7-1-1989

Digital Measurement of Partial Discharge

James Patrick Steiner
Purdue University

W. L. Weeks
Purdue University

Digital Measurement of Partial Discharge

J. P. Steiner
W. L. Weeks

TR-EE 89-40
July, 1989

School of Electrical Engineering
Purdue University
West Lafayette, Indiana 47907

DIGITAL MEASUREMENT OF PARTIAL DISCHARGE

A Thesis

Submitted to the Faculty

of

Purdue University

by

James Patrick Steiner

In Partial Fulfillment of the

Requirements for the Degree

of

Doctor of Philosophy

May 1988

This thesis is dedicated to my wife, Rebecca and my family.

ACKNOWLEDGMENTS

I would like to express my deepest appreciation to Prof. Walter L. Weeks for his support, encouragement and direction in preparing this thesis. Also I would like to thank the members of my committee, Michael Zoltowski, Edward Coyle and Vaughan Weston, for their invaluable comments, suggestions and discussions. Special thanks is due to Edward K. Duffy, Wilbur A. Starrett and William J. Schmidt (BF) of Essex (later Cablec) for their help in the cable experiments. I would also like to especially thank Eric S. Furgason for his help and suggestions. I would also like to acknowledge the support of Essex , Alcoa , Square D Company, Cablec and the Army Corps of Engineers (CERL).

TABLE OF CONTENTS

Page

LIST OF FIGURES

Figure Page

Figure Page

Figure Page

ABSTRACT

Steiner, James Patrick. Ph.D., Purdue University. May 1988. Digital Measurement of Partial Discharge. Major Professor: W. L. Weeks.

Various new measurement techniques have been developed for a high voltage phenomenon referred to as partial discharge. Partial discharge is a localized breakdown of the high voltage insulation system which is observed as low level, random emissions. Both electrical and acoustic emissions have been measured in underground power cables, solid cast power transformers and in lumped specimens. Typical problems complicating the measurements are the randomness of the emission, high levels of interference and extreme distortion of the signal by the propagation path.

Various signal processing techniques have been adapted to the measurement of partial discharge. The techniques investigated are capable of reducing noise in the measurements and have provided orders of magnitude improvement in sensitivity over ordinary methods. Some of the techniques studied are capable of providing information about the location of the partial discharge site.

CHAPTER ONE

PARTIAL DISCHARGE

1.1 Introduction

This study is concerned with the measurement of partial discharge in high voltage equipment. Partial discharge is a localized breakdown of the high voltage insulation system and is observed as extremely low level, random emissions that can be measured using electrical, acoustical, optical or thermal techniques. The most popular and, in most cases, the most sensitive methods are the electrical techniques. One common thread tying the different measurement methods together is that the observed signals are very low level, random emissions that are corrupted by noise. The randomness of the emissions complicates measurements but the noise is the primary difficulty associated with partial discharge measurements.

Various types of signal processing techniques are used to improve the sensitivity of the measurements and in some cases provide additional information that was not previously obtainable. The techniques that will be discussed are, for the most part, standard approaches that have been used for many years in different applications and have been successfully adapted to partial discharge measurements in this study.

Successful adaptation requires that the observed signals be modeled so that the necessary changes can be made in the techniques. Modeling of partial discharge is an extremely difficult task from the point of view of the observed signal properties. The difficulties arise because partial discharge is a complex physical process that is not completely understood at this time. Rather than trying to derive signal properties based on first principles, a phenomenological approach is taken because of the difficulties associated with developing a satisfactory physical model. The measurement techniques will use models based on experimentally observed behavior and in some cases the models will not be completely accurate but will lead to tractable results.

In other cases the models will be chosen so that existing solutions to other problems can be utilized.

There are different problems that are encountered in the measurement of partial discharge and each situation has its own problems. A lumped specimen was studied in a test cell that provides high integrity signals with little attenuation and minimum interference. In general, the attribute of lumped specimens is that the signal is attenuated and requires that a large amount of gain be used in the measurement system. Large gains always cause problems with the thermal noise in the amplifier, however, this is the simplest noise to deal with. Also associated with high gain is susceptibility to external interference. External interference can take many forms and in some cases it is easily dealt with in the same manner as the thermal noise but there are situations where the interference is difficult to eliminate.

Distributed systems, such as cables, suffer from the same type of difficulties with noise as lumped specimens. The difference in the measurement is that the propagation of the signals must now be taken into account. The presence of a well defined propagation path allows the use of estimation techniques so that the partial discharge site can be located. Various difficulties are encountered in estimating the position of the partial discharge site. These difficulties are associated with dispersion in the cable, reflections in the cable and the presence of multiple partial discharge sites.

Another type of distributed parameter system is the transformer and it is certainly the most difficult to measure and analyze. Transformer measurements suffer from the same type of noise and interference problems as mentioned above. The new element in the problem is that the propagation paths are more difficult to define and describe. From an electrical standpoint it has been possible to construct linear circuit models as discrete approximations to the distributed parameter system but these models appear to be valid only across narrow frequency bands and at lower frequencies. Acoustically, the propagation paths escape description except

for very localized simple models. This difficulty can be traced to the highly nonhomogeneous nature of the materials though which the acoustic signals must propagate.

1.2 Background

In the manufacture, installation, or operation of high voltage equipment small occlusions may develop in the insulation. These occlusions enclose a different dielectric material than that used for insulation, and it is often a gas whose dielectric constant is much less than material around it. This causes an abnormally high electrical stress in the void and partial discharges may occur. Discharges in voids, enclosed by dielectric or conductor and dielectric, cause degradation of the insulation. Damage resulting from these discharges may lead to total failure of the insulation system. Studies on time to failure have been made for various types of defects [1]. This study investigated the life time of dielectrics with cavities and Fig. 1.1 shows its results as a double logarithmic plot of field strength vs. life. Consider a typical cable system for which field strengths are on the order of 0.5 to 2 kV/mm. As can be seen, for relatively minor defects, the lifetime of a cable, under normal operating conditions, can extend into thousands of hours.

To relate the measurements made to the discharge process certain fundamentals need to be discussed. In a void with dielectric or conductor-dielectric boundaries two basic mechanisms are believed to cause breakdown. The first is the Townsend mechanism and the second type, which is believed to start initially as a Townsend mechanism, is the streamer mechanism [24,25,2,3,4].

The Townsend mechanism is a release of an avalanche of electrons. When the surface of a material is highly stressed and an energetic particle or photon of sufficient energy strikes this surface an electron is emitted. This electron collides with molecules in the void and releases further electrons. These electrons in turn release further electrons thus creating an avalanche. The molecules that released the electrons are now positive ions. As the positive ions accumulate, a positive space charge develops which enhances the field and accelerates the electrons causing further electron release. As the discharge continues the space charge migrates away from the surface. As the space charge migrates, the field, which was originally enhanced by the positive ion cloud, is reduced and this quenches the electron avalanche.

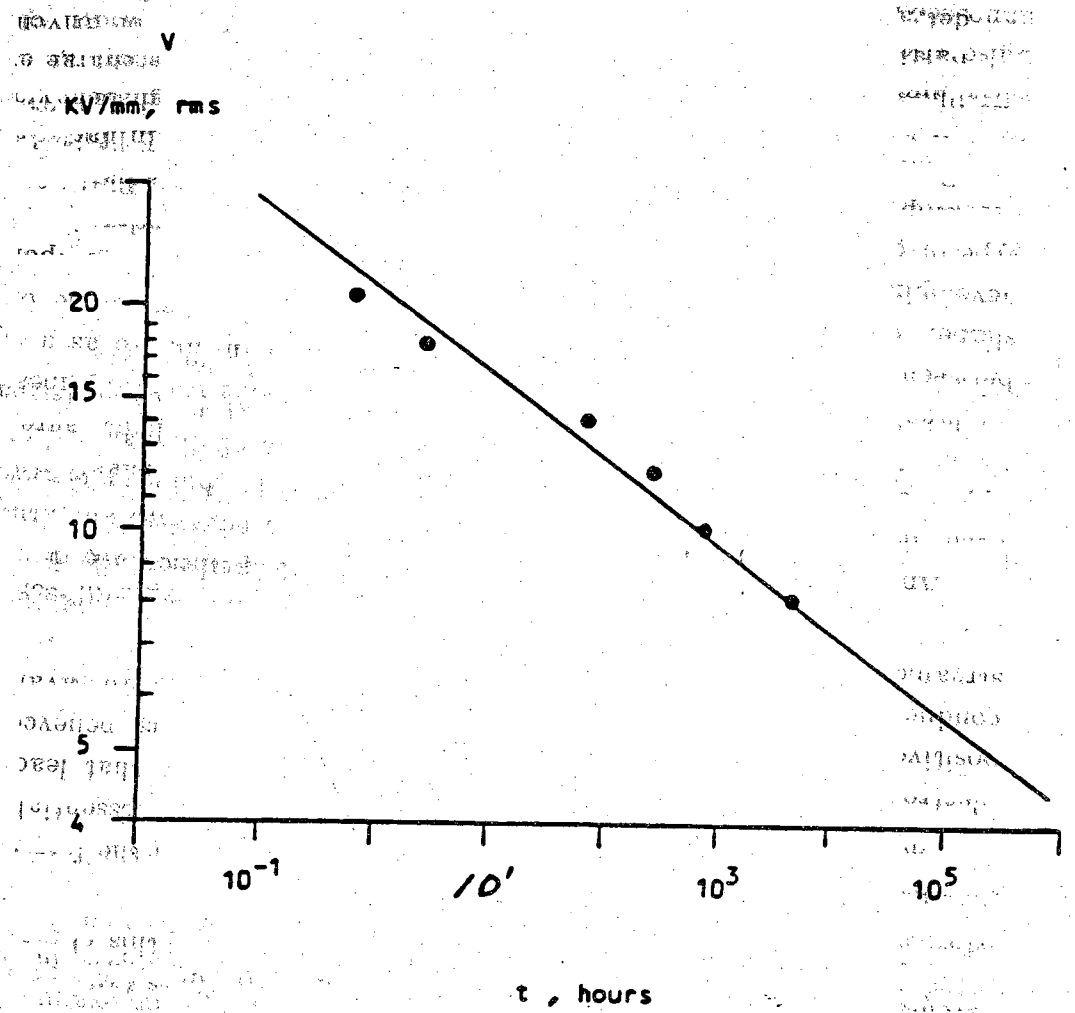


Figure 1.1. Voltage-life Characteristics of polyethylene.

The scenario above is greatly simplified and two points need emphasizing. The initializing event that starts the avalanche is random. For example, it can in be generated by background radiation which has events that are often modeled as a Poisson process. In fact radioactive illumination of the void drastically alters the observed behavior of the partial discharge [67]. The second point is that the space charge eventually detracts from the field and most of the positive ions must be driven into the opposite electrode before the probability for another discharge becomes appreciable. In other words there is a mechanism which inhibits further events from occurring. A point not brought out in the above discussion is that typical partial discharge pulses have random energies (amplitudes). It is not clear what mechanisms cause the randomness but this characteristic has been recognized for several decades.

The second type, the streamer mechanism, begins with the formation of electron avalanches, as in the Townsend mechanism. Essentially, the positive space charge causes auxiliary electron avalanches that leads to a conductive plasma channel that bridges the void. It is believed that streamer development is preceeded by a buildup of roughly 10^6 avalanches [3].

An elaborate and accurate analysis of partial discharges in voids has been developed in [24,25]. That work takes a rigorous mathematical approach using Maxwell's equations along with the continuity equations for the various charged species in the void to describe the propagation of the avalanche across the void. There appears to be excellent agreement between the theoretically predicted shape and the experimentally measured shape. One signal type observed in this study was referred to as a slowly developing partial discharge (Townsend mechanism). An example of this type of discharge is shown in Fig. 1.2. Another type of discharge that was observed was referred to as a rapidly developing partial discharge (streamer mechanism). An example of this type of discharge is shown in Fig. 1.3. The figures depicting these two types of partial discharge were band limited to 20 MHz and do not indicate the broad band nature of the signals. Typical pulse rise times can be less than 1 ns with pulse widths as short as a few nanoseconds. However, for large void dimensions the pulse width of the ionic component can be as large as thousands of nanoseconds.

In many typical measurement situations the the received signal is so distorted that the original pulse shape is inconsequential. For example in transformer measurements there are so many self resonances in the

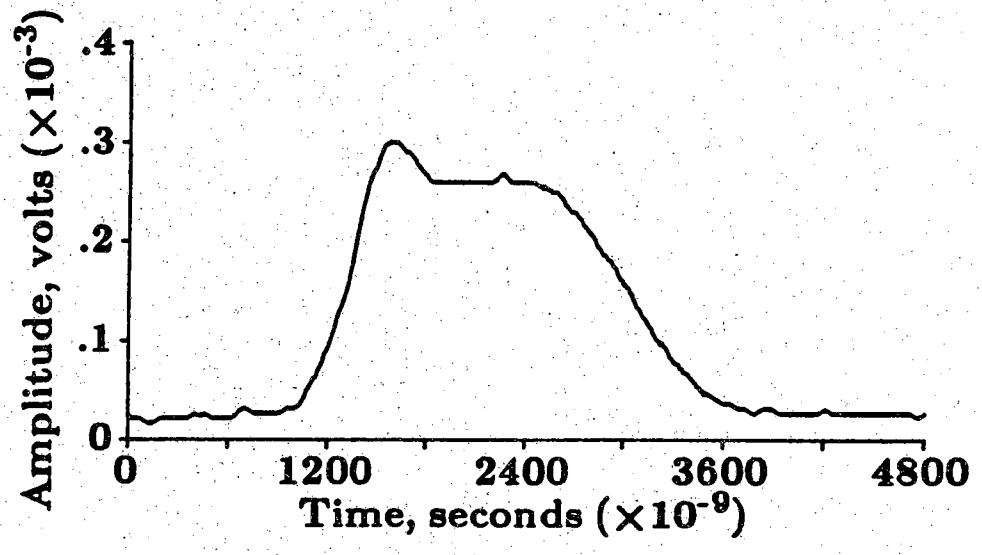


Figure 1.2. Slowly developing partial discharge.

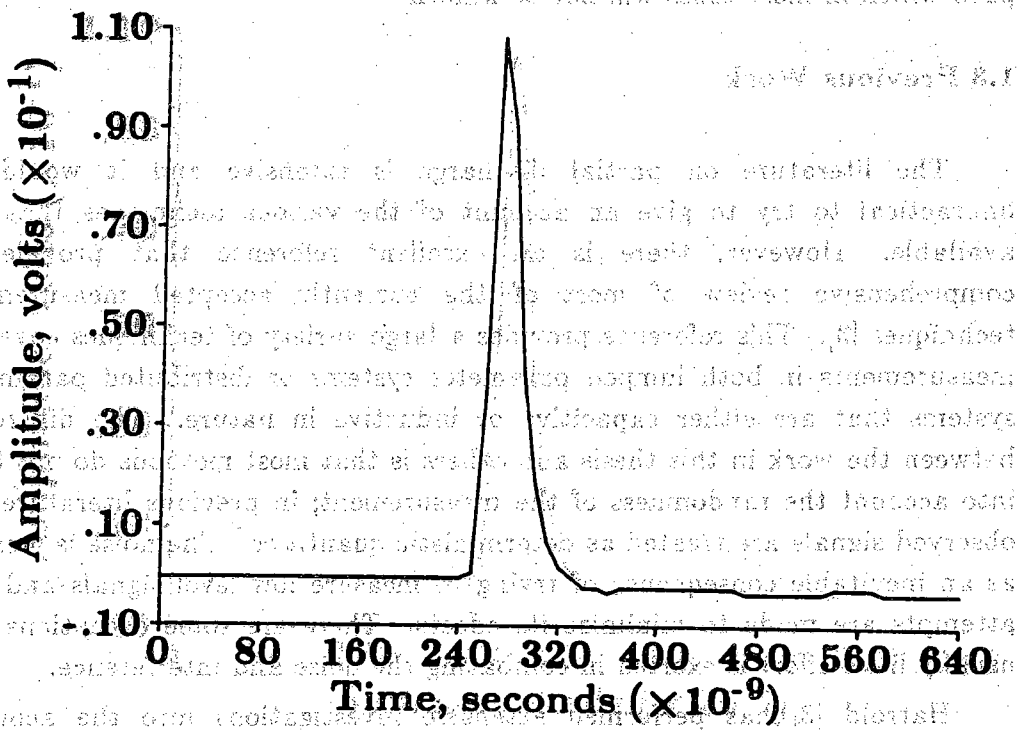


Figure 1.3. Rapidly developing partial discharge.

transformer that the observed signals bear no resemblance to the original pulse shape. Another example is given by URD cable which is so lossy that the maximum received bandwidth after propagation of only 500 meters will be a few MHz. In both of these cases the original waveform was completely obscured. In these cases the only partially distinguishable characteristics of the original waveform are the time of occurrence and the energy. The most important factor influencing the observed signal is then the propagation path which in most cases will not be known.

1.3 Previous Work

The literature on partial discharge is extensive and it would be impractical to try to give an account of the various techniques that are available. However, there is an excellent reference that provides a comprehensive review of most of the currently accepted measurement techniques [3]. This reference provides a large variety of techniques covering measurements in both lumped parameter systems or distributed parameter systems that are either capacitive or inductive in nature. The difference between the work in this thesis and others is that most methods do not take into account the randomness of the measurement; in previous literature the observed signals are treated as deterministic quantities. The noise is treated as an inevitable consequence of trying to measure low level signals and few attempts are made to minimize its effects. There are some exceptions but usually little effort is exerted in combating the noise and interference.

Harrold [3] has performed extensive investigations into the acoustic behavior of partial discharge. The references contained in [3] is not an exhaustive list of Harrold's work but does provide a reasonably complete list of investigations made into the acoustic behavior of partial discharge.

The thesis is different from others because acoustic methods are used for locating partial discharge in solid cast transformers. These efforts are detailed in chapter four where a rigorous development of the measurement techniques is presented. A slightly different approach was taken and the main similarity between this study and others is that acoustic signals are utilized.

A most notable work is [21] in which the fundamental limitations of partial discharge measurements are discussed. This paper discusses the matched filtering approach to amplitude estimation in white noise for ultra

wide band measurements. Results are given relating the sensitivity of the measurement to the bandwidth of the measuring system. Further results are given relating the minimum sensitivity of partial discharge measurements to the length of the cable through which the partial discharge is observed. This was derived using the bandwidth relationship and the measured transfer characteristics of a typical lossy power cable.

Another method, [9], by the same authors, describes a partial discharge location technique using a correlation based system. This method is intended for ultra wide band measurements and only uses one or two bits of amplitude information. As a consequence the results are highly nonlinear and it would appear that subsequent processing might be impeded by the presence of the nonlinearity. One of the attributes of correlation analysis as used in time delay estimation is that the correlation functions can be optimally filtered. This aspect of time delay estimation is discussed in chapter two.

Another method, [82], attempts to optimize the temporal resolution of the measurement using a Wiener filter. The formulation in [82] assumes that the signal is stationary which is not the case for transient pulses. Narrow band examples are used as illustrations of the technique and the estimate of the charge (amplitude) is provided by performing the integration

$$Q = \int_{-\infty}^{\infty} i_o(t)\, dt$$

where $i_o(t)$ is the optimally filtered current waveform. The problem with this solution is that the original observed waveform does not appear to have a DC component and therefore the optimaly filtered version cannot have a DC component. The integral used to estimate the charge therefore evaluates to zero.

Another paper, [10], discusses the advantages of signal processing for making partial discharge measurements. The authors state that full wave rectification of the acoustic signals and electrical signals can be used to locate the source of acoustic emissions in transformers. The paper is extremely vague and it is difficult to see what methods they are using and how the signals are modeled and utilized in the subsequent processing.

Acoustic location methods are put on solid mathematical foundation in chapter four of this thesis in which a similar method is used for locating acoustic emission sites.

A partial discharge location scheme is described in [8] but it is not completely clear how the location is determined. The method is described using a block diagram with the critical blocks stating what the block output is without any mention of the manner in which the output is determined. The critical element in the block diagram is the time of arrival section and it is never stated how the time of arrival is measured. The purported accuracy is described but there is no way of analytically verifying this.

An energy measurement scheme is described in [83] using a cross-correlator. This is a novel approach for partial discharge measurements that is closely related to some techniques described in this chapter. Unfortunately, the authors failed to recognize the tremendous power afforded using correlation techniques. Use of similar techniques can provide vast improvements in system sensitivity and experimental simulations have indicated that measurements sensitivities can be improved by at least three orders of magnitude.

An intriguing new approach to partial discharge measurements has been developed in [84]. This technique appears to have great promise if analytical methods can be found for interpreting the data. A short discussion concerning a mathematical framework for this technique is contained in chapter two. Measurements such as those found in [84] will probably lead to the next generation of partial discharge measurements techniques.

Another fascinating new method is contained in [13] in which X-rays are used to excite the partial discharge process. The technique involves illuminating the partial discharge site with a modulated X-ray source (via a chopper). The modulation waveform is then cross-correlated with the output of a standard partial discharge detection circuit using a lock-in amplifier. In essence the modulation source correlates with the changing statistics of the random point process describing the partial discharge. This technique is used to locate partial discharge site by sweeping the X-ray across the device under test. If no partial discharge site is encountered then the correlator does not respond. However, if the X-ray source illuminates a partial discharge site then the correlator responds which is the indication that there

is a site at that location.

1.4 Contributions

This thesis applies the theory of digital signal processing to partial discharge metrology. As seen in section 1.3, which is a limited search of current literature, very little work has been done in formulating partial discharge measurements in terms of modern signal processing techniques. This thesis develops a signal processing framework by presenting a signal model and adapting a variety of signal processing techniques to partial discharge measurements.

The title of this thesis (Digital Measurement of Partial Discharge) was chosen because the techniques that are presented measure well defined parameters associated with the partial discharge, such as its location. These parameters are fundamental to any method aimed at interpreting partial discharge. The information presented in this thesis might be thought of as the front end processing of any technique aimed at interpreting partial discharge because it provides accurate methods for measuring the data. Analysis of data requires that a model be used so that inferences about the partial discharge can be made in terms of the model. Analysis and interpretation of partial discharge is a difficult task and is only briefly investigated in chapter two. The model presented in chapter two and similar models are believed to provide the proper mathematical tools for any such investigation.

Paramount to the development and analysis of a signal processing algorithm is the use of an accurate signal model. A signal model is presented in chapter two which is believed to be new to partial discharge measurements. The model is very natural for this application and is widely used in many disciplines. In essence this model takes into account the random variations of both the time of occurrence and the amplitude of the partial discharge. The models presented are only first order models and more elaborate and accurate models could be developed with further work. However, experimental work shows the consistency of the measurements with the simple, first order models.

Some fundamental measurement techniques are presented in chapter two which can be of use in practice. The techniques applied are standard methods in signal processing but have not been used in partial discharge

measurements before. Also a more accurate lower bound on the measurability of partial discharge is discussed in terms of a well known result. Finally, a detection technique for measuring ultra low level partial discharge is discussed along with a technique which greatly reduces the amount of data processing required.

Chapter three discusses general techniques for estimating the locations of partial discharges. There are several contributions in this chapter related to the generalized cross-correlator. A detailed analysis of the bias of delay estimators is presented and this material is believed to be new. The bias expressions are developed from techniques that appeared recently in signal processing literature and are an extension of these results to practical measurement situations. Two new generalized cross-correlators are presented one of which is a minimax version. These generalized cross-correlators are used to enhance the resolution of the delay estimation measurements so that multiple partial discharge sites are more easily located.

Partial discharge measurements in cables are discussed in chapter four and several contributions are presented in this chapter. Several standard signal processing techniques are adapted to the partial discharge measurement in cables for the first time; system identification, equalization and generalized cross-correlation. Also a discrimination technique for eliminating impulsive interference in partial discharge measurements is used for the first time and is implemented using a generalized likelihood ratio test. A new application of the Gerchberg-Papoulis algorithm is presented which is capable of iteratively removing the effects of narrow band interference. Super-resolution is investigated using an ad hoc implementation of the MUSIC algorithm. The MUSIC algorithm is applied to the analysis of multiple partial discharge sites in short cables when data is collected with a low bandwidth system. Finally, the peculiarities of the signal model are exploited to provide a simple high resolution method for delay estimation in cables.

Partial discharge measurements in solid cast transformers are presented in chapter five. One technique that is investigated is the acoustic method which has been used for several years by the transformer industry; however, this is the first time it has been applied to solid cast transformers. Furthermore, the technique has been refined in a manner which is believed to be new. Current literature on acoustic measurement techniques lack the mathematical details presented in chapter five and this development is also

believed to be new. A new asymptotic maximum likelihood measurement of the attenuation of the acoustic energy is developed which provides a new method for locating partial discharge. A new, unique measurement technique, based on the rf magnetic fields in the transformer windings, is presented. Together with an asymptotic maximum likelihood algorithm, the magnetic measurements provide a new method to locate partial discharge in transformers. Finally, the same ad hoc formulation of MUSIC, used in chapter four, is applied to the signature analysis of partial discharge waveforms in transformers. This new technique provides another highly effective method for locating partial discharge in transformers.

CHAPTER TWO

LUMPED SPECIMENS AND MODELS

2.1 Introduction

Good models make it possible for digital signal processing to improve measurements of partial discharge. In the past, when partial discharge has been modeled, the approach has been to develop expressions that describe the physical behavior of the individual partial discharge pulse. While this is of great interest, it does not provide the information necessary for the signal analysis used in this thesis. In contrast, the work in this thesis requires statistical models of the entire observed signal process so that signal processing algorithms can be developed. Since the necessary physical models required to develop statistical descriptions of the observed signal process are not yet available the approach to modeling in this thesis is phenomenological.

The chief advantage of using signal models is that many results that are available in signal processing literature will be available for adaptation to partial discharge measurements. Depending upon the situation different types of models will be used. If the observed partial discharge is very active and the pulses are not easily separated then the appropriate model would be a stochastic process. In many cases there are only a few pulses and single events are easily observed. The appropriate model in this case is a nonstationary transient model in which only the single pulse is considered.

Individual partial discharges are random, time localized events. When random events are time localized they are often modeled as a stochastic point process. There are various forms which a point process can take and the general type which will be used for partial discharge can be called a shot noise process. To characterize this type of process the times of occurrence of the partial discharge and the amplitudes must be measured. Section 2.2 considers various estimates of these quantities.

In general, mathematical analysis of point processes is difficult unless very special models are used. Section 2.3 describes, in detail, a special point process which is referred to as the Poisson process. The reason for treating this case in detail is the amount of literature available for this model. Also, under suitable conditions or interpretation, the Poisson model appears to be a valid model for the analysis of partial discharge.

In more accurate analysis it may be desirable to employ more sophisticated models which use more of the available information. Section 2.4 will consider another special model referred to as the renewal model in which the memory of the process is restricted to the previous event and the intensity is homogeneous. The distribution of the inter-arrival times between partial discharge pulses will be the important parameter in this case.

