fluctuations of the current expressed by the variance.

$$E(\Delta I^2) = \left(\frac{q}{\tau}\right)^2 \cdot E(\Delta n_\tau)^2 = \frac{q^2 v}{\tau} = \frac{q}{\tau} \cdot E(I) \quad (2-4)$$

This result agrees with our intuition. The fluctuations of  $I$  decrease as  $\tau$  increases.

We will now study a more interesting case when this current is applied at the input of a linear system with an impulse response  $h(t)$ . Because the current is constituted from random pulses with a mean rate  $v$  we can apply the Campbell theorem:<sup>6</sup>

If the impulse response of a linear system is  $h(t)$  then the mean response of the system to a random series of pulses occurring at a mean rate  $v$  is

$$E(v) = v \int_{-\infty}^{\infty} h(t) dt$$

and variance  $E(\Delta v^2) = v \int_{-\infty}^{\infty} h^2(t) dt$  (2-5)

Proof of the theorem is left for the reader. Campbell's theorem is a very powerful tool for noise calculations in the time domain. If any kind of noise, at the input of a linear system, can be expressed as a sequence of uncorrelated impulses then the equation (2-5) gives directly the noise at the output of the system.

#### 4. Noise Power and Spectral Representation of Noise.

The majority of electronics systems are more easily analyzable in frequency domain. The frequency or spectral analysis of random noise signal differs somewhat from the standard analysis of deterministic signals.

Let's expand the random signal  $v(t)$  in Fourier series along an interval  $T$ . Thus,

$$v(t) = \frac{1}{T} \sum_{n=-\infty}^{\infty} c_n e^{j\omega_n t} \quad \omega_n = \frac{2\pi n}{T}$$

where  $c_n = \frac{1}{T} \int_{-T/2}^{T/2} v(t) e^{-j\omega_n t} dt$  (2-6)

Note that  $c_n$  is a random variable because  $v(t)$  is random. Because  $v(t)$  is a real,  $c_n^* = c_{-n}$ . If the process is stationary, the expectation value of  $v(t)$  must be time independent.

$$E(v(t)) = \sum_{n=-\infty}^{\infty} E(c_n) e^{j\omega_n t} \quad (2-7)$$

So all expectation values of the Fourier coefficient  $c_n$  except for  $E(c_0)$  must equal 0.

So we are now in the peculiar situation of having a Fourier-series representation of  $v(t)$  which is valid for the particular time interval  $T$ , but which varies statistically with the chosen member of the ensemble.

To obviate this difficulty let us calculate

$$E(v^2(t)) = \sum_m \sum_n E(c_m c_n^*) \cdot e^{j(\omega_m - \omega_n)t} \quad (2-8)$$

and again if  $v(t)$  is a stationary random process expression (2-8) must be independent of  $t$  and all expectation values  $E(c_m c_n^*)$  must vanish except when  $n = m$ , so

$$E(v^2(t)) = \sum_{n=-\infty}^{\infty} E(c_n c_n^*) = E(c_0^2) + 2 \sum_{n=1}^{\infty} E(c_n c_n^*)$$

From equation (2-7)  $E(c_0)$  is the time-average value of  $v(t)$  over the period  $T$ . If we take  $T$  large enough and assume that the system is ergodic  $c_0$  is the same for all systems in the statistical ensemble.

Thus  $(E(c_0))^2 = c_0^2 = E(c_0^2)$  and the expected fluctuations of  $v(t)$  from equation (2-8) are

$$E(\Delta v(t)^2) = 2 \sum_{n=1}^{\infty} E(c_n c_n^*) \quad (2-9)$$

As  $T$  is made "very long", the interval  $1/T$  between the frequencies of successive Fourier components becomes "very small". Provided that  $E(c_n c_n^*)$  varies smoothly with  $n$  we can rewrite equation (2-9) as

$$\begin{aligned} E(\Delta v(t)^2) &= \lim_{T \rightarrow \infty} 2 \sum_{n=1}^{\infty} T \cdot E(c_n c_n^*) \cdot \frac{1}{T} \\ &= \int 2T \cdot E(c(f) \cdot c^*(f)) df = \int w(f) df \end{aligned} \quad (2-10)$$

where  $w(f)df = \lim_{T \rightarrow \infty} TE(c_n c_n^*)$ ,  $f = \frac{n}{T}$

The function  $w(f)$  has dimension of square voltage for unit bandwidth. Divided by resistance it becomes the power density. It is called the power spectral density function or power spectrum. Let us show that knowledge of a power spectrum  $w(f)$  at the input of a linear system allows the determination of a power spectrum function  $w_o(t)$  at its output.

If  $v(t)$  is presented as the input to a linear system described by transfer function  $H(f)$  (Fourier transform of its impulse response  $h(t)$ ) the output can be written

$$v_o(t) = \sum_{n=-\infty}^{\infty} H\left(\frac{n}{T}\right) c_n e^{j\omega_n t}$$

so  $E(v_o(t)) = H(0) \cdot c_0$

and  $E(\Delta v(t)^2) = E(v_o(t)^2) - E^2(v_o(t))$

$$= 2 \sum_{n=1}^{\infty} H\left(\frac{n}{T}\right) \cdot H^*\left(\frac{n}{T}\right) E(c_n c_n^*)$$

which in the limit  $T \rightarrow \infty$  can be written as

$$E\left(\Delta v_o(f)^2\right) = \int_{-\infty}^{\infty} |H(f)|^2 w_i(f) df \quad (2-11)$$

or

$$w_o(f) = w_i(f) \cdot |H(f)|^2$$

The equation (2-11) is an extremely important relation for the frequency analysis of random signals passing through a linear system. The spectral analysis of random signals focuses on the power distribution. This is related to the integrated square of signals and accounts for the  $|H(f)|^2$  term appearing in the transfer relation (2-11). The band-limiting effect of a linear system on a random input signal is expressed by multiplication of the input spectral density by  $|H(f)|^2$ . This is a direct analogy to the multiplication of the input Fourier transform by the transfer function  $H(f)$  for deterministic signals.

Here we can ask a question. Can we carry this analogy also into the time domain? The answer is positive, but before going into detail let us return for a moment to our example of shot noise.

## 5. Power Spectrum of Shot Noise

Let us rewrite the Campbell theorem (2-5) using the Fourier transform  $H(f)$  of the impulse response  $h(t)$  of the linear system. ( $H(f)$  is a well-known system transfer function.)

$$E(\Delta v^2) = \nu \int_{-\infty}^{\infty} h^2(t) dt = \nu \int_{-\infty}^{\infty} H(f) \cdot H^*(f) df$$

We have used Parseval's theorem. Since  $h(t)$  is a real function of time  $H(-f) = H^*(f)$  and  $H(f) \cdot H^*(f)$  is an even function of frequency we can write

$$E(\Delta v^2) = 2\nu \int_{0}^{\infty} |H(f)|^2 df \quad (2-12)$$

The form of equation (2-12) is similar to (2-11). Let us substitute  $v$  with direct measured quantities. If a current  $I$  exhibiting only shot noise is presented at the input of the linear system ( $v = \frac{I}{q}$ ; individual pulses (electrons) have a charge  $q$ , so the linear system responds as  $q h(t)$ ,  $q^2 h^2(t)$ ) we can substitute for  $v$  in (2-12) and obtain

$$E(\Delta v^2) = 2 \frac{I}{q} \cdot \int_0^{\infty} q^2 |H(f)|^2 df = 2qI \int_0^{\infty} |H(f)|^2 df \quad (2-13)$$

This result can be interpreted according to the equation (2-11) as a power spectrum at the output of the linear system with the power spectrum  $w(f) = 2qI$  imposed at its input. We have so shown that shot noise (or more generally any Poisson-like fluctuations) has the power spectral density  $w_s(f) = 2Iq$  (2-14)

## 6. Autocorrelation Function <sup>7</sup>

Let us return to the question raised in Section 4 about analogy of the power density function in time domain. Basically, we require some measure of how the noise process varies with time. Specifically, if we consider the noise waveform  $v(t)$  of Fig. 1, we note that as  $t_2 \rightarrow t_1$ ,  $v(t_2)$ , regarded as a random variable, becomes more closely related to the  $v(t_1)$ . As  $t_2 - t_1$  increases we see less dependence of  $v(t_2)$  on  $v(t_1)$ . We make this concept more precise by defining the autocorrelation function

$$R_v(t_1, t_2) = E(v(t_1) \cdot v(t_2)) \quad (2-15)$$

It is apparent that if  $t_2 \rightarrow t_1$   $R_v \rightarrow E(v^2)$  or autocorrelation becomes the statistical second moment. If at some value of the difference  $(t_2 - t_1)$ ,  $v(t_1)$  and  $v(t_2)$

are statistically independent  $R_v \rightarrow E^2(v)$ . (0 if  $E(v)=0$ .)  $R_v(t_1, t_2)$  thus provides one possible measure of how rapidly  $v(t)$  changes in time.

For a stationary process  $R_v(t_1, t_2)$  can depend only on the time difference  $t_2 - t_1 = \tau$  so that we can write

$$R_v(\tau) = E[v(t) \cdot v(t + \tau)] \quad (2-15a)$$

For an ergodic process we can replace the ensemble average  $E$  by time average and write

$$R_v(\tau) = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T v(t) \cdot v(t + \tau) dt \quad (2-16)$$

The relationship between the power spectral density  $w(f)$  and the autocorrelation function is given by (Wiener-Khintchine theorem)<sup>8, 7</sup>

$$R(\tau) = \int_0^\infty w(f) \cdot e^{j\omega\tau} df$$
$$w(f) = 2 \int_{-\infty}^\infty R(\tau) \cdot e^{-j\omega\tau} d\tau ; \quad f \geq 0 \quad (2-17)$$

So that  $R(\tau)$  and  $w(f)$  are almost a Fourier transform pair. (If we define "mathematical power spectral density"  $w_m(f)$  as  $w_m(f) = \frac{1}{2}w(f)$  if  $f > 0$  and  $w_m(f) = -w_m(-f)$  for  $f < 0$ ;  $w_m(f)$  and  $R(\tau)$  are a Fourier transform pair.) The autocorrelation function is a means of characterizing the noise in the time domain as the power spectral density is in the frequency domain. The physics phenomena producing the noise may be more easily analyzable in either one domain or the other so relations (2-17) allows us to translate the noise description from one domain into another.

Let us return again to the example of a shot noise. A succession of pulses occurring randomly in time can be most conveniently described

in the time domain. In this case  $R(\tau)$  is zero for all values of  $\tau$  except  $\tau = 0$  and has the form of a "delta function." Its Fourier transform has all frequencies present in equal proportion. Calculation can be carried out and not surprisingly the result is identical with already known relation (2-14).

The thermal noise which we are going to investigate next has its origin most easily analyzable in the frequency domain.

#### 7. Johnson or Thermal Noise

Johnson noise arises from the thermal motion of the charge carriers in a resistor. In its equilibrium with the surroundings at the temperature  $T$  the average potential difference from one end of an isolated resistor to the other is zero. However, energy fluctuations of the order  $kT$  per degree of freedom manifest themselves as a power noise source.

Thermal noise was first studied experimentally by J. B. Johnson in 1928. A. Einstein had predicted in 1906 that Brownian motion of the charge carriers would lead to a fluctuating e.m.f. (electromotive force) across any resistance in thermal equilibrium. The power spectrum was calculated by H. Nyquist in 1928 based exclusively on thermodynamic reasoning. We will briefly repeat his treatment.

Let us consider a resistance  $R$  connected to one end of an ideal transmission line of characteristic impedance  $R_0$ , length  $l$  and wave velocity  $c$  that is shorted at the other end. After a while a state of equilibrium is reached. This means that power flowing from the resistance into the line is the same as the power flowing from the line into the

resistance. The energy inside the transmission line flows in two different directions; one half to the left and one half to the right. Power flowing into the resistor is, therefore, equal

$$P = \frac{1}{2} c \frac{E}{\lambda} \quad (2-18)$$

where  $E$  is the energy of electromagnetic waves inside the transmission line.

To calculate the energy we can, according to the rule of statistical thermodynamics, give to each normal mode of the line an energy

$$\epsilon = \frac{hf}{e^{hf/kT} - 1} = kT \quad \text{in the frequency range of interest} \quad (2-19)$$

where

$T$  - is the absolute temperature

$h$  - Plank constant

$k$  - Boltzmann constant

To calculate the number of normal modes we can short also the other end of the transmission line. (The line is still in equilibrium at  $T$  since, if we connect it back to the resistance no net power transfer to or from the line will occur.)

For the shorted transmission line modes are related with frequency according to

$$n \frac{\lambda}{2} = \frac{nc}{2f} = \ell \quad (2-20)$$

where  $n$  is an integer. Thus if  $\ell$  is "large" the number of modes in  $df$  is

$$dn = \frac{2\ell}{c} df$$

and the mean energy in frequency intervals  $df$ , trapped in the transmission line from equation (2-19) is

$$dE = \frac{2}{c} \cdot kT df$$

or using relation (2-18)

$$dP = kT df \quad (2-21)$$

We can see

- i) The noise power is independent of resistance
- ii) The noise power is white (independent of frequency)
- iii) The noise power is independent of the charge or the electron  
(q - does not enter into relation (2-21))

The relation (2-21) is usually written in one of two following forms.

$$w_i(f) = 4kT/R \quad (2-22)$$

or  $w_v(f) = 4kT \cdot R \quad (2-23)$

Equation (2-22) refers to Fig. 2a, where a physical resistance R was replaced by an idealised noiseless resistance R and a noise current generator in parallel. If the current power density of the generator is  $w_i(f) = 4kT/R$  the maximum power density available for an external circuit is

$$\frac{dP}{df} = R \cdot w_i(t)/4 = kT \quad \text{so relation}$$

(2-22) agrees with relation (2-21).

Similarly equation (2-23) refers to Fig. 2b where a physical resistance was replaced with a noiseless resistance and a noise voltage generator in series. It can be shown again that formulae (2-21) and (2-23) are identical.

### 8. Comparison between Shot and Thermal Noise

Let us comment about differences and similarities between shot and thermal noise. Shot noise depends on the number of the flowing charge carriers, thermal noise does not. Shot noise is a fluctuation in the number of charge carriers. It involves fundamentally a directed motion with random transit. Thermal noise is a fluctuation in the velocity (direction as well as magnitude) of the charge carriers.

Whether we have shot or thermal noise is essentially a question of whether or not the number of current carriers is proportional to the current. In a np junction or vacuum tube there are almost no current carriers normally present. Fluctuations must be associated with the number of current carriers that enter. (I in equation 2-14). In a resistor a very large and constant number of current carriers is always present. The net flow of current is a negligible perturbation of the random thermal motion of the carriers. Thus, we have the noise independently of the current, i.e. equations (2-22), (2-23).

Equations (2-22) and (2-23) are sometimes used to define equivalent noise resistance. It is simply a value of resistance which gives the same noise power as the noise source under study. For example, we can define the noise resistance corresponding to the shot noise due to the current I. From equations (2-14) and (2-22) we obtain

$$R_n = \frac{2kT}{qI} \left( = \frac{50 \text{ mV}}{I} \right) \quad (2-24)$$

at room temperature.

We can carry this parallel between shot and thermal noise even further. Because both have the same form of the frequency spectrum and because of Wiener-Khintchine theorem (Equations (2-17) they have also the same form of the autocorrelation function. Therefore, we can imagine the thermal noise in the time domain in the same way as the shot noise, that is, as a succession of pulses occurring randomly in time, and use

Campbell's theorem. (2-5)

More formally, if we connect a voltage noise source  $e_n$  with a voltage power density  $w_v(t) = 4kTR$  to the input of a linear system with the impulse response  $h(t)$  the fluctuations at the output are given (according to (2-11))

$$E(\Delta v_o^2) = w_v \int_0^{\infty} |H(f)|^2 df = \frac{1}{2} w_v \int_{-\infty}^{\infty} |H(f)|^2 df$$

where  $H(f)$  is the Fourier transform of  $h(t)$ . (Because  $h(t)$  is real  $H(-f) = H^*(f)$  and  $(H(f))^2$  is an even function of  $f$  we can extend the lower integration limit to  $-\infty$  and take one half of the integral.)

Using Parseval's theorem we can write

$$E(\Delta v_o^2) = \frac{1}{2} w_v \int_{-\infty}^{\infty} h^2(t) dt \quad (2-25)$$

which can be considered a form of the Campbell theorem (2-5) with a "rate" equal 1/2 of the noise power density.

We will conclude this part with a remark about the excess noise.

## 9. Flicker Noise.

In addition to shot noise and thermal noise which are white noise sources, there is often another noise source in many devices which predominates at lower frequencies. This flicker noise has a spectral density  $w(f)$  proportional to  $1/f^u$  where  $u$  is between .8 and 1.5. Origin of flicker noise is poorly understood. In frequencies of interest in signal processing it should not present a problem. Paradoxically, the devices intended for high frequency use exhibit the smallest amount of flicker noise, i.e. they have the smallest noise at lowest frequencies as well. This is probably related to the better care taken in production of high frequency devices.

## Part III

### MEASUREMENT OF SIGNAL CHARGES IN PRESENCE OF NOISE

#### Noise in Signal Amplification

We have seen in Part I that particle detection in high energy physics experiments requires the measurement of the charge delivered at the output electrode of the detector. To optimize the signal charge measurement under conditions imposed by the experiment is the main subject of the signal processing.

The detected charge has to be amplified early in the process to avoid possible additional contamination by other noise sources. From many different physical systems capable of providing gain mechanism, direct charge controlled amplifiers are the only amplifiers used in particle detection.

In this type of amplifier there is a charge in transit through the device (i.e., current is flowing), and its transport properties are modulated by charge on a separate control electrode. Examples are vacuum tubes (historical) and field effect transistors in which carrier transport is effected directly by the electric field of the grid or gate. Also, the bipolar transistor belongs to this category. Modulation of the charge in the base-emitter junction alters the base emitter voltage, hence the injection current. We stress here that bipolar transistors at the frequencies of interest are not current controlled devices. The noise analysis of all charged controlled devices is identical. The method consists of examining the noise in the current flowing through the device and then pretending that this noise is actually due to a noise voltage generator connected in series with the control electrode (Fig. 3). This is the series noise voltage and can be expressed as a series noise resistance. (According to relation [2-23]  $\overline{e_n^2} = 4 kTR$ , where we write  $\overline{e_n^2}$  instead of  $W_i(f)$ . The series noise voltage always appears as linearly added to the signal, therefore the signal-to-noise ratio is degraded.

The required connection between the magnitude of noise current fluctuation at the amplifying element output and the magnitude of the voltage source  $e_n$  at the control electrode (see Fig. 3) to produce such noise current fluctuations is given by the mutual conductance  $g_m$ .

Mutual conductance is defined as a ratio of a change in the controlled current through the device and a change in the control electrode voltage.

$$g_m = \frac{\Delta I_{DEV}}{\Delta V_{CON}} \quad . \quad (3-1)$$

(This is correct for all input voltages, not only for noise.)

Using Eq. (3-1), we can write

$$e_n = i_n / g_m \quad , \quad (3-2)$$

where  $i_n$  is the output current noise source of the device.

The next important source of noise is due to the fact that the input impedance of the amplifying element is not infinite. Consequently, there is some current  $i_g$  flowing into (or out) of the control electrode. This current exhibits fluctuations  $\Delta i_g$  and produces corresponding voltage fluctuations  $Z_s \Delta i_g$ , where  $Z_s$  is the impedance of the source. The control electrode "leakage" current noise can be equivalently represented by a parallel noise current generator  $i_{np}$  (Fig. 4a).

For detectors where  $Z_s$  is often capacitive and signal can be represented by a current source parallel to the capacitance, it is convenient to replace the series noise source  $e_{ns}$  with an equivalent current source  $i_{ns} = e_{ns}/Z_{in}$  parallel to the input impedance so signal and all noise sources are situated at the same place at the input and are directly comparable (Fig. 4b).

Summarizing, we have seen that the charge controlled amplifying device has two uncorrelated noise sources to be considered with it; the series noise source  $e_n$  and the parallel noise source  $i_n$ . Transforming the series noise as shown in Fig. 4b, we see that the effect of the parallel noise is independent of the detector impedance while the effect of the series noise is inversely proportional to the detector impedance. (There is some correlation, its effects are negligible in first order analysis.)

## 2. Series Noise in Bipolar Transistors

Series noise analysis of a bipolar transistor is particularly clean and instructive. The operation of a bipolar transistor is generally known,<sup>14</sup> but we have the feeling that in order to prevent some misunderstanding we will outline some of its properties important for our analysis here.