Point process models can be of use when trying to estimate various quantities that may be of interest. Section 2.5 gives the details of some experimental work which uses these models to estimate fundamental quantities associated with the geometry of the gap in which the partial discharge is occurring. Measurements are made which can be used to infer the size of the gap. Further measurements are made which estimate the conductivity of the surface of the gap for high conductivities. Similar types of arguments will be used in the analysis in later chapters when the locations of the partial discharge is inferred.

Techniques used to increase the signal to noise ratio are presented in section 2.6. One technique is similar to a detection technique in that a threshold is selected to provide the detection of an event. The mean number of level crossings of the signal process are used to approximate the increase in SNR. This technique is capable of detecting pulses buried in thermal noise and has provided gains in the SNR of the measurements. The advantages that this technique has over a rigorous detection approach is that this method can be easily implemented with standard equipment and, more importantly, the data rate is slow which reduces the computational load.

2.2 Amplitude and Occurrence Time Estimation

One of the most fundamental quantities that is of interest in partial discharge measurements is the charge transferred by the discharge current.

That charge is defined as

$$Q = \int_{-\infty}^{\infty} i(t)\, dt$$

where i(t) is the current in the partial discharge. It is easy to see that this integral is equivalent to evaluating $I(\omega)$ at $\omega = 0$ where $I(\omega)$ is the Fourier transform of the current pulse. Unfortunately, this equation is not of much use in practical situations because the measurement circuits used are almost never DC coupled to the partial discharge site. The usual manner in which this problem is circumvented is through calibration of the system under test. In this manner the voltage amplitude of the output from the detection circuit is compared with a calibrated reference value. The problem is then to estimate the amplitude of the voltage in the best manner possible and some solutions to this problem will be presented below.

Another fundamental quantity of interest is the occurrence time of the partial discharge pulse. If statistical analysis of the partial discharge process is desired then both the charge and time of occurrence are necessary pieces of information. If it is assumed that the pulse shape is known then the occurrence time is easily estimated using some very classical results. These results will be discussed in conjunction with the LMMSE estimator described below because the solutions are intimately related.

The signals to be discussed will be general in nature; they could be the output from either a narrow band or wide band partial discharge detector. The observed signal will be modeled as

$$x(t) = a_i\, s(\, t - \tau_i\,) + n(t)$$

where s(t) is the output of the particular partial discharge detector being used. The pulse shape, s(t), will be assumed known and is easily measured using the calibration signal. The amplitude, a_i and the time of occurrence, τ_i of the signal, s(t), will be considered to be random. The noise term, n(t), will be assumed to be stationary and have a power spectral density N(f). The noise term, n(t), will model the thermal noise in the amplifiers of the detector and the quantization noise of the digitizer. Any external interfering signal that is from a stationary random process can also be modeled by the

noise term, n(t). A small amount of white Gaussian (thermal) noise will be assumed to be present at the input to the quantizer and as a consequence the noise from the quantizer will be assumed to be uniformly distributed white noise. Nonstationary interference such as intermittent or impulsive signals not related to the desired signal, s(t), will not be considered in the ensuing analysis. Other techniques, discussed in later chapters, will be used to discriminate between the desired signal and interference.

The first estimator to be discussed is the LMMSE filter, [42], and takes the form of a linear filter, $H(\omega)$, which is given by

$$H(\omega) = \frac{\alpha\, S^*(\omega) \exp(-j2\pi\tau_i)}{N(\omega)}$$

where $S^*(\omega)$ is the complex conjugate of the Fourier transform of s(t). The constant α is given by

$$\alpha = \frac{A}{1 + A \int_{-\infty}^{\infty} \frac{|S(\omega)|^2}{N(\omega)}\, d\omega}$$

where A is the mean square value of a_i. It should be noted that the filter attenuates the observed signal at those frequencies at which the noise level is high and passes those frequencies where the noise is low. This could be particularly useful when combating narrow band interference, for example the interference due to an AM broadcast. The estimate of the amplitude is then given by

$$\hat{a}_i = \mathcal{F}^{-1}\{ X(\omega)H(\omega) \}\Big|_{t=0}$$

As can be seen, to implement this filter, it is necessary to know the arrival time, τ_i, of the partial discharge pulse. Since the optimal linear filter is in essence a correlation operation the estimate of the arrival time is the time at which the output of the filter achieves its maximum. This estimate provides the second piece of information necessary to statistically characterize the partial discharge process. Other pieces of information that are required are

the mean square value of the random amplitudes. It is impossible to know this quantity a priori but satisfactory results have been obtained by simply using the reference value from the calibration. It is also possible to adaptively adjust the mean square value as data becomes available, however, this aspect was not pursued.

There are other implementations of the LMMSE estimator which address different aspects of the amplitude estimation problem. One typical problem, encountered in practice, is the estimation of the amplitudes of closely spaced partial discharge pulses. When this occurs, superposition of the pulses can cause errors in the amplitude estimate. As mentioned in chapter one, a solution to this problem was given in [82]. However, the formulation was incorrect and the amplitude estimate for most cases (non-DC coupled) is zero. The proper solution can be found in [42] and is given by

$$H(\omega) = \frac{\beta\, \overline{a^2}\, S^*(\omega)}{\overline{a^2}\, \lambda\, |S(\omega)|^2 + N(\omega)} \quad ; \qquad H(0) = 0$$

where β is defined as

$$\beta = \left[1 + \int_{-\inf}^{\infty} \frac{\overline{a^2}\, |S^*(\omega)|2}{\overline{a^2}\, \lambda\, |S(\omega)|^2 + N(\omega)}\, d\omega \right]^{-1}$$

This solution assumes that the input signal process is a homogeneous Poisson shot process with intensity parameter λ and that the signal is not DC coupled. The estimate of the amplitude is given by

$$\hat{a}_i = \mathcal{F}^{-1}\{ X(\omega) H(\omega) \} \Big|_{t=t_i}$$

where t_i is determined as the peak of the filter's output. In essence, this filter performs a deconvolution of the input signal process. In some cases it will be necessary to constrain the solution to reduce the variance of the estimate. Some typical constraints for this type of problem are discussed in later chapters. In general, deconvolution schemes of this type provide increases in the pulse resolution on the order of 3 to 5 times before the

variance of the estimate becomes excessive.

The second method, the autocorrelation (AC) estimator, is an ad hoc technique and is very simple to implement. This technique is based upon a maximum likelihood argument used in a different context in chapter 5 but it should be noted that this estimator is not a maximum likelihood estimator for this case. The estimate is formed by evaluating the autocorrelation function at zero delay, removing the bias due to the noise, and taking the square root of the result. The value of the time autocorrelation function at zero delay is defined by

$$y = \int_{-T/2}^{T/2} x^2(t)\, dt$$

where the partial discharge pulse is assumed to be contained in the interval $(-T/2 , T/2)$. It was more convenient in this analysis to omit the $\frac{1}{T}$ factor from the usual definition of the time autocorrelation function. The resulting value will be an estimate of the sum of the noise energy and the energy in the partial discharge pulse. Taking the expected value one obtains

$$E\{ y \mid a_i \} = a_i^2 \mathscr{E} + T\sigma^2$$

where σ^2 is the variance of the noise and $\mathscr{E}$ is the energy in the signal pulse, s(t). The term, $T\sigma^2$, is an unwanted bias term and is removed from the calculation before further processing. The final step is to calculate the square root of the result after removal of the bias. The estimate of the amplitude is then

$$\hat{a}_i = \frac{\sqrt{y - T\sigma^2}}{a_{calib}}$$

where a_{calib} is the reference level obtained through calibration.

One interesting sidelight is that the AC method can provide meaningful results when the partial discharge rate is too high to resolve the individual pulses. Consider the case in which the observed signal is

$$x(t) = \sum_{i=1}^{N_T} a_i \, s(t - \tau_i) + n(t)$$

where there are now N_T pulses in the interval ($-T/2$, $T/2$). If the process is a homogeneous Poisson process with intensity parameter λ then the estimate of the the autocorrelation function at zero delay becomes

$$y = \lambda A \mathscr{E} + T\sigma^2$$

where A is the mean square value of the amplitude defined above. The output of the estimator is slightly different because it now gives an estimate of the square root of λA. We no longer get an estimate of the magnitude of the charge but an estimate of its average rms value.

Three estimators were compared experimentally; the peak value, LMMSE and AC estimators. Measurements were made using a Data 6000 waveform analyzer which has an 8 bit A/D converter and a general purpose microprocessor. The measurements were made using a precision calibration source, calibrated attenuators and a noise generator. Two different signals types where investigated: a broad band pulse and a narrow band pulse. These signals were chosen because they are representative of outputs from typical partial discharge detectors. The broad band pulse had a band width of 2 MHz while the narrow band pulse had a bandwidth of 10 KHz and a center frequency of 185 KHz. The signal was fed through an attenuator into one input of a summing amplifier and the noise source was fed through an attenuator into the other input of the summing amplifier. In this manner various input signal levels and noise levels could be measured.

The first measurement represents a typical calibration curve obtained by injecting a calibration pulse into the measurement system. The calibration signal was varied from 1% to 1000% of the full scale input level of the digitizer. The curve in Fig. 2.1 represents the nonlinear transfer function of the measurement system, with the noise bias removed, for various input levels using the AC estimator. The curve is linear until the full scale input level is reached and after this point the measurements begin being compressed. This calibration curve can be used to undo the nonlinearity by mapping the measured value into its actual value. The calibration curve will be unique to the measurement system, the calibrating

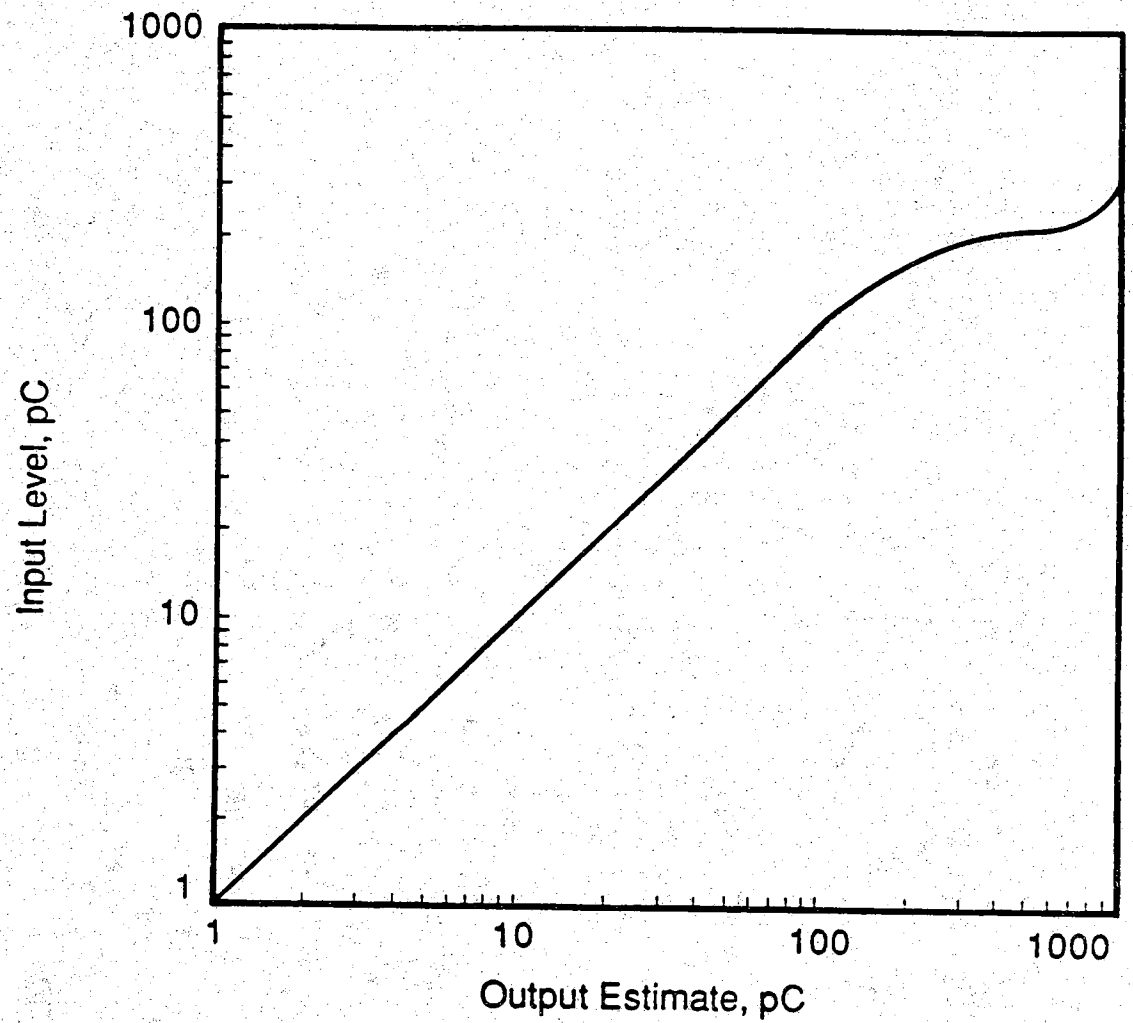


Figure 2.1. Nonlinear transfer function of measurement system.

signal and noise level. The removal of the bias, $T\sigma^2$, in the nonlinear region is an ad hoc approach which is not correct. The correct approach uses a measured calibration curve to perform the mapping, however, this requires a calibration curve for each noise level. An ad hoc approach of subtracting an equivalent noise bias for signals operating in the nonlinear region can be used if a calibration measurement is made. The calibration would measure a time, τ_{eq}, which would be the time that the digitizer operated in its linear region for a given input level. The bias that is removed is then $\tau_{eq}\sigma^2$ which still is not quite correct but leads to satisfactory results. This particular nonlinearity is referred to as a soft limiter and these results were incidental to the measurements that were made. There is an entire branch of signal processing devoted to optimal nonlinearities for use in nonlinear quantization schemes. These were not pursued but undoubtably would yield improved quantization schemes for partial discharge.

The second set of measurements illustrate the relative performance of the three different estimators obtained using an 8 bit quantizer. The measure of performance will be a signal to noise ratio which is defined as

$$\text{SNR} = 20 \log \left[\frac{\text{mean of estimate}}{\text{root mean square error of estimate}} \right]$$

For input signals less than the full scale input level of the digitizer the rms error is due to the variance of the estimator. When the signal exceeds the full scale input level of the digitizer the rms error for the peak value estimator includes both the bias due to the clipping of the signal and the variance of the estimate. This error term was included in the peak value estimator because there is not a method to compensate for measurements exceeding the full scale input level of the digitizer. However, since it is possible, corrections for the nonlinearity and noise bias were made for the LMMSE and AC estimators where appropriate. Fig. 2.2 illustrates the performance of the various estimators for different input levels to the digitizer. A sampling rate of 2.5 times the Nyquist rate was used and the waveforms contained 128 points. The rms input noise was approximately 1/2 lsb of the quantizer. As can be seen, in the linear region of the digitizer both the LMMSE and AC estimator perform almost 20 dB better than the peak value estimator. In the nonlinear region, the peak measurement has large errors because there is not a method to compensate for the

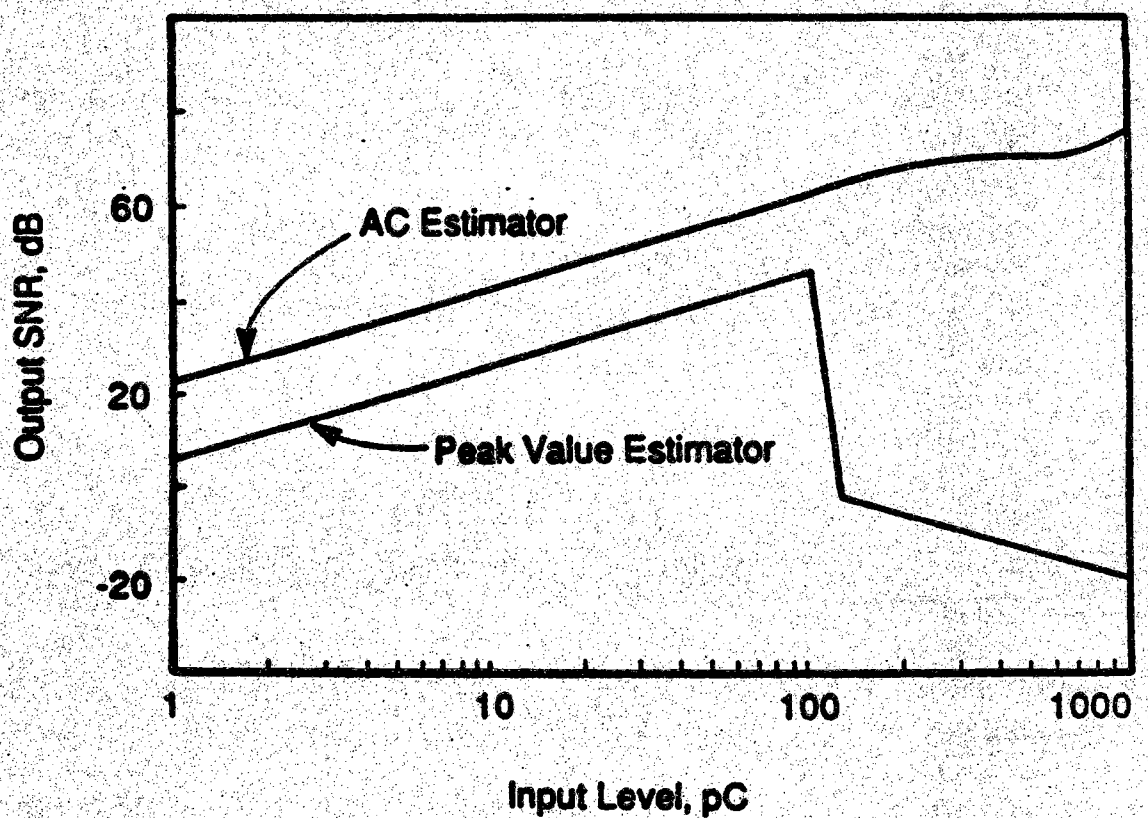


Figure 2.2. Performance of estimators versus digitizer input level.

nonlinearity. However, the LMMSE and AC estimator maintain their performance because there is a technique to map the measurements into their correct values.

The final set of measurements show the relative performance of the three estimators under varying noise conditions. In these measurements the rms error is due only to the variance of the estimator because the digitizer was operated in its linear region. The input signal was maintained at a fixed level of about 10 % of the full scale input level of the digitizer. The input noise level was adjustable so that the input SNR could be varied between -10 dB and 30 dB. Fig. 2.3 illustrates the effectiveness of both the LMMSE and AC estimators in low level measurements while the peak value estimator tracks the input SNR. The AC estimate performed better than the LMMSE estimate at low SNR's and it is believed that losses in the performance were observed because the arrival time also had to be estimated.

As was illustrated by the experimental measurements the LMMSE and AC estimators out perform the simple peak value estimator. However, implementation of the LMMSE technique requires relatively sophisticated computations that are time consuming. Special purpose hardware is available to accelerate the calculations, but is expensive. The AC method provides estimates, that in these measurements, were as good as the LMMSE estimate. However, the major advantage of the AC estimator is that it is much simpler to implement. The hardware implementation of the AC estimator is particularly simple and could easily operate in real time. The basic configuration could use an inexpensive 8 bit A/D flash converter. A pre-estimate could be calculated using a single chip multiplier-accumulator. After rapid collection, the data could then be properly rescaled by removing the bias due to the noise, correcting for the nonlinearity, if necessary, and taking the square root. This could be easily performed on any general purpose microprocessor in an expedient fashion for display purposes.

One final variation that should be mentioned is that a modification of these techniques would be useful in narrow band partial discharge detectors that use a diode envelope detector. This technique was not pursued but the implementation would amount to simply integrating the output of the envelope detector to develop the estimate. From some preliminary results it appears that similar gains in performance can be obtained.

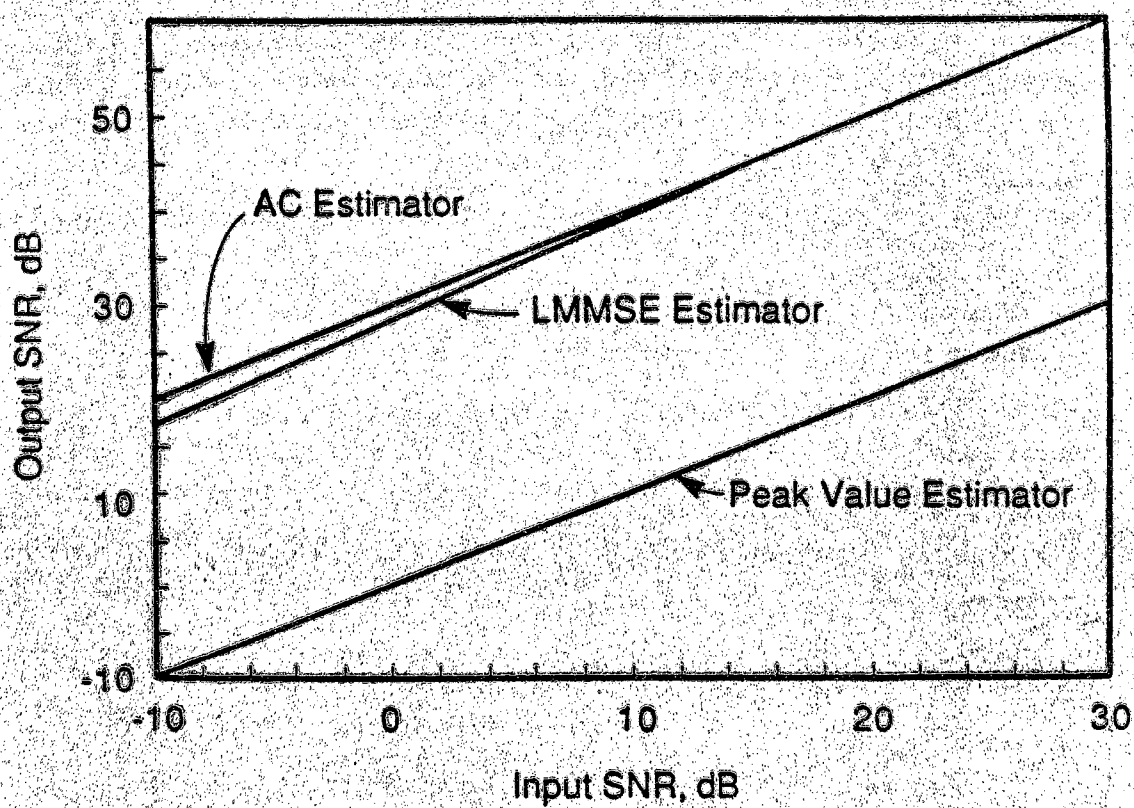


Figure 2.3. Performance of estimators for low level measurements.

The results discussed above provide solutions to the amplitude estimation problem when the second order statistics of the noise are known. The LMMSE solution is the optimal solution when the noise is Gaussian (assuming the arrival time is known) and is the optimal linear solution when the noise is from a wide sense stationary random process (not necessarily Gaussian). The important question that needs to be answered is what is the performance of the best estimator. This question can be answered in terms of a bound on the variance of the estimate. The bound is the Cramer-Rao lower bound and gives the minimum variance of the amplitude estimate. To calculate the bound assumptions about the statistical nature of the noise must be made. If the noise is from a stationary Gaussian random process with power spectral density, $N(\omega)$, then the bound on the variance of the amplitude estimate is given by [41]

$$\mathrm{var}(\,\hat{a} - a\,) \geqq \left[\int_{-\infty}^{\infty} \frac{|S(\omega)|^2}{N(\omega)}\, d\omega\right]$$

where it is assumed that the observation interval is infinite and that $S(\omega)$ is the Fourier transform of the unit amplitude pulse, $s(t)$.

The actual quantity of interest is the charge and these results need to be stated in terms of bounds on estimating the charge. Following Boggs [21] the pulse will be assumed to have a Gaussian shape

$$s(t) = a\, e^{-\frac{t^2}{2\tau_0^2}}$$

where a is the amplitude to be estimated. The rms duration of the pulse is given by $\frac{\tau_0}{\sqrt{2}}$ and the rms bandwidth is $\frac{1}{\sqrt{8\,\pi^2\,\tau_0^2}}$. The noise will be assumed to be white Gaussian noise with the variance $\frac{N_0}{2}$. For purposes of analysis and comparison to the charge, a DC coupled system will be used. Also, to simplify interpretation, a relatively simple case of a power cable will be considered. The charge measured from one end of the cable is $\frac{a}{2\,Z_0}\sqrt{2\,\pi\,\tau_0^2}$ where Z_0 is the characteristic impedance of the cable. The

output SNR of the estimator is given by

$$SNR_{out} = \frac{a^2 \mathscr{E}_s}{\frac{N_0}{2}}$$

where $\mathscr{E}_s$ is the energy of the unit amplitude pulse and is $\sqrt{\pi \tau_0^2}$. The noise, N_0, is defined as $k T_{eq}$ where k is Boltzmann's constant and T_{eq} is the effective noise temperature of the entire system. The effective noise temperature includes the contributions to the thermal noise from cable, measurement impedance and amplifier. The minimum measurable charge, with the estimate having the (subjectively) chosen SNR_{out}, is

$$Q_{min} = \left(\frac{8}{\pi}\right)^{1/4} \frac{1}{Z_0} \left[\frac{SNR_{out} \frac{N_0}{2}}{B}\right]^{\frac{1}{2}}$$

where B is the rms bandwidth of the Gaussian pulse and N_0 has the approximate value of 4×10^{-21} at the temperature 290° K. As an example consider a typical half kilometer cable with $Z_0 = 35$ and having an rms bandwidth of $B = 5$ MHz. If the system has a $NF = 4$ so that $\frac{N_0}{2}$ is 16×10^{-21} then $Q_{min} = 0.014$ pC for an output SNR of 20 dB. This particular number represents the theoretical minimum measurable charge if the optimal estimator were used. Since an optimal estimator is not available an additional loss will be incurred from the estimator used but will be small (on the order of a few dB).

2.3 Poisson Shot Noise Process

Observed partial discharge signals are random in nature and difficult to characterize. Current technology in partial discharge measurements tacitly assumes a deterministic model for partial discharge. However, to take full advantage of signal processing techniques a stochastic model should be used when modeling a signal that has random qualities. Unfortunately, present lack of understanding of the mechanisms of partial discharge makes a

stochastic model developed from first principles still beyond reach. Models for partial discharge signals can, however, be constructed based on experimental observations. The goal of this section will be to develop a first order theoretical model for partial discharge that will allow easy application of the theory to practical measurement situations. This development is taken directly from Snyder [61] but only the salient points are mentioned.

Typically, individual partial discharges are time localized and are randomly spaced in time with respect to each other which suggests modeling partial discharge in terms of a stochastic point process. The simplest example of a point process is the time homogeneous Poisson process for which each localized event is statistically independent of the others. There are several methods for deriving the Poisson process. The most informative approach is the constructive method in which the time intervals between partial discharges are considered to be exponentially distributed with mean $1/\lambda$. The resultant process is then, by definition, a time homogeneous Poisson process with the intensity function equal to a constant, λ. The intensity function λ has the interpretation of being the mean number of partial discharges per unit time. For small Δt the probability of having a single partial discharge in the interval $(t, t + \Delta t)$ is then

$$\Pr(N_{t,\, t+\Delta t} = 1) = \lambda \Delta t + o(\Delta t)$$

where $\frac{o(\Delta t)}{\Delta t}$ approaches zero as Δt approaches zero.

Work by Devins, [12], showed that a model for the statistical time lag associated with the epochs of the individual discharges under DC conditions is exponential. If the probability density of the time between partial discharges is truly exponential, then by definition, it must be a Poisson process.

Most of the experimental results that will be discussed were made under AC conditions and the voltage across the defect is therefore a function of time. Consequently, the epochs of the partial discharges will be related to the phase of the high voltage excitation (e.g., 60 Hz). If the process is still modeled as Poisson, this variation can be taken into account by introducing a time varying intensity function, λ_t, which is now a function of the excitation voltage. The introduction of a time varying intensity function means that the process is now nonstationary. The interpretation of this

nonstationarity is that the statistical quantities describing the process are now functions of time. For example, the average number of partial discharges occurring in the time interval $(t_1, t_1 + \Delta t)$ will not be the same as the average number of partial discharges occurring in a different time interval $(t_2, t_2 + \Delta t)$ because these averages are functions of the times t_1 and t_2. For small Δt the probability of having a single partial discharge in the interval $(t, t + \Delta t)$ is

$$\Pr(N_{t,\, t+\Delta t} = 1) = \lambda_t \, \Delta t + o(\Delta t)$$

where $o(\Delta t)$ is defined above. This nonstationarity will complicate the analysis if a further simplification cannot be found.

To simplify the problems associated with the above nonstationarity, further assumptions about the nature of the nonstationarity will be made. The excitation voltage is a sinusoid and therefore it will be expected that the intensity function, λ_t, will also be periodic with the same period as the excitation voltage. Based on the assumption of a periodically varying intensity function, the process is now called periodic nonstationary. This interpretation is appropriate if observations are made during a short time interval, T, located at a particular phase of the excitation voltage. If the process is observed for many cycles of the excitation voltage and the process is periodic nonstationary, then it is expected that on the average there will be $N = \lambda_\phi T$ partial discharges per cycle, where $\lambda_\phi = \lambda_t$ is now approximately a constant. The intensity function will be approximately a constant for any chosen, small interval at a given fixed phase, however, if the positioning of the measurement interval is moved to another phase of the excitation voltage this constant can be expected to change.

To complicate the matter further, the partial discharge process undergoes a slow metamorphosis. For example, Luczynski [25] observed that in his specimens the process initially consisted mainly of partial discharge pulses with very short durations and large energies. These high energy partial discharges are referred to as rapidly developing partial discharges. Rapidly developing partial discharges occurred relatively infrequently with only a few partial discharges per cycle of excitation voltage and therefore having relatively small values of λ_t . Luczynski [25] also observed that after a period of many hours the process changed to one with a high incidence of slowly developing partial discharges. Slowly developing partial discharges

have longer durations and low energy. The number of slowly developing partial discharges per cycle of excitation voltage was large implying a large λ_t . This change represents another type of nonstationarity and is referred to as a slowly varying nonstationarity. If the measurement interval is confined to a small enough time period in the entire history of the process, then it can still be considered to be periodic nonstationary. In the following description the measurements intervals will be assumed to be sufficiently short so that the slowly varying nonstationarity can be ignored.

To develop the model further, some measure of the energy must be associated with the occurrence times of the partial discharges. The manner in which this is included is to use a marked point process [61]. Now each epoch has associated with it a mark, U_i, indicating its energy, which also can be described statistically. The assumptions about the stationarity of the process with intensity function λ_t can also be made about the time varying nature of the parameter(s) describing the mark distribution. For the Poisson model the mark process will be considered to be independent of the occurrence times and that the individual marks, U_i, will be independent of each other.

The reason for discussing the Poisson process in detail is that this model is the easiest to manipulate mathematically. Literature abounds with results concerning the Poisson process and several results will be presented which apply to the partial discharge measurements made in this thesis. The assumptions required for the partial discharge process to be a Poisson process are very restrictive and this model may not apply in all cases. However, when a more complete model is developed for partial discharge a limit theorem will be presented. Using this limit theorem, a method to sample the partial discharge process will be discussed which allows one to interpret it as a Poisson process.

There are various assumptions which can be used to define a marked Poisson process and the following conditions are an example of one set of conditions which have the advantage that they can be easily modified to admit non-Poisson models. The first assumption that is made is that the process is conditionally orderly which means that multiple events can not occur simultaneously. This is a reasonable assumption because it is highly unlikely that two partial discharges can occur simultaneously (the pulses could be separated by a pico-second or femto-second) The second assumption that is made is that the process develops without aftereffects

which means that the past history of the partial discharges is independent of the future behavior of the partial discharges. In other words all previous partial discharges will not influence the behavior of any partial discharges in the future. Also a finite, integrable function, λ_t, called the intensity of the process, is assumed to exist. The third assumption that is made is that $Pr(N_{t_0} = 0) = 1$ where t_0 is the beginning of the interval on which the Poisson process is defined. This simply means that the number of partial discharge pulses counted at the beginning instance of observation is zero. The final assumption is that the marks are mutually independent and independent of the Poisson process.

As mentioned these conditions are restrictive and it is informative to consider the appropriateness of the assumptions. The conditional orderliness restriction can be removed by treating multiple occurrences as a mark. This might be useful if a single discharge were considered to be a multiplicity of simultaneous events, for example the number of electrons in the partial discharge. Evolution without aftereffects is an assumption that could be difficult to justify in some situations. In various cases there appears to be an influence of previous partial discharge behavior of future behavior. The independence of the marks is also an assumption that may not always be accurate. In some cases there appears to be a dependency between the counting process and the marks. If these cases are considered then highly complex models are required.

The simplest model that can be used is the time homogeneous Poisson shot noise process [61]. In the case of a time homogeneous Poisson model the time between the occurrences of the partial discharges will be exponentially distributed with density

$$p(\tau) = \frac{1}{\lambda} e^{-\frac{1}{\lambda}\tau} u(\tau)$$

where the average number of events per second is λ. Each of these events can be considered to be delta functions occurring at random times, with random areas passing through a linear filter, h(t), as depicted in Fig. 2.4. In this case the observed signal, s(t), is given by

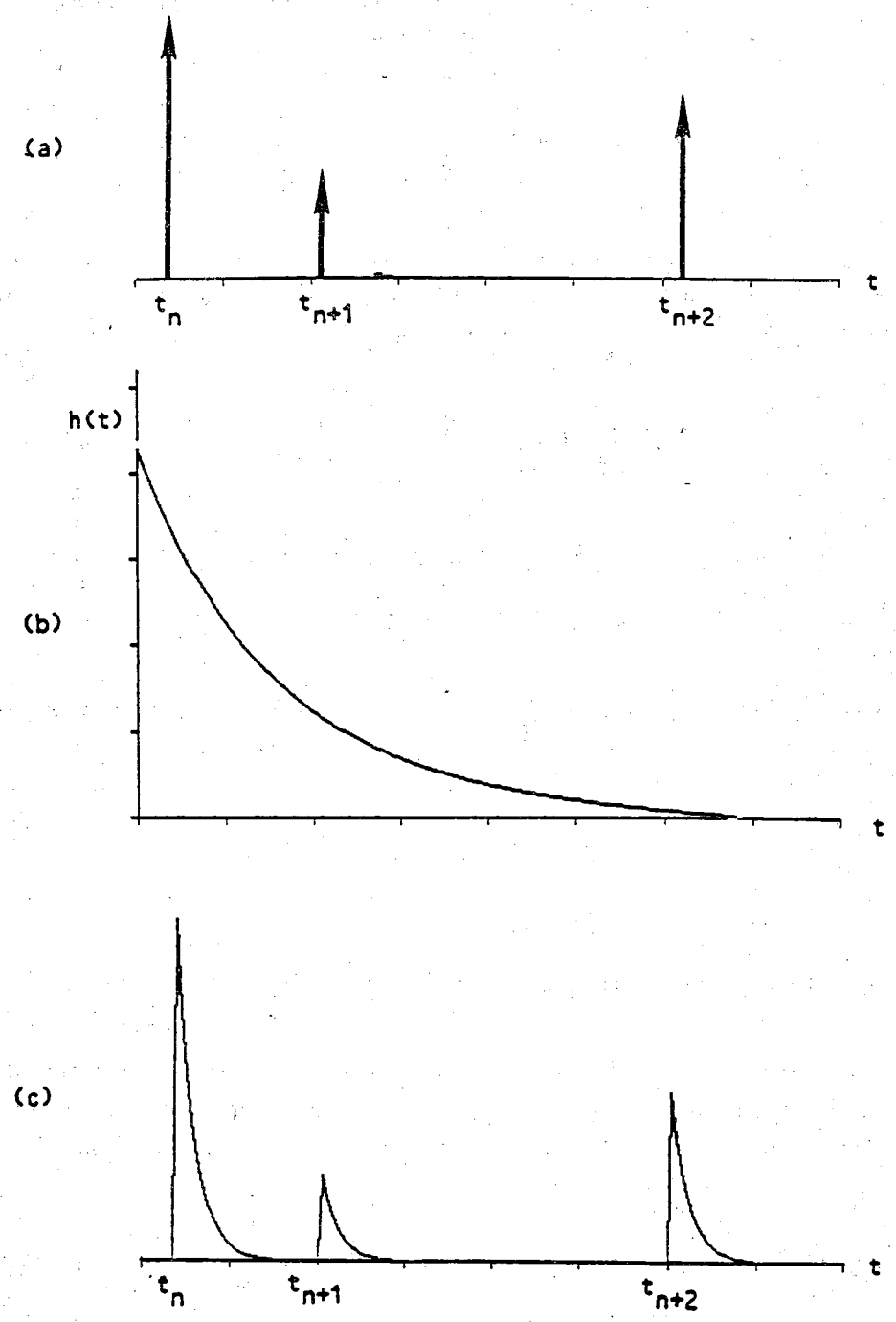


Figure 2.4. Generation of signal using the model. a) Marked point process. b) Linear filter. c) Resulting filtered process.

$$s(t) = \sum_{i=1}^{N_T} a_i \, \delta(t - t_i) * h(t)$$

where a_i is the random area of the delta function (the mark) and * is the convolution operator. The linear filter, for most practical purposes, can be considered to be the cascade of the impulse responses of the propagation path and the detection circuit. In general this will not be true because the physical behavior of the partial discharge itself determines the shape of the pulse. An excellent example of such a case is provided by the slowly developing partial discharge whose shape is determined by the physical and geometrical properties of the void. However, in most cases the true shape will be masked by the impulse response of the filters (medium) through which it is observed. A more complete model could include the random variation of the impulse response. This random variation could arise in a number of ways, for example, the randomness of the individual discharge itself could cause fluctuations in the the pulse shape. Another manner in which the impulse shape could become random is if the partial discharges are emanating from different positions in the device under test. In this case the random variation will be due to the impulse response of the different propagation paths.

As mentioned above a further complication can be included by assigning a mark to each random event. These marks need not be restricted to just the random amplitude, charge or energy of the individual partial discharge pulse. In general, a mark is any auxiliary random variable associated with the epoch of the partial discharge. Suppose that there are two separate partial discharge sites then the mark could be the position of the observed partial discharge interior to the device under test. Both of the amplitude and position marks can be combined into a single vector mark. Another example would be to use the partial discharge type as the mark, it could be either a slowly or rapidly developing partial discharge. The signature of a particular partial discharge might be used if extremely complex structures are being measured (e.g., a transformer). This case would also correspond to partial discharges located at different positions.

There are various quantities that are of interest which can be obtained from the modeling. For complex (non-Poisson) models it is sometimes only possible to calculate some of the moments but for the Poisson shot noise model it is possible to derive a complete statistical description. This

description is in terms of the characteristic function of a marked, filtered Poisson process and is given by [61]

$$M_t(j\,\omega) = \exp\left\{ \int_{t_0}^{t} \lambda_t\, E\,[\, \exp(\, j\,\omega h(\, t,\, \tau : \mathbf{u}\,)) - 1\,]\, d\tau \right\}$$

where $\mathbf{u}$ is the mark vector and h is the pulse shape. A similar expression for the joint characteristic function for y_t at times t_1, and t_2 is given by

$$M_{t_1,\, t_2}(j\,\omega_1,\, j\,\omega_2) =$$

$$\exp\left\{ \int_{t_0}^{t} \lambda_t\, E\, \left(\exp[\, j\,\omega_1 h(\, t_1,\, \tau : \mathbf{u}\,) + j\,\omega_2 h(\, t_2,\, \tau : \mathbf{u}\,)] - 1 \right) d\tau \right\}$$

These expressions are useful for calculating the moments of the process for example the average charge is given by

$$Q_{avg} = \int_{-\infty}^{\infty} E\{\, a(t)\}\, \lambda(t)\, h(t)\, dt$$

where both the amplitude and intensity are considered to be periodic functions of time.

In later chapters higher order moments of the filtered process will be used in the analysis of different techniques. Second order moments are the basis of many signal processing algorithms because these moments are easily utilized in the analysis. Furthermore, second order moments are easily measured using standard (highly developed) spectral estimation techniques. The autocorrelation function of the shot noise is given by

$$R(t_1, t_2) = \int_{t_0}^{\min(t_1, t_2)} \lambda_\tau E[h(t_1, \tau : u)\, h(t_2, \tau : u)]\, d\tau$$

$$+ \int_{t_0}^{t_1} \lambda_\tau E[h(t_1, \tau : u)]\, d\tau \int_{t_0}^{t_2} \lambda_\tau E[h(t_2, \tau : u)]\, d\tau$$

where the expectation is taken with respect to the mark. The case that is of most interest in the following chapters is the cross-correlation function of the shot noise passing through two dissimilar, time-invariant, linear filters. Typical observation windows are short with respect to the period of the process (5 μs vs. 16.66 ms). In this respect the intensity function will be considered to be a constant. For two different filters, $h_1(t)$ and $h_2(t)$ the cross-correlation function is

$$R_\phi(\tau) = (\overline{a_\phi})^2 \lambda_\phi^2 \int_{-\infty}^{\infty} h_1(t)\, dt \int_{-\infty}^{\infty} h_2(t)\, dt$$

$$+ \overline{a_\phi^2}\, \lambda_\phi \int_{-\infty}^{\infty} h_1(\gamma)\, h_2(\tau + \gamma)\, d\gamma$$

where it is assumed that the statistics of the amplitude are functions of ϕ but constant in the observed window. The subscript ϕ is used to indicate that these results apply only to the small window located at phase ϕ. The power spectral density of the partial discharge is then easily found by letting $h_1 = h_2$ and taking the Fourier transform of $R_\phi(\tau)$ yielding $S_\phi(\omega)$.

Other disciplines have found the Poisson model to be of great use in statistical inference of the underlying nature of the mechanisms generating the process. One application of the Poisson model is for making a statistically based decision about the type of process being observed. For example a hypothesis testing procedure could be used to determine what type of partial discharge process is being observed. Something similar to this is presently being performed on a regular basis by operators of commercial partial discharge detection equipment. A CIGRE report, [87], gives list of expected detector outputs for different partial discharge types. Two papers, [88,89], were recently presented which describe the use of expert systems, based on the CIGRE report, for determining the type of partial

discharge being observed. Using a hypothesis testing procedure, based on the Poisson model, could lead to a similar system. Another prominent application in statistical inference is the estimation of various parameters associated with factors influencing the behavior of the point process. This particular application could be used for investigating the influence of physical changes on the partial discharge process. If the situation had been catalogued then the reverse problem is possible; determine the values of the variables controlling the partial discharge. Some work, other than that contained in this thesis, has been performed on this problem. In the past few years a new measurement technique has been developed, [84] which in essence makes the measurements of partial discharge in the manner to be presented below. However, the measurements are not analyzed using this framework.

Both of the applications, discussed above, require the use of a likelihood ratio for the observed process. The few results presented thus far were concerned with the integrated response observed through a linear filter. In some cases this is the desired approach for analyzing partial discharge. Now suppose that the estimation procedures discussed in section 2.2 are used to estimate both the occurrence time and mark of each event. The result is a sequence of points, (t_i, u_i) which can be used in the analysis of the point process via the likelihood function. In the analysis that follows it will be assumed that only the sequence of points is available for use.

An important function associated with the analysis of the Poisson process is a function referred to as the sample function density [61]. The sample function density can be roughly described as the probability of observing a particular realization of the point process. For a univariate Poisson process (without marks) the sample function density is given by

$$p\left[\{ N_\sigma; t_0 \leqq \sigma < t \} \mid \alpha \right] =$$

$$\exp\left\{ - \int_{t_0}^{t} \lambda_\sigma(\alpha)\, d\sigma + \int_{t_0}^{t} \ln[\lambda_\sigma(\alpha)]\, dN_\sigma \right\}$$

where the second integral has the evaluation

$$\int_{t_0}^{t} \ln[\lambda_\sigma(\alpha)]\, dN_\sigma = \sum_{i=1}^{N_t} \ln[\lambda_{t_i}]$$

The variable, α, is a non-random parameter that is associated with (and may control in some sense) the particular process being observed. In the estimation problem α will be the variable that is under study and will be estimated. In the hypothesis testing problem α can be thought of as an index to the process being observed and will be an integer; in the binary problem $\alpha = 0, 1$.

For a marked Poisson process the sample function density is given by

$$p\left[\{ x_\sigma;\ t_0 \leqq \sigma < t \} \mid \alpha \right] =$$

$$\exp\left\{ -\int_{t_0}^{t} \lambda_\sigma(\alpha)\, d\sigma + \int_{t_0}^{t} \ln[\lambda_\sigma(\alpha)]\, dN_\sigma + \sum_{i=1}^{\infty} \ln[\, P(\, U_i \mid \alpha\,)\,]N_t(U_i) \right\}$$

For a marked Poisson process the sample function density can be also be written as

$$p\left[\{ x_\sigma;\ t_0 \leqq \sigma < t \} \mid \alpha \right] = \prod_{i=1}^{\infty} p\left[\{ N_\sigma(U_i);\ t_0 \leqq \sigma < t \} \mid \alpha \right]$$

where the process has been quantized and split along the level U_i so that there are now sub-processes each having it's own sample function density given by

$$p\left[\{ N_\sigma(U_i);\ t_0 \leqq \sigma < t \} \mid \alpha \right] =$$

$$\exp\left\{ -\int_{t_0}^{t} P(U_i \mid \alpha)\lambda_\sigma(\alpha)\, d\sigma + \int_{t_0}^{t} \ln[P(\, U_i \mid \alpha\,)\lambda_\sigma(\alpha)]\, dN_\sigma(U_i) \right\}$$

This representation follows from an extremely important theorem discussed in [61] in which a marked Poisson process can be represented as a superposition of sub-processes. Each of these sub-processes is a Poisson process with it's intensity function given by

$$\lambda_\sigma(\alpha;\, i\,) = P(\, U_i \mid \alpha\,)\lambda_\sigma(\alpha)$$

where there are now M sub-processes for (i = 1, 2, .M) for M quantization levels. This characterization can also be extended to the continuous case, however, measurements by nature are quantized so it seems that the denumerable case is the most practical to discuss. The factorization of the marked Poisson process into M separate processes is an extremely important and valuable tool for analysis.

Another type of analysis can be performed if a count rate histogram is utilized. The above formulations assume that all information is available; the exact time of occurrence of each point in order of occurrence. Collecting data can be difficult in situations where the data rate is high and the observation interval long. However, there are instrumentation techniques available that allow high data rates and long observation intervals while remaining cost effective. Using these techniques, only the number of counts in a particular quantization segment, (ϕ_i, u_i) would be recorded in terms of a histogram using a histograming memory (which are commercially available). The data collection system would collect the number of counts in each bin where each bin would be assigned to a small phase window and charge window. Analysis of the measurements would then proceed in terms of the number of counts in each bin. The histogram will be at least a two dimensional histogram with the number of dimensions being equal to p+1 where p is the dimension of the mark space. For a one dimensional mark space the probability of observing a particular histogram contour (a contour along u_i) is given by

$$Pr\{\, N_{t_{j-1},\, t_j}(U_i) = n_j;\, j = 1, .., k \mid \alpha\,\} =$$

$$\prod_{j=1}^{k} \left\{ [\,(n_j)\,!\,]^{-1} \left[\int_{t_{j-1}}^{t_j} P(U_i \mid \alpha)\lambda_\sigma(\alpha)\, d\sigma \right]^{n_j} \exp\left[- \int_{t_{j-1}}^{t_j} P(U_i \mid \alpha)\lambda_\sigma(\alpha)\, d\sigma \right] \right\}$$

where it is assumed that the counts n_j were collected in k sub-intervals of the interval (t_0, T). The interval, (t_0, T) would be an integral number of periods of the excitation voltage and the sub-intervals, (t_{j-1}, t_j), would be the phase intervals (ϕ_{j-1}, ϕ_j) observed over the integral number of periods used. Since the marks are assumed to be independent of the counting process and independent from each other the probability of observing the number of counts in the sub-intervals (t_{j-1}, t_j) (the phase bin) becomes

$$\Pr\{ N_{t_{j-1}, t_j} = n_j;\ j = 1, .., k \mid \alpha \} =$$

$$\prod_{i=1}^{M} \Pr\{ N_{t_{j-1}, t_j}(U_i) = n_j;\ j = 1, .., k \mid \alpha \}$$

Using these equations the log likelihood function is easily seen to be

$$\ell\,(\alpha) = \sum_{i=1}^{M} \ell_{\,i}(\alpha)$$

where

$$\ell_{\,i}(\alpha) = - \int_{t_0}^{T} P(U_i \mid \alpha)\lambda_\sigma(\alpha)\, d\sigma + \sum_{j=1}^{k} \int_{t_{j-1}}^{t_j} \ln[\, P(U_i \mid \alpha)\lambda_\sigma(\alpha)\, d\sigma \,]$$

It is interesting to note that this likelihood function is the sum of the likelihood functions for each charge bin.

Now that the basic formulas have been presented their use will be described for different situations. The sample function density is the more useful of the two characterizations because it uses all of the information available. The likelihood function for the count rate data is more restrictive because it ignores potentially useful data (the occurrence times). However,

as pointed out, data collection requirements are greatly reduced using the count rate approach. Furthermore, the amount of numerical analysis is reduced because of the compression in data. For these reasons, emphasis will be placed on the count rate analysis but the techniques are completely analogous for the sample function density analysis. In the following, reference will be made to the likelihood function which can be either the count rate likelihood function or the sample function density because the sample function density is the likelihood function for the time data.

The first case to be discussed is the estimation problem in which a variable, α, influencing the partial discharge process, is estimated. The value of α which maximizes the likelihood function is the value of α which maximizes the probability of observing the recorded data. In other words maximization of the likelihood function (or it's logarithm) with respect to α gives the maximum likelihood estimate, $\hat{\alpha}_{ml}$, of α. Solution of the likelihood equation will yield an estimate of α

$$\left.\frac{\partial \ell(\alpha)}{\partial \alpha}\right|_{\alpha=\hat{\alpha}_{ml}} = 0$$

The solution can easily be extended to a parameter vector α. As an example, consider the estimation of the intensity of the process in terms of the count rate data.

$$\left.\frac{\partial \ell\,[\lambda(\phi_j, u_i)]}{\partial \lambda(\phi_j, u_i)}\right|_{\lambda=\lambda_{ml}} = 0$$

The maximum likelihood estimate is then

$$\lambda_{ml}(\phi_j, u_i) = \frac{n_j(u_i)}{T/k}$$

where it is assumed that each phase bin was observed an equal amount of time and $t_0 = 0$

Estimation of parameters influencing the partial discharge process requires a priori knowledge of the functional form of $\lambda(\alpha)$. These functional forms could be theoretically predicted but as already mentioned sufficient

understanding of the physics of the process is lacking for this approach to be fruitful. The alternative procedure is the phenomenological approach in which particular situations are measured while carefully controlling the variable to be examined. The desired end result is a functional form for the dependence of the intensity on the variable in question. Two examples of this procedure are investigated in section 2.5. Once this parameterization has been accomplished, then the next time that this situation is encountered, the variable can be estimated based upon the previously measured functional form. The measurements made in [84] attempt to parameterize the behavior of partial discharge histograms similar to the type developed here for voids with different geometries.

The proposed procedure presents an insurmountable task in attempting to parameterize all partial discharge types for all possible variables. The results from this impossible task would certainly be a desirable data base, however, the purpose of this exposition is not to propose the development of such a data base. Rather, it is to present a mathematical framework, under which partial discharge can be studied. Study of partial discharge in this manner may lead to a greater understanding of the mechanisms that influence it's behavior.

2.4 Higher Order Models

The Poisson model contained in the previous section assumes that all of the events are statistically independent of each other with exponentially distributed inter-occurrence times and in certain situations this may be a reasonable assumption. There are instances in which the inter-occurrence times are clearly not exponentially distributed and the Poisson model will be incorrect. Furthermore, there are situations where there appears to be a dependence between events and in these case it is more appropriate to describe the partial discharge process with more elaborate models. Use of more elaborate models for statistical inference will give better results (assuming the models are accurate) because the information is being used correctly. However, when the models for a point process become more elaborate the analysis soon becomes intractable. There are some models and special cases in which results can be obtained. One point process which models the dependence between the events and is somewhat tractable is the self exciting point process. Self exciting processes model dependencies

between events in the process and this model seems well suited for modeling partial discharge. Referring back to the conditions for the partial discharge process to be Poisson the condition that is relaxed for this case is evolution without aftereffects.

A self-exciting point process is one in which the intensity of the process is now allowed to be a function of the past events. The intensity function for the process is given by

$$\lambda_{t,\,U_i} = \begin{cases} P(U_i)\ \mu_t(0), & t_0 < t \leqq w_1 \\ P(U_i)\ \mu_t(\ N_t;\ w_1,\ ,..w_{N_T}), & w_{N_T} < t \leqq w_{N_{T+1}} \end{cases}$$

where the marks are assumed to be independent of the counting process and of each other. The variables, w_j, are the occurrence times of the partial discharge pulses. The equation states that the present intensity at time, t, is a function of the number of previous events as well as the times at which they occurred. In other words the probability of observing a partial discharge in (t, t + Δt) is

$$P(\ N_{t,t+\Delta t} = 1 \mid N_t;\ w_1,\ ,..w_{N_T};u_i\ ) = \lambda_{t,U_i}\ \Delta t + o(\Delta t)$$

for small Δt. The sample function density can be written down for this case and can be found in [61]. In general this process is still difficult to analyze and further simplifications can be used.

A widely used model which is similar to the Poisson model is the renewal model. In this case the interarrival times are still considered to be independent, identically distributed random variables but they no longer have an exponential distribution. The renewal process has inter-occurrence times that can have an arbitrary distribution. The renewal process is a special case of the self exciting point process in which the intensity function is homogeneous and dependent on only the previous event. In this case the the process is described as a homogeneous, memory one, self exciting point process with

$$\mu_t(N_t; w_1, \ldots w_{N_t}) = h(t - w_{N_t})$$

for some function h(t) which is referred to as the hazard function.