Let us start by considering a forward biased n<sup>+</sup>-p junction (i.e., very heavily doped n-type on lightly doped p-type semiconductor semiconductor). For an idealized case under consideration, the forward current is given by

$$i_f = i_o (e \frac{qV}{kT} - 1) \approx i_o e \frac{qV}{kT} \quad , \quad (3-3)$$

where  $i_o$  is the reverse saturation current

V - voltage applied to the junction

q - charge of an electron

k - Boltzmann constant

T - absolute temperature.

In the vicinity of the junction in the p-type part of the device, the bulk of the current is carried by free electrons and their density falls off exponentially with the distance x (see Fig. 5a)

$$n_e(x) = n_o \cdot e^{-\frac{x}{l_e}}, \quad (3-4)$$

where  $l_e$  is the electron diffusion length in the p-type material. If the p region is made extremely thin (compared to  $l_e$ ), and it is followed by a reverse biased, lightly doped n region (Fig. 5b), then electrons injected from the emitter into the base have a high probability of reaching the edge of the collector-to-base depletion layer and being collected.

In the usual way, we will define  $\alpha$  and  $\beta$  of a transistor. If the emitter current is  $I_E$ , then some fraction  $\alpha$  ( $0 < \alpha < 1$ ) of it will be collected and will appear in the collector circuit as the collector current  $I_c$ . The difference between  $I_E$  and  $I_c$  constitutes the base current.

$$I_c = \alpha I_E$$

$$I_B = (1-\alpha) I_E$$

$$\Rightarrow I_c = \left( \frac{\alpha}{1-\alpha} \right) I_B = \beta I_B \quad (3-5)$$

The quantity  $\beta$  relating  $I_c$  and  $I_B$  is called the current gain of the transistor and can be made large (up to  $\sim 1000$  in today's transistors).

Let us calculate  $g_m$  of a transistor in grounded emitter configuration, i.e., the transistor's base is the control electrode and the modulated device current is the collector current.

If the base voltage is  $V_B$ , Eqs. (3-3) and (3-5) give

$$I_c = \alpha I_o e^{\frac{qV_B}{kT}},$$

or by differentiation with respect to  $V_B$

$$\frac{dI_c}{dV_B} = \frac{q}{kT} I_c$$

By using Eq. (3-1),  $g_m = \frac{q}{kT} \cdot I_c (= \frac{I_c}{25m \text{ Volts}})$  (3-6)  
at room temperature.

The value of mutual conductance is directly traceable to the Boltzmann factor of Eq. (3-3), i.e., it is controlled by the thermal energy distribution of carriers at junction (ambient) temperature. It follows that no charge controlled amplifier in which the carriers are at the same temperature as the transistor can have higher  $g_m$  than a transistor operating at the same current.

We can now easily derive the formula for the series noise of a transistor. The collector current  $I_c$  exhibits full shot noise (since the carriers constituting this current are minority carriers in the transistor base and at the collector depletion layer are being collected independently).

Recalling (2-14), we can write the power current density as

$$w_i = \overline{i^2} = 2qI_c , \quad (3-7)$$

or by using (3-2) and (3-6), we can express series noise power density

$$\overline{e_n^2} = \frac{1}{g_m} \cdot 2kT . \quad (3-8)$$

It is convenient to express the noise  $e_n$  as being generated by a resistor  $R_s$ . Comparing (3-8) with formula (2-22), we obtain

$$R_s = \frac{1}{2} \frac{1}{g_m} . \quad (3-9)$$

In practice, formula (3-9) holds down to  $R_s \approx 50\Omega$ . Below that value, the other series noise sources became important. The most evident is the actual series resistance in the base structure and the connecting leads. In the absence

of excess noise their noise resistance is equal to the ohmic resistance, so we can write

$$R_s = \frac{1}{2} \frac{1}{g_m} + R_{ex} \quad (3-10)$$

Before concluding this section, let us study the base diffusion capacitance of the bipolar transistor. This capacitance is the dominant base-to-emitter capacitance and plays an important role in the charge detection by bipolar transistors. It is not a geometrical capacitance, but a consequence of the transistor amplification mechanism.

The concentration of the minority carriers in the base is proportional to the collector current. The constant of proportionality is just the time  $\tau$  necessary for carriers to diffuse across the base.

$$Q_B = \tau \cdot I_c \quad (3-11)$$

The same amount of charge of the opposite polarity has to be supplied to the base to be electrically neutral. This charge of base majority carriers must be supplied through the base lead. That means that the change of the charge stored in the base can be related to the change in the flow of the collector current and via mutual conductance back to the change of the base voltage. We can write

$$\Delta Q_B = \tau \cdot \Delta I_c = \tau \cdot g_m \cdot \Delta V_B$$

or  $\Delta Q_B = C_D \cdot \Delta V_B$ ,

where  $C_{DIF} = \tau g_m$  . (3-12)

The total base-to-emitter capacitance in a real bipolar transistor is the sum of the dominant diffusion capacitance (2-12) and all geometrical capacitances.

### 3. Series Noise in Field Effect Transistors

The amplification action of a field effect transistor is based on voltage control of majority carriers flow through a semiconductor channel. The effective width, and with it the conductance of the channel is varied by reverse potential of the gate-channel junction (Fig. 6). The gate is the control electrode. The two channel ends are called drain and source. The drain is that channel-end which is at higher reverse voltage with respect to the gate. There is usually no structural difference between drain and source.

The basic noise source in the field effect transistor is the thermal noise of the conducting channel. It can be shown that this noise is equivalent to a noise current generator connected between the source and drain of magnitude  $^{15}$

$$\overline{i_n^2} \cong 0.7 \times 4 kT g_m \quad . \quad (3-13)$$

From this relation, we can write the series voltage noise and the series noise resistor

$$R_s \cong \frac{.7}{g_m} \quad . \quad (3-14)$$

Comparing (3-14) and (3-9), we see that the series noise of junction field effect transistors and bipolar transistors are comparable.

Vacuum tubes (mentioned here mainly for historical reasons) have higher series noise. For vacuum triode we have

$$R_s = \frac{a}{g_m} \quad \text{with } 2.5 < a < 4 \quad . \quad (3-15)$$

#### 4. Parallel Noise

As has been mentioned in Section 1, any current flowing to or from the control electrode produces the noise. This was represented in Fig. 4. Assuming the full shot noise of the central electrode current, the noise current power density is given by

$$\overline{i_n^2} = 2qI_g$$

where  $I_g$  denotes the current flowing through the control electrode. For junction field effect transistors, the  $I_g$  values are in  $10^{-10} - 10^{-14}$  A range. The bipolar transistor has a much higher base current  $I_g = I_c/3$ , at least in the  $10^{-6}$  A region.

The parallel noise is therefore much smaller for a FET. In many cases the parallel noise is important and then a field effect transistor is the best amplifying device.

We will return to the question of the most suitable amplifiers after determining the relative importance of noise sources for the charge measurement.

#### 5. Optimum Filtering for Charge Measurement <sup>2,16</sup>

Now we are going to answer one of the fundamental questions we asked in the Introduction. What is the minimal charge that we can detect on the output electrode of a detector? As a criterion for this minimal charge, we will adopt equivalent noise charge (ENC). ENC is the quantity of charge which, appearing on the detector, would give an output pulse of height equivalent to rms (root mean square) of the noise. (To see a signal on the screen of an oscilloscope, it has to be at least three ENC if the scope is independently triggered. If we use the oscilloscope in the self-triggered mode, the signal should be at least 5 ENC.)

The basic equivalent circuit of a detector and amplifier is shown in Fig. 7. The signal is represented as a current source in parallel with the total input capacitance. Noise sources were represented by the current generators also in parallel with the total input impedance. The amplifier is shown as an ideal noiseless amplifier followed by additional amplifiers which includes the filtering or "pulse shaping". It is important that filtering is performed after amplification so that no noise is added by dissipative filter components.

The whole linear system from the detector to the output of the filter is described by its impulse response  $h(t)$  in the time domain or equivalently by its transfer function  $H(f)$  in the frequency domain. We assume that the shape of the signal  $s(t)$  is known and we are measuring the amplitude  $A$  of the signal  $A \cdot s(t)$  in the presence of noise. Our problem is to find the pulse response of the system  $h(t)$  which gives the best measurement of  $A$ , i.e., we have to find a filter which maximizes the signal-to-noise ratio at the output of the system at some appropriate "readout" time  $t_o$ .

The filter output at  $t_o$ , due to the signal  $A \cdot s(t)$ , is determined by the Fourier transform of the filter output frequency spectrum  $A S(f) \cdot H(f)$ . [ $S(f)$  is the Fourier transform of the signal  $s(t)$ ].

$$g(t_o) = A \int_{-\infty}^{\infty} S(f) H(f) e^{j\omega t_o} df \quad (3-16)$$

The fluctuations at the output due to the noise power density

$$2w(f) = \overline{i_{pn}^2} + \overline{i_{sn}^2} \quad \text{at the input according to Eq. (2-11) are equal}$$

$$\overline{(\Delta v)^2} = \int_0^{\infty} 2w(f) \cdot |H(f)|^2 df = \int_0^{\infty} w(f) \cdot |H(f)|^2 df \quad (3-17)$$

The signal-to-noise ratio  $\rho$  to be maximized is

$$\rho = (g(t_0))^2 / (\Delta v^2) \quad (3-18)$$

Let us recall the Schwarz inequality for complex functions

$$|\int u \cdot v \, df|^2 \leq \int |u|^2 \, df \cdot \int |v|^2 \, df \quad (3-19)$$

The equality holds only if

$$u(f) = a \cdot v^*(f) \quad (3-20)$$

If we substitute  $u(f) = H(f) \cdot w^{\frac{1}{2}}(f)$  and

$$v(f) = A(S(f)/w^{\frac{1}{2}}(f)) \cdot e^{j\omega t_0},$$

and apply inequality (3-19) to the ratio (3-18), we obtain

$$\rho \leq A^2 \int_{-\infty}^{\infty} \frac{|S(f)|^2}{w(f)} \, df, \quad (3-21)$$

and  $\rho$  is maximum, i.e., equality sign holds if

$$H(f) = a \frac{S^*(f)}{w(f)} \cdot e^{-j\omega t_0} \quad (3-22)$$

It is convenient to split the optimum filter  $H(f)$  into two parts. The first part  $H_1(f)$  transfers the noise spectrum  $w(f)$  into a white spectrum  $w_0$  and the second part  $H_0(f)$  completes the maximization of signal-to-noise ratio. If we denote by  $S_1(f) = S(f) \cdot H_1(f)$  the signal Fourier transform modified by the filter  $H_1(f)$ , we can directly apply relation (3-22) to find the optimizing filter  $H_0(f)$ .

$$H_0(f) = S_1^*(f) \cdot e^{-j\omega t_0} \quad (3-23)$$

where we have chosen  $a = w_0$ .

The filter pulse response, determined as the Fourier transform of the transfer function (3-23), is

$$h_o(t) = s_1(t_o - t) \quad , \quad (3-24)$$

which is the mirror image delayed by  $t_o$  of the signal  $s_1(t)$  at the input of the optimizing filter. The filter is called a matched filter, since it is matched to a particular waveform.

The maximum signal-to-noise ratio for the matched filter with a white noise source at the input obtained from Eq. (3-22) is

$$\sigma_{\max} = \frac{A^2 \int_{-\infty}^{\infty} |s_1(f)|^2 df}{w_o} = \frac{A^2 \int_{-\infty}^{\infty} (s_1(t))^2 dt}{w_o} = \frac{E_s}{w_o} \quad . \quad (3-25)$$

Because the numerator is equal to the energy  $E_s$  of the input signal. Eq. (3-25) shows that the matched filter uses the full signal energy. (It is not the energy of a detecting particle nor the energy loss of a particle in the detector.) Eq. (3-25) allows us to define the noise equivalent energy (NEE) =  $w_o$ . This quantity is directly related to the above defined noise equivalent charge.

The optimum filter pulse response (Eq. [3-24]) can be found also directly by a similar analysis in the time domain. Let us show that the output waveform of the matched filter is symmetrical with respect to measured time  $t_o$  so any practical output waveform has a maximum at that time.

The output of the filter can be obtained as the convolution integral of the signal and the filter pulse response (3-24).

$$g(t) = \int_{-\infty}^{\infty} s_1(\lambda) h_o(t-\lambda) d\lambda = \int_{-\infty}^{\infty} s_1(\lambda) s_1(t_o - t + \lambda) d\lambda$$

or for  $t = t_o + \tau$ ,

$$g(t_o + \tau) = \int_{-\infty}^{\infty} s_1(\lambda) s_1(\lambda - \tau) d\lambda = \int_{-\infty}^{\infty} s_1(\lambda' + \tau) \cdot s_1(\lambda') d\lambda' = g(t_o - \tau) \quad . \quad (3-26)$$

Before applying the derived formulae to the practical detection, let us study a closely related problem of extracting the time information from a signal in the presence of noise.

#### 6. Optimum Filtering for Time Measurement

Let us consider the same basic equivalent circuit (shown in Fig. 7) as in the previous section. The input signal  $A \cdot s(t)$  has a known shape  $s(t)$  with an unknown amplitude  $A$ . To measure the time related with the time of occurrence of the signal, we have to maximize the signal slope to noise ratio. We decided to perform the measurement at some later time with respect to the occurrence of the signal.

The filter output at  $t_o$  is again given by relation (3-16) so the slope of the output waveform at the time  $t_o$  is

$$g'(t_o) = \int_{-\infty}^{\infty} j\omega s(f) H(f) e^{j\omega t_o} df \quad . \quad (3-27)$$

The fluctuations at the output are given by (3-17), so the signal slope-to-noise ratio to be maximized is,

$$\left( \frac{1}{\delta t} \right)^2 = \rho = \frac{(g'(t))^2}{(\Delta v^2)} \quad . \quad (3-28)$$

Applying the Schwarz inequality we obtain,

$$\rho \leq A^2 \int_{-\infty}^{\infty} \frac{|s(f)|^2}{w(f)} \quad , \quad (3-29)$$

and the equality sign holds if

$$H(f) = a \frac{-j\omega s^*(f)}{w(f)} \cdot e^{-j\omega t_o} \quad . \quad (3-30)$$

We can again split the optimum filter into two parts; first the prewhitening filter  $H_1(f)$  and the second filter  $H_o(f)$  for which we can apply the relation (3-30), put  $a = -w_o$  and obtain

$$H_o(f) = juS_1^*(f) \cdot e^{-jw_o t_o} \quad . \quad (3-31)$$

The filter impulse response is

$$h_o(t) = \int_{-\infty}^{\infty} H_o(f) \cdot e^{jw_o t} df = \frac{d}{dt} \int_{-\infty}^{\infty} S_1^*(f) \cdot e^{jw_o(t_o-t)} df \quad ,$$

or taking complex conjugated equation we can write

$$h_o(t) = \frac{d}{dt} S_1(t_o-t) \quad . \quad (3-32)$$

Comparing Eqs. (3-32) and (3-24), we see that the pulse response of the optimum filter for timing measurements in the presence of white noise is the derivative of the pulse response of the optimum filter for the amplitude measurement. This is true for any noise power spectrum. (It follows from relations (3-30) and (3-22).)

For a matched filter, we can find the  $\rho_{\max}$  which corresponds to the minimized error in the timing  $\delta t$ . From Eq. (3-29), written for the case of the white noise and Eq. (3-31), we obtain

$$\left( \frac{1}{\delta t} \right)^2 = \rho_{\max} = \frac{A_o^2}{w_o} \int_{-\infty}^{\infty} |juS_1(f)|^2 df = \frac{A_o^2}{w_o} \int_{-\infty}^{\infty} (S_1'(t))^2 dt \quad . \quad (3-33)$$

We can see that the matched filter for the best timing uses in an optimal way the derivative of the signal waveform to extract the time occurrence of an event. We will show that at the measured time  $t_o$ , the resulting waveform  $g(t)$  crosses 0. Hence, the best timing results from a zero crossing method which has the well known antiwalk properties as well. Let us show something slightly more

general, namely that the output waveform  $g(t)$  is an odd function of time around the time  $t_0$ . By analogy with the previous section, we can write

$$g(t) = \int_{-\infty}^{\infty} s_1(t-\lambda) \cdot h_0(\lambda) d\lambda = \int_{-\infty}^{\infty} s_1(t-\lambda) \frac{d}{d\lambda} s_1(\lambda) d\lambda =$$
$$[s_1(t-\lambda) \cdot s_1(\lambda)]_{-\infty}^{\infty} - \int_{-\infty}^{\infty} \frac{d}{d\lambda} s_1(t-\lambda) \cdot s_1(\lambda) d\lambda .$$

Where we used integration per parts. The signal has to vanish for  $t \rightarrow \pm \infty$  and by simple manipulation similar to those for obtaining relation (3-26) we can show that

$$g(t_0 - \tau) = -g(t_0 + \tau) . \quad (3-34)$$

#### 7. Feasibility of Detecting the Macroscopic Electromagnetic Field of a Relativistic Particle

As an illustration of the fundamental character of the previous sections, let us analyze the detectability of a relativistic particle via its macroscopic electromagnetic field.

The basic motivation for wishing to directly detect the electromagnetic field of a relativistic particle is the fact that the shape of the field depends strongly on the  $\gamma$  ( $= \frac{E}{m}$ ) of the particle. Let us assume that with a suitable probe we can extract all electromagnetic field energy  $U$  contained outside a cylinder of radius  $r_0$  around the trajectory of the particle. Let us also assume that the full extracted energy is brought as a signal to the input of an amplifier. Electromagnetic energy  $U$  can be expressed as

$$U = \frac{q^2 \gamma}{32 \epsilon_0 r_0} \quad (3-35)$$

where  $r_0$  and  $\gamma$  were defined above and

$q$  - is the charge of the particle

$\epsilon_0$  - is the absolute dielectric constant ( $\frac{1}{4\pi\epsilon_0} = c^2 \times 10^{-7}$ ) .

We see that the extracted energy is proportional to the particle energy  $\gamma$ .

We can calculate the value of  $\gamma$  for which the signal energy equals the noise power of the amplifier. Without going into detail, we know that the noise power must be according to the Eq. (2-21) of the order of  $kT$ . (This time we deal with the real physical power and energy instead of square of current or voltage as usual.) Calculating  $\gamma$  we obtain

$$\gamma_{\min} = \frac{32\epsilon_0 r_0 kT}{q^2} \quad (3-36)$$

for  $r_0 = 1$  cm,  $kT = 25$  meV, we obtain  $\gamma_{\min} \approx 5 \times 10^5$  which is too large for the method to be practical. (It can be shown that the length  $\ell$  of the "extracting probe" is  $\ell \approx \gamma \cdot r_0$ , which makes the detector several km long.)

Part IV. Charge Measurement for Practical Detection Systems.

1. Charge measurement under additional constraints. 15-19

Let us return to the measurement of the charge signal. The input circuit with the parallel and series noise sources is shown in Fig. 7. Let us calculate the impulse response of an amplitude matched filter for the detector signal being an impulse of current containing total signal charge  $Q$ .

$$s(t) = Q \cdot \delta(t) \quad (4-1)$$

The total noise current power density  $2w_i(f)$  at the input is

$$2w_i(f) = \frac{4kT}{R_p} + \frac{4kTR_s}{|Z|^2} = 4kT \left( \frac{1}{R_p} + \omega^2 C_t^2 R_s \right) \quad (4-2)$$

where  $R_p$  - is the equivalent parallel noise resistance

$R_s$  - is the equivalent series noise resistance

$Z = \frac{1}{j\omega C_t}$  is the input impedance

$C_t$  - total input capacitance

We can rewrite equation (4-2) using the "noise corner time constant"

$$\tau_c = C_t \cdot \sqrt{R_s R_p} \quad \text{as}$$

$$2w_i(f) = \frac{4kT}{R_p} (1 + \omega^2 \tau_c^2) \quad (4-3)$$

To obtain white noise we have to pass the noise and signal through a prewhitening filter with the transfer function  $H(f) = \frac{1}{1+j\omega\tau_c}$ .  
The noise then becomes white:

$$2w_o = \frac{4kT}{R_p} \quad (4-4)$$

and the signal becomes

$$s_1(t) = \begin{cases} \frac{Q}{\tau_c} \cdot e^{-\frac{|t|}{\tau_c}} & t > 0 \\ 0 & t < 0 \end{cases} \quad (4-5)$$

Now we use equation (3-18) to calculate the square signal to noise ratio for the matched filter.

$$\sigma_{\text{max}}^2 = \frac{Q^2}{4kTC_t \sqrt{\frac{R_s}{R_p}}} \quad (4-6)$$

From equation (4-6) we obtain the minimal equivalent noise charge

$$(\text{ENC})_{\text{opt}}^2 = 4kTC_t \cdot \sqrt{\frac{R_s}{R_p}} \quad (4-7)$$

Under the most favorable conditions (usually not present in particle physics experiments) ENC can be below 30 electron charges (as low as  $\sim 10$ ). Equation (4-5) shows that the signal waveform  $s_1(t)$  has an infinite duration, so the matched filter requires an infinite delay to respond in a symmetrical way. Such a filter can never be built in practice.