To use this model for delay estimation it will be necessary to investigate the second order moments of the integrated process. For a point process with independent marks the correlation function can be written as

$$R(t_1,t_2) = \overline{a}^2 \int_0^{t_1} d\tau_1 \int_0^{t_2} d\tau_2\, g(t_1-\tau_1)\, g(t_2-\tau_2)\, f_2(\tau_1,\tau_2)$$

$$+ \overline{a^2} \int_0^{\min(t_1,t_2)} d\tau\, g(t_1-\tau)\, g(t_2-\tau)\, f_1(\tau)$$

where the functions $f_1(\tau)$ and $f_2(\tau_1,\tau_2)$ are referred to as the first and second order product densities, respectively [62]. The function g(t) is the transfer characteristic of the linear filter through which the process is observed. First order product densities have the interpretation that $f_1(t)\Delta t$ is the probability of observing an event in the infinitesimal interval $(t, t + \Delta t)$ regardless of the points occurring elsewhere. Second order product densities have the interpretation that $f_2(t_1, t_2)\Delta t_1 \Delta t_2$ is the probability of simultaneously observing events in the infinitesimal intervals $(t_1, t_1 + \Delta t_1)$ and $(t_2, t_2 + \Delta t_2)$ regardless of the points occurring elsewhere. These product densities are usually difficult to determine but there are a few cases for which they are easily found. For a stationary Poisson process the product densities are

$$f_1(\tau) = \lambda$$

and

$$f_2(\tau_1,\tau_2) = \lambda^2$$

so the correlation function becomes those presented in the previous section.

It is also possible to write down the correlation functions for a renewal process and the form depends on the actual description of the time between

events. The second order product density for a renewal process takes the form

$$f_2(\tau_1,\tau_2) = f_1(\tau_1)f_1(\tau_2 - \tau_1)$$

so if the first order product density can be found the second order product density can be easily calculated. The first order product density is found from the solution of the integral equation

$$f_1(t) = p(t) + \int_0^t f_1(\alpha)p(t - \alpha)\, d\alpha$$

where p(t) is the probability density function of the interarrival times of the events. This integral equation is easily solved using the Laplace transform and the solution is the inverse Laplace transform of

$$F_1(s) = \frac{P(s)}{1 - P(s)}$$

This transform is particularly easy to calculate for the gamma density function. This solution will apply for the memory one self exciting point process (non-homogeneous renewal process) if the periodic nonstationarity assumption is made. In this case the observations would be made in a small phase window for which the inter-occurrence times would be homogeneous.

There is a further assumption that can be made in many cases when the intensity function is small enough. The correlation function for a renewal process takes into account the correlation existing between subsequent events, (e.g., partial discharges). Typically, the small time intervals, in which the measurement is made, contain either one partial discharge or none. Therefore, when calculating the correlation function from the data, there are no subsequent partial discharge events entering into the calculation. Since there are not any subsequent events, the correlation function is the same as for a Poisson process, justifying the Poisson assumption for the case in which the intensity is small enough. This particular case is actually very common considering the small observation windows used in the implementation of the delay estimation schemes to be

discussed.

There is a more general model which encompasses all the cases discussed and is referred to as the general marked point process. In this case the intensity of the process can now be a function of the marks and dependencies between the counting process and the marks is allowed. The form of the intensity will be

$$\lambda_{t,\,U_i} = \begin{cases} \mu_{t,\,U_i}(0), & t_0 < t \leqq w_1 \\ \mu_{t,\,U_i}(\,N_t;\, w_1, \ldots w_{N_t},\, u_1, \ldots u_{N_t}) & w_{N_t} < t \leqq w_{N_t+1} \end{cases}$$

Consideration of a complex model such as this leads to intractable equations which are of little use in the analysis. There is method by which such a process can be analyzed using a very simple approach. The approach throws out a tremendous amount of information that could be potentially useful but the resulting simplicity justifies the method to be proposed. The method is a technique by which the general point process is sampled so that the resulting point process can be interpreted as a Poisson process. Current statistical research on point processes is concerned with an operation on point processes referred to as pooling. Pooling of point processes means that individual processes are overlaid one on top of another and the statistics of the resulting process are considered. Under the proper conditions the pooled process converges to a Poisson process [61]. This result can be shown to hold for univariate point processes as well as marked point processes. The drawback of using the pooling operation to simplify data collection and analysis is that this technique ignores potentially useful information contained in the unpooled sequence of events.

The method for sampling partial discharge involves pooling the point processes from small phase windows on different cycles of the excitation. The assumption of periodic nonstationarity is still used and each component of the pooled process comes from a phase window located at the same phase but from a different cycle. The phase windows are of an extent such that each component process pooled into the result contributes no more than one point. The basic assumption that is made for this technique to work is that the partial discharges occurring at sufficiently separated times are statistically independent. It is presently not clear if separation of the

processes by only a single cycle is sufficient for independence. The resulting pooled process in each window will be approximately a marked Poisson process with intensity λ_ϕ. While there is no mathematical proof that true partial discharge sampled in this manner converges to a marked Poisson process, if the assumptions made about the process are true then it will be approximately a marked Poisson process. Using this sampling method and assuming that the resultant pooled process is Poisson then the analysis in the previous section applies. Questions about the residual dependencies between adjacent phase bins are still unanswered. If statistical analysis indicates the existence of these dependencies then the sampling method can be modified. The modification would involve sampling in a manner which would make the constituent point processes to be pooled more sparse. This could be accomplished, for example, by sampling only a few phase bins per cycle which would make the (per cycle) point processes more sparse.

In actual practice when partial discharge with large intensities are observed it is likely that the large intensities are due to multiple discharge sites that are simultaneously active. When this is the case then these multiple sites will be independent if they are sufficiently separated. Sufficient separation means that there is little physical communication of charged species and photons between the sites. In this case the process is already pooled and it is likely that a Poisson model may be accurate.

2.5 Experimental Study

The following experiments demonstrate the consistency of the point process model discussed in the previous sections. The results are not definitive in the sense that they prove that the model is correct but rather the results demonstrate that the model is consistent with the observations. Only a limited number of experiments were performed for this particular aspect of the thesis.

High voltage systems, by necessity, are large to prevent unwanted discharge and breakdown. Typical high voltage systems also have complex structures, in which the signals of interest may propagate through many poorly defined paths. Subtle details of the signals may then be masked. The structures considered in the following sections, concerned with estimating the location of the partial discharge site, are too large for a controlled study of the characteristics of the partial discharge process. To

properly study the partial discharge process with as many variables as possible under control, some type of test cell must be used.

The test cell used in this work is a slight modification of the test cell used by Luczynski [25]. The principle modification was the inclusion of an acoustic transducer to permit the acoustic emission to be measured simultaneously with the electric signal. The test cell, depicted in Fig. 2.5, provides high fidelity signals with high sensitivity and allows the control of parameters influencing the discharge process. In particular the variables of interest are the materials surrounding the discharge gap and the actual discharge gap geometry. When a high voltage is applied to the high voltage electrode a discharge occurs in the region labeled E in Fig. 2.5. The acoustic transducer is a Physical Acoustics integral preamp transducer with a center frequency of 150 KHz. The insulating material used on both faces of the discharge gap was Plexiglas.

The test cell is small compared to the electrical wavelengths studied, (bandwidth of 20 MHz), however, it is still large with respect to the acoustic wavelengths. Unwanted multipath effects will be present for the acoustic signals and therefore a characterization of the typical acoustic signal will be necessary in order to extract information about the waveforms. The acoustic signature of the test cell was studied by fracturing a pencil lead in the center of the test cell. A typical response to this impulsive force is shown in Fig. 2.6. The peaks in the envelope of the response are due to reverberations in the test cell. The first peak is the desired signal, the second is due to a reflection from the outer edge of the central measuring electrode and the third peak is from a reflection from the outer edge of the guard electrode.

The signal model discussed above makes assumptions about the time varying nature of the process and the type of stationarity that will be used. The relative energy vs. the phase of the excitation voltage is shown in Fig. 2.7. The relationship of the energy to the phase of the excitation voltage was developed by averaging over many cycles of the excitation voltage a quantity referred to as the mark accumulator process [61] which for this application will be defined as

$$\mathcal{E}(\phi_k) = \sum_{i=1}^{N_{\phi_k}} a_i^2(\phi_k)$$

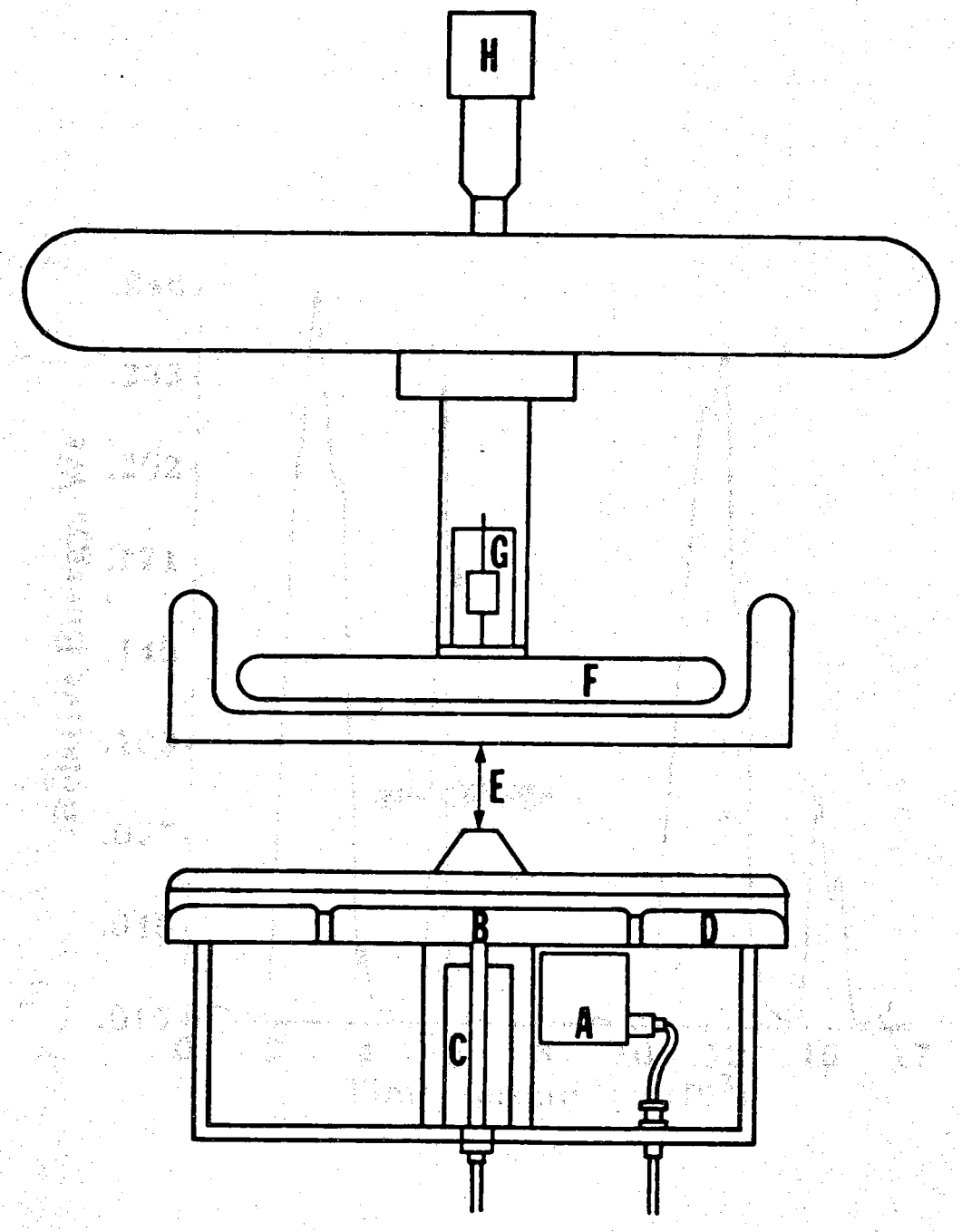


Figure 2.5. Test cell. A: acoustic transducer; B: central measuring electrode; C: coaxial connection; D: guard electrode; E: discharge region; F: high voltage electrode; G: isolation resistor; H: micrometer adjustment.

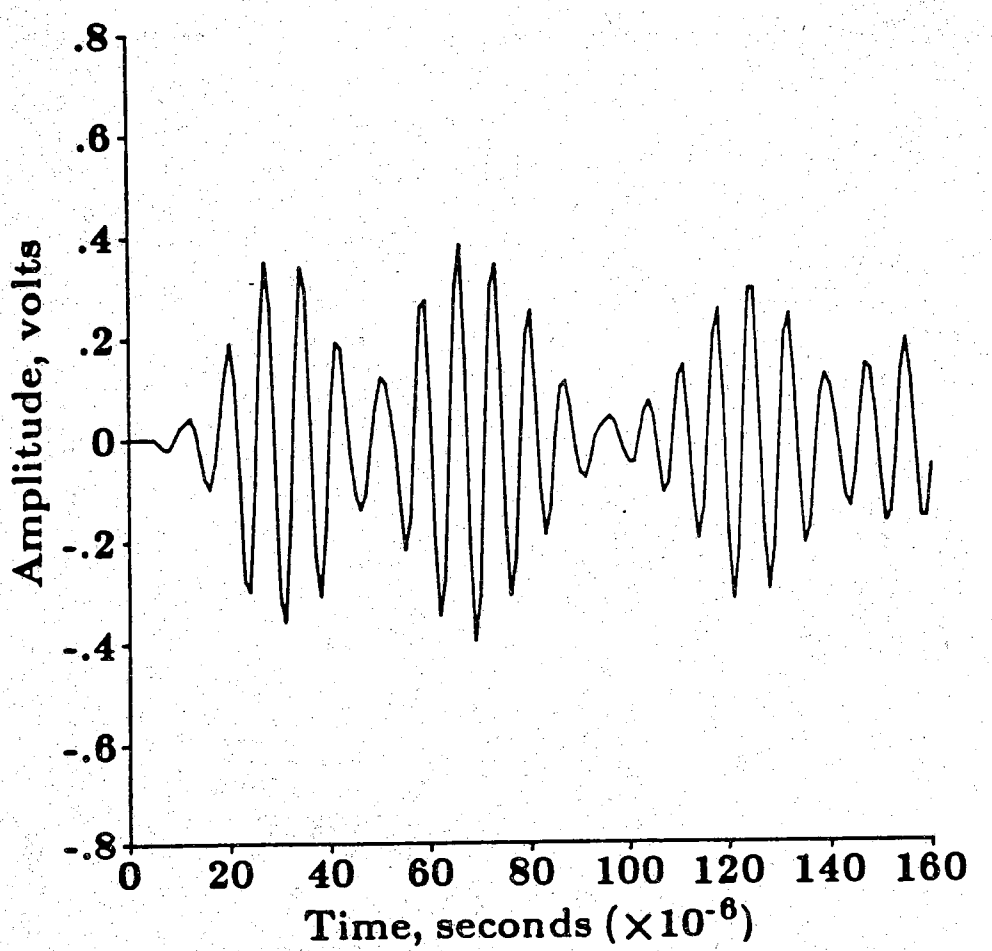


Figure 2.6. Acoustic response of test cell.

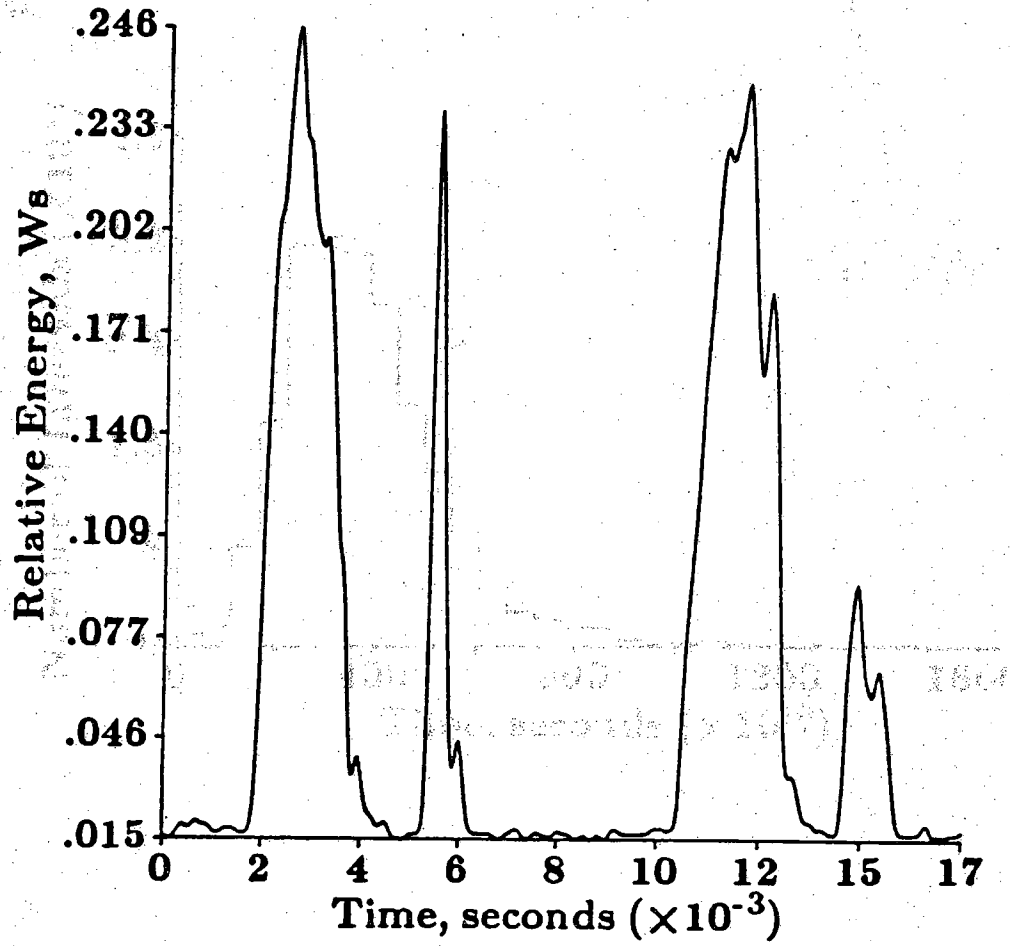


Figure 2.7. Relative energy versus excitation voltage phase.

which is defined slightly different than in [61]. In this case the mark is the energy and the mark accumulator function is defined for each phase bin. As can be seen from Fig. 2.7 the energy is distributed across the excitation voltage in a cyclic manner.

A model was developed in which assumptions about the stationarity of the process were made. To verify the assumptions about the stationarity of the process, it is necessary to test the stationarity of the components of the model. It was stated above in the section about the signal model that if the measurements were confined to a small periodic interval, (with the same period as the excitation voltage), then the parameters describing the process would be constant. In particular it is necessary to test the stationarity of the times between partial discharges and their energies. The method that was used to test the hypothesis of stationarity was a nonparametric test called the run test, which can be used to test for trends in time, [15]. The test used the mean and variance of the samples over 20 second intervals and checked the stationarity over a half hour period at a significance level of 0.05. The run test was applied to similar sample paths to test for trends in time. Results of the test indicate whether or not the sample paths came from a stationary process; usually, the data was accepted as stationary.

A typical histogram of the time between discharges is shown in Fig. 2.8. The histogram is not exponential, ruling out a Poisson process. The next order approximation would be a renewal process where the times between partial discharges can assume an arbitrary distribution. A typical histogram of the energies for the acoustic emission is depicted in Fig. 2.9 for a fixed phase window. An interesting relation between the gap size, x, and the energy emerged from these measurements. Fig. 2.10 shows the relationship between the mean energy of the discharge and the gap size for a fixed phase window. When plotted on a semilog scale the resulting curve suggests using

$$b(x) = k\exp(mx)\ ; \quad k > 0,\ m > 0,\ x > 0$$

as a model for the energy/gap relationship, where b(x) is the parameter controlling the energy distribution. The constants k and m can be determined by performing a simple least squares fit on the data in Fig. 2.10. Combining this with the histograms of the energy, a simple model can be constructed by assuming the histograms can be modeled as gamma density functions. This yields a model for the mark distribution as a function of gap

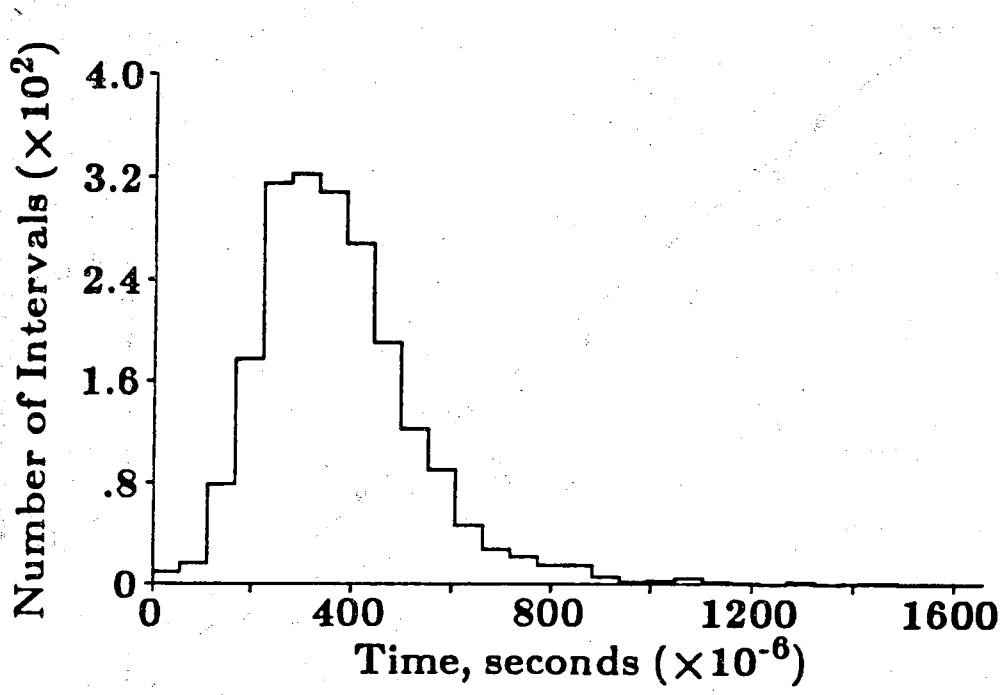


Figure 2.8. Histogram of time between partial discharges.

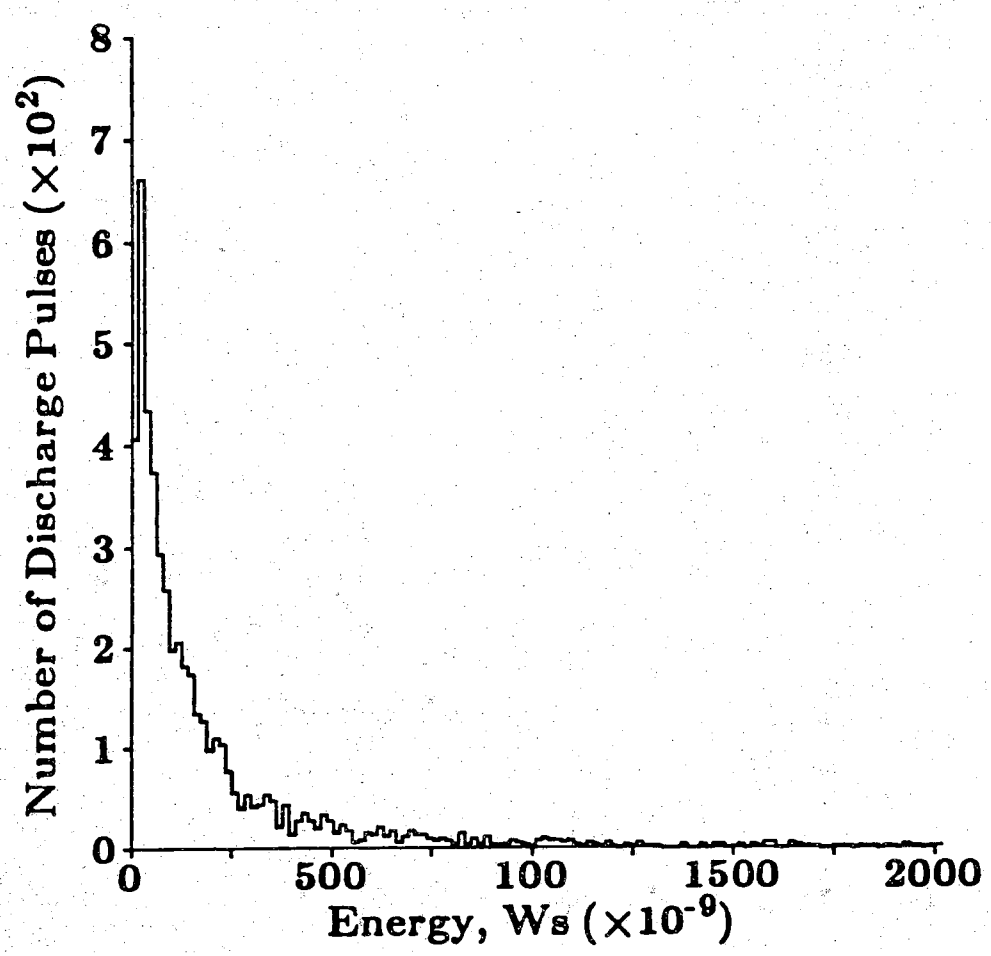


Figure 2.9. Histogram of acoustic energy from partial discharges.

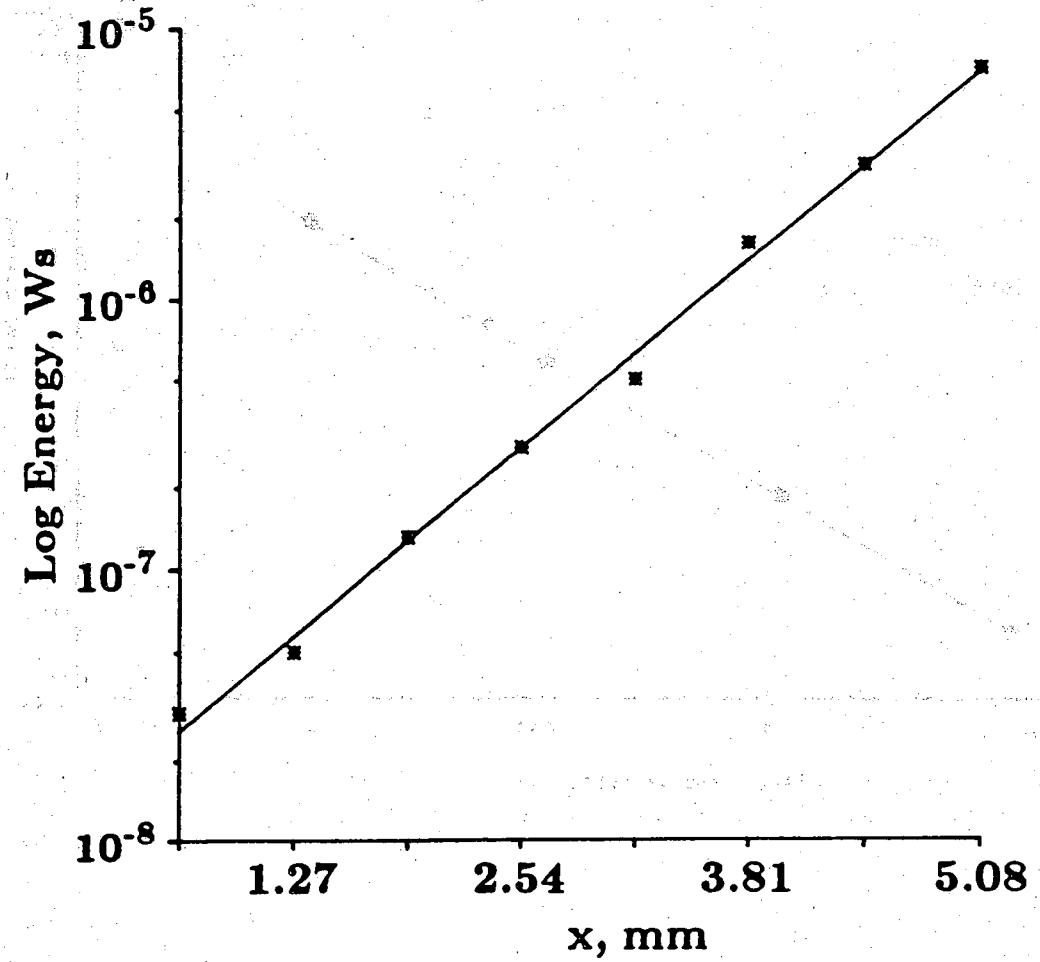


Figure 2.10. Mean energy versus gap size.

size

$$p(u;x) = \frac{b^{n+1}(x)}{n!} u^n \exp(-b(x)u)$$

This particular representation has a mean value of

$$m(x) = \frac{n+1}{b(x)}$$

and a variance of

$$\sigma^2(x) = \frac{n+1}{b^2(x)}$$

It should be remarked that the constants k and m are actually time varying quantities. As mentioned above, the parameters controlling the mark distribution will be assumed to be periodic nonstationary. Therefore, for any small time interval located at a fixed phase of the excitation voltage, the quantities k and m will be constants, (which will change as the time interval is moved to another phase of the excitation voltage).

Another experiment which was performed involved the study of partial discharge when the surface conductivity, ρ_s, is changed. The same type of electrode configuration was used as in the energy/gap experiment except that the surface conductivity of the discharge region was controlled. The gap separation was held constant and the surface conductivity was controlled by coating the faces of the partial discharge region with resistive paint. The results of the experiment are illustrated in Fig. 2.11 where the mean partial discharge energy versus surface conductivity is plotted. Only a few points are plotted because of the difficulty of obtaining controlled surface conductivities on the samples. On a log-log scale there appears to be a linear relationship between the mean energy and the surface conductivity. Further study of the energy relation to the surface conductivity is needed to validate this model.

A classic example of partial discharge with non-exponential occurrences times and a complex dependency structure is the DC point to plane process

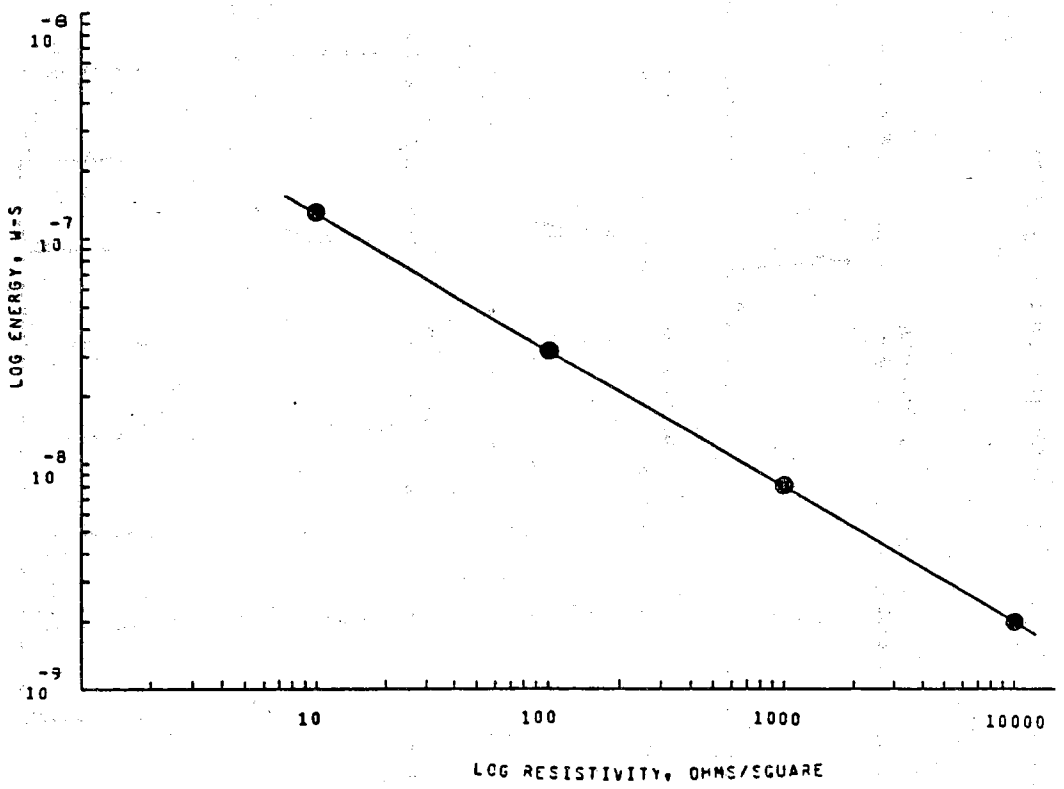


Figure 2.11. Mean partial discharge energy vs. surface conductivity.

referred to as Trichel pulses. Fig. 2.12 illustrates a typical sample function from a DC point to plane partial discharge in which the events seem to form a periodic sequence. Trichel pulses seem very regular but there is a small randomness in the behavior of discharge processes like these. Referring to Fig. 2.13 a histogram of the time between events and a histogram of the relative charge in the events clearly show that even a seemingly stable partial discharge process like DC Trichel pulses have random characteristics.

There are several reasons for considering DC Trichel pulses as an experimental beginning for detailed statistical analysis under a point process framework. The most important reason is that this partial discharge has been extensively studied and is probably the best understood partial discharge type. Also, the experimental conditions are relatively simple to reproduce without having to resort to exotic sample preparation techniques. Another reason is that the excitation voltage is constant removing one source of nonstationarity in the process. There are similar types of partial discharge that can occur under AC conditions and this beginning point would lend understanding to the AC case. Finally, Trichel pulses appear to be so regular that if a reasonable model can be found for this situation, which appears to have a highly complex dependency structure, then modeling of other more random appearing partial discharges would seem, intuitively, a more tractable task.

A single model for the simple case of Trichel pulses is difficult to develop because the statistical characteristics change as the applied potential is changed. At lower potentials for which the repetition rate remains below several hundred KHz a renewal model appears to be an accurate model for the occurrence times. At higher repetition rates a complex dependency structure begins to become apparent with the memory of the process extending over large numbers of events in the series. The following analysis used observation times that ranged from 300 μs to 15 ms. Slowly varying nonstationarities from changes in the materials due to aging were observed; the measurements changed slightly when repeated with a gap that had been allowed to age.

A simple experiment was performed using the same test cell as used in the other experiments except that the high voltage electrode was replaced by a phonograph needle. The experiments were performed at 766 mm Hg, 22 ° C and 60% relative humidity. The spacing of the point to plane was 16 mm and the applied voltage was varied between 5 KV and 8.5 KV with the point being kept at the most negative potential. A single recording provided the

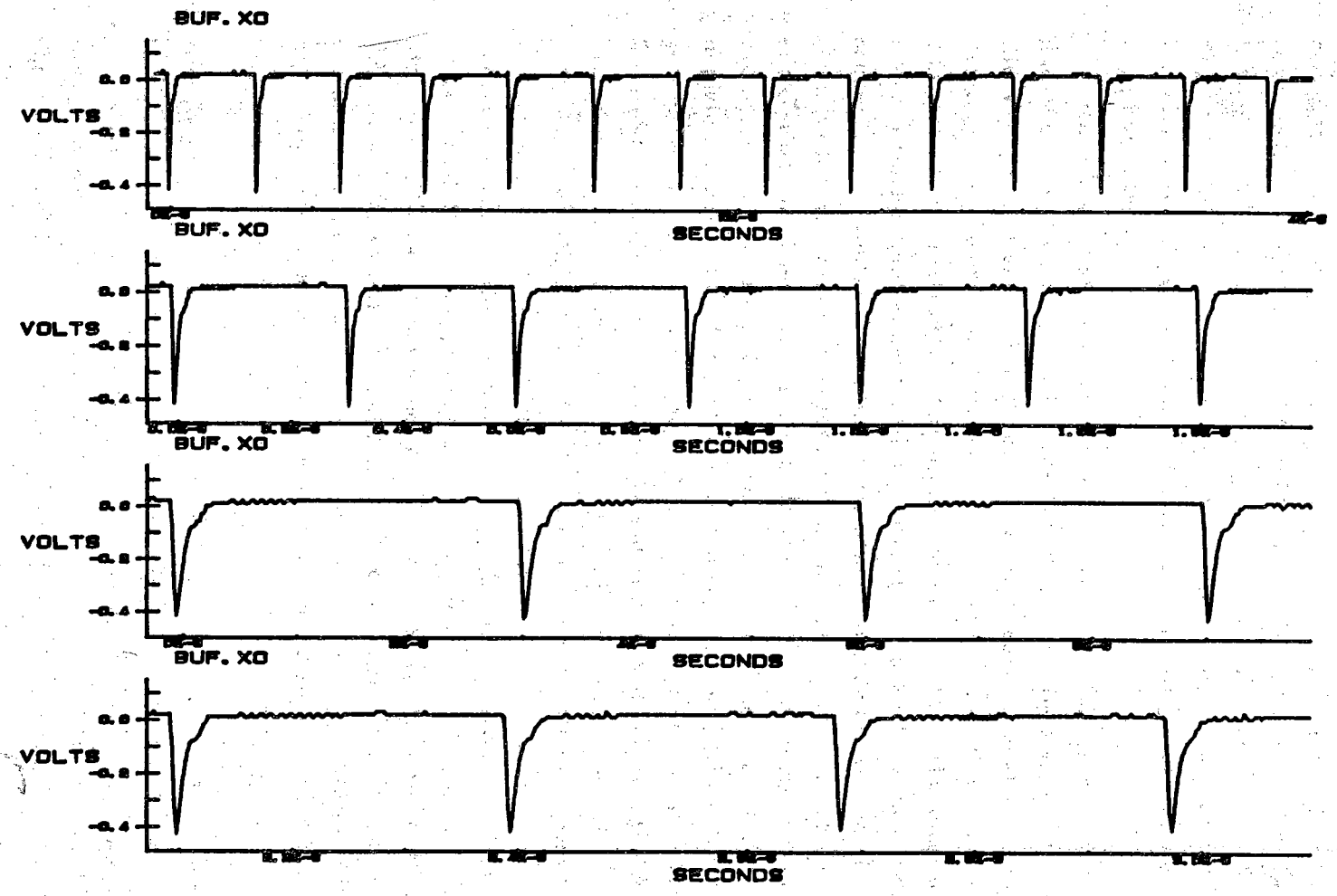


Figure 2.12. Typical sample function of Trichel pulses.

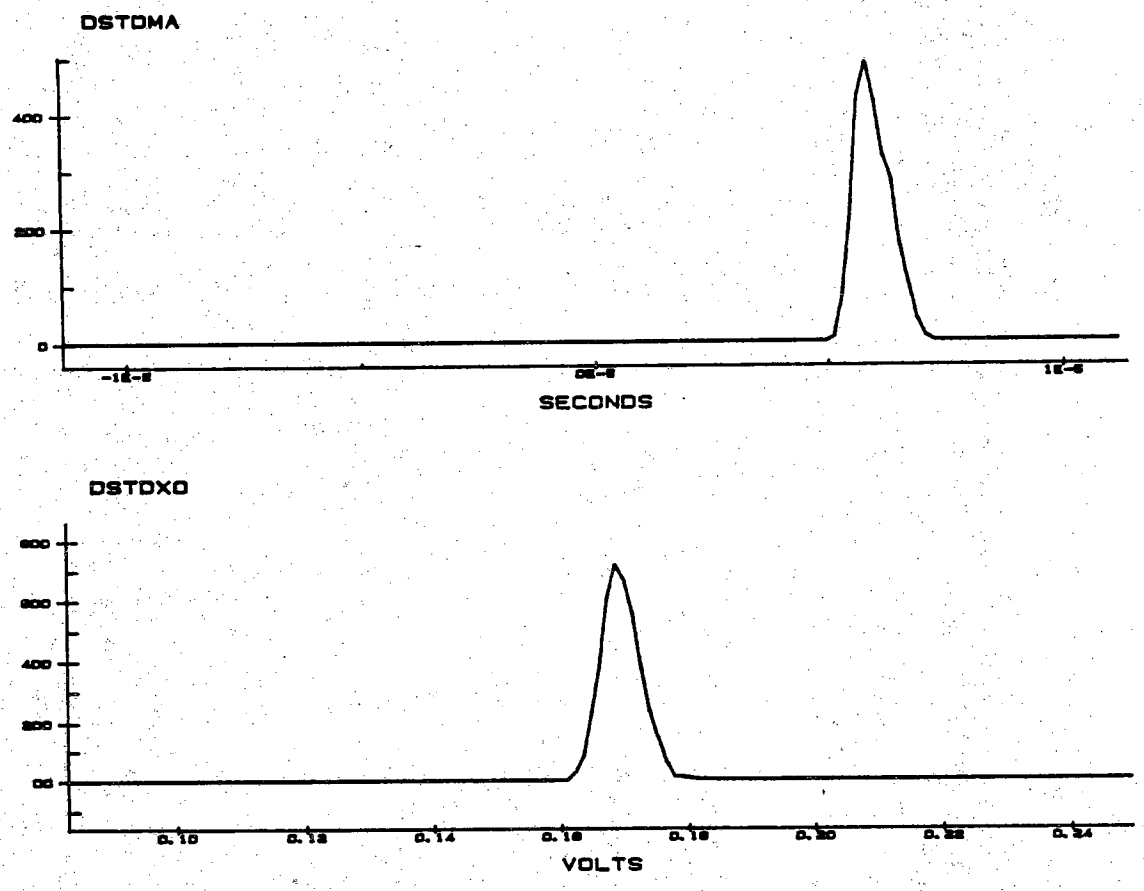


Figure 2.13. Statistical description of Trichel pulses. a) Histogram of time between partial discharges. b) Histogram of partial discharge energies.

data (30,000 points) and the techniques in section 2.2 were used to measure the data points for the sequence. Two sequences were formed; an inter-occurrence time sequence and an amplitude sequence. Typical sequences contained between 100 and 400 data points. The sampling rate was varied between 2 MHz and 100 MHz so that an adequate number of data points could be collected for different repetition rates. The input bandwidth to the recorder was varied between 500 KHz and 10 MHz depending upon the sampling rate.

Statistical analysis of the series of events is difficult without a specific model to use. Since the inter-occurrence time histogram is not exponential the Poisson process is ruled out. The next easiest model is the renewal model where the inter-occurrence times can have an arbitrary distribution. The general characteristic of renewal processes is that the inter-occurrence times form an independent series of events. Since the series forms an iid sequence of events the serial correlation coefficient should be be zero (ideally) for lags other than zero. A natural first step is to check the serial correlation coefficient to ensure that there is not correlation between the events; a typical example of the correlation coefficient for different lags is depicted in Fig. 2.14. A hypothesis test was performed which tests for serial correlation in the sequence [85]. This test was performed at a significance level of 0.05 and the data always passed this test if the repetition rate was low enough. As can be seen from Fig. 2.14 the amplitude data has a serial correlation which is similar to the inter-occurrence time's. Another aspect that is apparent is that there is a strong correlation between the inter-occurrence times and the amplitudes. The assumption that the counting process and the mark process are independent is ruled out by this result. More sophisticated models need to be considered to handle this type of dependence for example a mutually self exciting process. It is interesting to examine similar sample serial correlation functions obtained from high repetition rate Trichel pulses. Fig. 2.15 and 2.16 illustrate two different situations; one in which the serial correlation is just beginning to become apparent and the other in which there is clearly a significant memory reaching far into the past of the process. It is interesting to note the highly structured form of the serial correlation coefficient for the highest repetition rate data. It seems that this serial correlation function might be easily described with a simple parametric model. Also note the similarity between the correlation functions for the data set; even the cross-correlation has the same form.

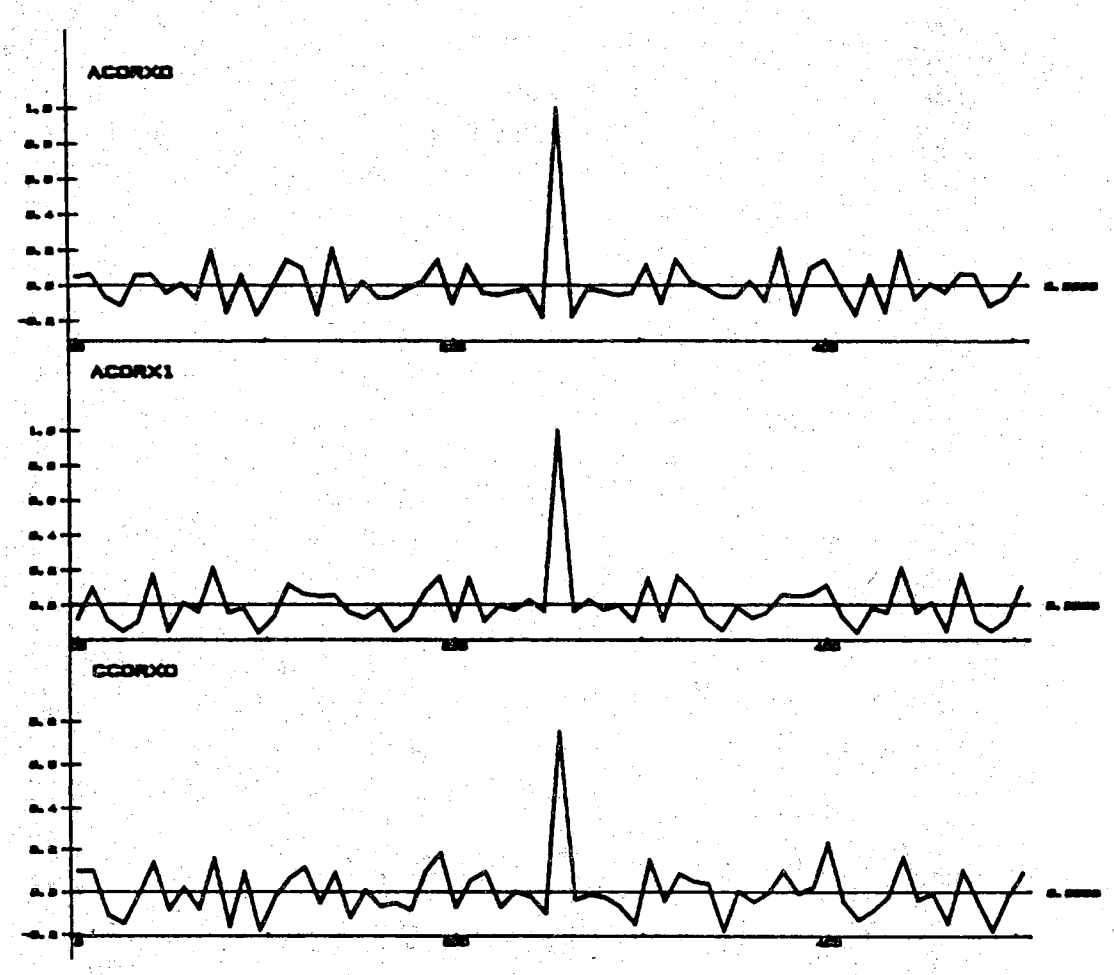


Figure 2.14. Serial correlation functions, $\lambda = 5.7 \times 10^3$. a) Autocorrelation of time sequence. b) Autocorrelation of energy sequence. c) Cross-correlation of time and energy sequences.

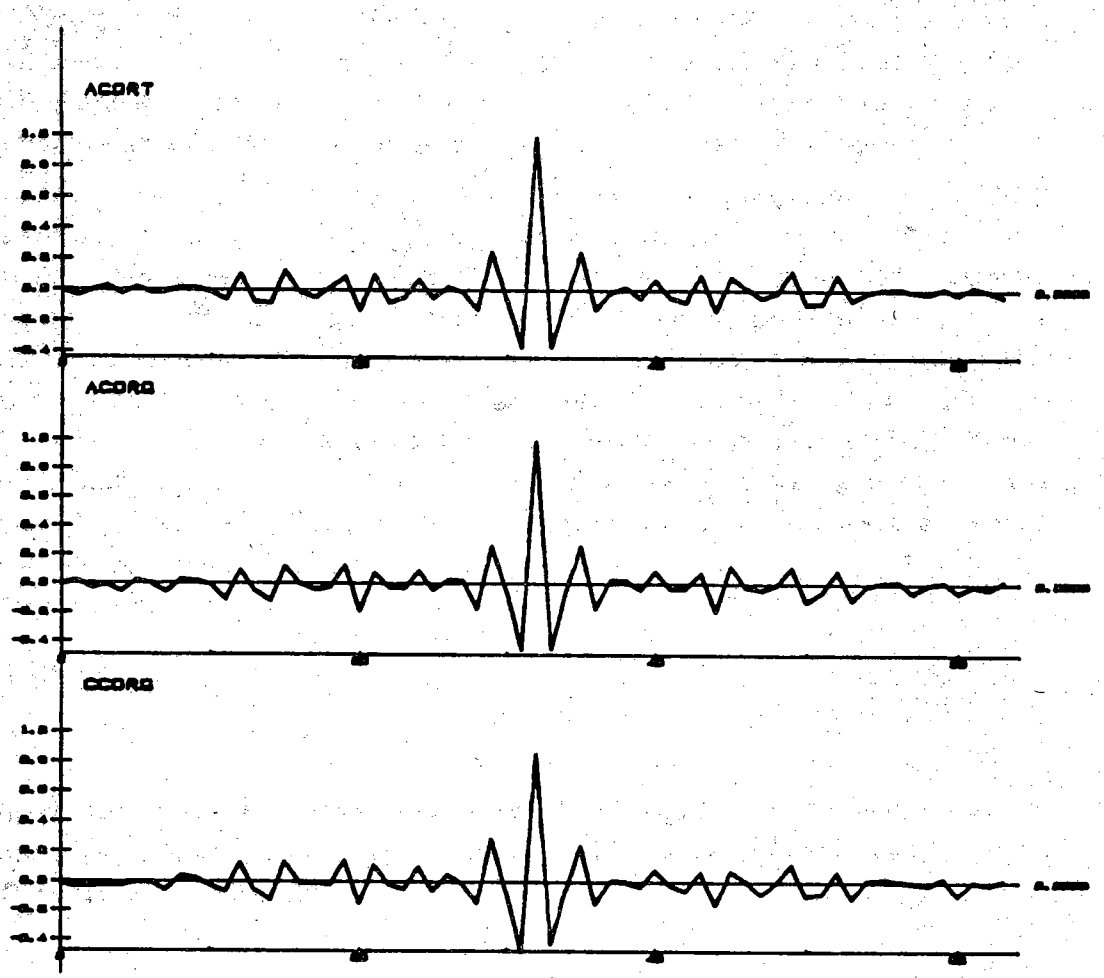


Figure 2.15. Serial correlation functions, $\lambda = 8.0 \times 10^5$ a) Autocorrelation of time sequence. b) Autocorrelation of energy sequence. c) Cross-correlation of time and energy sequences.

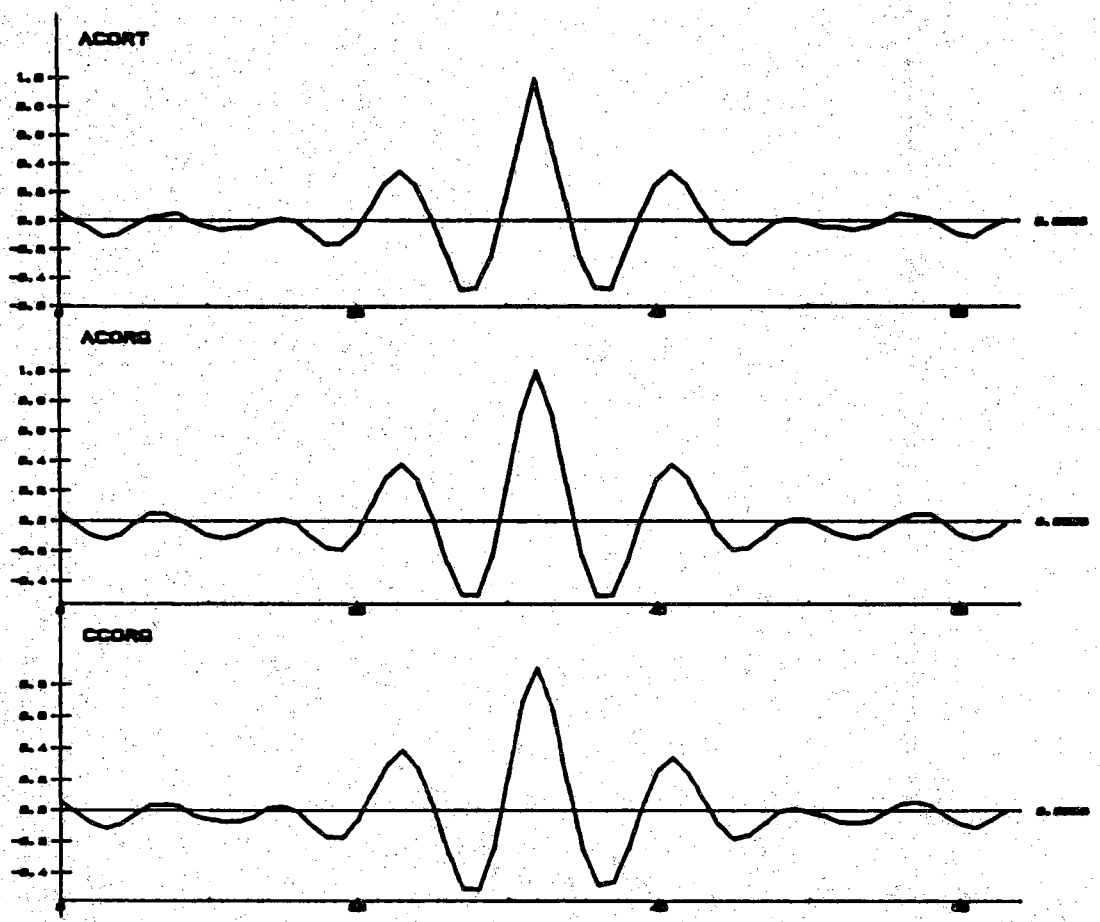


Figure 2.16. Serial correlation functions, $\lambda = 1.43 \times 10^6$ a) Autocorrelation of time sequence. b) Autocorrelation of energy sequence. c) Cross-correlation of time and energy sequences.