The impulse response of an ideally matched filter with the time origin shifted in such a way that the waveform peaks at zero is shown in Fig. 8.  $\left( h(t) = e^{-\frac{|t|}{\tau_c}} \right)$

We notice that for  $|t| > 2\tau_c$  the ideal pulse response is not very different from 0. Since the ideal response corresponds to a stationary function a small change in the shape of the function should not increase the ENC too much. Intuitively we expect that the exact shape of the impulse response is not very critical to obtain low noise performance of the system. For any filter we can define the ratio  $\eta$  of the ENC obtained with that filter and the ENC of the matched filter. The ratio  $\eta$  measures how much of the noise performance we are sacrificing with a given filter.

$$\eta_{\text{Filter}} = \frac{(\text{ENC})_{\text{Filter}}}{(\text{ENC})_{\text{Matched}}_{\text{OPT}}} \quad (4-8)$$

Let us calculate this ratio for a filter with the triangular impulse response shown in Fig 8.

$$h(t) = 1 - \frac{|t|}{t_m} \quad |t| \leq t_m$$

$$= 0 \text{ otherwise} \quad (4-9)$$

The calculation of the ENC can be conveniently done in the time domain. For the parallel noise we can use Campbell's theorem which is directly applicable because the noise source can be thought as a sequence of random impulses. (See equation (2-5) )

The contribution of the parallel noise into ENC can be written as

$$\overline{q_{np}^2} = \frac{2kT}{R_p} \cdot \int_{-\infty}^{\infty} (h(t))^2 dt \quad (4-10)$$

To write the expression for the contribution of the series noise source  $e_n$  we have to take into account that the voltage source is placed differently at the input and therefore its impulse response is different from the impulse response of the signal. It is easy to see that the impulse response of the voltage source is  $C_t \frac{d}{dt} h(t)$  where  $h(t)$  is the impulse response due to a unit signal charge. So substituting for  $h(t)$  into equation (2-5) we can express the series noise as:

$$\overline{q_{ns}^2} = 2kTR_s C_t^2 \int_{-\infty}^{\infty} \left( \frac{d}{dt} h(t) \right)^2 dt \quad (4-11)$$

Straight forward evaluation of integrals gives the following results.

$$\overline{q_{np}^2} = \frac{2kT}{R_p} \cdot \frac{2}{3} t_m \quad (4-12)$$

$$\overline{q_{ns}^2} = 2kT \cdot R_s \cdot C_t^2 \cdot \frac{2}{t_m} \quad (4-13)$$

The ENC ratio  $n$  of the triangular filter has a minimum for  $t_m = \sqrt{3} \tau_c$  and its minimal value is about 1.08. We see that the noise of the system is only about 8% above the theoretical limit.

We note also that for  $t_m = \sqrt{3} \tau_c$ , parallel and series noise contribute the same amount to the total noise. From the equations (4-12) and (4-13) we see that the contributions of the parallel noise increases with the time  $t_m$  while the contribution of series noise decreases with the shaping time  $t_m$ .

This general behavior is independent of the precise shape of the impulse response of the system and can be intuitively understood. The parallel noise source feeds the same impedance as the detector signal. Since the noise can be thought of as a sequence of random impulses a longer shaping time allows more noise pulses to contribute to the output. This process is similar to the random walk which also gives a deviation proportional to the square root of the processing time.

Random voltage pulses from the series noise source are equivalent to current "doublet" pulses when transformed to the same point at the filter input as the detector signal. Current "doublet" pulses can be visualized as derivatives of the Dirac delta function  $\delta(t)$ , i.e., as a pair of impulses of opposite polarity separated by  $\Delta t$ , where the product of the pulse area and  $\Delta t$  equals unity. A doublet passing through a flat part of the impulse response function gives zero contribution to the output because the contribution from the first pulse of the doublet is exactly canceled by the contribution from the second pulse. Only when impulse response function has a steep part two pulses of the doublet produce different responses at the measuring time giving

a non zero contribution to the noise.

The triangular pulse response can be realized in practice. It has the final time duration  $2t_m$ . The detection rate of the system is limited to  $1/2t_m$  level. The choice of  $t_m \approx \tau_c$  which gives the best noise performance is often impossible in high energy physics experiments because the noise corner time constant is too long ( $\sim 10^{-4}$  to  $10^{-5}$  sec) and the resulting rate limitation unacceptable.

If the duration of the system impulse response is much shorter than the noise corner time constant the contribution of the parallel noise (equations 4-10, 4-12) is negligible and the ENC is given by equation (4-11) and (4-13).

We would like to stress that the length of the filter pulse response  $t_m$  has nothing to do with the speed of charge removal from the detector capacitance. We will return to this point later in the section dealing with amplifiers.

Up to now we have assumed the detector input is a current impulse  $\delta(t)$ , or practically a pulse of a duration much shorter than the filtering time  $t_m$ . If the signal has a duration which can fluctuate, the system impulse response has to have a flat top of the duration about equal to the longest signal charge collection time. If the signal duration does not fluctuate and the total charge must be measured the total length of the system impulse response has to be longer than the duration of the signal. This can limit the rate performance of the system. If the charge collection time is too long, it is possible to have a system in which the duration of the output signal is shorter than the duration of the signal at the input.<sup>20</sup> In these cases only a part

of the signal is used so ENC is usually much larger.

## 2. Optimal Matching between Detector and Amplifier

We are going to answer another important question asked in the Introduction. What is the "best" amplifying device to be used with a given detector?

Let us assume that the duration of the output waveform is limited by the rate considerations so the total noise is due to the series noise source of an amplifying device. Also we will assume a triangular shaping of duration  $2t_m$  at the base (this assumption is not essential, we could do the analysis with any suitable shaping function) so the equivalent noise charge (ENC) can be written as

$$(ENC)^2 = 4kTR_s (C_A + C_D)^2 / t_m \quad (4-14)$$

The relation (4-14) is exactly the relation (4-13) where we have written explicitly the total input capacitance as a sum of detector and amplifier capacitance. From section III-3

know that the noise series resistance  $R_s$  can be written as

$$R_s = \frac{a}{g_m} \quad (4-15)$$

where  $a = .5$  and  $.7$  for bipolar and junction field effect transistors respectively. ( $a = 2.5-4$  for vacuum triodes).

To reduce the noise series resistance  $R_s$  we have to increase the mutual conductance of the amplifier. This can be in principle achieved by paralleling several amplifying devices. If we call the number of amplifying devices  $n_A$  the equation (4-15) can be written as

$$(ENC)^2 = 4kT \frac{a}{n_A g_m} \left( n_A C_A + C_D \right)^2 / t_m \quad (4-16)$$

(We must not forget that paralleling of amplifiers increases the total input capacitance.) Equation (4-16) has a minimum for

$$n_A = \frac{C_D}{C_A} \quad (4-17)$$

which gives the  $(ENC)_{opt}$  as

$$(ENC_{opt})^2 = 16kTC_D \frac{a C_A}{g_m} \cdot \frac{1}{t_m} \quad (4-18)$$

The equation (4-17) says that the lowest noise can be achieved when the capacitance of the amplifier equals the detector capacitance. It is easy to see that under this condition the maximum of the signal energy from the detector electrode is transferred to the preamplifier so the condition (4-17) also maximizes the signal energy  $E_s$  as defined in (3-25). (To be really precise, if we consider a broader class of connections between the detector and the preamplifier by using an "ideal" inductance we can increase the signal energy  $E_s$  at the preamplifier and hence improve ENC up to the factor  $\sqrt{1.5}$ . The stray capacitances of real inductors make this solution impractical in cases of low amplifier capacitance.)

For the optimal capacitance matching between the detector and the preamplifier the  $ENC_{opt}$  as given by (4-18) depends on the amplifying device only via the ratio  $C_A/g_m$  ( $a$  - is practically the same for bipolar and FET transistors)

Discussing the bipolar transistor (in Part II.3) we have seen that the ratio of the predominant, i.e., diffusion capacitance  $C_{diff}$  and the

mutual conductance  $g_m$  is the transit time  $\tau_A$  of the minority carriers across the base of a bipolar transistor. For a junction field effect transistor we can find a very similar relation. The  $C_A/g_m$  ratio is practically equal to the transit time  $\tau_A$  of the majority carriers through the FET channel. The transit time  $\tau_A$  is a measure of the reciprocal of the unity gain frequency  $f_T$  of the amplifying device.

Using  $f_T$  we can rewrite the equation (4-18) as

$$(ENC_{opt})^2 = 16kTC_D \frac{a}{f_T t_m} \quad (4-19)$$

We have found that the unity gain frequency  $f_T$  of an amplifying device is the only parameter defining the noise performance of the signal processing system under all the above assumptions. Devices intended for the use at the highest frequencies, which have the highest  $f_T$ , give also the best noise performance. The state of the art silicon junction FET have  $f_T$  around 1GHz, bipolar transistors up to 5GHz and the  $f_T$  of GaAs FETs can be as high as 20 GHz. The extremely small  $C_A$  ( $\sim 1\text{pF}$ ) which makes the capacitance matching difficult particularly at higher frequencies and the excess flicker noise up to 100 MHz have prevented the wide use of GaAs FETs as amplifying devices so far. The shot noise of the base current (source of the parallel noise) limits the use of bipolar transistors for the applications with shaping time shorter than few  $10^{-8}$  sec. Silicon junction FETs give the best performance for  $t_m > 10^{-7}$  sec. The situation is summarized in Fig. 9, where the ENC versus  $t_m$  is plotted for several detectors and amplifying devices.

Optimal capacitance matching can be obtained by several means. There is a wide scale of FETs with different  $C_A$ . If necessary several FETs are

usually used at the preamplifier input. A bipolar transistor offers a range of  $C_A$  as a function of its standing current  $I_C$  (equation 3-6).

For a detector with a very high capacitance the only practical matching is by transformer matching. Liquid argon calorimeters provide a very important example of a detector for which the transformer matching is a necessity.

### 3. Signal Processing for Liquid Argon Calorimeters

#### 3.1 Signal Formation in Liquid Argon Calorimeters <sup>25</sup>

As it was mentioned in Part I calorimeters measure the energy and position of particles through the process of total absorption. In this process, secondary particles are generated, which themselves react inside the detector. Finally, almost all of the incident energy of the particle will be converted into heat (hence the name "calorimeter"). The increase of the temperature is obviously too small to be measured and ionization or excitation in a detection medium are used to produce a measurable signal. The use of calorimeters is increasing in high energy physics experiment for the following reasons (i) it is the only way to detect and measure the energy of the neutral particles, (ii) the accuracy of the energy measurement due to detector factors improves as the square root of the particle energy, (iii) signal from calorimeter which carries the information about the energy deposition is available at very short time scale for the event selection (triggering).

A liquid argon calorimeter uses liquid argon as a detection medium and measures total energy via the ionization sampled in the liquid argon.

This technique is now very popular mainly for electromagnetic calorimeters. Roughly speaking, the liquid argon calorimeter consists of parallel conducting plates (electrodes) a few mm thick and a few mm apart immersed in liquid argon. Advantages of using ion chamber approach for the energy sampling are:

- (i) Simple high-accuracy charge calibration is possible.
- (ii) The detector has uniform sensitivity over the whole detector volume.
- (iii) The detector can be practically arbitrarily subdivided and so provide position resolution.
- (iv) Large detectors of arbitrary shape can be constructed.

On the negative side we have the rate limits given by the mobility of electrons in the liquid argon and the requirements of the operation at cryogenic temperatures.

To study the signal from a liquid argon ion chamber we have to take into account that the liquid argon is a single carrier medium as far as charge collection is concerned. The positive ions do not contribute to the signal charge due to their very low mobility. The basic relations for the current and charge waveform for planar electrode geometry are illustrated in Fig. 10. Part a of Fig. 10 shows the current and charge in an external circuit for one ion pair. The current, due to one carrier (electron), is determined as  $q/t_d$ , by the drift time  $t_d$  across the gap. (See also Appendix 1). The charge measured in the external circuit is determined by the ratio of the distance traversed by electron ( $d-x$ ) and the electrode spacing  $d$ ,

$$Q_s(x) = q \frac{d-x}{d} \quad (4-20)$$

If ionization is uniformly distributed across the inter-electrode gap as in the case of ionization resulting from a passage of a charged particle the induced signal equals one half of the ionization charge on the average. (Fig. 10b). The resulting current and charge waveforms are different for localized and for uniform ionization across the gap. We can note that for uniform ionization three quarters of the observable charge is collected in one half of the drift time across the gap. The saturated drift velocity of the electron in liquid argon is around 5 mm/  $\mu$ s as so for a typical gap width of 2 mm the collection time is about 400 ns. Standard rate considerations in high energy physics experiments limits the resolving time to the 1  $\mu$ s region. The signal current cannot be considered as an impulse compared with the duration of the filter, so the output amplitude due to a unit charge signal will be smaller than the amplitude due to the unit charge impulse.

Let us estimate the amount of charge produced by absorption of, let us say, an electron of total energy E. Liquid argon samples typically between (10-20) % of the total energy. The energy loss to produce 1 ion pair in liquid argon is .25eV. Taking into account that we see only one half of charge the energy loss per observed electron is 50 eV. If we sample 10% of the energy we need effectively 500 eV to produce an electron charge. In other unit, one GeV of energy deposited inside the calorimeter gives a signal of .32 pC.

### 3.2 Detector-Preamplifier Matching for Liquid Argon Calorimeters.

In the preceding section we have determined the quantity of charge in typical sampling liquid argon calorimeters. Calorimeters involve a

117/2

large number of plates and, therefore, a large capacitance. The capacitance of a bigger section of a calorimeter  $C_d$  can reach up to  $10^{-7}$  F. Connected directly to the gate of a junction FET with an input capacitance of  $10\text{pF}$  the amplifier would see only one part in  $10^4$  of the total charge. Paralleling of  $10^4$  input FETs is not a practical solution; therefore, the capacitance matching has to be accomplished by a transformer.

The detector-amplifier circuit configuration with a matching transformer, provisions for detector biasing, and charge calibration is shown in Fig. 11. Due to the large detector capacitance the blocking capacitor  $C_b$  cannot be much bigger than the detector capacitance  $C_d$  so it has to be taken into account. An equivalent circuit showing the essential elements for noise analysis is given in Fig. 12. The shape of the impulse response is also shown in Fig. 12. Bipolar shape of the total duration  $\lambda$  is assumed. Bipolar shape is a necessity if the shift in the energy spectrum due to the pile-up has to be avoided.

We are going to calculate ENC of the configuration shown first assuming that the signal is a current impulse. We will transform the signal charge and the detector capacitance and the blocking capacitance to the secondary of the matching transformer and use equations (4-10) and (4-11) with an impulse response as shown in Fig. 12.

We can write the following equations for the transformed signal charge and capacitancies:

$$Q^{(T)} = \frac{Q}{n}$$

$$C^{(T)} = \frac{C}{n^2} \quad (4-21)$$

where  $n$  is the transformation ratio.

The transformed blocking capacitance  $C_b^{(T)}$  acts as a capacitive divider between transformed detector capacitance  $C_d^{(T)}$  and amplifier capacitance  $C_A$ . Hence, the relation between the transformed charge at the detector  $Q^{(T)}$  and the voltage at the preamplifier input  $V_{PR}$  can be written as

$$Q^{(T)} = \left( C_A \left( 1 + \frac{C_d}{C_b} \right) + C_d^{(T)} \right) \cdot V_{PR} \quad (4-22)$$

which allows us to write equation (4-11) in the following way:

$$\overline{q_{ns}^{(T)2}} = 2kT R_s \left( C_A \left( 1 + \frac{C_d}{C_b} \right) + C_d^{(T)} \right) \int_{\infty}^{\infty} \left( \frac{d}{dt} h(t) \right)^2 dt \quad (4-23)$$

or after substituting for transformed quantities from (4-21) and evaluating the integral we can express the equivalent noise charge (not transformed) of the series noise source as

$$(ENC)_s^2 = 32kT R_s \frac{1}{\lambda} \left( n C_A \left( 1 + \frac{C_d}{C_b} \right) + \frac{C_d}{n} \right)^2 \quad (4-24)$$

In a similar way we can find the expression for the equivalent noise charge due to parallel noise sources.

$$(ENC)_p^2 = \frac{2}{3} kT \frac{\lambda}{R_p} \cdot n^2 \left( 1 + \frac{C_d}{C_b} \right)^2 \quad (4-25)$$

For the short duration of the filter impulse response  $\lambda$  as used in practical calorimeters the parallel noise can be kept small compared with the series noise and can be neglected. The total ENC is thus given by (4-24). There is an optimum transformer ratio for which the ENC has a minimum.

$$n_{OPT}^2 = \frac{C_d}{C_A \left( 1 + \frac{C_d}{C_b} \right)} \quad (4-26)$$

From equations (4-21), (4-22) and (4-26) we see that the optimal noise situation is obtained with the transformer ratio which equals transform detector capacitance to the capacitance of the preamplifier. (Modified slightly by the effect of a finite blocking capacitance  $C_b$ .) The condition (4-26) maximizes again the transfer of the signal energy from the detector to the preamplifier.

To keep the parallel noise small requires certain care. A liquid argon ion chamber does not have significant leakage current. The dominant source of parallel noise in the input circuit are the losses in the ferrite core of the matching transformer. We will not analyze the noise performance of the transformer here. With a correct choice of the ferrite core with the right number of transformer turns wound in a correct way the parallel noise due to the transformer losses can be kept sufficiently small.

The leakage inductance of the transformer plus the inductance in the connection between the detector and amplifier requires a damping to prevent oscillations in the input circuit. An amplifier input impedance of  $500-1000\Omega$  is required for an aperiodical waveform. The damping has to be implemented with no resistive component connected to the input. We will describe an amplifying configuration satisfying the above condition ("cool resistor") in some detail later.

Substituting  $n_{OPT}$  from equation (4-26) into (4-24) we obtain the minimum equivalent noise charge due to the series noise

$$ENC_{MIN}^2 = 128 kTR_s \frac{1}{\lambda} C_d C_A \left( 1 + \frac{C_d}{C_b} \right) \quad (4-27)$$

Equation (4-27) shows that for an optimal matching the ENC increases with the square root of detector capacitance. (We can assume  $\frac{C_d}{C_b} = \text{constant}$ )

Let us consider two sections of a calorimeter each matched to its preamplifier. Noise at the output of one section is uncorrelated with the noise at the output of the other section. The square of the noise of their linear sum thus equals the sum of noise squares from individual sections which is the same (according equation 4-27) as the noise coming from a section with the capacitance equal to the sum of two capacitances. We see that for an optimal matching the noise is given by the capacitance of the considered section independently of the number of subdivisions and other read-out details. A fine subdivision can improve the signal to noise ratio at the later stage of the analysis, because for a given event we can exclude the subsections without the signal and so decrease the effective capacitance of the detector. This approach is not practical in triggering when the signal from the calorimeter is used to provide the total energy for the event selection. In this case we have to take the signal from a major part of the calorimeter which has the equivalent noise charge given by equation (4-27) independently of all details of the detector segmentation.

### 3.3 Optimization of Electrode Distance in Liquid Argon Calorimeter

All noise relations for the liquid argon calorimeter were derived up to now under the assumption of an impulse (delta function) signal. We have already seen that real signal due to uniform ionization across the gap has a "sawtooth" waveform (shown in Fig. 10b)

$$i = \frac{Nq}{t_d} \left(1 - \frac{t}{t_d}\right) \quad 0 \leq t \leq t_d \quad (4-28)$$

If this waveform is presented at the input of the preamplifier-filter-amplifier system whose overall impulse response  $h(t)$  is our well-known bipolar function with triangular lobes the output waveform is the convolution of these two functions. Fig. 13 shows the result of the convolution. Two effects are apparent. i) The width of the output waveform, i.e., the resolving time  $\lambda_m$ , is the sum of the electron drift time across the gap and the width  $\lambda$  of the system impulse response;

$$\lambda_m = \lambda + t_d \quad (4-29)$$

ii) The amplitude of the output is reduced compared to that of  $h(t)$  for impulse excitation of the same charge.