Checking for serial correlation is important as a first step but this does not guarantee independence. More powerful tests which test for serial independence must be used on the sequence. The preferable manner for testing a series of events for the general renewal hypothesis would be to use a nonparametric test for serial independence in the sequence. The run test [15] is a nonparametric test which can be used for testing the hypothesis of serial independence. There are other tests which can also be used and can be found in [85]. The results from applying the run test to the data, which did not exhibit serial correlation, were that the sequences consist of serially independent events at the 0.05 significance level. These results are not as definitive as would be desired because the run test has a low power.

The concept of pooling the data to form a Poisson process was also investigated using a Trichel pulse process. In this experiment multiple records were superposed to form the pooled process. Figure 2.17 illustrates the results of pooling pulses from a Trichel pulse process which had a repetition rate of 28130 pulses per second. The pooled record length was 20.44 μs and contained the superposition of sixteen records. On the average, the pooled records contain about 9 pulses each. There are several different tests that can be used to test the data to see if it was generated by a Poisson process [85]. The statistical test that was used is called a dispersion test [85] and uses a test statistic, d, which is given by

$$d = \sum_{i=1}^{k} \frac{(n_i - \bar{n})^2}{\bar{n}}$$

where there are k observation intervals of equal length with n_i events contained in the i^{th} observation interval. The term, $\bar{n}$ is the mean number of events and is given by

$$\bar{n} = \frac{1}{k} \sum_{i=1}^{k} n_i$$

For large enough k a chi-squared distribution with k-1 degrees of freedom is a good approximation to the distribution of d under the null hypothesis. Data similar to that shown in Fig. 2.17 was tested using 40 pooled records and the Poisson hypothesis was accepted at the 0.05 significance level. A

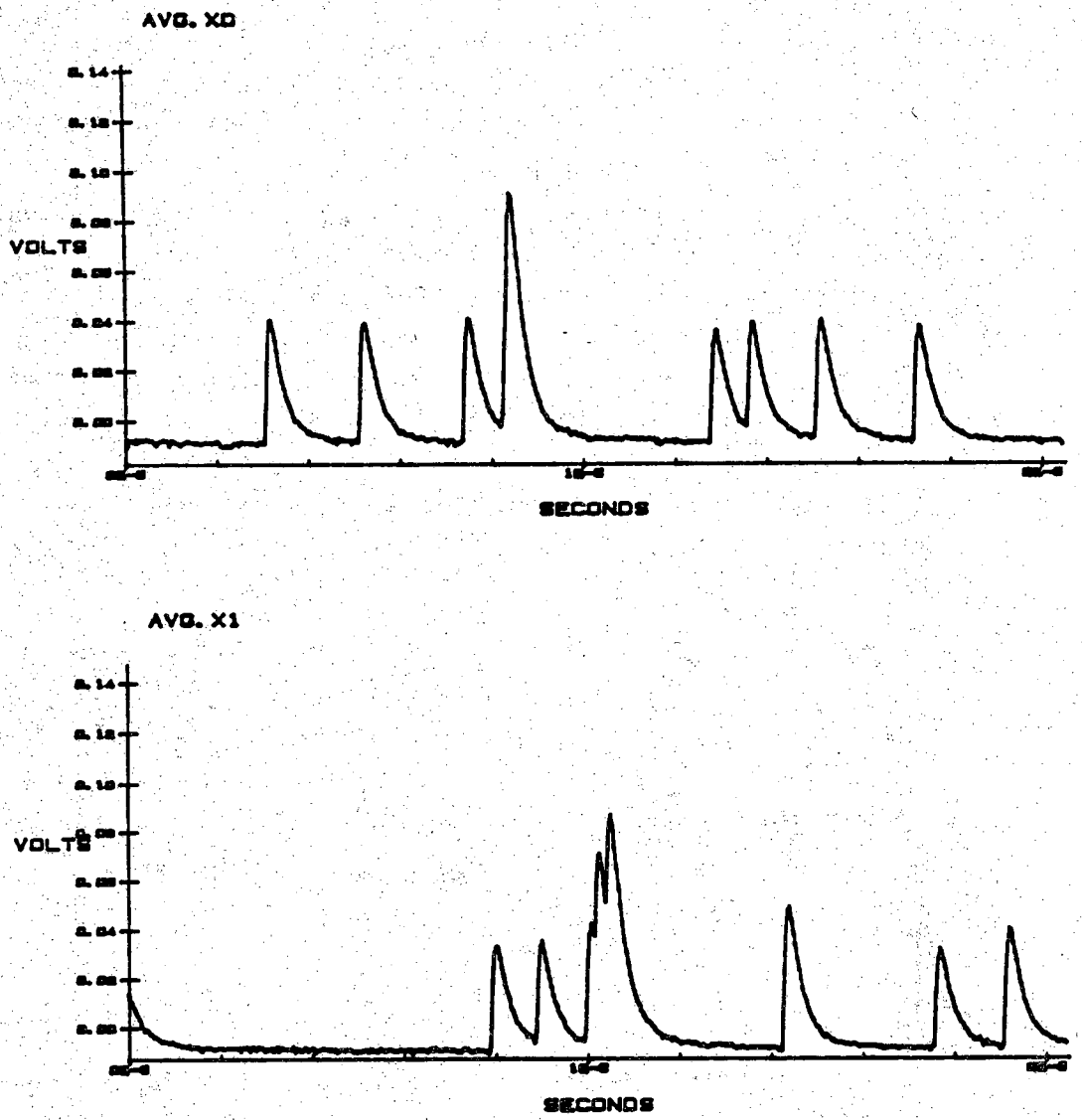


Figure 2.17. Typical pooled records for Trichel pulses.

typical histogram for the times between events in a pooled process is shown in Fig. 2.18.

2.6 Ultra Low Level Measurements

It is of great interest to measure partial discharge signals with extremely low levels of charge. This section will deal with a technique which can aid in the measurement of low level pulses buried in noise with small repetition rates. The basic approach will use cross-correlation and averaging to reduce the noise level in the estimate. Cross-correlation will be used to provide a coherent averaging method. To increase the output signal to noise ratio an ad hoc detection technique based on the level crossings of a Gaussian random process will be used. The Gaussian noise assumption will be used because it is an excellent model for the thermal noise in the amplifiers and is the dominant source of noise in the application to be discussed. The final estimate will not be the charge in a single pulse but will be an average value that is difficult to interpret. This technique applies to partial discharge pulses that have peak values either greater than or less than the standard deviation of the Gaussian noise. The most interesting case is that in which the peak value of the pulse is on the order of the standard deviation of the noise. In this case the SNR is too low and the individual partial discharge pulse is not directly observable.

The most desirable approach would be to setup up a detection procedure to test the hypothesis that the signal is present. However, there are several unwanted parameters in the formulation for which the a priori probability densities are unknown. Implementation of such a scheme is difficult in this case and approximate procedures are needed. Another reason for using the method to be discussed is that is provides a simple, real time alternative to a true detection algorithm which can become numerically intensive. If a digitally implemented detection scheme were used and the signals were observed for a one minute interval using an eight bit quantizer with a 100 MHz sampling rate then over 10 giga-bytes of data would have to be processed. This would translate into an excessive amount of data processing for even a super-mini computer to perform. However, once the preprocessing has been performed by the algorithm to be presented then it is desirable to perform a hypothesis test using, for example, the generalized likelihood ratio test discussed in chapter four. Follow up processing by more exact tests will further increase the output SNR of the measurements.

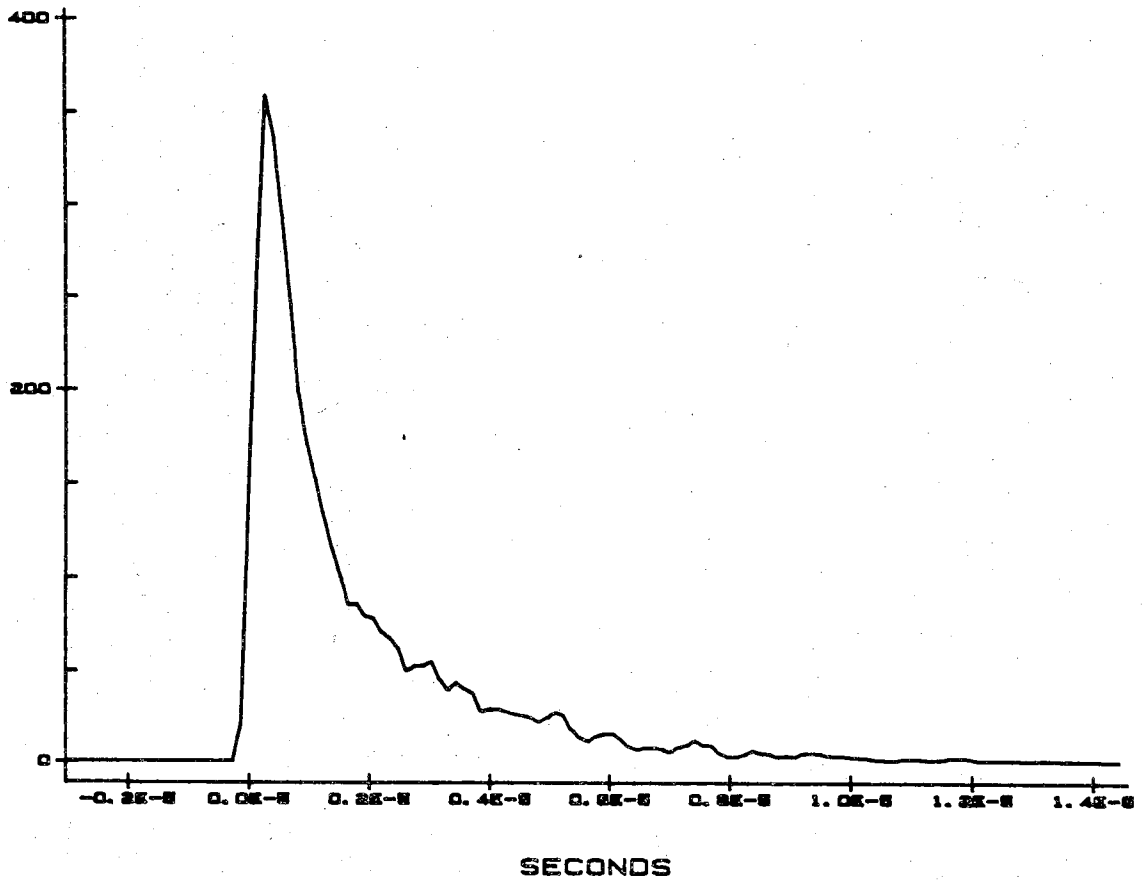


Figure 2.18. Typical histogram of times between events for pooled records.

The measuring system will consist of two inputs; these two inputs can be at physically separated locations or at a single location (two amplifiers connected to the same point). The input signals are modeled as

$$x(t)=\sum_{i=1}^{N_T} a_i s(t - t_i) + n_1(t)$$

and

$$y(t)=\sum_{i=1}^{N_T} a_i s(t - t_i - D) + n_2(t)$$

where D is the delay between the signal s(t) at the two input points and is zero if the amplifiers are connected to the same point. The number of events in the interval, T, is N_T and will be assumed to be Poisson distributed. The noise terms $n_1(t)$ and $n_2(t)$ will be assumed to be Gaussian and independent from each other and from the signal s(t). The input SNR for either channel is easily shown to be

$$SNR_{in} = \frac{\lambda \overline{a^2} \int_{-\infty}^{\infty} |S(\omega)|^2 \, d\omega}{\int_{-\infty}^{\infty} N(\omega) \, d\omega}$$

where $S(\omega)$ is the Fourier transform of the partial discharge pulse, $N(\omega)$ is the noise power spectral density, the partial discharge process is assumed to be homogeneous with intensity parameter λ, the pulse is not DC coupled and the mean square value of a_i is $\overline{a^2}$. The Poisson process is assumed to be homogeneous because no a priori knowledge is available about the intensity function λ_t. The homogeneous intensity, λ, will have the interpretation that it is the average number of partial discharge pulses per second even though the true process will not be homogeneous.

One approach to the problem of measuring low level partial discharge would be brute force averaging of the cross-correlation function. This approach has limitations for low repetition rate (small λ) partial discharge

because the output SNR is still low when λ is small. There are tremendous difficulties associated with retrieving these pulses from the noise if there are only a few pulses per cycle of the excitation voltage because the variance of the estimate will be large compared to the partial discharge energy. An obvious modification would be to limit the processing to a small phase window, of extent ϕ, to a region where the probability of measuring a partial discharge is the highest. The intensity, λ was assumed to be constant because there is not a method for knowing the true function λ_t before measuring it. In actual practice some information is known; the discharges are most likely to occur near the voltage peaks. If the correct phase window is chosen which contains all of the partial discharge pulses then the equivalent homogeneous intensity will be

$$\lambda_{eq} = \frac{360}{\phi} \lambda$$

That is to say that the intensity parameter increases giving a higher input SNR which yields an improved output SNR. Furthermore, if an accurate trigger is available so that the processing can be limited to regions in the near vicinity of the pulse then the output SNR increases dramatically. The problem with using only the regions in the near vicinity of the pulse is that these regions are difficult to determine unless the pulse has sufficient amplitude. In this case the pulse itself would be required to provide the trigger.

There is an alternative which can provide increased performance by locating the regions (on the average) in the vicinity of the partial discharge pulse. The algorithm locates data segments that have a significant probability of containing a partial discharge pulse and then processes only those segments. This approach yields better results because the data has an effective higher input SNR. This technique is not perfect because it will miss pulses and indicate that a pulse is present when it is not. These deficiencies amount to a loss in the output SNR but the improvement of the output SNR can be almost as great as if the region in the vicinity of the pulse were actually known. This technique relies on the level crossings of a random process and works by using the trigger on the digital recorder as the detection level. If the noise level is high then a trigger set at a proper level will initiate recording a few times per second (when the noise crosses the

trigger threshold). The basic idea lies in the fact that when a small signal is added to the noise then the recorder will trigger more often than if the noise alone were present. Proper selection of the trigger level will improve the output SNR because there is an increased probability that there is a partial discharge and not noise alone.

To analyze the algorithm consider what happens when noise alone is present at the input to the recorder. The noise will be Gaussian with zero mean and having a power spectral density, $N(\omega)$. The trigger level will be set at a level u so that any upcrossing of the level u by the noise will initiate recording. The mean number of triggers (upcrossings) in the interval T is given by ,[86]

$$E\{ C_u(0,T) \} = \frac{T}{2\pi} \left[\frac{\alpha_2}{\alpha_0} \right]^{1/2} e^{-\frac{u^2}{2\alpha_0}}$$

where the α_j terms are the spectral moments of the noise and are given by

$$\alpha_0 = R_n(0)$$

$$\alpha_2 = - R_n''(0)$$

where R_n is the autocorrelation function of the noise and the prime refers to differentiation. The second spectral moment, α_2, is required to be finite; white noise is not allowed. When a signal is added to the noise then the mean number of triggers is given by [86]

$$E\{ C_{m(t)}(0,T) \} = \frac{1}{\sqrt{\alpha_0}} \int_0^T \left\{ \phi\left(\frac{m(t)}{\sqrt{\alpha_0}}\right) \left[2\sqrt{\alpha_2}\,\phi\left(\frac{m'(t)}{\sqrt{\alpha_2}}\right) + m'(t) \left[2\Phi\left(\frac{m'(t)}{\sqrt{\alpha_2}}\right) - 1 \right] \right] \right\} dt$$

where ϕ and Φ refer to the Gaussian probability density and error function respectively. The expression above gives the crossings of a curve m(t) by the Gaussian noise. The particular curve that is of interest is the level shifted signal $m(t) = a\,s(t) - u$ where s(t) is the partial discharge pulse with amplitude a and u is the trigger level of the transient recorder. The integrals defining the expected number of triggers, for realistic signals, are difficult to calculate and further simplifications need to be found. Furthermore, the signal is not deterministic and is actually a random process which further complicates the analysis.

Fortunately, some assumptions can lead to a usable implementation of this technique. One assumption is that the point process has a small intensity so that the probability of two partial discharge pulses having significant overlap is small. The next assumption is that the partial discharge pulse will be modeled as a rectangular pulse with the pulse width being equal to the duration of the true pulse. These two assumptions lead to a simple approximate solution in which the level crossings of the noise can be used. It should be noted that even if the two assumptions were true then the solution to be presented would still be approximate. The solution to be presented is based on a relative frequency approach to the probabilities and is a very simple approach to an extremely complex problem. However, the approach works and provides a simple technique to measure low level partial discharge.

Several definitions need to be stated before the solution can be given. These definitions are

ΔT = recorded observation interval for a single trigger

N_{pd} = number of triggers due partial discharge plus noise

N_n = number of triggers due noise only

D = duration of a single partial discharge pulse

T = total observation interval

N_t = total number of triggers

λ = average number of partial discharge pulses per second

By assumption the process is a homogeneous Poisson process so the probability that an event is observed in $(t, t + \Delta t)$ is approximately $\lambda\Delta t$. Since it was assumed that the partial discharge pulses do not overlap the coverage of the interval, (0,T), by the square pulses is $\lambda T D$ where D is the width of the square pulse. Also, the coverage of (0,T) by noise only is $T - \lambda T D$. The mean number of triggers in the interval (0,T) due to the noise only is approximately given by

$$E\{ C_u(0, T - \lambda T D) \} = \frac{(T - \lambda T D)}{2\pi} \left[\frac{\alpha_2}{\alpha_0} \right]^{1/2} e^{-\frac{u^2}{2\alpha_0}}$$

The mean number of triggers in the interval (0,T) due to partial discharge pulses plus noise is approximately given by

$$E\{ C_{u-a}(0, \lambda T D) \} = \frac{\lambda T D}{2\pi} \left[\frac{\alpha_2}{\alpha_0} \right]^{1/2} e^{-\frac{u^2}{2\alpha_0}} \int_0^\infty e^{-\frac{a^2 - 2au}{2\alpha_0}} p(a)\, da$$

where p(a) is the probability density of the partial discharge amplitude. The amplitude, a, is assumed to be of one polarity and if the discharge has a negative polarity then the same expression applies if a negative trigger is used (downcrossing). The integral in the above expression is easily evaluated for special probability densities such as the gamma density. If the amplitudes are gamma distributed with density given by

$$p(a) = \frac{\beta^{n+1}}{\Gamma(n+1)} a^n e^{-\beta a} \; ; \qquad a \geqq 0$$

then the mean number of triggers is approximately given by

$$E\{ C_{u-a}(0, \lambda T D) \} = \frac{(\lambda T D)}{2\pi} \left[\frac{\alpha_2}{\alpha_0} \right]^{\frac{1}{2}} \mu_n [2\pi \alpha_0]^{\frac{1}{2}} \frac{\beta^{n+1}}{\Gamma(n+1)} \left[\frac{e^{-\frac{u^2}{2\alpha_0}}}{e^{-\frac{(u - \alpha_0\beta)^2}{2\alpha_0}}} \right]$$

The constant, μ_n, is found from evaluating the expression

$$\mu_n = \int_0^\infty a^n \frac{1}{\sqrt{2\pi\alpha_0}} e^{-\frac{[a - (u - \alpha_0\beta)]^2}{2\alpha_0}} da$$

Using these formula, N_n, N_{pd} and N_t are approximately given by

$$N_n = E\{C_u(0, T - \lambda T \Delta T \}$$

$$N_{pd} = E\{C_{u-a}(0, \lambda T \Delta T \}$$

and

$$N_t = N_n + N_{pd}$$

The input SNR can also be written as

$$SNR_{in,1} = \frac{\lambda T \mathscr{E}_{pd}}{T \sigma_n^2}$$

where $\mathscr{E}_{pd}$ and $T\sigma_n^2$ are the average partial discharge pulse energy and noise energy respectively. The level, u, is chosen so that $N_t \Delta T < T$ which means that no more than one trigger due to noise occurs in ΔT. If this condition is not met then the following approximations for the increase in SNR will be pessimistic. Also the noise bandwidth is assumed to be comparable to the pulse bandwidth so that $C_{u-a}(0, D) < 1$. This means that each partial

discharge pulse will provide only one trigger rather than multiple triggers. As a practical matter, this choice of bandwidth is preferred because it provides the best input SNR. In fact, to achieve better performance, an analog matched filter should be used on the incoming data which would force the signal and noise to have comparable bandwidths. Provided that these conditions are met then the input SNR for the reduced data set is

$$\mathrm{SNRin,2} = \frac{N_{pd}\,\mathscr{E}_{pd}}{N_t\,\Delta T\,\sigma_n^2}$$

The gain in the input SNR is

$$G_{SNR} = \frac{SNR_{in,2}}{SNR_{in,1}} = \frac{N_{pd}}{N_t}\,\frac{T}{\Delta T}\,\frac{1}{\lambda T}$$

In the noiseless case where $N_{pd} = N_t = \lambda T$ the gain in the SNR is $\frac{1}{\lambda \Delta T}$ for which it is assumed that the only triggers generated are generated by the partial discharge pulses themselves. The SNR gain can be put into a more suggestive form using the definitions $\lambda_{pd} T = N_{pd}$ and $\lambda_n T = N_n$ which gives

$$G_{SNR} = \frac{1}{\lambda \Delta T}\left[\frac{\lambda_{pd}}{\lambda_{pd} + \lambda_n}\right]$$

where the first term represents the ideal gain and the term in brackets represents the loss in gain due to the nonideal behavior of the algorithm.

The above formulas represent the gain in the input SNR using the algorithm and it is of interest to find a similar expression for the output SNR. The estimate being used is the cross-correlation between the two inputs and expressions for the variance of a cross-correlation estimate can be found in [15]. Ignoring the contribution to the variance from the randomness of the partial discharge amplitude and considering only the contribution from the noise gives the output SNR for direct averaging

$$SNR_{out,1} = \frac{(\lambda T \mathscr{E}_{pd})^2}{T \mathscr{P}}$$

The term, $\mathscr{P}$, is defined by

$$\mathscr{P} = \int_{-\infty}^{\infty} R_{n_1 n_1}(\tau) R_{n_2 n_2}(\tau) \, d\tau$$

where $R_{n_1 n_1}(\tau)$ and $R_{n_2 n_2}(\tau)$ are the autocorrelation functions of the noise in the two channels. Using the level crossing algorithm gives an output SNR given by

$$SNR_{out,2} = \frac{(N_{pd} \mathscr{E}_{pd})^2}{N_t \Delta T \mathscr{P}}$$

The gain in the output SNR is given by

$$G_{SNR} = \frac{1}{\lambda \Delta T} \left[\left(\frac{\lambda_{pd}}{\lambda} \right) \frac{\lambda_{pd}}{\lambda_{pd} + \lambda_n} \right]$$

where the first term is the ideal gain in the SNR and the term in brackets is the loss due to the algorithm.

It should be pointed out that this analysis is only approximate and gives a simple closed form expression for a highly complex problem. These expressions might best be treated as upper bounds on the possible increase in the SNR. Even though these expressions are approximate and the technique ad hoc, the results of using this technique are excellent. The following experimental simulation illustrates the effectiveness of the algorithm for measuring partial discharge buried in noise. The rms level of the noise was 97 mv with a peak pulse height of 200mv. Fig. 2.19 illustrates a typical data record with a pulse buried in noise. The noise had a bandwidth of 2MHz, the pulse had a duration of 500 ns and λ was 60 . The observation time, T, was 600 seconds, the sampling rate was 10 MHz and

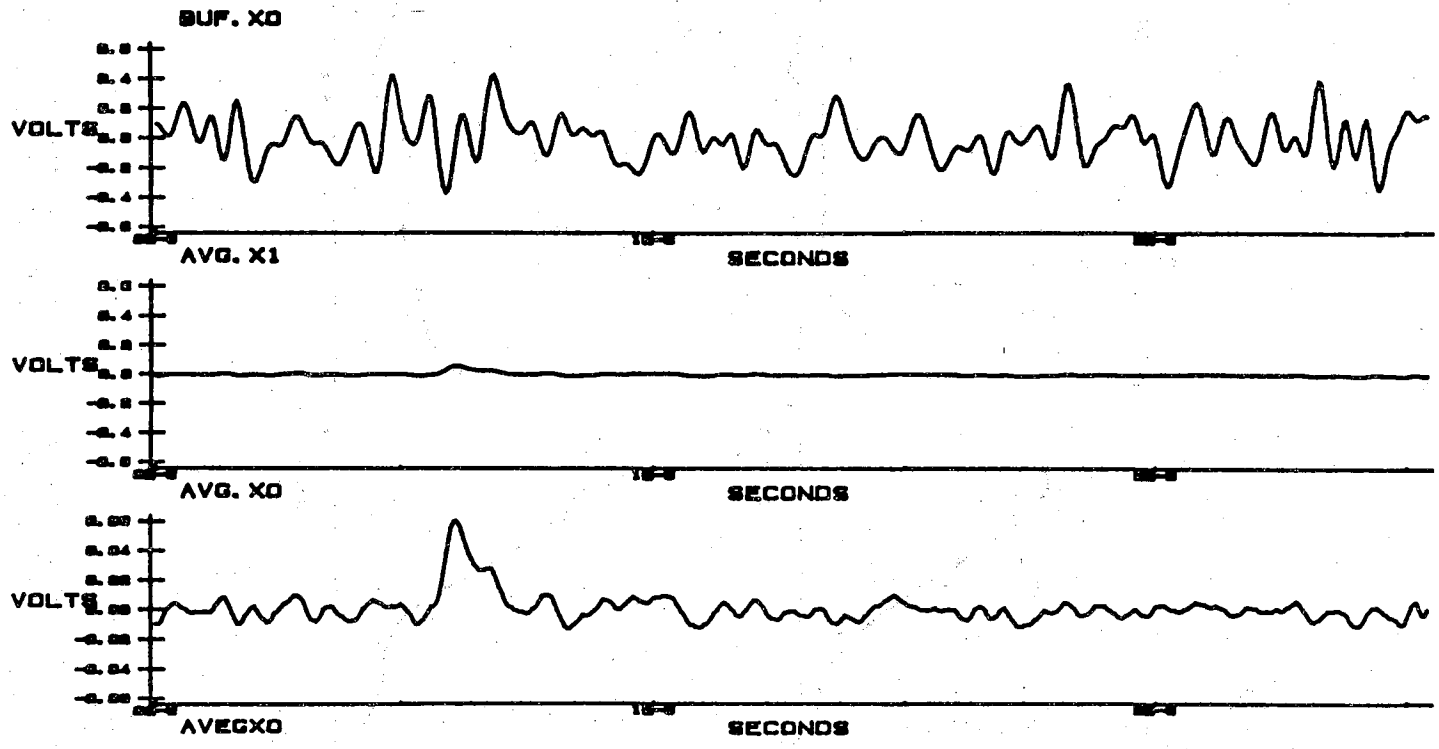


Figure 2.19. Low level partial discharge pulse. a) Pulse buried in noise. b) Actual pulse shown to scale of a). c) Magnified (X10) pulse.

ΔT was 6.4 μs. The input SNR for a single data segment containing a pulse was about -6 dB while the input SNR for the entire process was about -74 dB. The trigger level was 500 mv and the residual mean at the digitizer input was 1.8 mv giving an effective trigger level of 498.2 mv. The measured value of N_n was 2230 while the theoretical value was 2244. The measured value of N_{pd} was 335 while the theoretical value was 319. The results of the processing are shown in Fig. 2.20. The theoretical gain in the input SNR was 25 dB, the theoretical gain in the output SNR was 5 dB and the theoretical output SNR was 15 dB. The average energy in the partial discharge pulses is calculated from

$$\mathcal{E}_{pd} = \frac{N_t}{N_t - N_n^{theor}} \hat{R}(0)\,\Delta T$$

where $\hat{R}$ is the estimated cross-correlation function. The estimate for the partial discharge was 16.9 nv^2 and the actual value was 17.5 nv^2, which is within the expected accuracy based on the theoretical output SNR. This result corresponds to measuring partial pulses with a charge of about 0.0045 pC that occurs only once per cycle at 60 Hz. The output had a modest gain in SNR as compared to direct averaging but this figure does not reflect the true performance of the algorithm. To accurately assess the algorithm both the gain in output SNR and compression of the data need to be examined. For direct averaging 12 giga-bytes of data would need to have been collected and processed. Using the level crossing algorithm only 328 kilo-bytes were processed which is a compression factor of 36550. One final comment is in order about the algorithm. The trigger level is in the exponent of the formulas and consequently the algorithm is very sensitive to small changes in this level.

To improve the output SNR of the measurement further, more sophisticated processing could be performed. The generalized likelihood ratio test could be used as a second screening algorithm before the data is included in the estimate. This test is discussed in chapter four where an approximation is presented for the tests implementation. Using some definitions, the performance gain from the second test can be stated in terms of the previous results. These definitions are

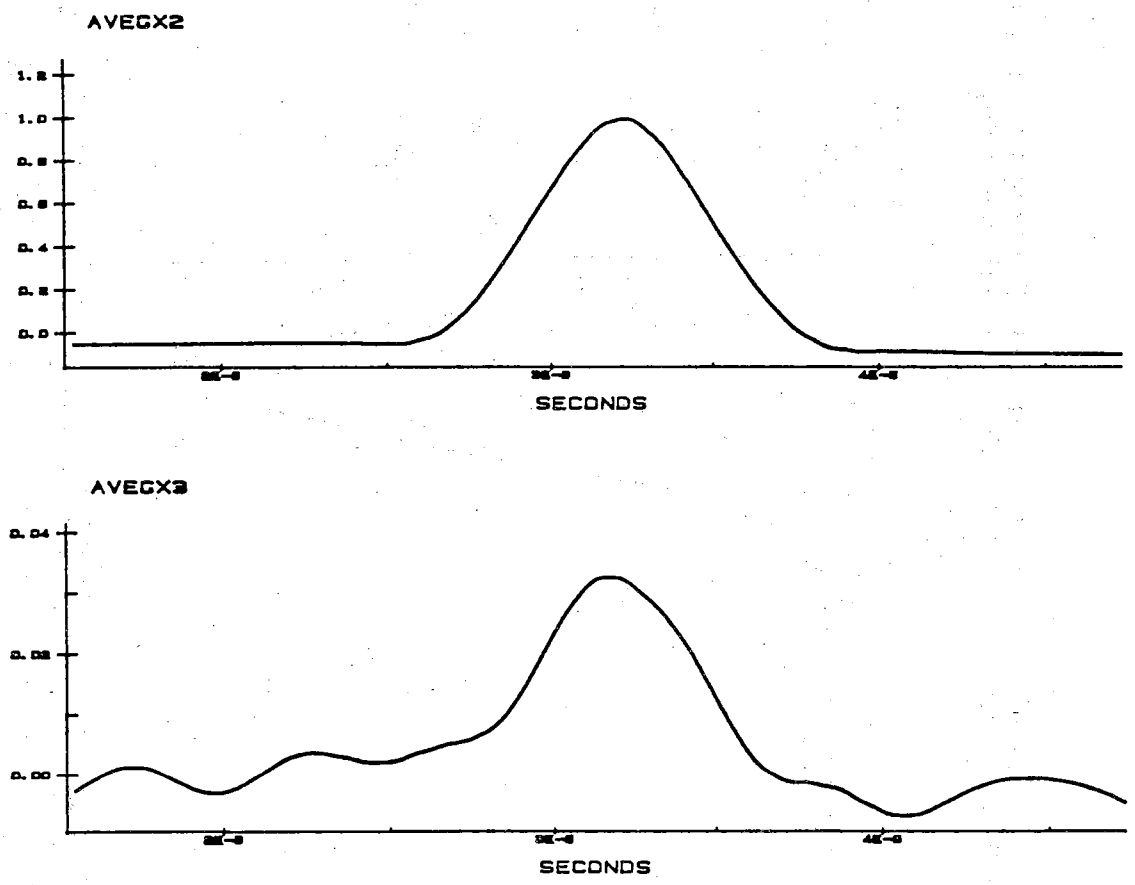


Figure 2.20. Results of level crossing algorithm. a) Ideal output of algorithm. b) Output of algorithm for experimental data.

$$P_D = \Pr(\text{deciding a pulse is present} \mid \text{pulse is present})$$

$$P_F = \Pr(\text{deciding a pulse is present} \mid \text{pulse is not present})$$

When a hypothesis test is performed on the reduced data set containing a total of N_t records then new values for the counts of detected and falsely detected partial discharge pulses must be used. These new numbers are given by

$$M_t = N_t \, (P_D + P_F)$$

$$M_{pd} = N_t \, P_D$$

Replacing N_t, N_{pd} and N_n with M_t, M_{pd} and M_n in the above expressions then gives the performance of the the two algorithms used in conjunction. The only motivation for using these algorithms in conjunction is to reduce the amount of data that needs to be processed. If the processing resources were available then only a detection algorithm would be necessary (and desirable). An ad hoc detection scheme, based on the cross-correlation function, can be used and will respond to any correlated input (both partial discharge or interference). The test would be

$$\int_{\Delta T} x(t)y(t)\, dt \underset{H_1}{\overset{H_0}{\gtrless}} \gamma$$

where γ would be the threshold of the test. Histograms under the two alternatives using this cross-correlation based test are shown in Fig. 2.21 for the same data discussed above. The H_0 alternative is signal plus noise and the H_1 alternative is noise only. It is easy to see from Fig. 2.21 that a decision region can be chosen such that P_F is small which will improve the output SNR if the detection algorithm is used.

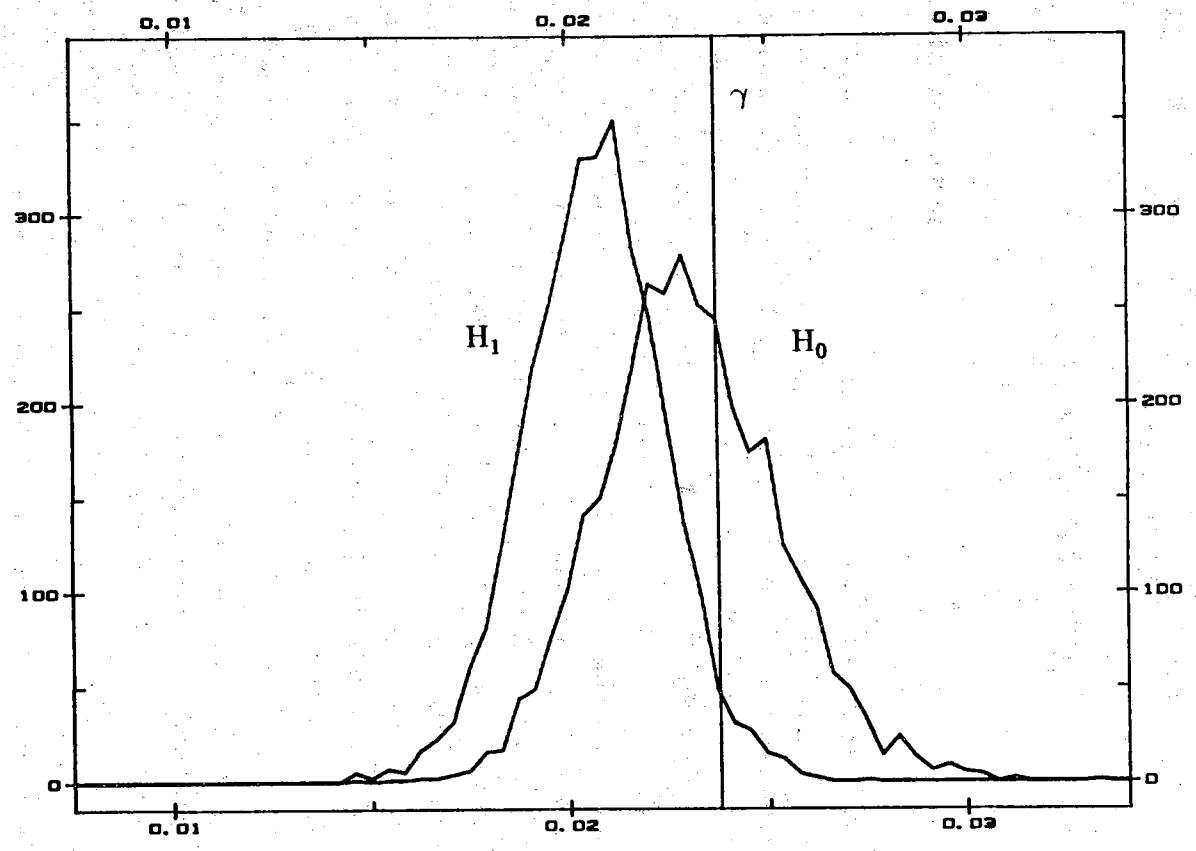


Figure 2.21. Histograms of the processor output under the two hypothesis. H_0 is to the right of the threshold and H_1 is to the left of the threshold.

CHAPTER THREE

DELAY ESTIMATION

3.1 Introduction

This chapter discusses techniques for locating partial discharge sites in high voltage equipment using digital signal processing techniques to estimate the time of arrival of the signals. There is a vast body of literature pertinent to delay estimation including passive delay estimation techniques [16]. Passive delay estimation is required to locate a signal source, (signal emitter), when the measuring device has no knowledge about the absolute temporal characteristics of the signal, (ie; its absolute time origin). For example, passive sonar systems use the sound emitted by enemy submarines to deduce their locations without any information about the absolute time at which the sound was produced. The time delay information is deduced using the relative time delays between the signals, (multiple receiver approach), or the signal and its reflection, (single receiver approach). There are so many delay estimation techniques that it would be impractical to discuss each in detail so several will be outlined but only a few specific approaches will be consider in detail.

Cross-correlation techniques have several advantages over other techniques that are used to determine delay differences. The primary advantage that cross-correlation has is that it maps a signal with a random time variation into a time (time delay) stable function. This allows the use of a 100% coherent averaging. In other words by averaging cross-correlation functions one is assured that each element of the average is perfectly lined up with the others. Direct averaging of the signals can never provide this perfect alignment. The next advantage is that uncorrelated noise averages to zero so that signals buried in thermal noise can be easily extracted. The final advantage that cross-correlation has for this application is that physically separated partial discharge sites are uncorrelated with each other.

This says that a cross-correlation function for a device with multiple discharge sites will simply be the superposition of the individual cross-correlation functions of each site.

A non-correlation based technique for locating the partial discharge sites is discussed in section 3.2. This method seems to predominate the current techniques for locating partial discharge even though this method has been abandoned by other disciplines because of its inaccuracies. This method is included for comparison and an error analysis is performed later in the chapter.

The generalized cross-correlation (GCC) method for locating a passive emitter is discussed in section 3.3. This technique is the preferred method for source localization because of the accuracy that it provides. This technique uses prefilters that are optimal in some prescribed sense. The different variations that will be discussed are useful in solving some of the problems encountered in measuring time delay in high voltage systems. High accuracy results are possible using the maximum likelihood (ML) weighting function and this method is the preferred method for small data records. If the measurements are made in a system with narrow band interference the Wiener processor (WP) implementation seems to provide the best results.

Bias errors that can occur when making delay estimates are discussed in section 3.4. In the ideal case many delay estimators are unbiased, however, in the general measurement situation various types of bias errors can corrupt the estimate. These errors can arise because of the lack of temporal resolution when overlapping pulses are encountered. Other errors occur when the channel contains a phase distortion which distorts the correlation function causing a shift in it's maximum. Bias errors can occur through the processing methods that are chosen to estimate the delay. For example, in certain situations autocorrelation estimates can be highly biased. Section 3.5 discusses the variance of delay estimates and illustrates the advantage of correlation based estimators over time of arrival estimators. The Cramer-Rao lower bound is discussed along with a local analysis of the variance for wide sense stationary processes and transient nonstationary processes.

The methods that are used to reduce the delay bias due to overlapping correlation functions are referred to as deconvolution. Section 3.6 discusses an ad hoc generalized cross-correlator, referred to as the modified PHAT (PHAse Transform) processor, which provides greater range resolution in the

measurements. The filter is a combination of the PHAT GCC and the ML GCC which uses a naive approach to deconvolution for the frequencies with high signal to noise ratios and uses the optimal ML solution for the frequency regions with a low signal to noise ratio. Section 3.7 discusses a new GCC based on a minimax equalizer that recently appeared in literature. This particular implementation provides improved range resolution at the cost of a greater variance of the estimate. The unusual aspect of this processor is that the necessary signal spectral characteristics are estimated and the estimation error is incorporated into the solution.

3.2 A Non-correlation Based Location Technique

The first method uses only the estimated arrival times at a single receiver and there are several approaches to estimate the time of arrival of a signal. The simplest method is to use the time at which the signal crosses some predetermined threshold (the threshold is often chosen to be the signal's maximum). This approach is used in TDR systems and has limited application because of the difficulty in determining the appropriate threshold for distorted, noisy or unknown waveforms. Furthermore, if the discharge site is extremely active or there are multiple discharge sites, active at nearly the same time, the number of discharges present on the TDR screen prevents a correct delay estimate from being made. This technique has been used in factory testing in attempts to locate partial discharge sites. Typically, a wide band coupling device is placed at one end of the cable and the pulses are observed on an oscilloscope. The observed pattern consists of a directly arriving pulse and a pulse reflected back from the opposite end of the cable. There are other terms due to further reflections as the pulse rings back and forth on the cable but these are of little interest. The location is deduced from an estimate of the time difference between the first two pulses appearing on the screen. The time difference will be $\tau = \mathrm{T} + \mathrm{b} - \mathrm{a}$, where T is the total electrical length of the cable, a is the propagation time from the site to the cable end with the coupler and b is the propagation time from the site to the other end of the cable. The position of the site, ℓ, is then given by

$$\ell = \tfrac{1}{2} [\, \mathrm{L} - (\, \tau - \mathrm{T}\,)\nu\,]$$

where L is the length of the cable and ν the propagation velocity.

As mentioned above, there are several difficulties with this approach. The cable is dispersive and consequently neither the front edge of the pulse or the pulse shape is preserved. The second pulse is less in amplitude than the first. The amplitudes are random, causing a jitter on the screen. All these factors contribute to the difficulty of obtaining a delay estimate using either the peak or some predetermined amplitude threshold as a reference on the two pulse. These problems can easily cause large errors in cables of substantial length. Another problem occurs when the discharge is low level, or if there is a large amount of interference because it is very difficult to obtain a stable enough display in order to make a measurement. A further problem occurs when there is a large amount of discharge activity present in the cable. In the case where there is only one site, which is very active, confusing results can be obtained. The problem occurs when the reciprocal of the count rate intensity function is on the order of the electrical length of the cable. In this case the second pulse on the screen is not necessarily due to the reflection from the opposite end of the cable but may be due to a second discharge occurring. The same phenomena happens when the reciprocal of the combined count rate intensity function for multiple partial discharge sites is on the order of the electrical length of the cable.

An improvement in this technique can be made by using matched filtering. There are numerous articles on this subject, most of which apply to active delay estimation schemes. This technique cross-correlates a known waveform with the incoming signal; the times of arrival are then the times at which the output correlation function attains its maximums. Mathematically this operation is represented by

$$R(\tau) = \frac{1}{T} \int_0^T x(t)\, s(t-\tau)\, dt$$

$$x(t) = s(t) + n(t)$$

where $x(t)$ is the signal received at a single location on the cable , $n(t)$ is a white noise process, and $s(t)$ is the known partial discharge waveform. This approach would measure the times of arrival of the signal and its reflections. Matched filtering yields excellent delay estimates provided that the signal shape is accurately known before hand. The same approach used for the

TDR would then be used to locate the site, the only difference being the method of choosing the arrival time.

This approach has fewer problems than the TDR approach but it also fails in certain situations. The problems associated with selecting a proper threshold are now eliminated and this technique performs very well in situations where there is high levels of noise. However, there is now the problem of selecting a proper known waveform to use for the matched filter. A calibration experiment can provide an adequate waveform for this purpose. However, because the signal will travel though an unknown amount of cable the shape will be slightly different than the calibrating signal potentially causing an error in the delay estimate. The major problem is due multiple discharge waveforms arriving in close temporal proximity to each other. In this case the matched filter will not provide any better result than the TDR did.

There are generalizations to the matched filter approach which take into account non-white noise case [42]. The resulting correlation function will then become

$$R(\tau) = \mathcal{F}^{-1}\left\{ \frac{X(\omega)\, S^*(\omega)}{N(\omega)} \right\}$$

where $X(\omega)$ and $S(\omega)$ are the Fourier transforms of the received data and known waveform respectively. The term, $N(\omega)$, is the power spectral density of the non-white wide sense stationary noise corrupting the measurement. This particular implementation is effective in eliminating narrow band interference from the measurements which is a particularly troublesome problem in these measurements. Another problem that can be addressed using this approach is the problem associated with closely spaced pulses. If the spacings are exponentially distributed random variables (a Poisson process) with a mean value $\frac{1}{\lambda}$ then the matched filter can be modified to [42]

$$R(\tau) = \mathcal{F}^{-1}\left\{ \frac{X(\omega)\, S^*(\omega)}{\lambda |S(\omega)|^2 + N(\omega)} \right\}$$

This particular approach essentially deconvolves the input waveform, however, the variance rapidly increases and typical resolution performance can not be expected to exceed a factor of two if a reliable estimate is desired.

3.3 Generalized Cross-correlation

Correlation methods are widely used in delay estimation problems and there is vast literature on various correlation techniques. In the past decade a new type of correlation scheme emerged and is referred to as generalized correlation, [18,39]. The basic idea is to prefilter the signals in some optimal manner and then cross-correlate. The optimization criteria is dictated by the needs and constraints of the particular problem at hand.

Once a measurement model is developed the correlation techniques can be easily implemented on a computer using an FFT approach. The received signals are as

$$x(t) = s(t) + n_1(t) + i_1(t)$$

$$y(t) = s(t-D) + n_2(t) + i_2(t)$$

where the signal s(t) and a delayed version, s(t-D), are the signals from the partial discharge site to be located. The parameter to be estimated is the relative delay, D, between the two signals. For simplicity, the distortion due to the propagation through different path lengths will not be included. The signals will be generated by either a wide sense stationary random process or by random transient events. The thermal noise from the amplifiers is modeled as wide sense stationary noise with power spectral density $N(\omega)$. The quantization noise from the digitizer, will be assumed to be zero mean, independent, identically distributed, (iid), random variables. The measurement noise, $n_1(t)$ and $n_2(t)$, is a combination of both the thermal noise and quantization noise. The signals $i_1(t)$ and $i_2(t)$ are associated with signals from other sources not presently of interest and interfering with the signals from the current site under investigation. These signals can come from a number of different sources for example reflections of the signal from discontinuities in the device, other partial discharge sites, signal sources exterior to the device under test. With the exception of the reflected signals

the interfering signals, $i_1(t)$ and $i_2(t)$, will be considered to be independent of $s(t)$ but correlated with each other.

The delay is determined from the cross-correlation function which is estimated as

$$\hat{R}(\tau) = \frac{1}{T} \int_{-T/2}^{T/2} x(t)\, y(t+\tau)\, dt$$

The delay estimate, $\hat{D}$, between the two signals is the value of τ at which the function, $\hat{R}(\tau)$, has its maximum

$$\hat{D} = \arg \left\{ \max \hat{R}(\tau) \right\}$$

Cross-correlation is an effective manner for determining the delay between signals and has been used extensively for this purpose by many disciplines.

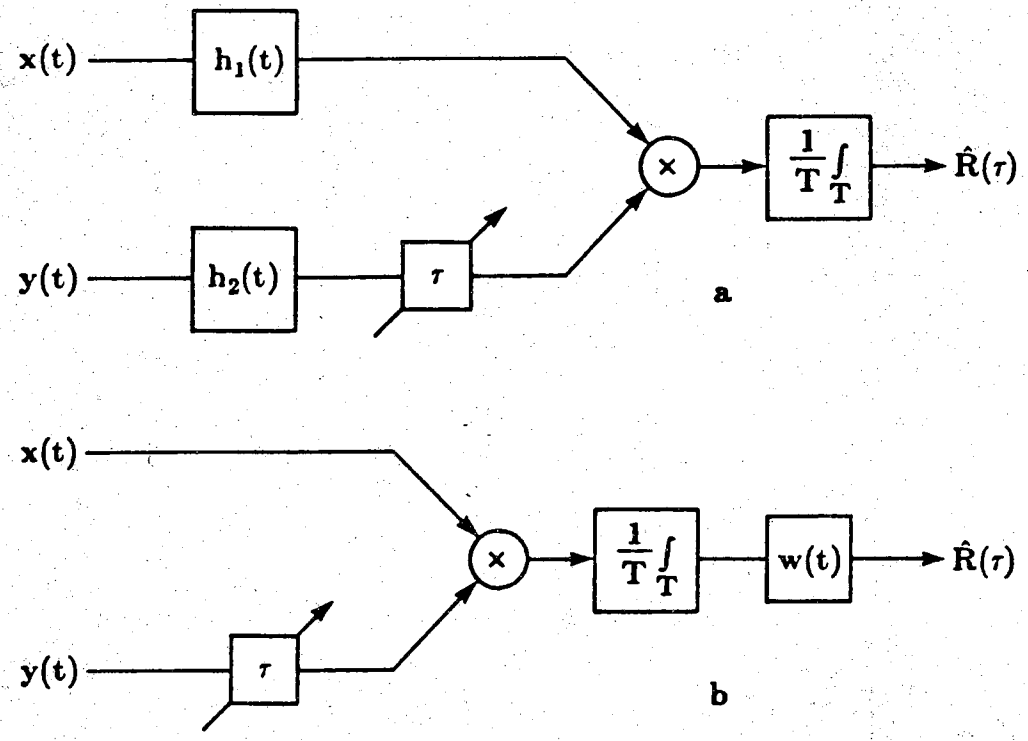
The performance of the cross-correlator can be improved upon by optimally prefiltering the input signals before cross-correlating. This operation is referred to as generalized cross-correlation [18,37,38,39,59] and its implementation is shown in Fig. 3.1. The filters $h_1(t)$ and $h_2(t)$, or equivalently, a weighting function, $w(t)$, are chosen to optimize the measurement of D in some chosen manner. The resultant correlation function, is given by

$$R(\tau) = \mathcal{F}^{-1}\{ H_1(\omega)\, H_2^*(\omega)\, S_{XY}(\omega) \}$$

$$= \mathcal{F}^{-1}\{ W(\omega)\, S_{XY}(\omega) \}$$

The term $S_{XY}(\omega)$ is the cross-spectral density of the signals $x(t)$ and $y(t)$ and can be expressed as

$$S_{XY}(\omega) = G_{p1}(\omega)\, G_{t1}(\omega)\, G_{p2}^*(\omega)\, G_{t2}^*(\omega)\, S(\omega)$$



GENERALIZED CROSS CORRELATOR

$$x(t) = s(t) + n_1(t)$$

$$y(t) = s(t-D) + n_2(t)$$

$$\hat{R}_{xy}(\tau) = \mathcal{F}^{-1}\{\hat{G}_{xy}(f)\}$$

$$\hat{R}_{xy}^{GCC}(\tau) = \mathcal{F}^{-1}\{\hat{G}_{xy}(f)W(f)\}$$

Figure 3.1. Block diagram of generalized cross-correlator.

where $S(\omega)$ is the spectral density of the signal before being filtered. The filters acting on the signal will be the cascade of the filter, $G_p(\omega)$, describing the propagation paths through the device and the filter, $G_t(\omega)$, describing the transfer characteristics of the measuring circuits. Once the relative delay, D, between the two signals is determined, then along with the known length, L, between measuring points and a measured velocity of propagation, ν, the partial discharge site can be located. The location, ℓ, will be given by

$$\ell = \tfrac{1}{2}(L - \nu D)$$

Some of the generalized cross-correlation techniques available can be adapted to autocorrelation processing. In this case the correlator would use the form of the generalized correlator that utilizes the weighting function, $W(\omega)$. The inputs x(t) and y(t) would be the same signal and would come from a single receiver. In this case, the interfering signals, due to reflections, would be necessary to determine the relative delay. The correct delay would be the second largest peak in the correlation function, (since the autocorrelation function, by definition, always has its maximum at the origin). The location, ℓ, will be given by

$$\ell = L - \frac{\nu D}{2}$$

A few of the more useful approaches are mentioned below. The Hannan-Thompson processor, (HT), is the maximum likelihood implementation of the generalized cross-correlator [39]. The HT processor attains the theoretically best performance of any delay estimator, (it achieves the Cramer-Rao lower bound on the variance). The constrained deconvolution processor, (CD), solves the deconvolution problem with the constraint that the output noise be held to a prescribed level or that the gain bandwidth of the weighting be held at a prescribed level. This filter is then used to prefilter the signals before correlating. This approach yields the best resolution of the correlation techniques for a given output SNR. This particular processor is developed in section 3.7 along with a minimax version. The Wiener processor, (WP), makes a best, (linearly constrained),

estimate of the input signals before correlating [38]. The WP approach can be useful in situations where there is a large amount of interference or distortion. There are other generalized cross-correlators but these are the most useful in the present study.

3.4 Bias of Delay Estimates

When processing signals to determine where the sources of the emissions are located one is naturally concerned about the validity of the estimate. For example, assume that the measurements will be made on a cable of length L and that the cable has matched impedances at both ends. A measurement is then made of a signal arriving at both ends of the cable. Furthermore, assume that a perfect parametric model of both the signal and noise are available. The question is how accurately, based on a record of length T, can the position of the signal be located. Typical measures of performance are the variance, bias and probability of error of the delay estimate. These calculations will depend on the parametric model of the signal and the noise and the technique used to determine the delay. Many times these calculations are intractable so certain assumptions and approximations may be necessary.

One factor affecting the accuracy is the error due to the bias of the estimate. Typically, estimates are sought which are either unbiased or asymptotically unbiased, (as the the record length increases the bias goes to zero). Fortunately, for two input delay estimators the bias of the delay estimate is zero in the ideal case. However, when a situation develops in which the peaks are not highly resolved there is the possibility that this lack of resolution will cause the estimate to be biased. The bias is due to the overlapping of adjacent individual correlation functions constituting the entire correlation function. Asymmetry in the differential channel will also contribute to the bias by skewing the correlation functions.