We are going to find an "optimal" gap width under two constraints. The first is the resolving time  $\lambda_m$  and the second is the constant ratio of the energy sampled in liquid argon. The first constraint is given by the rate requirement. The second constraint defines the total volume of the detector. Let us change the detector gap  $d$ . If  $d$  is very small, the capacitance of the detector is high and according to equation (4-27) equivalent noise charge rather high. If we increase the gap width, the detector capacitance decreases, but so also the amplitude of the output waveform which reaches 0 when the electron drift time  $t_d$  is equal to the resolving time  $\lambda_m$  ( $\lambda = \lambda_m - t_d = 0$ ). Intuitively there is an optimal gap width which gives the lowest ENC.

(There are many other important parameters entering into the choice of the gap width and hence the granularity of sampling for this kind of detector. The present optimization is only with respect to the electronic noise which may not be the limiting factor for a particular calorimeter

under consideration.)

Let us make our intuitive argument more quantitative. If we set a constraint on the resolving time  $\lambda_m$ , and vary the ratio of  $t_d$  to  $\lambda_m$  from zero to one, maintaining the total charge constant, then the amplitude of the output decreases from unity to zero as shown in Fig. 14. Calling this amplitude  $a (t_d/\lambda_m)$  we can obtain the equivalent noise charge for the case of the "sawtooth" signal from equation (4-27) by dividing the ENC for the impulse case by  $a (t_d/\lambda_m)$

$$\text{ENC}^2 = 128kT R_s C_A \left(1 + \frac{C_d}{C_b}\right) \frac{C_d}{a^2 \left(\frac{t_d}{\lambda_m}\right)} \cdot \lambda \quad (4-30)$$

The capacitance of the detector  $C_d$  depends on the gap width. If we use the second constraint of constant ratio of liquid argon and plate material the detector capacitance  $C_d$  is inversely proportional to the square of gap width  $C_d = \frac{a}{d^2}$ . (Capacitance of one gap goes as  $1/d$  and the number of gaps in the detector has also  $1/d$  dependence.) Substituting  $d = v_e \cdot t_d$  and neglecting the term  $C_d/C_b$  we can write equation (4-30) as

$$\text{ENC} \propto \sqrt{kT R_s C_A} \frac{1}{v_e \lambda_m^{3/2}} \left[ \frac{t_d}{\lambda_m} \left( \frac{\lambda}{\lambda_m} \right)^{1/2} \cdot a \left( \frac{t_d}{\lambda_m} \right) \right]^{-1} \quad (4-31)$$

The dimensionless term in the square bracket is plotted in Fig. 15. It has a minimum value of 6 at  $t_d/\lambda_m \approx 1/3$ .

Thus, there is an optimum value of the interelectrode gap determined by the required resolving time  $\lambda_m$  and by the electron drift velocity  $v_e$

$$d = \frac{1}{3} v_e \cdot \lambda_m \quad (4-32)$$

To conclude this part we would like to stress that the event resolving time  $\lambda_m$  does not define the limit on event timing in the liquid argon calorimeter. The best possible time resolution of the occurrence of an event in liquid argon calorimeter is given by equation (3-33). For a bipolar shaping considered here and the "sawtooth" form of the input current, the time jitter of an event with the total charge  $Q$  at the zero crossing is

$$\delta t = \frac{\text{ENC}}{Q 4a \left( \frac{t_d}{\lambda + t_d} \right)} \quad (4-33)$$

where ENC is given by equation (4-27) and  $a$  is plotted in Fig. 14.

As an example for  $C_d = 10^{-8}$  F;  $\lambda = 1 \mu\text{s}$ ,  $a = .6$  and the 1 GeV of the deposited energy in the calorimeter with 10% of energy sampled by liquid argon the time jitter  $\delta t$  is 6 ns.

Part V.

## POSITION SENSITIVE GAS DETECTORS

### 1. Fundamentals of Particle Detection in Proportional Gas Counters

Today the position-sensitive detectors in high energy physics experiments are dominated by multiwire gas counters operated in the proportional mode.

Detectors come in a large variety of sizes and arrangements, but the underlying physical principles of detection are the same. Gas counters have an appealing initial simplicity with a great design flexibility, but the details of their operation are surprisingly complicated when closely examined. In these notes, the emphasis is on the signal processing, so only very brief descriptions of the fundamentals of gas counter operation are presented. <sup>24, 25</sup>

The physical processes involved in detection by gas proportional counters are: i) Primary ionization of the gas atoms and molecules by the electromagnetic field of high energy particles; ii) secondary ionization by some electrons produced by the primary ionization; iii) drift of electrons toward the anode wire; iv) impact ionization of "hot" electrons in the high electric field near the anode wire producing the charge multiplication (avalanche gain mechanism); v) collection of electrons by the anode and an exchange of remaining positive charges among various components of the gas mixture; vi) drift of positive ions toward the cathode and signal formation on the surrounding electrodes; and vii) signal processing and readout.

As mentioned above, the main purpose of wire chambers in high energy physics experiments is to provide position sensing detection. In some applications, the measurement of the ionization due to the high energy particle (or the accompanied x-rays) is also required. Examples are: gas proportional chambers used as sampling detectors in calorimeters, chambers intended for particle identification via ionization in relativistic rise region, and detection of transition radiation x-rays, etc.

A proportional gas counter consists of a thin wire stretched at the axis of a conducting cylinder filled with a suitable gas. The electrons liberated in the gas by an ionizing radiation drift toward the central wire at a positive potential relative to the cylinder. Electrons arriving in the high field region in the vicinity of the wire experience inelastic ionizing collisions, forming a multiplication avalanche. The pulse induced in an external circuit between the anode and the cathode is mainly produced by the motion of positive ions. (Appendix 1)

Multiwire proportional chambers usually consist of a plane of equally spaced anode wires, sandwiched between two cathode planes. In the major part of the chamber volume the electric field is uniform, but close to the wire it is almost identical to the field of a cylindrical proportional counter. The simplest localization is done by identifying the wire which has produced an avalanche.

The relation between the position of the ionizing column created by a particle and the avalanching process, due to the arrival of electrons at the anode, can be used for an accurate position measurement. A gas detector designed to operate this way is called a drift chamber. A possible configuration of drift chamber electrodes is shown in Fig. 16a. (26) Electrons produced at time  $t_0$  by incoming charged particles drift against the electric field and reach the anode wire where avalanche multiplication occurs at a time  $t_1$ . The distance of the track from the anode wire is therefore given by

$$x = \int_{t_0}^{t_1} w_D t \quad , \quad (5-1)$$

where  $w_D$  is the drift velocity of electrons. Before studying the main limitation of the position accuracy obtainable by drift chambers, let us say a few words about the charge multiplication in a avalanche process.

A passage of a charged particle through a 1 cm thick layer of gas produces, typically, about 100 electron-positive ion pairs. We already know that this amount of charge is not large enough for direct detection. The amplification in an avalanche brings the signal into a detectable level. However, there is a penalty to pay. The avalanche process is a stochastic (random) process itself and introduces additional fluctuation into the detection. (Its name is "linear birth process".) If we study avalanche processes originated by a single electron; we see that resulting total charges have a broad distribution. This kind of distribution is very common in detection and is called the Furry distribution. In the case of the avalanche multiplication process in gas detectors, the probability density function of the gain distribution can be approximated by a simple exponential

$$P(g) = \frac{1}{g_0} \cdot e^{-\frac{g}{g_0}} \quad (5-2)$$

where  $P(g) \cdot \Delta g$  is the probability for the gain to be between  $g$  and  $g + \Delta g$  and  $g_0$  in the mean gain.

If we detect  $n$  primary electron charges having the mean value  $E(n)$  and the variance  $\text{var}(n)$  via an avalanche process described by Eq. (5-2), the mean value  $E(N)$  and the variance of the total number of charges  $\text{var}(N)$  after multiplication can be expressed as <sup>17</sup>

$$\begin{aligned} E(N) &= g_0 \cdot E(n) \\ \text{var}(N) &= g_0^2 \cdot \text{var}(n) + E(n) \cdot \text{var}(g) \\ (\text{var}(x) &= E(x^2) - E(x)^2) \end{aligned} \quad (5-3)$$

The variance of  $g$  is directly calculable from the probability density function (5-2), and its value is  $g_0^2$ . For the square of the relative width we can obtain from Eq. (5-3)

$$\frac{\text{var}(N)}{E^2(N)} = \frac{\text{var}(n)}{E^2(n)} + \frac{1}{E(n)} \quad (5-4)$$

If the distribution of the number of primary electrons  $n$  has Poisson statistics,  $\text{var}(n)/E^2(n) = 1/E(n)$ , we see that the relative width of the number of electrons after multiplication in the avalanche process (5-2) is degraded by a factor of  $\sqrt{2}$ . In the measurement of x-ray energy, the distribution of the number of primary electrons is narrower than the Poisson distribution with the same mean number of electrons. (Fano factor) Therefore, the relative importance of the second term in Eq. (5-4) is greater and the energy resolution degrades typically by a factor of 2. (Broadening of the resolution given by Eq. (5-4) does not depend on the mean gain  $g_0$ . It can be shown, however, that the stability of the mean gain decreases with the value of mean  $g_0$ .)

The loss of the energy resolution is of little importance for chambers where only position information is required. Low gas gain has, however, enormous practical advantages for the chamber operation. The lifetime of the chamber is extended, external and internal high voltage discharges are easier to prevent. Positive ions produced by an avalanche at the anode of a drift chamber move slowly toward the cathode across the drift volume. The presence of the positive charges inside the drift volume can seriously affect the electric field and hence the drift velocity of the primary electrons. The velocity in Eq. (5-1) can become dependent on the past chamber history and the method stops to work. The effect depends on the total charge in the drift volume which is product of the detection rate and the mean gain. Smaller mean gain allows operation of the chamber at higher rates.

## 2. Position Resolution in Drift Chambers <sup>22</sup>

The space resolution in drift chambers is mainly limited by two factors:  
i) diffusion of electron during their drift toward the anode and ii) discontinuity of the primary ionization.

The resolution for a track close to the anode wire is limited by the fact that the ionization left by a fast particle, instead of being a "continuous" line, is concentrated in separated clusters. The effect is shown in Fig. 16b. A track with drift path 0 (crossing the anode noise) gives a reading which corresponds to the distance of the closest ionization cluster to the anode. If the mean number of primary clusters per cm is  $N$ , the probability of zero clusters within the distance  $d$  from the anode is

$$P_0(d) = e^{-2dN} \quad . \quad (5-5)$$

From this probability we can derive the mean distance of the closest cluster  $E(d_{\min}) = 1/2N$  and the variance  $\text{var}(d_{\min}) = 1/(2N)^2$ . We see that the drift distance measurement has a systematic error  $1/2N$  (we cannot measure negative distances) and also the spread  $\sigma = 1/2N$ . We can take their sum in quadrature as the contribution of the primary ionization effect to the space resolution (for a particle passing through the anode).

$$\delta x_i(o) = \frac{1}{\sqrt{2}N} \quad . \quad (5-6)$$

For gases used in drift chambers,  $N$  is about  $25/\text{cm}^2$ , so  $\delta x_i(o)$  is about  $300 \mu$  (for a track crossing the anode wire). The importance of the fluctuation in the ionization decreases rapidly with the track distance from the anode wire  $x$ . Assuming a constant drift velocity, we can write

$$\delta x_i(x) = \delta x_i(o) \cdot (\sqrt{x^2 + \delta_i x^2(o)} - x) \approx \delta_i x(o) \cdot \frac{\delta x_i(o)}{x} \quad (5-7)$$

so already 3 mm from the anode wire the fluctuations in the primary ionization contributes only  $30 \mu$  to the total resolution.

The effect of the electron diffusion is more important in the major part of the drift cell. For a single electron drifting during the time  $t$  the rms displacement  $\delta x_D$ , with respect to the mean drift path  $x$ , is

$$\delta x_D = \sqrt{2Dt} = \sqrt{\frac{2Dx}{\mu E}} \quad (5-8)$$

where  $D$  is the diffusion constant,  $\mu$  is the electron mobility and  $E$  the electric field. The diffusion constant  $D$  is given by the Einstein relationship

$$D = \frac{\zeta kT}{q} \cdot \mu \quad (5-9)$$

where  $k$ ,  $T$ ,  $q$  have their usual meaning and  $\zeta$  is the ratio of the effective electron temperature and the gas temperature  $T$ . The reason for having the electron agitation ratio  $\zeta$  in the Eq. (5-9) is that free electrons in gas under the influence of an external electric field are, in general, not in thermal equilibrium with the gas. From Eqs. (5-8) and (5-9) it follows that the mixtures used in the drift chambers should have the electron agitation ratio close to one. (We can notice that the diffusion of electrons in gases and thermal noise in resistors are basically the same effects.)

Combining Eqs. (5-8) and (5-9), we can write

$$\delta x_D = \sqrt{\frac{2\zeta kT}{q}} \cdot \sqrt{\frac{x}{E}} \quad . \quad (5-10)$$

It looks as if we could improve the diffusion by simply increasing the electric field  $E$ . This is correct for low values of  $E$ . Physically it means that the drift velocity increases with the electric field and therefore the electron drift time decreases. At the some value of  $E$ , the drift velocity stops to increase (for some gas mixtures it even starts to decrease!) and  $\delta x_D$  stops to improve with

the electric field. It means that electrons are heated and the increase of the agitation electron velocity decreases the electron drift velocity. The majority of the drift chamber works in a plateau region of the drift velocity as a function of the electric field. The independence of the drift velocity of the electric field removes the constraint of the uniform electric field inside the drift cell and allows a much wider variety of electrode configuration in a drift chamber design.

Since in a drift chamber a swarm of electrons are drifting from the particle track to the anode wire, the diffusion effect on the position resolution is smaller than  $\delta x_D$  for one electron given by Eq. (5-10). The rms of the total resolution can be written as

$$\delta_{TOT} = \sqrt{\delta x_i^2(x) + A \cdot \delta x_D^2(x) + \delta_{el}^2} \quad (5-11)$$

where A is the factor which accounts for the statistical effects in the diffusion of the electron swarm and  $\delta_{el}^2$  are the fluctuations in the drift time measurement. A typical resolution versus drift distance curve is shown in Fig. 16c. We see that increasing the drift path from zero (anode wire), the resolution at small x improves due to the decreasing importance of the ionization density effect but at large x increases as the total error starts to be dominated by diffusion effects.

The drift time measurement includes the measurement of the particle passage time through the chamber (usually provided by a scintillation counter) and the time of the avalanche at the anode wire. Two examples of signals at the output of the preamplifier connected to the drift chamber anode are shown in Fig. 17b and c. Output waveforms are very irregular as can be expected from the ionization randomly clustered along the particle path. (This is in contrast to a regular waveform due to the point-like ionization of an  $Fe^{55}$  6 keV x-ray.) The irregularity of the waveform

prevents us from using any matched filter for a optimal timing. The best results for a standard drift chamber, where the information about the drift path is contained only in the time of arrival of the first cluster, were obtained by a leading edge timing method. From the signal processing point of view, standard drift chambers can use only very small portions of the signal to derive the time and hence the position information. Any drift chamber electrode configuration capable of equalizing the drift time for all clusters produced by ionizing particles would improve the chamber performance. Not only the available primary signal would be much larger, but also the effect of the diffusion would be decreased by full use of the statistics of all primary clusters.

Let us return to the Eq. (5-11) describing the resolution of drift chambers. The contribution of fluctuations in drift time measurement  $\delta_{el}$  to the total drift chamber resolution depends on the signal from the chamber and on the noise of the amplifier. We will calculate the minimal gas gain of the chamber, i.e., the gain which makes the anode signal sufficiently large so the time jitter in the timing circuit, due to the amplifier noise and hence  $\delta_{el}$ , is negligible compared with other terms in Eq. (5-11).

We will use results of the noise analysis from Section IV.2. A typical capacitance of the anode wire is 25 pF, ( $\approx 10$  pF/m, plus additional capacitance in chamber feed-throughs). The measurement time  $t_m$  is very small. The best amplifying device under these conditions is a bipolar transistor with high  $f_T$ . Taking  $f_T = 3$  GHz,  $t_m = 10$  ns, the ENC calculated from Eq. (4-19) is equal to about 1000 electron charges. The time jitter can be written as

$$\Delta t = \frac{t_m \cdot ENC}{Q_s} \quad (5-12)$$

where  $Q_s$  is the part of the wire chamber signal utilized for timing and  $t_m$  is the half width at the base of the assumed filter impulse response. The product

of the time jitter and the drift velocity should be small compared with the diffusion limit of the chamber which allows us to write

$$Q_s \approx \frac{t_m \cdot ENC \cdot w_{\text{drift}}}{\delta x_D} \quad . \quad (5-13)$$

Taking  $\delta x_D = .1 \text{ mm}$ ,  $w_D = 5 \text{ cm}/\mu\text{s}$ , the required signal charge for timing  $Q_s$  is about 5000 electrons. In the first 5 ns only about 10% of the avalanche charge is induced on external electrodes (see Appendix1). If we assume an average of 5 electrons contributing to the signal, the minimal gas gain is about  $10^4$ .

The calculation was done for the optimal capacitance matching and for a transistor with high  $f_T$ . The choice of the "best first transistor" is sometimes comprised by an additional requirement for the device to survive possible instabilities and discharges in the chamber. If the electronic noise is increased, the gas gain has to be increased by the same factor resulting in rate limitation and other disadvantages as described in the previous section.

To conclude, we would like to stress that drift chambers are detectors still in a development stage and substantial improvements in their performance are very likely to occur.

### 3. Second Coordinate Readout in Drift Chambers

We have seen that position resolution in a drift direction of the drift chamber can be around .1 mm. Such a fine resolution is difficult to achieve by other position detection means for the comparable active areas. A stereo view is the most popular way to obtain particle coordinates in space. There is, however, a very wide class of experiments performed at colliding beam facilities where the implementation of stereo views presents considerable problems.

In a typical colliding beam experiment, the anode wires are oriented parallel to the beam axis and to the magnetic field. The position in the direction of the magnetic deflection is measured by the drift time method. The problem arises in determining the second coordinate along the anode wire. Second coordinates are used to define the polar angle of particle tracks and to establish the consistency of the track recognition. Usually the position accuracy sufficient for this direction is lower than the position accuracy perpendicular to the anode wire.

There are many second coordinate readout methods. These methods can be divided into two groups. Methods within the first group utilize the signal induced on the cathode for the second coordinate measurement. The cathode strips and delay line parallel to the anode wire are the methods most frequently used.

The cathode strips method (we can include here, also, a graded density cathode method) can achieve the greatest position accuracy for a given gas gain of the chamber, but it presents the well known ambiguity problems and therefore reconstruction difficulties for multiparticle events. The delay line parallel to the anode wire can limit the rate capability of the chamber and its physical presence in the drift volume can interfere with the drift field requirements. All methods within the first group require additional material within the drift chamber volume. This usually complicates, considerably, the mechanical construction of the chamber and introduces multiple scattering which may finally limit the precision of the momentum measurement.

Methods within the second group utilize the anode signal for the second coordinate measurement and so do not require any additional material inside the chamber, i.e., in the particle path. The position determination by charge division<sup>32</sup> and the second coordinate measurement by timing of the electromagnetic wave propagation along the anode wire (direct timing)<sup>33</sup> belongs to the second group. The first method requires the

resistive anode and gives a constant relative error  $\Delta t/t$  for a given signal charge. The direct timing method requires a high conductivity anode and gives the resolution  $\Delta t$  largely independent of the length of the wire  $l$ , for a given signal charge. We will analyze a few of the methods in detail.

#### 4. Charge Division Readout in Drift Chambers 34, 35

The principle of the method is shown in Fig. 18. The resistive anode wire of a proportional detector is terminated with low impedance ( $Z=0$ ) amplifiers at both ends. If the signal charge is injected at the distance  $x_0$  from the A end of the wire, the ratio of the charges collected at the two ends is the inverse ratio of the wire resistances between the injection point and the wire ends. For the usual case of the uniform wire resistivity, we can write

$$\frac{q_A}{q_B} = \frac{l-x_0}{x_0} \quad (5-14)$$

where  $q_A$ ,  $q_B$  are charges collected at A and B ends of the wire and  $l$  the length of the wire.

The Eq. (5-14) is evidently correct for the DC current for which the presence of distributed capacitance  $C(x)$  and inductance  $L(x)$  as shown in Fig. 18 is irrelevant. For the real chamber signal, which contains components at very high frequencies, the presence of  $C(x)$  and  $L(x)$ , together with the wire resistance  $R(x)$ , defines the signal propagation from the injection point to the wire ends. The first question to ask is the following. Is Eq. (5-14) correct also for an impulse injection into the anode wire? Is the charge division method correct also in cases where the signal propagation along the anode wire is mainly electromagnetic wave propagation? The second question considers the compatibility of the drift time

measurement with the charge division method. What is the degradation of the signal rise time due to the resistive anode? How long does it take for the injected charge to be collected by the amplifiers at the wire end and does this collection time limit the rate of the chamber?

The answer to the first question is positive. The charge divides between the two ends according to Eq. (5-14) independently of the distribution of inductance capacitance along the line. Reference 35 contains a general proof of relation (5-14). This intrinsic linearity of the charge division method is particularly important for a large detector system with many wires where differences in  $L(x)$  and  $C(x)$  from wire to wire can be expected. The linearity also justifies a relatively simple calibration at two points along the wire.

To answer the second question, we note that the anode wire plus surrounding electrodes form a transmission line and we are going to study the signal propagation along this line. Let us assume a special case of a homogenous line where all three line parameters (resistance  $R$ , inductance  $L$ , and capacitance  $C$  per unit length) are constant. The differential equation describing electromagnetic behavior of the line is the well known telegrapher's equation

$$\frac{\partial^2 u(x,t)}{\partial x^2} = CL \frac{\partial^2 u(x,t)}{\partial t^2} + RC \frac{\partial u(x,t)}{\partial t} \quad (5-15)$$

$u(x,t)$  - voltage

$x$  - position coordinate of the line

$t$  - time

$R, C, L$  - resistance, inductance, and capacitance of the line per unit length

$l$  - length of the line

The first term in Eq. (5-15) represents electromagnetic wave propagation and the second term diffusion. The boundary conditions due to low impedance pre-amplifiers at the ends of the wire are

$$u(0,t) = u(l,t) = 0 \quad (5-16)$$

Eq. (5-15), with the boundary condition (5-16), can be solved by the Fourier series method. Let us assume the solution in the form

$$u(x,t) = \sum_{n=1}^{\infty} X_n(x) \cdot T_n(t) \quad (5-17)$$

After standard manipulation, we can write for  $X_n(x)$

$$X_n(x) = A_n \cdot \sin \lambda_n x ; \lambda_n = \frac{n \cdot \pi}{l} \quad (5-18)$$

where all  $A_n$  have to be defined from the initial conditions.

Time equations then become

$$CL \frac{\partial^2 T_n}{\partial t^2} + RC \frac{\partial T_n}{\partial t} + \lambda_n^2 T_n = 0 \quad (5-19)$$

which are simple linear second order differential equations with constant coefficients. The complete solution can be carried out by the reader or found in ( ). We will concentrate here on the fundamental harmonic ( $n=1$ ) which has the slowest decay rate and, therefore, it defines the discharge time of the line. Substituting  $\pi/l$  for  $\lambda_1$ , the character of the solution of the main harmonic time Eq. (5-19) can be periodical, aperiodical, or critically damped. Thus we can define an RCL line to be

- i) underdamped line for  $\lambda R < 2\pi \sqrt{L/C}$
- ii) overdamped line for  $\lambda R > 2\pi \sqrt{L/C}$  or
- iii) critically damped line for  $\lambda R = 2\pi \sqrt{L/C}$

(Because  $\sqrt{L/C}$  is about  $350\Omega$  for the majority of chambers,  $\lambda R_{CRIT} \approx 2k\Omega$ .)

As is well known, critical damping gives the most rapid decay  $\sim e^{-\frac{R}{2L}t} = e^{-\frac{4\pi\cdot t}{\sqrt{L\cdot C}}}$ . For an underdamped line, all harmonics are also decaying as  $e^{-\frac{R}{2L}t}$ , but the R is smaller. An example of the voltage along a 2 m long anode wire at different times after an infinitely sharp charge was injected at the wire center for an underdamped line is shown in Fig. 19. We see that for this line with a resistance equal to  $\frac{1}{2}$  of the critical resistance the electromagnetic wave propagation is more important than the electric diffusion as the line propagation mechanism. The reflections of the wavefront are well visible and wavefronts contain more charge than the total amount of the diffused charged in the line.

For an underdamped line, the diffusive propagation dominates. The first harmonic has an aperiodical solution and slower decay time than a critically-damped line.

It can be calculated that for a critical damped line the rise time of the current pulse at the line end is about 1 ns (for an impulse injection), and about 40% of the total charge is delivered to the preamplifier within this time. We see that the deterioration of the rise time, due to the resistive anode, is negligible for the drift time measurement. Let us calculate the amount of time needed to discharge the line and to compare it with the time spread of the primary ionization arriving to the anode wire in a typical drift chamber (Fig. 17b,c).

For a critically damped line, the time constant of the decay is

$$\tau_C = \frac{2L}{R_{CRIT}} = \frac{2Lc}{2\pi\sqrt{L/C}} = \frac{L}{\pi C} = \frac{\tau_0}{\pi} \quad (5-20)$$

where  $c = 1/\sqrt{LC}$  is the speed of electromagnetic wave propagation along the line (vacuum) and  $\tau_0 = L/c$  is the time for the signal to travel the single length of the line. Let us take six decay time constants as a criterion for the line discharge time (.2% of the charge left). This discharge time is equal to the time needed for light to travel twice the length of the line ( $\approx 15$  ns).

Thus, in practical drift chambers, the charge collection time at the wire end is limited by the time spread of ionization arriving at the wire and the use of wires with higher resistivity does not limit the rate capability of the chamber. If the wire resistance is large, compared with critical resistance, we can neglect  $L$  in Eq. (5-15) and obtain the well known diffusion equation. The first term in the time equation (5-19) becomes zero and we can easily obtain the line discharge time constant  $\tau_D$  for a diffusive RC line.

$$\tau_D = \frac{RC}{\pi^2} \quad (5-20)$$

### 5. Position Resolution of Charge Division Method <sup>36</sup>

We will use the results of the noise analysis from Section IV.1 to calculate the resolution of the method. Since the resistive line represents a dissipative position sensing medium, its noise is inherently present with the signal. The basic equivalent circuit is shown in Fig. 20a. It is, essentially, Fig. 7 with a physical resistance  $R_p$  shorting the input. (We have shown only one end of the line; the other is connected to the ground for our analysis.) Equation (4-10) is directly applicable, so parallel noise can be written

$$\overline{q_{np}^2} = \frac{2kT}{R_p} \int_{-\infty}^{\infty} (h(t))^2 dt \quad (5-21)$$

To write down the contribution from the series noise  $e_n$ , we have to take into account that now we have a real resistance connected to the input, so the input admittance is  $1/R_p + j\omega C_{in}$ . The impulse response of the voltage source is  $(1/R_p + C_{in} d/dt)h(t)$ , where  $h(t)$  is the impulse response due to a unit signal charge. Following the same arguments as in Section IV.1, we can express the series noise as

$$\overline{q_{ns}^2} = 2kTR_s \int_{-\infty}^{\infty} (1/R_p h(t) + C_{in} d/dt h(t))^2 dt = 2kTR_s C_{in}^2 \cdot \int_{-\infty}^{\infty} [1/\tau_{in}^2 h^2(t) + (d/dt h(t))^2] dt \quad (5-22)$$

where  $\tau_{in} = R_p \cdot C_{in}$  is the input circuit time constant. (The cross term in integral (5-22) is zero as can be seen by per partes integration. It means that impulses and doublets with the same time origin are uncorrelated and the total series noise is the sum of their mean square contribution (Fig. 20b).

Let us now study the requirements on the pulse response of the overall filter  $h(t)$ . In order to preserve the intrinsic linearity of the charge division, we have to collect practically all charge. This means that the flat part of the impulse response has to be at least five time constants long (.2% nonlinearity), plus the time spread in the arrival of the primary ionization to the anode wire. Let us assume: i) the wire resistance is sufficiently high so the line discharge time constant  $\tau_D$  is given by Eq. (5-20) and ii) the time spread of the arriving ionization is small compared with  $\tau_D$ .

Because the line is "shorted" at the other end, a part of the line capacitance seen at the input circuit is "missing". It can be shown that the effective line admittance in frequencies of interest is approximately <sup>37</sup>

$$Y = \frac{1}{jR} + j\omega \frac{4C}{3} \quad . \quad (5-23)$$

To evaluate the relative importance of the parallel and the series noise, let us calculate the noise corner time constant  $\tau_c$  (see Eq. 4-3).

$$\tau_c = (C_A + 4C/3) \cdot (jR \cdot R_s)^{1/2} (\approx 30 \text{ ns}) \quad (5-24)$$

For  $C_A = 10 \text{ pF}$ ,  $4C = 30 \text{ pF}$ ,  $jR = 10 \text{ k}\Omega$  and  $R_s = 50\Omega$ . The value of the noise corner time constant  $\tau_c$  has to be compared with the duration of the flat part of the filter impulse response  $\tau_F$ . According above assumptions

$$\tau_F = 5 \cdot \tau_D = \frac{5 \cdot R C j^2}{\pi^2} (\approx 150 \text{ ns}) \quad . \quad (5-25)$$

for the same values of line parameters. We see that  $\tau_F \gg \tau_c$ , so the parallel noise dominates.

Referring to Fig. 21 for the definition of the filter impulse response, we can evaluate  $\overline{q_{np}^2}$  from Eq. (5-21),

$$\overline{q_{np}^2} = \frac{2kT}{iR} \times 1.2 \tau_F \approx 1.2 kT \cdot (C\ell) \quad (5-26)$$

where we substituted for  $\tau_F$  from Eq. (5-25). This noise adds to the position signal  $q_A$ . (The noise in  $q_A + q_B$  is much smaller since for  $\tau_F \geq 5\tau_D$ , it is anti-correlated at the two ends of the line.) The position resolution is then,

$$\frac{\Delta\ell}{\ell} = \frac{\sqrt{\overline{q_{np}^2}}}{q_A + q_B} = 1.1 \frac{\sqrt{kT(C\ell)}}{q_A + q_B} \quad (5-27)$$

The position resolution for the charge division as given by Eq. (5-27) is determined only by the total line capacitance ( $C\ell$ ), and is independent of the line-resistance. This is a direct consequence of the fact that the dominant noise is a thermal noise of the resistive anode in parallel to the input. If we try to decrease the anode noise by increasing the anode resistance, the filtering time has to be increased by the same ratio (Eq. 5-25), and the noise remains the same.

The result (5-27) holds only when the duration of the flat top of the filter impulse response is dominated by the time constant of the diffusive anode line. If the filter response is determined from the time spread of the arrival of the primary ionization to the wire  $\tau_F$  is not determined by (5-25), and the noise is given by the first part of the Eq. (5-26).

The amount of signal charge required to obtain 1% resolution by charge division method is few time  $10^6$  electrons for a typical drift chamber. The method uses total

ionization charge a. the anode wire, so the gas gain required for the charge division method and the gas gain for drift time measurement are comparable.

## 6. POSITION SENSING WITH DELAY LINE

Position sensing with delay lines is based on conversion of position information into a time difference between two signals from delay-line ends (Fig. 22a). The delay line is in principle a non-dissipative position-sensing medium. The noise in this system is generated in the line terminations and in the amplifiers. The position resolution is ultimately limited by the fluctuations in the shape of the avalanche around the anode wire. The electronic noise determines the magnitude of the signal required to achieve this resolution.

We are going to show that the delay-line termination can be realized without connecting any dissipative elements to the preamplifier input. In this way no additional noise source is added into the system, and the noise at the output should be smaller than the noise of a system with a conventional resistive termination.<sup>36</sup>

Figure 23 shows a basic circuit configuration of a charge-sensitive preamplifier with a capacitance  $C_0$  at the collector of a cascade transistor. Similar configurations are commonly used in low-noise preamplifiers because they do not contain any dissipative element at the input. (Resistor  $R_f$  can be made sufficiently large so its contribution to the total noise is negligible.) Let us calculate the input independence of this preamplified configuration.

The voltage at the output of the preamplifier is the current from the  $g_m$  element integrated on the capacitance  $C_0$ :

$$U_{\text{out}} = g_m \frac{1}{j\omega C_0 + 1/R_0} \cdot U_{\text{in}}, \quad (5-28)$$

and the apparent input current flowing through the feedback impedance  $Z_f$  is given by

$$I_{in} = U_{out}/Z_f . \quad (5-29)$$

From equations (5-28) and (5-29) we can calculate the input independence

$$Z_{in} = \frac{U_{in}}{I_{in}} = \frac{1}{g_m} \cdot Z_f (j\omega C_0 + \frac{1}{R_0}) . \quad (5-30)$$

At frequencies where  $I/R_0 \ll j\omega C_0$  and for a pure capacitive feedback ( $Z_f = 1/j\omega C_f$ ), equation (5-30) becomes

$$Z_{in} = \frac{1}{g_m} \cdot \frac{C_0}{C_f} . \quad (5-31)$$

We see that  $Z_{in}$  in (5-31) is real, i.e. a resistive termination was achieved by capacitance in feedback. Thus, termination can be realized without connecting any dissipative element to the input. The termination, however, cannot be noiseless since it is realized by a physical amplifier. Because of the feedback, the equivalent series noise voltage  $\overline{e_n^2} = 4kT R_s$  appears at the input terminal.

If we now terminate a delay line of a characteristic impedance  $Z_0$  with a preamplifier having  $Z_{in}$  adjusted to equal  $Z_0$ , we have only noise from the preamplifier in the system (Fig. 22c). The noise source  $e_n$  sees the input impedance  $2Z_0$  ( $Z_0$  of the preamplifier in series with the line impedance  $Z_0$ ), so the spectral density of an equivalent parallel (across the input terminal of the amplifier) current noise-source generator is

$$\overline{i_{np}^2} = 4kT \frac{R_s}{(Z_0)^2} \times 2 , \quad (5-32)$$

where we have multiplied by a factor of two because of the same amount of noise generated at the other line-end.

### References

1. C.W. Fabjan and H.G. Fisher, Rep. Prog. Phys., 43, 1003 (1980).
2. P.W. Nicholson, Nucl. Electronics (John Wiley and Sons 1974).
3. A. van der Ziel, Noise in Measurements (John Wiley and Sons 1976).
4. F.N.H. Robinson, Noise and Fluctuation (Oxford University Press 1974).
5. M. Schwartz, Information Transmission, Modulation and Noise (McGraw - Hill 1980).
6. N.R. Campbell, Proc. Camb. Phil. Soc., 15, 117 (1909).
7. A. Papoulis, Probability, Random Variables and Stochastic Processes (McGraw - Hill 1965).
8. N. Wiener, Acta Math. Stockh, 55, 117 (1930).
9. A.I. Khintchine, Math. Ann. 109, 604 (1934).
10. J.B. Johnson, Phys. Rev. 32, 97 (1928).
11. A. Einstein, Ann Phys. 19, 298, 371 (1906).
12. H. Nyquist, Phys. Rev. 32, 110 (1928).
13. V. Radeka, IEEE Trans. Nuc. Sci. NS-11, 358 (1964).
14. A.S. Grove, Physics and Technology of Semiconductor Devices (John Wiley and Sons, New York 1967).
15. A. van der Ziel, Proc. I.R.E., 50, 1808 (1962) and 51, 461 (1963).
16. V. Radeka and N. Karlovac, Nucl. Instr. & Meth., 52, 86 (1967).
17. J. Sandweiss, Proc. of the 1977 ISABELLE Summer Workshop, BNL 50721, UC-28 p. 146.
18. V. Radeka, Nucl. Instr. & Meth., 99, 525 (1972).
19. F.T. Arecchi, G. Cavalleri, E. Gatti and V. Svelto, Energia Nucleare, 7, 691 (1960).
20. R.A. Boie, A.T. Hrisoho and P. Rehak, IEEE Trans. on Nucl. Sci., NS-28, 603 (1981).
21. C. Bussolati, S. Cova and E. Gatti, Zeitsch. fur Angewandte Mathematik und Physik, 22, 14 (1971).

22. W.J. Willis and V. Radeka, Nucl. Instr. & Meth., 120, 221 (1974).
23. V. Radeka, IEEE Trans. Nucl. Sci., NS-24, 293 (1977).
24. G.F. Knoll, Radiation Detection and Measurement (John Wiley and Sons 1979).
25. G. Charpak, F. Sauli, Nucl. Instr. & Meth., 162, 405 (1979).
26. A. Breskin, G. Charpak, F. Sauli, M. Atkinson and G. Schultz, Nucl. Instr. & Meth., 124, 189 (1975).
27. A.T. Bharucha-Reid, Elements of Theory of Markov Processes and their Applications (McGraw - Hill 1960) pages 439-442.
28. W. Farr, J. Heintze, H.K. Hellenbrand and A.H. Walenta, Nucl. Instr. & Meth., 154, 175 (1978).
29. P. Rehak and A.H. Walenta, IEEE Trans. on Nuclear Science, NS-27, 54 (1980).
30. E. Gatti, A. Longerio and H. Okuno, Nucl. Instr. & Meth. 163, 83 (1979).
31. A. Breskin, G. Charpak, F. Sauli and J.C. Santiard, Nucl. Instr. & Meth. 119, I (1974).
32. R.A. Boie, V. Radeka, P. Rehak and D.M. Xi, IEEE Trans. Nucl. Sci., NS-28, 71 (1981).
33. J.L. Alberti and V. Radeka, IEEE Trans. Nucl. Sci. NS-23, 251 (1976).
34. V. Radeka and P. Rehak, IEEE Trans. Nucl. Sci., NS-25, 46 (1978).
35. V. Radeka and P. Rehak, IEEE Trans. Nucl. Sci., NS-26, 225 (1979).
36. V. Radeka, IEEE Trans. Nucl. Sci., NS-21, 51 (1974).
37. R.B. Owen and M.L. Awcock, IEEE Trans. Nucl. Sci., NS-15, 290 (1968).
38. E. Durand, Electrostatique, Tome 2 (Masson et Cie, Paris, 1966) p. 196.

## APPENDIX

### Signal Induced by a Moving Charge

Ionization detectors use the ionization produced by a fast moving charged particle to detect and measure some quantities associated with a particle. The signal seen on the external electrodes of a detector, however, is due to the movement of the charged carriers inside the detector volume. The charged carriers move in an external electrical field which must be provided (by applying a bias voltage) in order to induce a signal.

To calculate the amount of the induced charge on outside electrodes, we will assume that the velocity of the charged carrier is much smaller than the light velocity so the system can be considered to be in an electrostatic equilibrium at any time. Expression for the induced charges can thus be obtained based only on the electrostatics.

Let us start with the Green's theorem<sup>(38)</sup> (called also the Gauss identity in German and French literature) which considers a set of conductors in two electrostatically possible states characterized by total charges  $Q_1, Q'_1$  and voltages  $V_1, V'_1$ , respectively. Green's theorem can be written:

$$\sum_i Q_i V'_i = \sum_i Q'_i V_i \quad (A-1)$$

where summation has to be performed over all conductors of the considered set.  $Q$ -s and  $V$ -s are the charges and voltages of the first electrostatic state.

Let us apply the Green's (Gauss') theorem to the conductors 1 and 2 shown on Fig. 24. The two electrostatically possible states are

$$q, Q_1, Q_2; V_{(P)}, V_1 = 0, V_2 = 0$$

and

$$q = 0, Q'_1, Q'_2; V'_{(P)}, V'_1, V'_2 = 0 \quad .$$

The first state has the charge,  $Q$ , at the point,  $P$ , and both conductors are grounded. The second state is free of the charge,  $q$ , and has only the second conductor grounded. Equation (A-1) can be written

$$q \cdot V'(P) + Q_1 V'_1 + Q_2 \cdot 0 = 0 \cdot V(P) + Q'_1 \cdot 0 + Q'_2 \cdot 0 \quad (A-2)$$

or

$$Q_1 = - \frac{V'(P)}{V'_1} \cdot q$$

Equation (A-2) gives the induced charge,  $Q$ , on the conductor, 1, connected to the ground due to the presence of the charge,  $q$ , at the point,  $P$ . The ratio  $V'(P)/V'_1$  is the voltage at the point,  $P$ , due to the unit voltage applied to the consider electrode (conductor #1) while all other conductors are grounded and the space is free of the charges. Equation (A-2) is always correct and applicable to the multielectrode detectors (cathode readout for example). The identical result can be derived from the energy balance equation. It holds also for semiconductor detectors where we have a fixed space charge present in the active (depletion) region of detector. The fixed charge modifies the real electrical field inside the junction, but because it does not move the induced charge is the same as without it (given by A-2) assuming the same motion of the free carriers. (This is a direct consequence of the linearity of the electrostatics.) We will stress again that  $V'(P)$  is a fictitious field and not the real field inside the detector. The real field is responsible for the carrier motion, the fictitious field,  $V'(P)$ , is a tool for the signal calculation.

Now we are ready to calculate the induced signal in ionization detectors. First, let us consider a parallel plate ion chamber having the gap distance,  $d$  (Fig. 10). We will assume that an electron-positive ion pair was created at the point,  $P$ , at the distance,  $x$  from the cathode (Fig. 10a). The net charge induced on the electrode (for example, anode) is the sum of charge induced by the

electron and the positive ion.

$$Q_{A,el} = -q \frac{V'_A(P)}{V'_A} = -q \frac{x}{d}$$

$$Q_{A,ion} = q \frac{V_A(P)}{V'_A} = q \frac{x}{d}$$

We can see that at the time of the ionization the induced charge on the anode due to the electron is canceled by an equal amount of the induced charge of the opposite polarity due to the positive ion. If we assume that electron drifts immediately towards the anode with the velocity  $v_d = d/t_d$  ( $t_d$  is the electron drift time across the gap) a constant induced current flows in the external circuit for which we can write

$$i = \frac{dQ_{A,el}}{dt} = -\frac{q}{d} \frac{dx}{dt} = -\frac{q}{t_d} . \quad (A-3)$$

This current flows for  $t = \frac{d-x}{v_d}$  so the total induced charge is

$$Q_s = -q \frac{(d-x)}{d}$$

which is the equation (4-20).

In cylindrical proportional chambers charge multiplication process takes place close to the anode wire. About 98% of the charge is created in less than 5 mean free paths from the anode. We will neglect this small motion of electrons before reaching the anode and consider the proportional chamber signal to be formed by a cloud of positive ions leaving the anode.

The ratio of  $V'(P)/V'_{anode}$  for the case of a cylindrical geometry can be easily found from Gauss' theorem and is equal.

$$\frac{V'(P)}{V'_{anode}} = \frac{\ln\left(\frac{r_c}{r}\right)}{\ln\left(\frac{c}{r_a}\right)} \quad (A-4)$$

where  $r$  - is the distance of the point P from the center

$r_a$  - is the anode radius

$r_c$  - is the cathode radius .

The induce current on the anode due to the total (multiplied) charge  $Q_t$  is

$$i = -Q_t \cdot \frac{d}{dt} \left( \frac{V'(P)}{V'_{\text{anode}}} \right) = Q_t \cdot \frac{1}{r_a \ln \left( \frac{r_c}{r_a} \right)} \cdot \frac{1}{r} \cdot \frac{dr}{dt} \quad (\text{A-5})$$

$dr/dt$  is the radial drift velocity of positive ions in the real electrical field of the proportional chamber. Assuming that the drift velocity is given by the external field,  $E$ , and by the ion mobility,  $\mu$ , we can write

$$\frac{dr}{dt} = E(r) \cdot \mu = \frac{E_a r_a}{r} \cdot \mu \quad (\text{A-6})$$

where  $E_a$  is the real electrical field on the anode surface. Equation (A-6) is the equation of motion for the positive ion which has an elementary solution

$$r^2(t) = 2E_a r_a \mu t + r_a^2 \quad (\text{A-7})$$

where we have assumed  $r(t=0) = r_a$ . Substituting the equation (A-7) (and its derivative) into (A-5) after simple manipulation we obtain

$$i(t) = \frac{i_0}{1 + t/t_0} \quad 0 < t < t_{\text{max}} \quad (\text{A-8})$$

where

$$i_0 = \frac{Q_t E_a \mu}{r_a \ln \left( \frac{r_c}{r_a} \right)}$$

$$t_0 = \frac{r_a}{2E_a \mu}$$

$$t_{\text{max}} = (r_b^2 - r_a^2) / 2E_a r_a \mu.$$

We have referred to equation (A-8) in Part V. The induced charge is given by:

$$Q_s(t) = i_0 \cdot t_0 \ln (1 + t/t_0) \quad (\text{A-9})$$

The signal current stops at  $t = t_{\text{max}}$ , that is, at the arrival of the positive anode to the cathode. It can be shown that  $Q_s(t = t_{\text{max}}) = Q_t$ .

This research supported by the U.S. Department of Energy under Contract

DE-AC02-76CH00016.

Figure Captions

Fig. 1. : A possible oscillogram picture of noise. The displayed voltage is a random process. If we consider the predictivity of  $v(t_2)$  from the knowledge of  $v(t_1)$ , we arrive at the concept of the autocorrelation function.

Fig. 2. : a) Physical resistance  $R$  represented by a noise current source parallel to an idealized noiseless resistance  $R$ .  
b) Physical resistance  $R$  represented by a noise voltage source in series with an idealized noiseless resistance  $R$ .

Fig. 3. : Series noise in charge-controlled amplifiers. Fluctuations in controlled output current are thought to be caused by a noise voltage generator in series with the control electrode.

Fig. 4. : a) Fluctuations in current flowing into (or out of) the control electrode represented by a parallel noise current generator  $i_{np}$ .  
b) For a current signal the series noise voltage source  $e_{ns}$  replaced by an equivalent current source  $i_{ns} = e_{ns}/Z_{in}$  parallel to the input.

Fig. 5. : a) Forward-biased  $n^+p$  junction. The density of current-carrying electrons in the p part of the junction falls off exponentially.  
b) Schematic of an npn transistor. The P-region is made extremely thin and is followed by a reverse-biased lightly doped n region.

Fig. 6. : Schematic representation of an n-channel junction field-effect transistor.

Fig. 7. : The basic equivalent circuit of the detector and amplifier. The system is described by the over-all impulse response  $h(t)$ . The series voltage noise source was replaced by an equivalent current noise source parallel to the detector.

Fig. 8. : a) The impulse response of an ideally matched filter with the time origin shifted in such a way that the waveform peaks at zero. The filter requires an infinite time to respond and is unfeasible.  
b) Triangular impulse response of a realizable filter. The noise is increased by only 8% compared with the noise of an ideal filter.

Fig. 9. : Equivalent noise charge (ENC) versus detection time  $t_m$  for different detectors and different amplifying devices.

Fig. 10. : Charge collection in the liquid argon ionization calorimeter:  
a) charge due to a single ion pair formed at a distance  $x$  from the negatively biased plate (solid line) and for an ion pair formed very close to the negatively biased plate (dashed line).  
b) charge due to uniform ionization.

Fig. 11. : Detector/preamplifier circuit configuration.

Fig. 12. : Equivalent circuit of detector and preamplifier for noise analysis. System pulse response of duration  $\lambda$  is shown at the bottom.

Fig. 13. : Top: current input signal  $i(t)$  from the liquid argon ion chamber. Bottom: impulse response of the system  $h(t)$  and the system response to the current signal at the input.

Fig. 14. : Relative amplitude of the system response as a function of the ratio of the input pulse duration  $t_d$  and the duration of the output pulse  $\lambda_m$ .

Fig. 15. : Equivalent noise charge (ENC) at the output of the calorimeter as a function of the ratio of the input pulse duration  $t_d$  and the output pulse duration  $\lambda_m$ .

Fig. 16. : a) Electrode configuration of multiwire drift chambers with relatively long drift path. Cathode wires are connected to uniformly decreasing potentials starting from ground in front of the anode. Field wires define the field at boundaries of two drift cells.  
b) Effect of the discontinuity of the primary ionization to the drift chamber resolution.  
c) Resolution of drift chamber versus drift distance.

Fig. 17. : "Current" waveforms at the preamplifier output:  
a) Signal due to the point-like ionization of an  $^{55}\text{Fe}$  6 keV X-ray.  
b,c) Signal due to the ionization of a minimum ionizing particle.

Fig. 18. : Principle of the charge division method. Charge is injected at the distance  $x_0$  from the A-end of the anode wire. The anode wire plus the surrounding electrodes form a distributed RCL line.

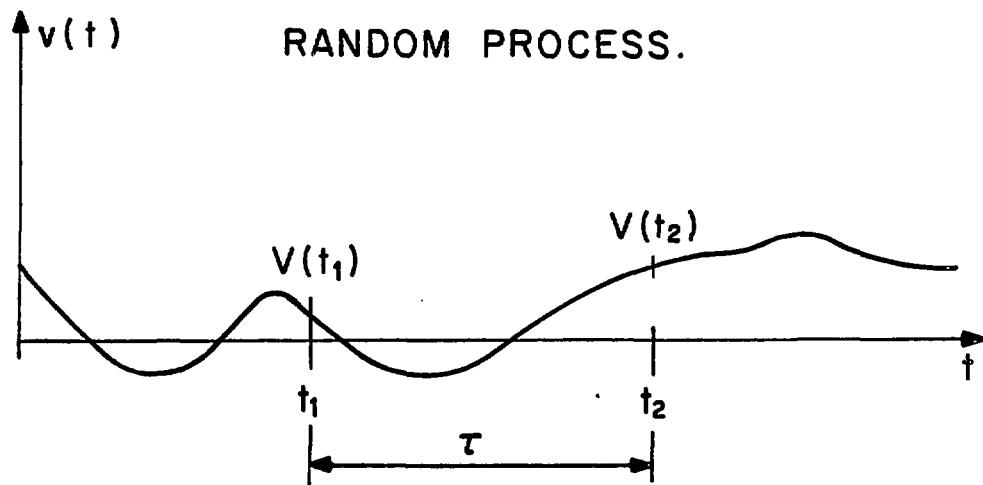


FIG. 1.

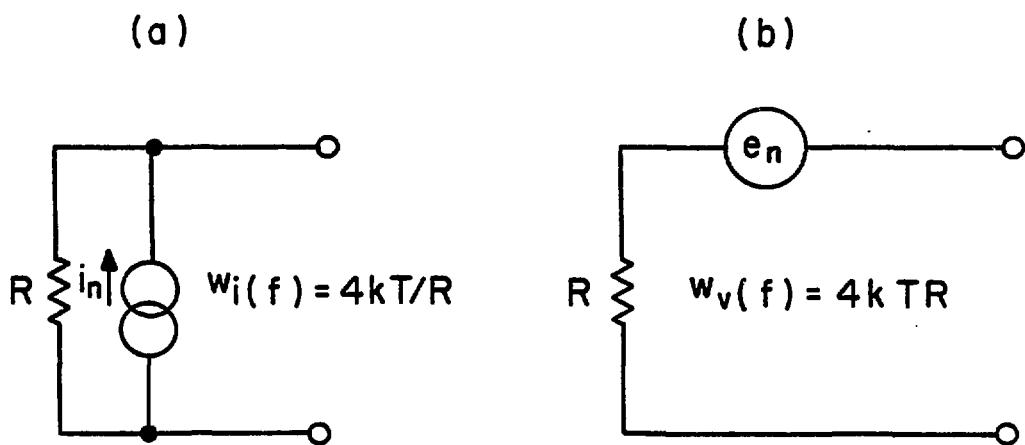


FIG. 2.

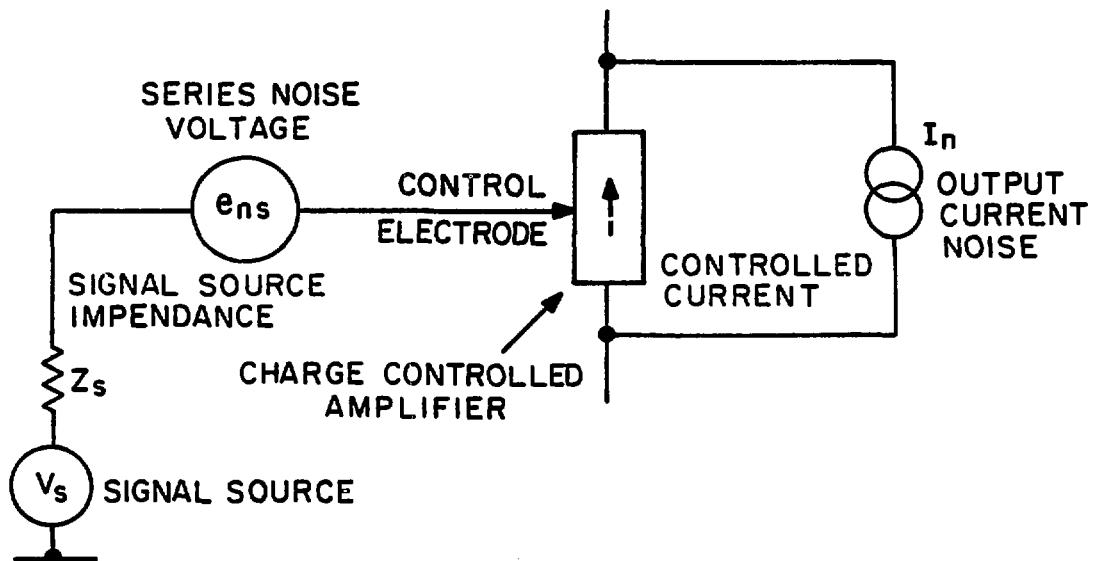
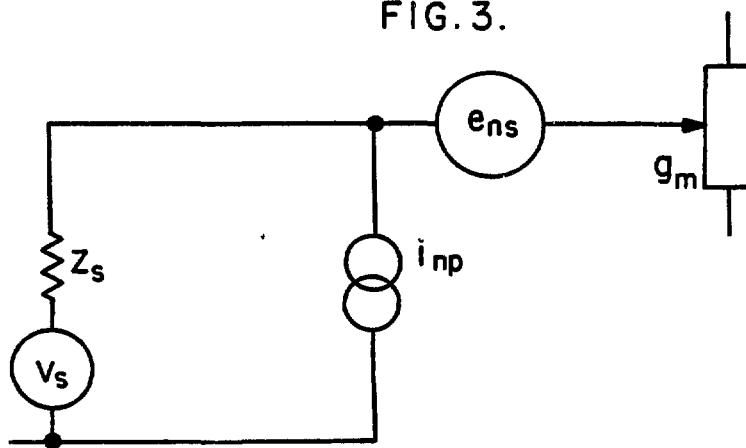
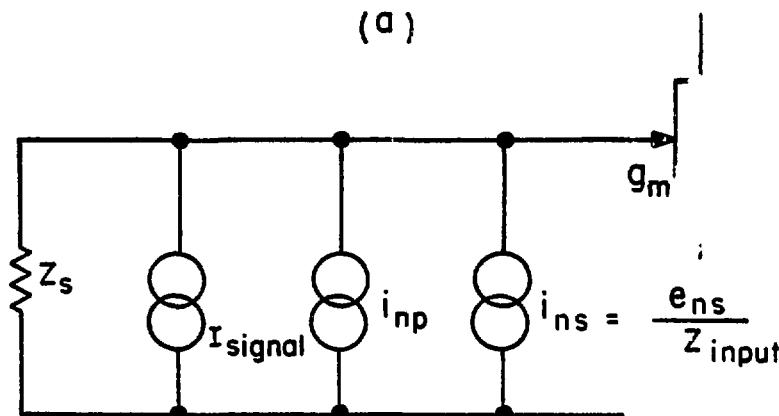


FIG. 3.

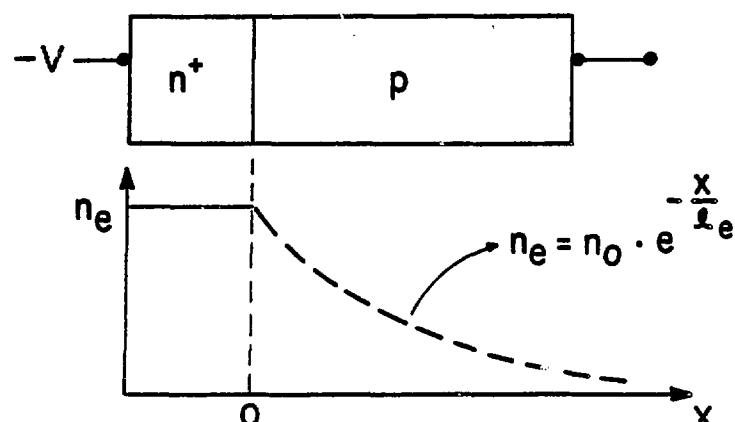


(a)

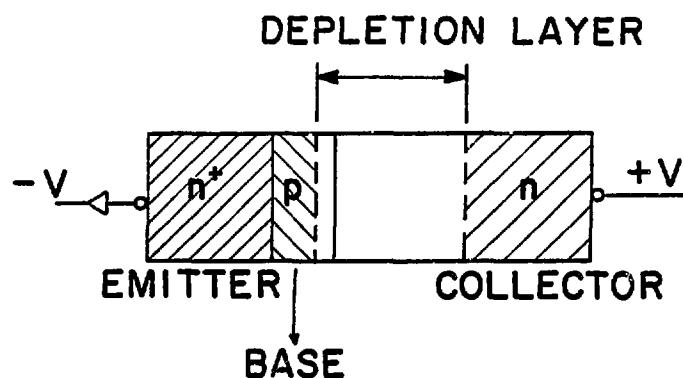


(b)

FIG. 4.



(a)



(b)

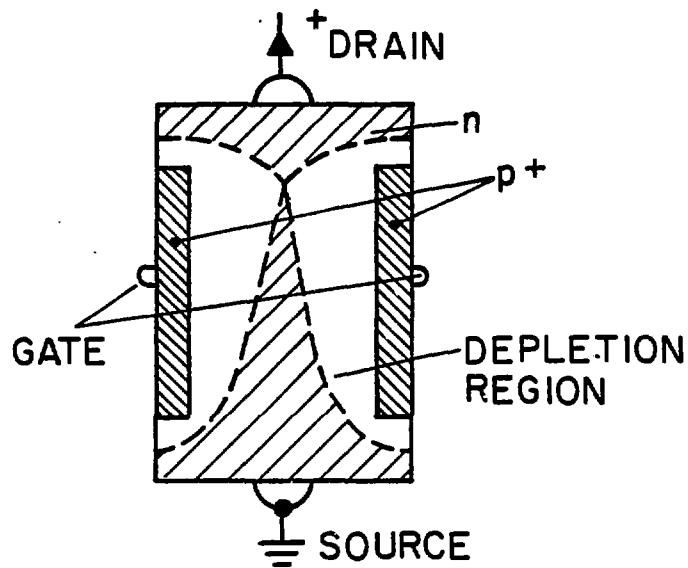


FIG. 6.

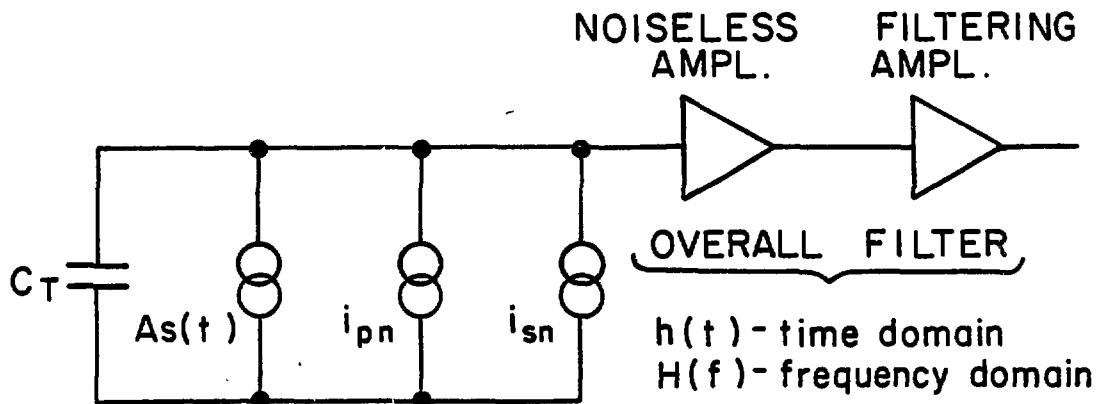
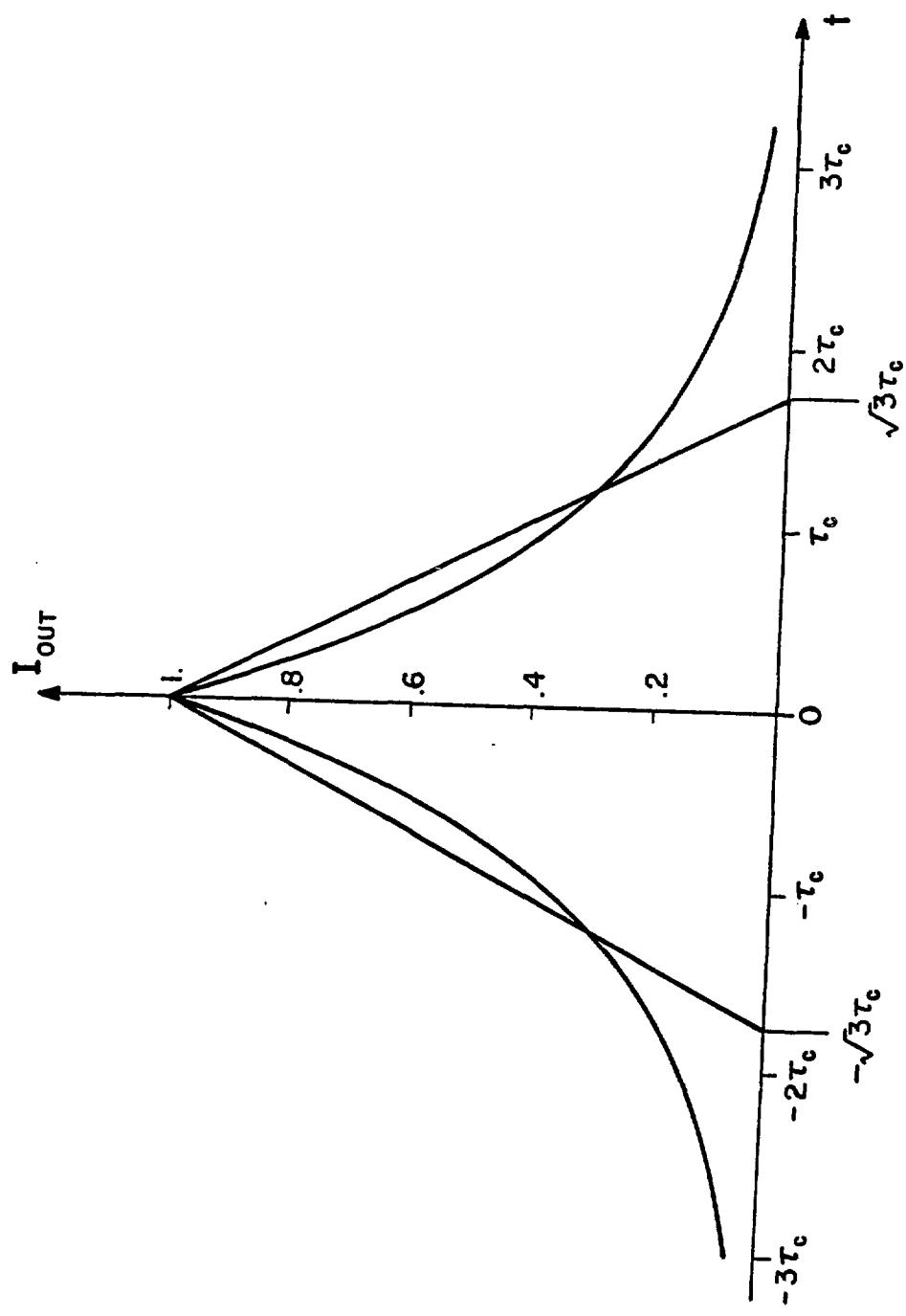


FIG. 7.



$F_{ij} \neq 0$

## NOISE CHARGE FOR DETECTORS vs. MEASUREMENT TIME

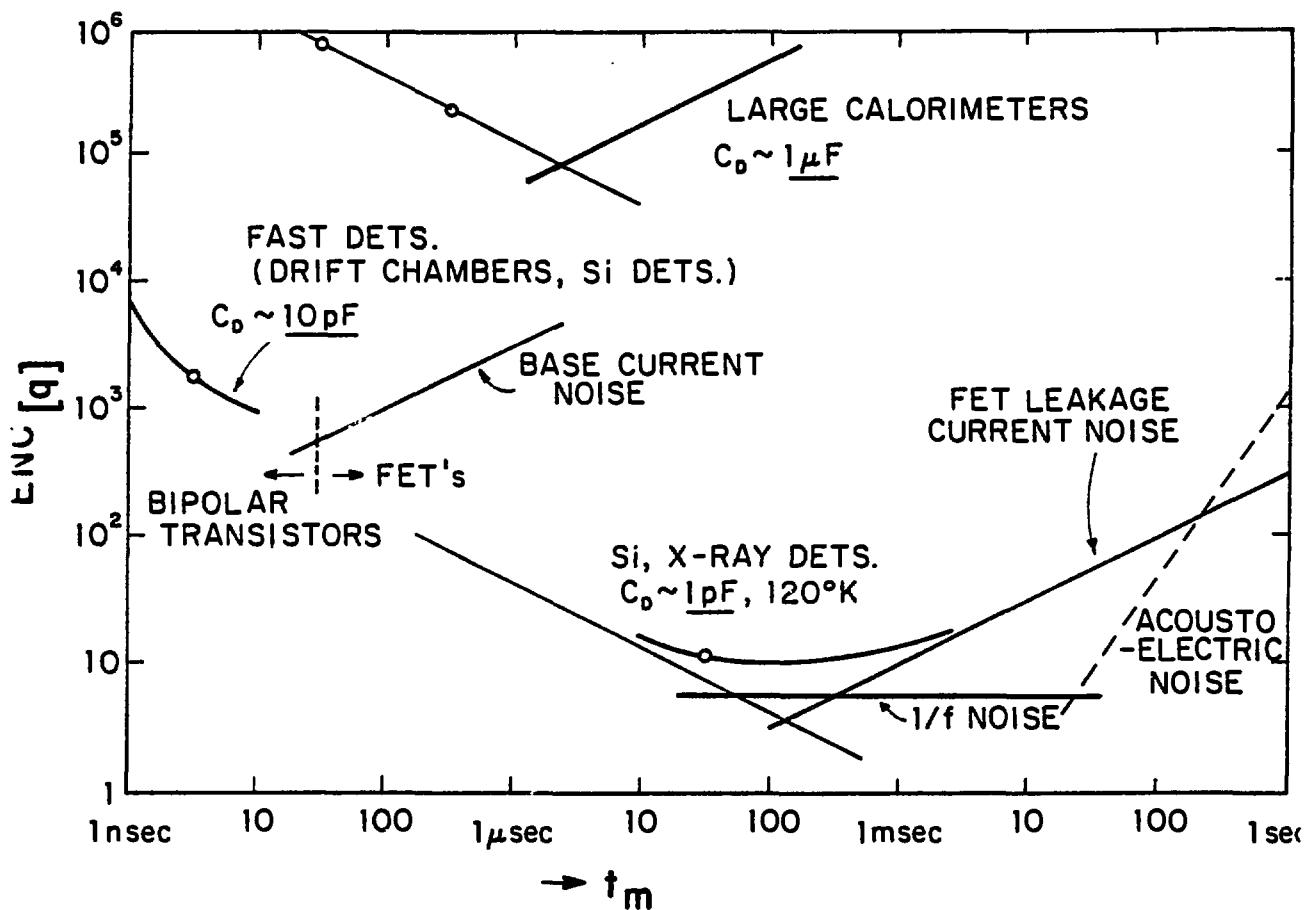
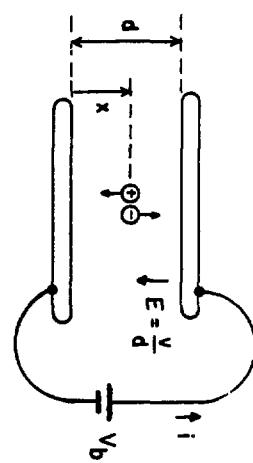
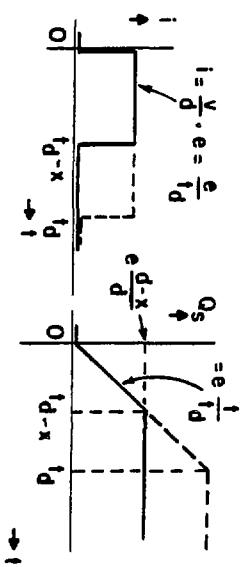


Fig. 9.

(a)



CURRENT



(b)

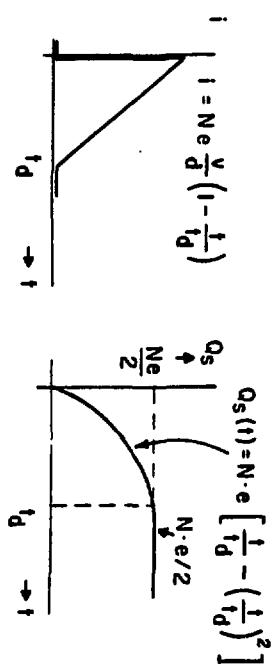
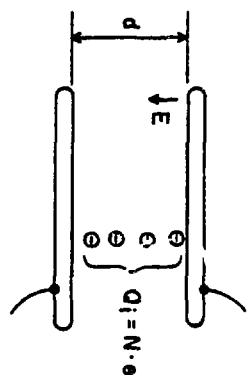


Fig 10.

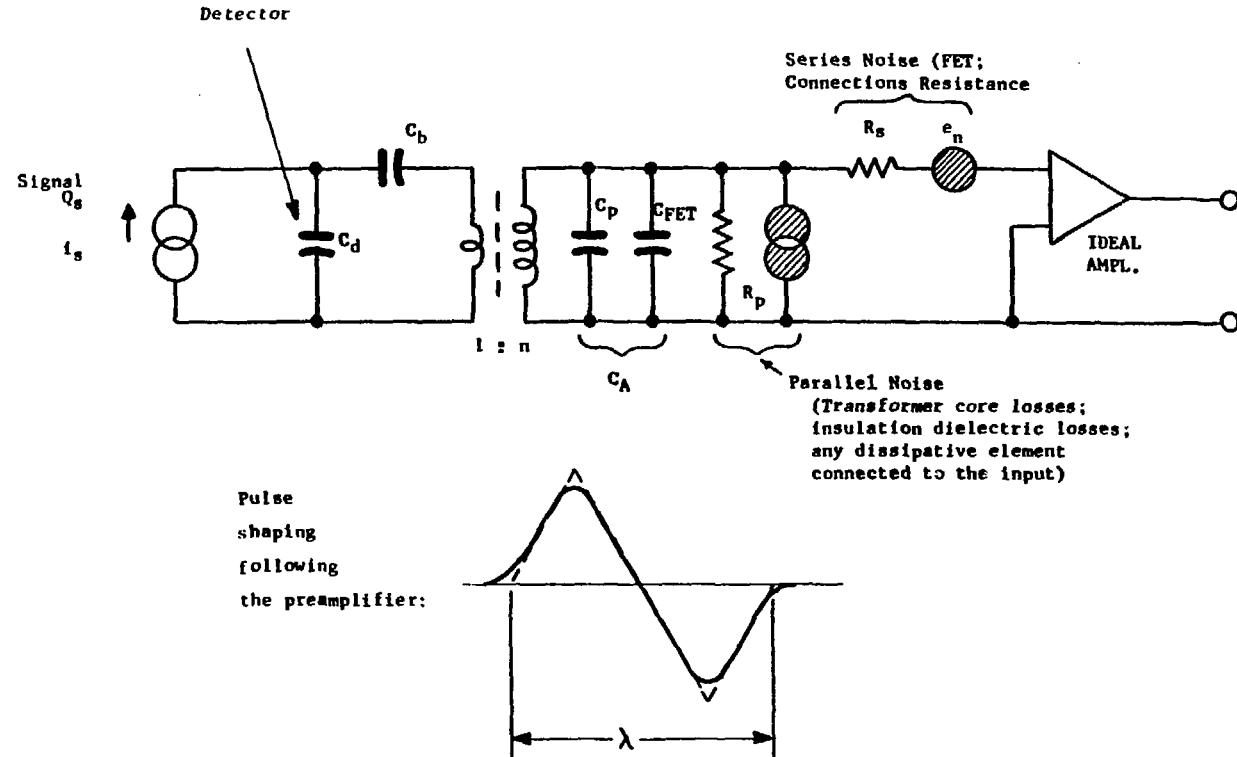


Fig. 1.

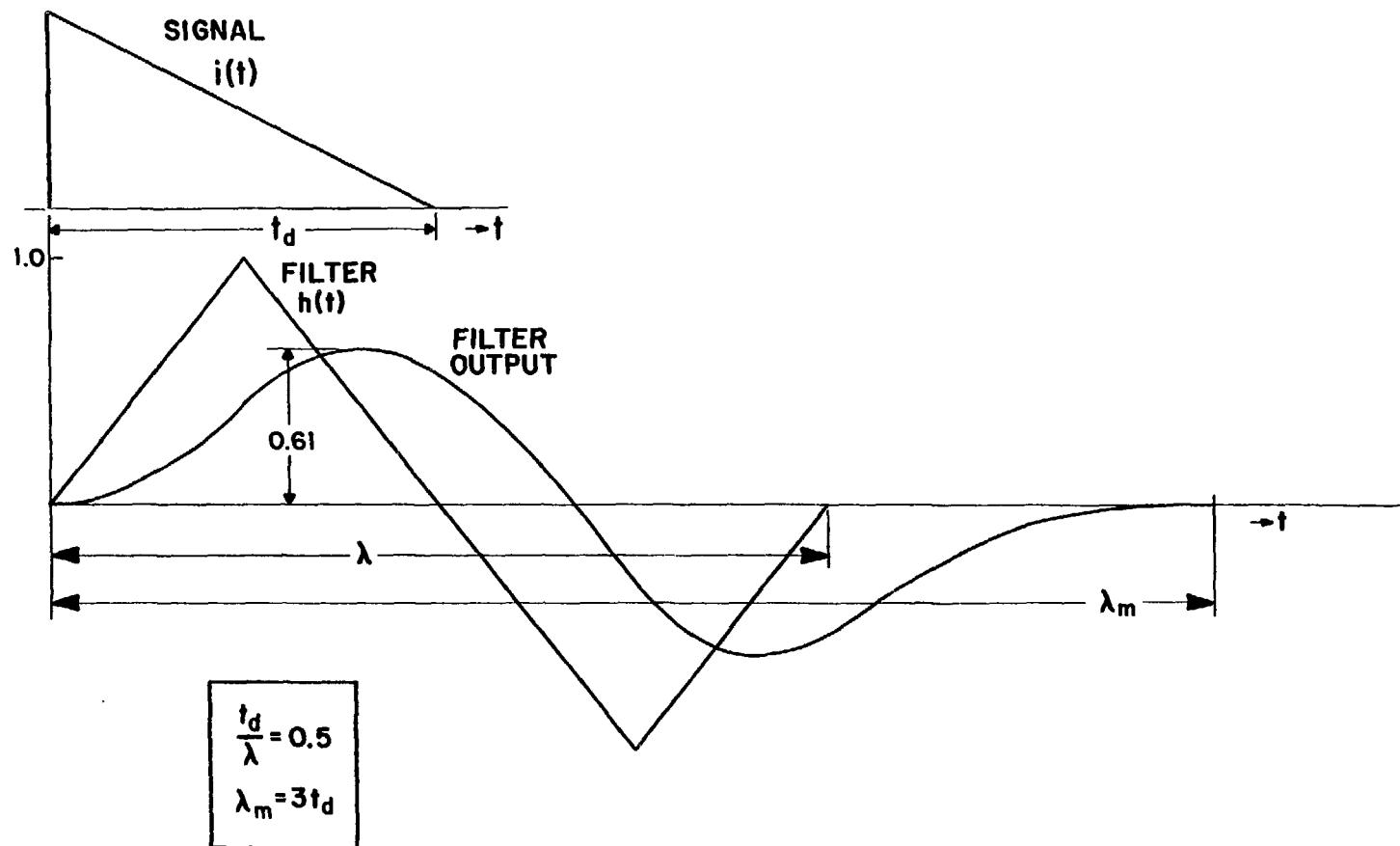
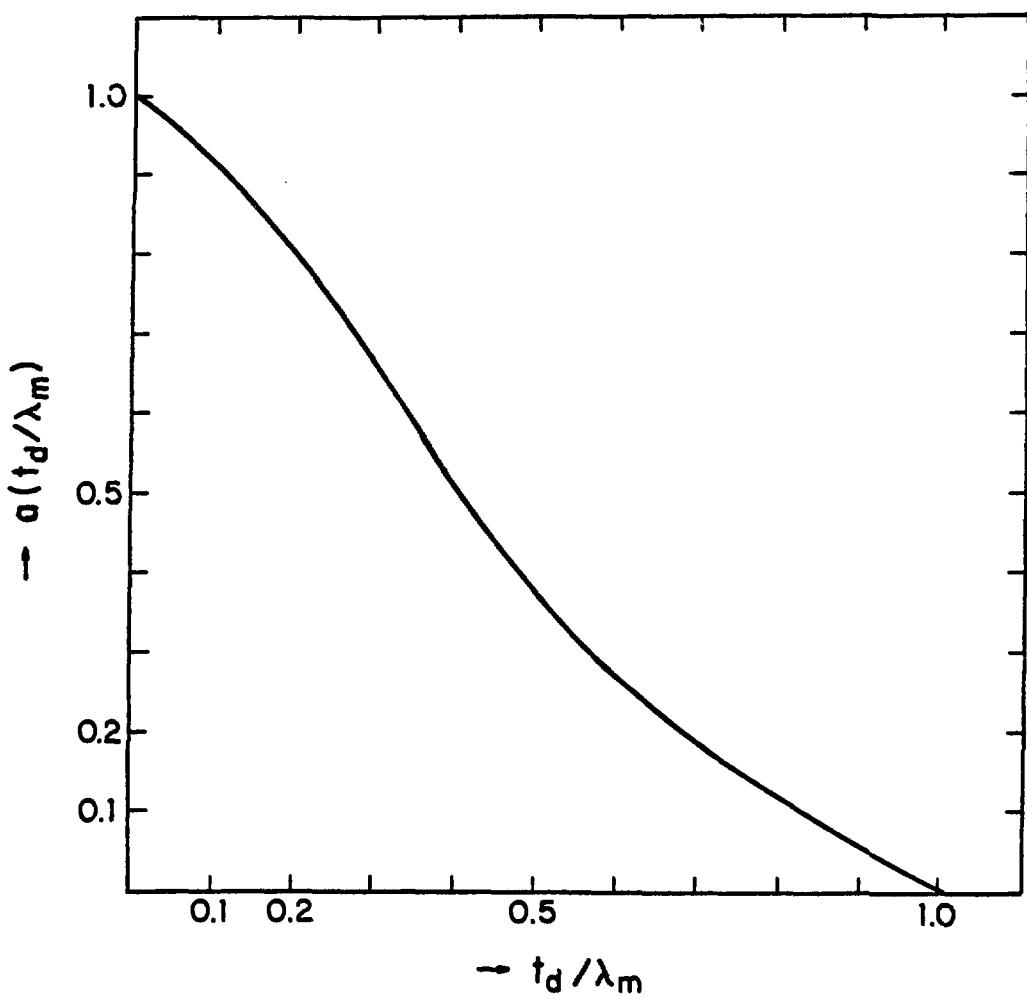


Fig. 13.



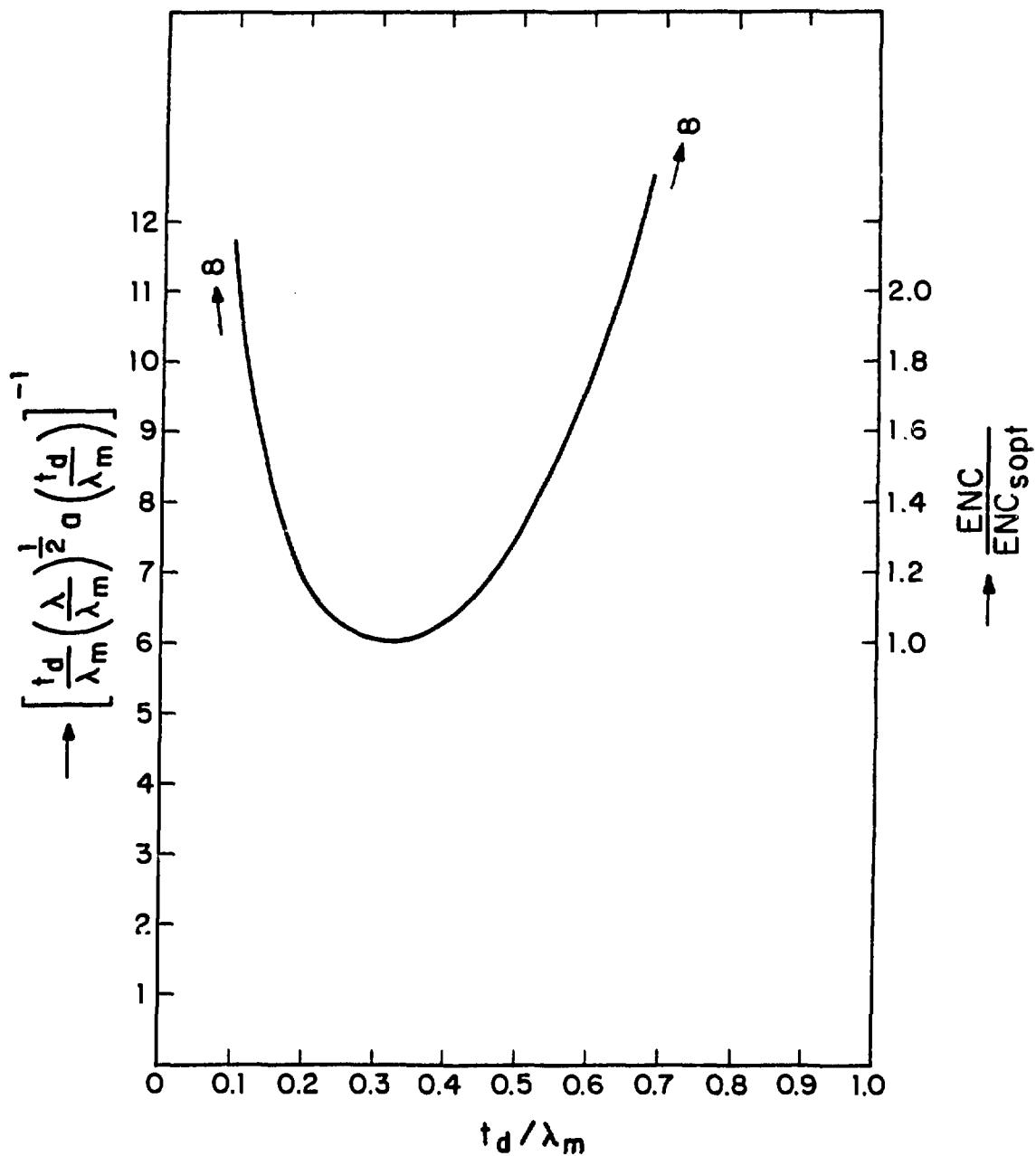


Fig. 11.

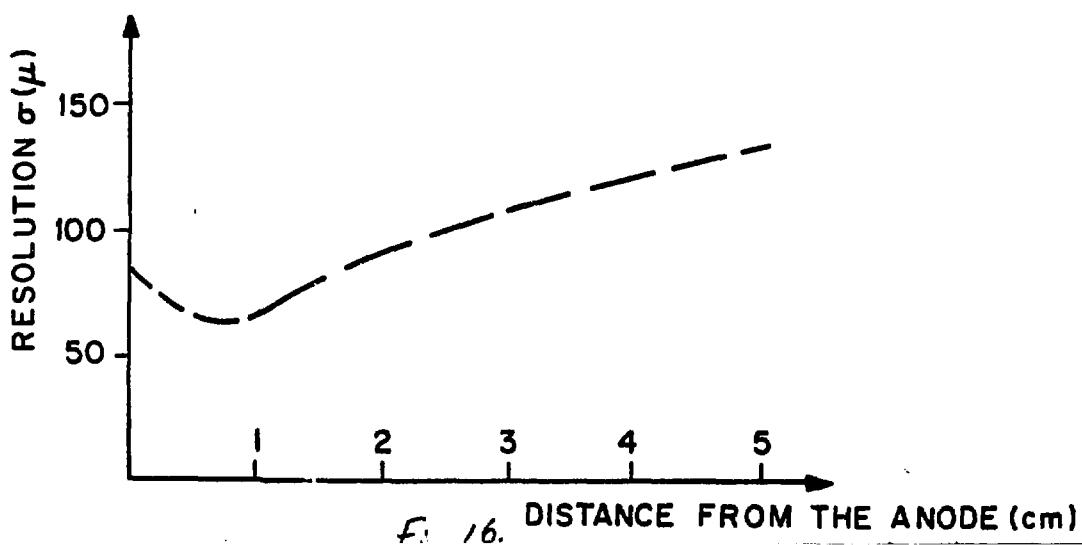
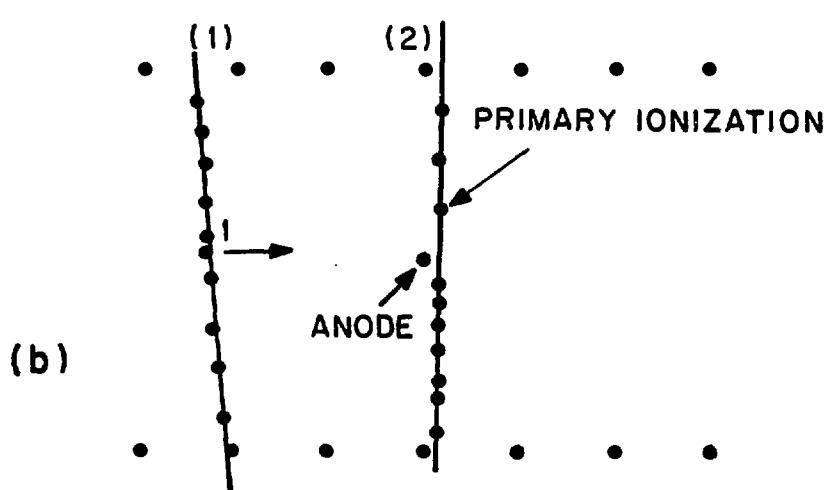
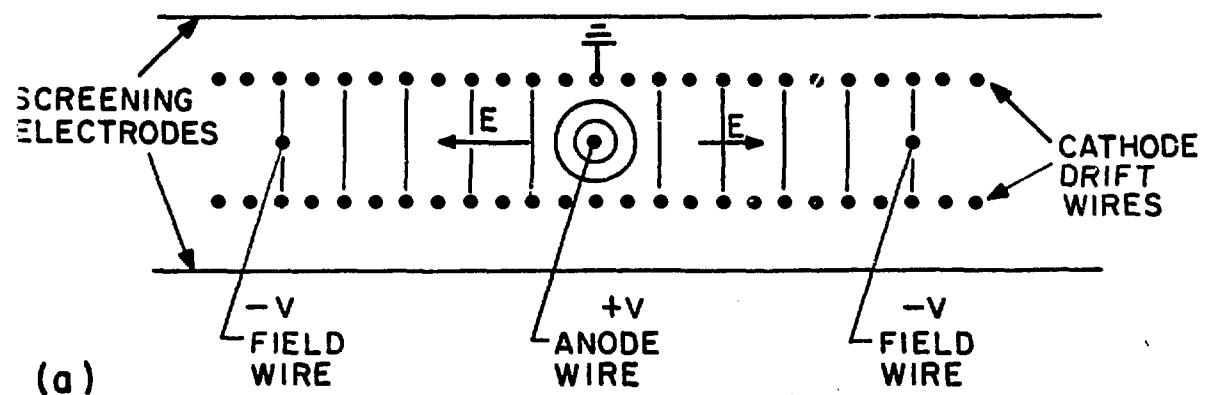
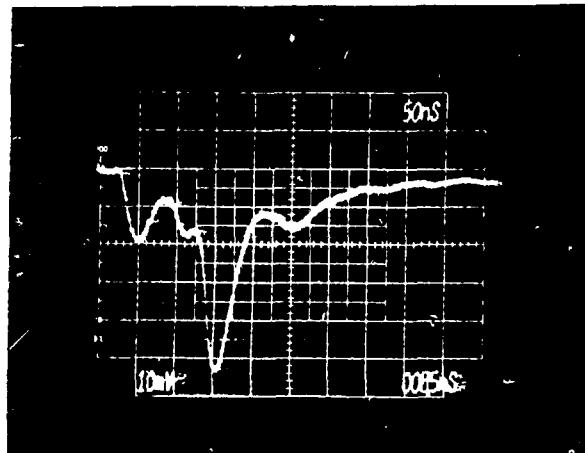
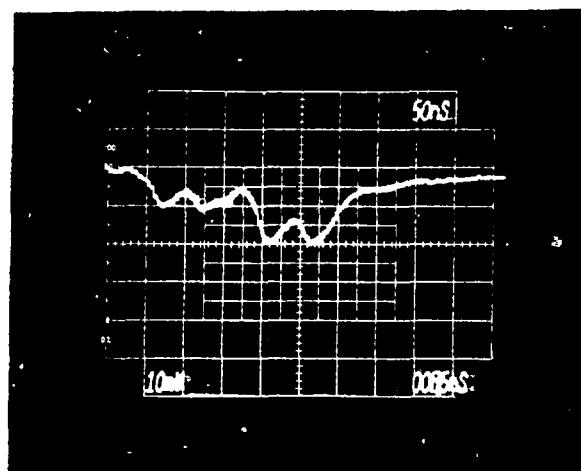
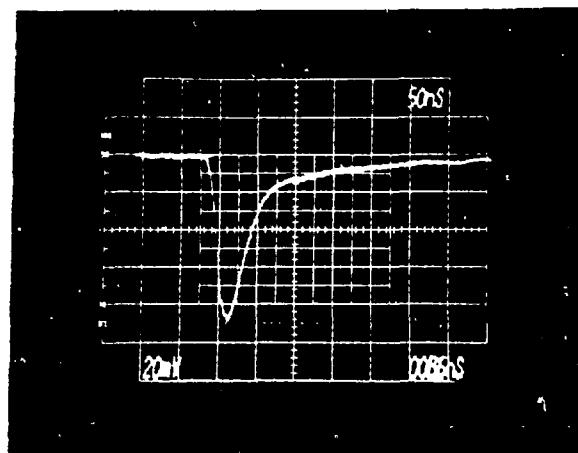


Fig. 16. DISTANCE FROM THE ANODE (cm)



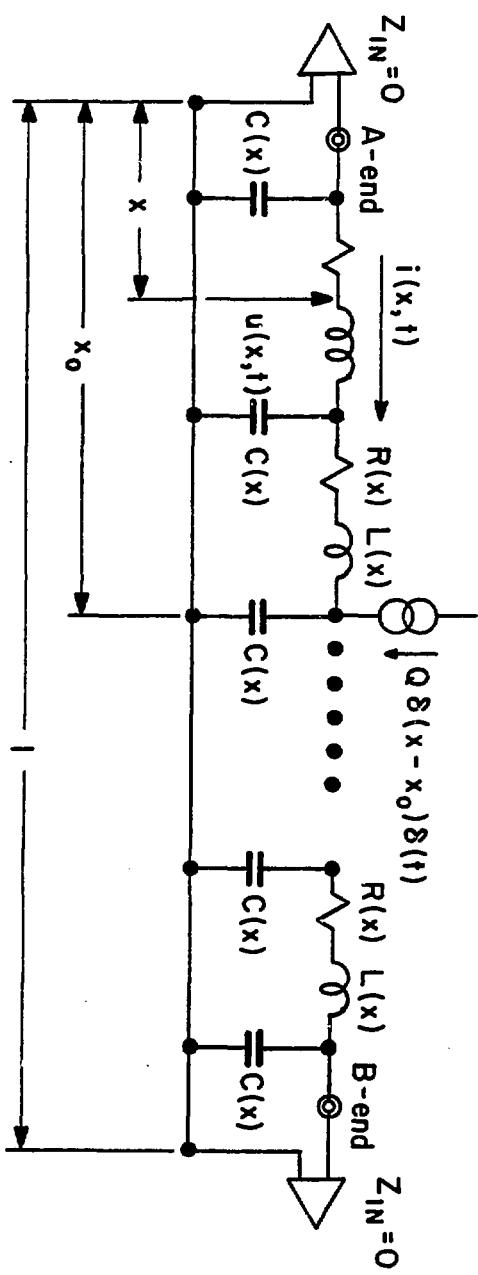


Fig. 18.

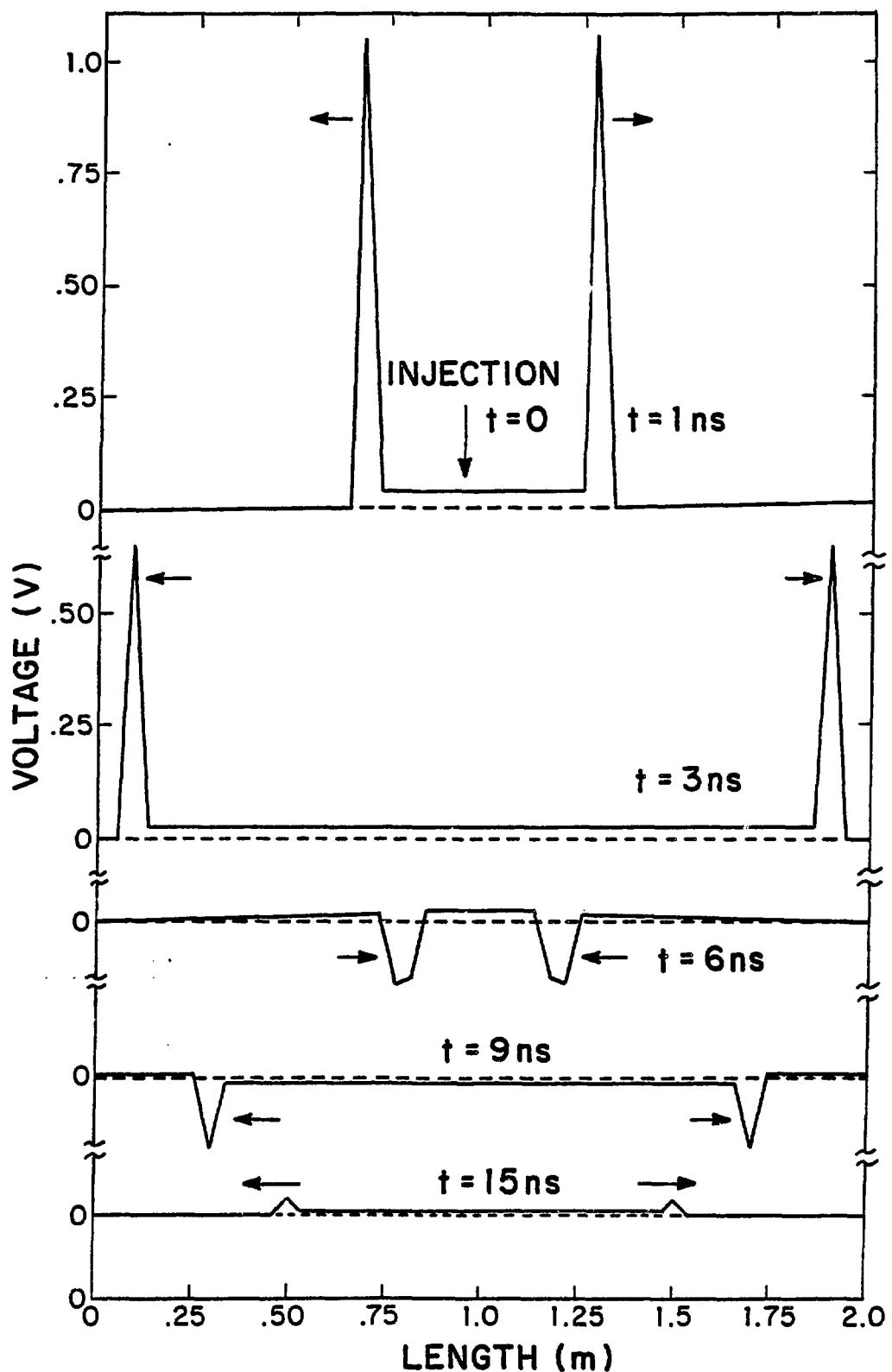
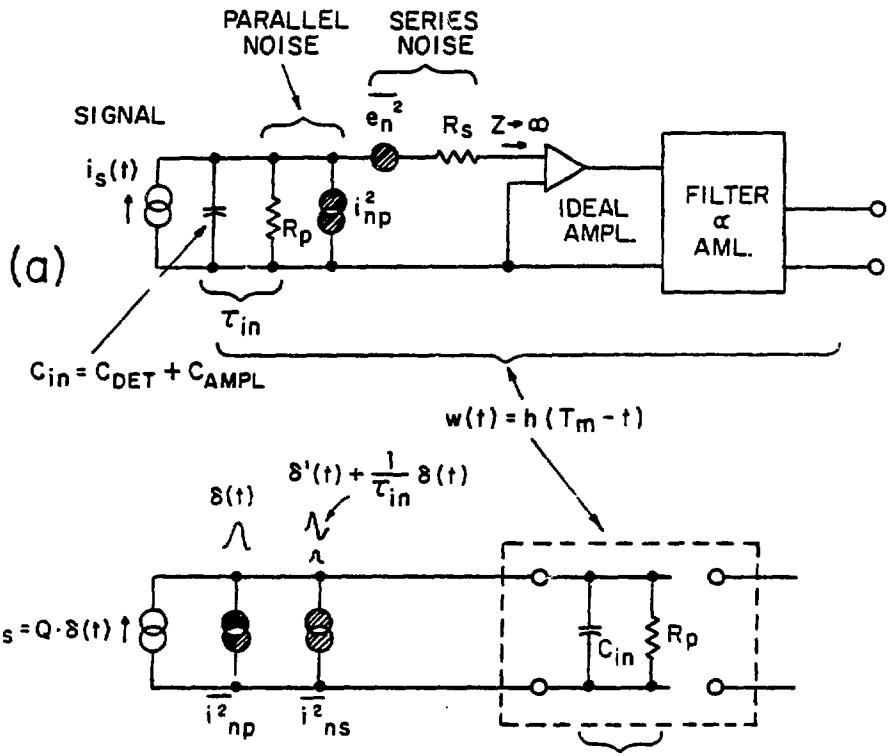
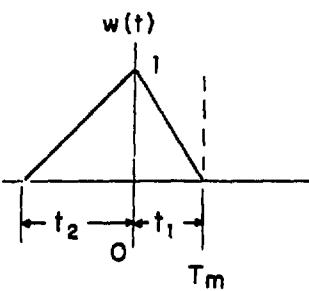
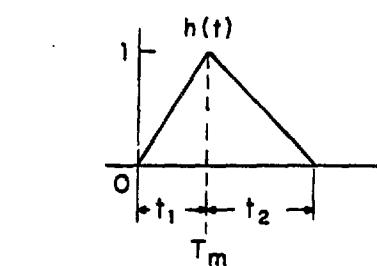


Fig. 19.



(b)



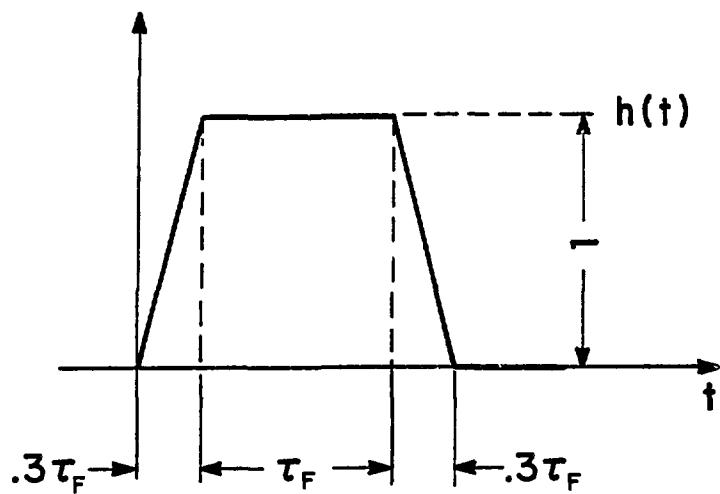


Fig 21

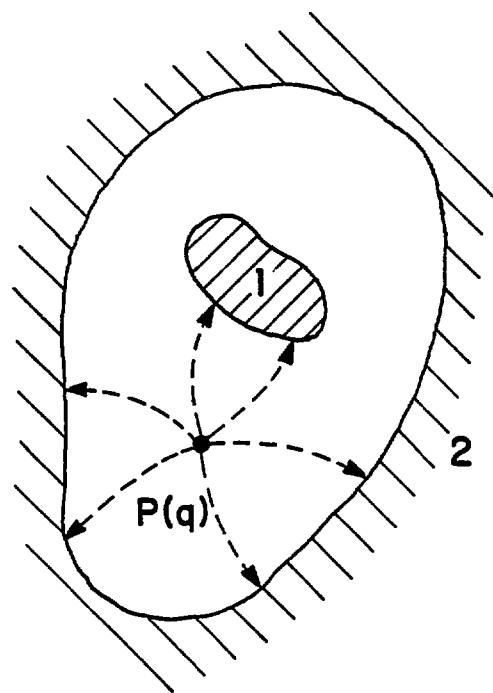


Fig 24