The following discussion on bias is an extension of an analysis given in [37] in which assumptions about the integration length are used as justification for some of the derivation. The purpose of the analysis in [37] was to show that the delay bias was zero but as will be seen below this only happens in special cases. To develop expressions for the bias it is necessary to consider a particular signal model. The measurements from two locations will be modeled as

$$x(t) = \sum_{i=1}^{N} s_i(t) + n_1(t) \tag{3.4.1a}$$

$$y(t) = \sum_{i=1}^{N} h_i(t) * s_i(t - D_i) + n_2(t) \tag{3.4.1b}$$

The signals, s_i, will be considered to be either wide sense stationary or random transient events with relative delays, D_i. If the signals are considered to be wide sense stationary then the expectation will be taken with respect to the appropriate random process. If random transient events are considered then the expectations will be taken with respect to the random amplitude. The transient events will be considered to be entirely contained within the observation interval. If necessary the pulse shapes of the transients could be considered random but this will not be done. The filter, h_i, is a differential filter describing the difference in the channel characteristic due to factors other than the linear delay. These filters will describe the nonlinear phase distortion and the spectral shaping of the signals. The two noise terms, n_1 and n_2 are zero mean, stationary random processes, (or wide sense stationary if desired). The noise terms are independent of each other and independent of the signals, s_i. Depending upon the situation, the signal terms, s_i, may or may not be independent of each other. To estimate the delays, D_i , the function $\phi(\tau)$ is used

$$\phi(\tau) = \int_{-T/2}^{T/2} [\, g_1(t) * x(t)\,]\,[\, g_2(t) * y(t+\tau)\,]\, dt$$

where the filters, $g_1(t)$ and $g_2(t)$, are the optimal filters used in the generalized cross-correlator. For simplicity, these filters will not be included in the following analysis, (it is a simple matter to include them if desired). The generalized cross-correlation prefilters are zero phase and it will be seen later that zero phase filters have no effect on the bias. The estimate of the delays, $\hat{D}_i$, are the values of τ at which $\phi(\tau)$ has its local maximums. This function consists of two terms

$$\phi(\tau) = \phi_S(\tau) + \phi_N(\tau) \tag{3.4.2}$$

The first term is comprised of the cross products of the signals and is given by

$$\phi_S(\tau) = \int_{-T/2}^{T/2} \left\{ \left[\sum_{i=1}^{N} s_i(t) \right] \cdot \left[\sum_{j=1}^{N} h_j(t) * s_j(t+\tau-D_j) \right] \right\} dt \quad (3.4.3a)$$

The second term is comprised of cross products of the signals and and noises and is given by

$$\phi_N(\tau) = \int_{-T/2}^{T/2} n_1(t) \left[\sum_{i=1}^{N} s_i(t+\tau) \right] dt \quad (3.4.3b)$$

$$+ \int_{-T/2}^{T/2} n_2(t) \left[\sum_{i=1}^{N} h_i(t) * s_i(t+\tau-D_i) \right] dt$$

$$+ \int_{-T/2}^{T/2} n_1(t)\, n_2(t+\tau)\, dt$$

Different assumptions about the signals, s_i, will be made depending on which effect is under consideration. However, for each case the bias of the estimate will be defined as

$$B(\hat{D}) = E\{\hat{D} - D\}$$

where $\hat{D}$ will refer to the estimate and D will refer to the true value and will not be subscripted unless necessary. To arrive at a form amenable to calculation of the bias it will be necessary to express $[\tau - D]$ in terms of the functions in Eq. 3.4.2. To arrive at this form it is necessary to differentiate Eq. 3.4.2 which yields

$$\left(\frac{\partial \phi(\tau)}{\partial \tau} \right) \Bigg|_{\tau=\hat{D}} = \left(\frac{\partial \phi_S(\tau)}{\partial \tau} \right) \Bigg|_{\tau=\hat{D}} + \left(\frac{\partial \phi_N(\tau)}{\partial \tau} \right) \Bigg|_{\tau=\hat{D}} \qquad (3.4.4)$$

$$= 0$$

Expanding $\phi_S(\tau)$ into a Taylor series about the true delay, D, yields

$$\phi_S(\tau) = \sum_{n=0}^{\infty} \frac{\phi_S^{(n)}(D)}{n!} [\tau - D]^n \qquad (3.4.5)$$

where $\phi_S^{(n)}(D)$ refers to the n^{th} derivative of ϕ_S with respect to τ evaluated at the delay D. Taking the derivative of Eq. 3.4.5 and inserting it into Eq. 3.4.4 yields

$$\phi_N'(\tau) + \phi_S'(D) = - \sum_{n=1}^{\infty} \frac{\phi_S^{(n+1)}(D)}{n!} [\tau - D]^{n+1}$$

To arrive at the desired relation it is necessary to use a reversion of series [75] which yields

$$[\tau - D] = \sum_{n=1}^{\infty} A_n [\phi_N'(\tau) + \phi_S'(D)]^n$$

where the coefficients A_n are the standard coefficients in a reversion of series. Retaining only the terms in the Taylor series up to the second derivative and performing the reversion yields

$$(\hat{D} - D) = - \frac{\phi_S'(D) + \phi_N'(\hat{D})}{\phi_S''(D)} \qquad (3.4.6)$$

where τ was evaluated at $\hat{D}$. This analysis is a local analysis of the bias whose validity holds near the point of expansion. If the range of validity needs to be extended higher order terms in the Taylor expansion can be retained however the analysis is extremely tedious. Taking the expectation of Eq. 3.4.6 results in

$$E\{\hat{D}-D\} = E\left\{\frac{\phi_S'(D)}{\phi_S''(D)}\right\} + E\left\{\frac{\phi_N'(\hat{D})}{\phi_S''(D)}\right\} \tag{3.4.7}$$

where the second expectation evaluates to zero. This reduces the bias to the expression

$$B(\hat{D}) = E\left\{\frac{\phi_S'(D)}{\phi_S''(D)}\right\}$$

This expectation is still difficult to evaluate so another Taylor expansion will be used. The functions, $\phi_S'(D)$ and $\phi_S''(D)$ are both random variables, defining these as

$$V = \phi_S'(D)$$

and

$$Z = \phi_S''(D)$$

These random variables will have mean values

$$E\{V\} = \mu_v = -T\left.\frac{\partial R_{SS}(\tau)}{\partial \tau}\right|_{\tau = D}$$

and

$$E\{Z\} = \mu_z = -T\left.\frac{\partial^2 R_{SS}(\tau)}{\partial \tau^2}\right|_{\tau = D}$$

Expanding the function

$$F(V,Z) = \frac{V}{Z}$$

into a Taylor series about the mean values μ_v and μ_z ,[6], retaining only the terms up to and including the second order partial derivatives, yields

$$F(V,Z) \simeq \frac{\mu_v}{\mu_z} + \frac{1}{\mu_z} V + \frac{\mu_v}{\mu_z^3} Z^2 + \frac{1}{\mu_z^2} V Z \tag{3.4.8}$$

Taking the expected value of Eq. 3.4.8 and inserting the expressions in for V and Z gives

$$E\left\{ \frac{\phi_S'(D)}{\phi_S''(D)} \right\} \simeq 2 \frac{\mu_v}{\mu_z} \tag{3.4.9}$$

$$+ \frac{\mu_v}{\mu_z^3} \frac{\partial^4}{\partial \tau^4} \int_{-T/2}^{T/2} dx \int_{-T/2}^{T/2} dt\, E\{ s(t)s(t+\tau-D)s(x)s(x+\tau-D) \} \Big|_{\tau=D}$$

$$+ \frac{1}{\mu_z^2} \frac{\partial^3}{\partial \tau^3} \int_{-T/2}^{T/2} dx \int_{-T/2}^{T/2} dt\, E\{ s(t)s(t+\tau-D)s(x)s(x+\tau-D) \} \Big|_{\tau=D}$$

To simplify these expressions any further will involve the calculation of the fourth order moments inside the integrals which can be calculated if specific models are considered.

The first case to be considered is the bias due to the measurement itself. To separate the individual effects the only signal that will be used in this case will be $s_1(t)$. Furthermore, the filter, $h_1(t)$ will be assumed to be zero phase so that only the spectral magnitude is altered. In most analysis the first derivative of ϕ_S is approximately zero because ϕ_S is assumed to be equal to R_{SS}, the autocorrelation function, for large record lengths, T. The bias is also zero for the shorter record lengths due to the symmetry of the moments used in the calculations. In fact one can continue the Taylor expansion and the resulting terms will either have an odd number of derivatives with respect to τ or be multiplied by μ_v which evaluates to zero. Only taking an odd number of derivatives will give a result of zero provided that the higher order moments are symmetric about the true delay, D. The filter has no effect because it is a zero phase filter and will always be an even function.

The second case to be considered is the bias due to nonlinear phase distortion in the channel. This distortion can come from a variety of sources and in a practical sense can never be eliminated. One example is the nonlinear phase of the measuring devices; it is difficult to exactly match the two receivers which will give a small nonlinear distortion. Another example is the dispersion in the propagation path which in the case of a cable is a small effect. Another example is given by reflected waves which reflect from discontinuities that have complex impedances. To separate the individual affects the only signal that will be used in this case will be $s_1(t)$. In most analysis the first derivative of ϕ_S is approximately zero. This is due to the fact that ϕ_S is an even function about the delay D. However, in nonlinear phase systems ϕ_S is no longer symmetric about D and consequently the derivative is no longer zero. It should be noted that the definition of the delay, D, for this analysis refers to the delay due to the linear phase shift only. One of the purposes of including $h_i(t)$ in the $y(t)$ channel was to easily include the effect of a nonsymmetric correlation function. The filter, $h(t)$,is a filter describing the difference between the two channels and can be decomposed into its even and odd components.

$$h(t) = h_e(t) + h_o(t)$$

$$= \frac{h(t) + h(-t)}{2} + \frac{h(t) - h(-t)}{2}$$

Considering only the first term of Eq. 3.4.9 and expressing it in terms of its Fourier transform gives

$$\frac{\mu_v}{\mu_z} = \frac{\int_{-\infty}^{\infty} \omega H(\omega) S_S(\omega) d\omega}{\int_{-\infty}^{\infty} \omega^2 H(\omega) S_S(\omega) d\omega}$$

$$= \frac{\int_0^\infty -j\,\omega H_o(\omega) S(\omega)\, d\omega}{\int_0^\infty \omega^2 H_e(\omega) S(\omega)\, d\omega}$$

It should be noted that H_o is purely imaginary and as a consequence the term $-j$ cancels. The magnitude of this bias is a function of the degree of asymmetry of the correlation function about its true linear phase delay, D.

The third case will develop the bias error due to overlapping correlation functions. In this case the signal will consist of the sum of two individual signals with relative delays, D_1 and D_2. The true delay being sought will be D_1. The correlation functions due to each individual signal will be assumed to be symmetric about its corresponding delay. This case is slightly different than the preceding cases and the change is incorporated by modifying Eq. 3.4.6 to be

$$(\hat{D} - D) = - \frac{\phi'_{S_1}(D_1) + \phi'_{S_2}(\hat{D}_1) + \phi'_N(\hat{D}_1)}{\phi''_S(D_1)}$$

A similar equation can be developed that replaces Eq. 3.4.9 and it's first terms are given by

$$E\left\{ \frac{\phi'_{S_1}(D_1) + \phi'_{S_2}(\hat{D}_1)}{\phi''_{S_1}(D_1)} \right\} \simeq 2\,\frac{\mu_V + \mu_I}{\mu_Z} \qquad (3.4.10)$$

where the terms μ_V and μ_Z are defined as above. The term μ_I is the interfering term and is given by the expression

$$\mu_I = - T \left. \frac{\partial R_{S_2}(\tau)}{\partial \tau} \right|_{\tau = \hat{D}_1}$$

Using Eq 3.4.10 the delay bias can be expressed as

$$2\,\frac{\mu_V + \mu_I}{\mu_Z} = \frac{2\int_0^\infty -j\,\omega\,S_2^o(\omega)\,d\omega}{\int_0^\infty \omega^2\,S_1(\omega)\,d\omega}$$

where $S_2^o(\omega)$ is the odd portion of the power spectral density of $s_2(t)$ shifted by the delay term

$$\exp[-j\,\omega\,(\hat{D}_1 - D_2)]$$

If the correlation function has a finite duration less than the delay difference, $D_1 - D_2$, then as the delay difference becomes larger the first and second derivatives of the correlation function at that point approach zero. In other words as the functions become more separated the bias due to the overlap goes to zero. This result can also be expressed in the time domain in terms of the autocorrelation functions R_1 and R_2 of the signals s_1 and s_2. The delay bias for overlapping correlation functions is given by

$$B(\hat{D}_1) = \frac{-2R_2'(D_2 - \hat{D}_1)}{R_1''(0)}$$

The interpretation of this bias is that the interfering correlation function tilts the desired correlation in such a manner that it's maximum is shifted. The tilt is affected by the slope of the interfering term at the location of the desired peak. Other factors affecting the bias are the bandwidth and power of the desired correlation function. The term in the denominator is the rms bandwidth squared (for low pass signals) times the power in the correlation function. If the correlation function has a large power or has a large bandwidth then the bias is reduced.

This bias expression has other practical implications other than giving the bias for overlapping correlation functions from adjacent (but desired) correlation functions. If a narrow band signal is interfering with the measurements then the resulting correlation function will be the superposition of the desired correlation functions and a damped sinusoid. Depending on the exact relationship between the phase of the sinusoid and

the desired peaks there can be significant bias in the desired estimates due the superposition. Another problem can be caused by low frequency high energy impulsive interference. This is a common interfering signal in partial discharge measurements and is caused by SCR noise, switching surges and the like. The superposition of these pulses into the measurement has an adverse effect because the correlation time of this interference is typically in excess of the total propagation time of the device under test. In other words the presence of this interference will bias all the delay measurements.

The next case will be concerned with the inherent bias in autocorrelation processing. The above analysis assumed that the noise in the two channels was independent. In contrast now the noise terms, n_1 and n_2 will be assumed to be equal. Furthermore, the signal model has to be adjusted so that reflections of the signal are included. Making these changes gives the new model

$$x(t) = \sum_{i=1}^{N} s_i(t-D_i) + n_1(t)$$

$$y(t) = \sum_{i=1}^{N} s_i(t-D_i) + n_1(t)$$

The new terms for the correlation output are

$$\phi_S(\tau) = \int_{-T/2}^{T/2} \left\{ \left[\sum_{i=1}^{N} s_i(t-D_i) \right] \cdot \left[\sum_{j=1}^{N} s_j(t+\tau-D_j) \right] \right\} dt$$

and

$$\phi_N(\tau) = \int_{-T/2}^{T/2} n_1(t) \left[\sum_{i=1}^{N} s_i(t+\tau-D) \right] dt$$

$$+\int_{-T/2}^{T/2} n_1(t)\left[\sum_{i=1}^{N} s_i(t+\tau-D_i)\right]dt$$

$$+\int_{-T/2}^{T/2} n_1(t)n_1(t+\tau)\,dt$$

To simplify the problem and make the effect clear consider the simple case in which the observed signal consists of only the direct signal and one reflection. Furthermore, assume that the channel is distortion free and that the reflected signal is delayed from the direct signal by a delay D. The signal model then becomes

$$x(t) = s(t) + s(t-D) + n(t)$$

In this case the autocorrelation function $R(\tau)$ is given by

$$R(\tau) = 2R_{SS}(\tau) + R_{SS}(\tau - D) + R_{SS}(\tau + D) + R_{NN}(\tau)$$

where R_{SS} and R_{NN} are the autocorrelation function of the signal and noise. In essence this problem is identical to the problem of overlapping echos discussed above. Suppose that interest is focused on the peak located at D, then the other terms are overlapping pulses and comprise the interfering term $I(\tau)$ where

$$I(\tau) = 2R_{SS}(\tau) + R_{SS}(\tau + D) + R_{NN}(\tau)$$

The same solution applies as for overlapping correlation functions and in this case the bias in the time domain is given by

$$B(\hat{D}) = \frac{2I'(\hat{D})}{R''_{SS}(0)}$$

This bias can cause difficulties in the location of the partial discharge site if the delay, D, is small. Consider for example the case of an underground

power cable. If the received signal has a bandwidth of 4 MHz then if D is less than 250 ns then the estimate can have significant bias. This correspond to a partial discharge site located within 125 feet of either end of a typical 1500 foot URD cable. In fact, if located too close to the end of the cable the desired secondary peak in the cable may not be discernible.

The analysis given in this section was concerned with the bias errors associated with correlation analysis. A similar analysis can be carried out for time of arrival measurement schemes for determining the bias. The results are the same as given above except that the autocorrelation functions (spectral densities) are replaced by the signal (magnitude spectrum). In other words the same errors associated with the time delay bias estimates obtained from correlation functions will affect time of arrival methods.

3.5 Variance of Delay Estimates

Another commonly used measure of the performance of an estimate is its variance which is defined as

$$\mathrm{Var}[\,\hat{D}\,] = E\{\,(\,\hat{D} - D\,)^2\,\} - \left[E\{\,\hat{D} - D\,\}\right]^2$$

This is a difficult quantity to calculate so a similar approach will be used as for the bias. However, before calculating the variance, it is interesting to ask what is the best possible performance that could be attained. This question is answered in terms of a lower bound on the variance of all possible delay estimators. To obtain a bound it is necessary to consider specific signal and noise models and for this type of problem the only tractable model is the Gaussian model. The bound is referred to as the Cramer-Rao lower bound [41] and is given by

$$\sigma_{\hat{D}}^2 \geqq \left. \frac{-1}{E\left\{ \dfrac{\partial^2 \ell n p(\mathbf{X} \mid Q, \tau)}{\partial \tau^2} \right\}} \right|_{\tau = D}$$

where $\mathbf{X}$ is the observation vector containing both x(t) and y(t), Q is spectral density matrix containing all the spectral densities describing the

Gaussian probability density function $p(X \mid Q, \tau)$. For Gaussian signals and noise this quantity is calculated in [39] and is given by

$$\text{minimum var}[\,\hat{D}\,] = \left[T \int_{-\infty}^{\infty} (2\pi f)^2 \frac{|\gamma_{12}(f)|^2}{1 - |\gamma_{12}(f)|^2}\, df \right]^{-1}$$

where $|\gamma_{12}(f)|^2$ is the magnitude squared coherence function. This bound is achieved by the ML generalized cross-correlator discussed in [37]. In general, bounds like this simply state the best performance and give no indication of how to achieve them.

For the general wide sense stationary case or for random transient pulses contained within the observation interval the following analysis gives the local variance of the delay estimate, $\hat{D}$. Only the unbiased case will be considered so using Eq. 3.4.6 as a starting point the variance is given by

$$\text{var}(\,\hat{D}\,) = E\left\{ \left(\frac{\phi_S'(D) + \phi_N'(\hat{D})}{\phi_S''(D)} \right)^2 \right\}$$

This can be treated as function of two variables V and Z

$$F(V,Z) = \frac{V^2}{Z^2}$$

where V and Z are given by

$$V = \phi_S'(D) + \phi_N'(\hat{D})$$

and

$$Z = \phi_S''(D)$$

Following the same line of reasoning as in section 3.4 a Taylor series expansion about the mean values μ_V and μ_Z will be used. Retaining all terms up to the second partial derivatives and deleting the terms whose expectations evaluate to zero yields

$$F(V,Z) \simeq \frac{V^2}{\mu_Z^2}$$

where

$$\mu_V = 0$$

and

$$\mu_Z = E\{\phi_S''(D)\}$$

This is the same result given in [37] except that the random variation of the signal term was not ignored in this analysis. The analysis given in [37] assumes a large observation time, T, and argues that the random variations of the signal can be ignored; this analysis does not make that assumption. The variance can be shown to be [37]

$$\text{var}[\,\hat{D}\,] = \frac{1}{T} \frac{1}{\left[\int_{-\infty}^{\infty} \omega^2 S(\omega)\, d\omega\right]^2} \cdot \left[\int_{-\infty}^{\infty} \omega^2 \left[S(\omega) N_1(\omega) + S(\omega) N_2(\omega) + N_1(\omega) N_2(\omega)\right] d\omega\right]$$

where $S(\omega)$, $N_1(\omega)$ and $N_2(\omega)$ are the spectral densities of the signal and noise respectively. If the generalized cross-correlator is used then the variance becomes

$$\text{var}[\,\hat{D}\,] = \frac{1}{T}\,\frac{1}{\left[\int_{-\infty}^{\infty} \omega^2\, W(\omega)\, S(\omega)\, d\omega\right]^2}$$

$$\cdot\left[\int_{-\infty}^{\infty} \omega^2\, W^2(\omega)\,\left[S(\omega)\,N_1(\omega) + S(\omega)\,N_2(\omega) + N_1(\omega)\,N_2(\omega)\right] d\omega\right]$$

where $W(\omega)$ is the optimal weighting function for the generalized cross-correlator's implementation.

Since a popular method of locating partial discharge is to use the peak value of the received waveforms a similar analysis will be carried out for this case. Consider the two received signals that are modeled as

$$x(t) = s_1(t - D_1) + n_1(t)$$

and

$$y(t) = s_2(t - D_2) + n_2(t)$$

where it will be assumed that the delay estimate will be made from transient type signals that are corrupted by additive wide sense stationary noise. The signals may be different from each other because of frequency dependent attenuation but the phases will be assumed to be undistorted. The desired delay difference is $D_1 - D_2$ and the variance of the delay estimate is given by $\text{var}[\hat{D}_1 - \hat{D}_2]$. Using the same reasoning that resulted in Eq. 3.4.6 the delay difference $\hat{D}_1 - D_1$ is given by

$$\hat{D}_1 - D_1 = \frac{-\,n_1'(\hat{D}_1)}{s_1''(0)}$$

where the signal $s_1(t - D_{1)}$ has it's maximum at $t = D_1$. The variance is found by calculating the expected value of

$$\text{var}[\hat{D}_1 - \hat{D}_2] = E\{(\hat{D}_1 - D_2)(\hat{D}_2 - D_2)\}$$

Expanding this expression gives

$$\text{var}[\hat{D}_1 - \hat{D}_2] = E\left\{\left[\frac{-n_1'(\hat{D}_1)}{s_1''(0)}\right]^2 + \left[\frac{-n_2'(\hat{D}_1)}{s_2''(0)}\right]^2\right\}$$

where the cross terms were ignored because the noise terms $n_1(t)$ and $n_2(t)$ are considered to be independent. The resulting variance expression is then given by

$$\text{var}[\hat{D}_1 - \hat{D}_2] = \frac{-R_1''(0)}{[s_1''(0)]^2} + \frac{-R_2''(0)}{[s_2''(0)]^2}$$

where the terms R_1 and R_2 refer to the autocorrelation functions of the noise $n_1(t)$ and $n_2(t)$ respectively. The important difference to notice between the correlation based delay estimator and the peak location based delay estimator is the scaling factor $\frac{1}{T}$. As the observation interval is increased the variance of the correlation based estimator decreases. There is no method by which the peak location estimator can be improved. In essence what this result says is that collecting additional data will improve the cross-correlation based estimate substantially. However, it should be noted that excellent results can be obtained from a single record with one partial discharge pulse of sufficient energy using cross-correlation. A typical partial discharge with a received bandwidth of 10 MHz and a charge of 0.025 pC will have an ENR of 14 dB for a record length of 2.56 μs when measured with an amplifier with an 8 dB noise figure. In this particular case the delay estimate has a standard deviation of about 7 ns.

3.6 Modified PHAT Processor

In situations where multipath signals are present it is necessary to increase the resolution of the resulting correlation functions. The following GCC is an ad hoc type of filtering which increases the bandwidth of the measurement [51]. The following filter is based upon ML and PHAT [39] generalized cross-correlators. The maximum likelihood, ML, implementation [39,18] of the generalized cross-correlator uses a weighting function of the form

$$W(\omega) = \frac{1}{|S_{XY}(\omega)|} K(\omega)$$

$$K(\omega) = \frac{C(\omega)}{1 - C(\omega)}$$

$$C(\omega) = \frac{|S_{XY}(\omega)|^2}{S_{XX}(\omega) S_{YY}(\omega)}$$

where $C(\omega)$ is the magnitude squared coherence function. As can be seen from the above equation, the first operation involved in the weighting is to set the cross-spectral magnitude equal to one. This is accomplished by dividing the cross-spectral density by its magnitude and if only this operation is performed then the GCC is referred to as the PHAT processor. If only this operation is performed, (assuming perfect knowledge of the cross-spectral density), the correlation function would have a single impulse at the delay D. In practice this leads to a correlation function which consists of a random collection of impulses with little relation to the true value D. The basic problem with this technique is that only estimates of the cross-spectral density are available. Furthermore, those frequencies where the cross-spectral estimates are poor have been given the same weight as those regions where the estimate is good. The ML weighting attenuates those frequencies at which the estimate of the phase of the cross-spectrum is poor while accentuating those frequencies where the estimate is good.

To interpret the ML weighting function, $K(\omega)$ will be rewritten as

$$K(\omega) = \frac{SNR_X(\omega)\, SNR_Y(\omega)}{SNR_X(\omega) + SNR_Y(\omega) + 1}$$

The shaping of the cross-spectrum provided by $K(\omega)$ can be described by considering the cases where there is large amount of signal energy at the frequencies of interest and where there is not. When there is a large amount of signal energy the SNR is high and $K(\omega)$ weights the cross-spectrum according to the combined SNR in both channels. If the noise in each channel has a constant spectral density across those frequencies then the cross-spectrum essentially retains its original shape. For frequencies where there is little signal energy and the SNR is low the weighting, $K(\omega)$, attenuates the cross spectrum as the product of the SNR in each channel.

The acoustic measurements that were made in this study used transducers that were narrow band and consequently the time resolution of the cross-correlation function is poor. The time resolution refers to the ability to unambiguously resolve individual peaks in the correlation function. In the structures studied, multiple propagation paths were present and consequently the correlation functions exhibit multiple peaks. It is therefore necessary to maintain an adequate time resolution in order to identify the individual peaks and measure the correct delay. Unfortunately, when using narrow band transducers, the resulting resolution of the measurements using a ML weighting are not adequate to separate the multiple peaks that are usually present in these measurements. The reason for this problem is that as mentioned above the ML weighting essentially retains the spectral shape at those frequencies where the SNR is high.

Using only the phase information by setting the magnitude of the cross-spectrum equal to one would provide the greatest time resolution, but as mentioned above this approach has other deficiencies. However, if the frequency range where the magnitude is set equal to one is limited to those frequencies where the SNR is high enough, then there is a resulting increase in resolution without the noise problems that cause the spurious peaks. The question remains how to choose those frequencies where the signal energy is large enough to prevent the occurrence of the spurious peaks. Intuitively it makes sense to use some measure of the SNR at the various frequencies to decide which frequencies to include. The function, $K(\omega)$, described above provides a convenient measure on which to base the decisions. So as a purely ad hoc method the weighting function, $W(\omega)$, is used to filter the

waveforms [53] and is given by the expression

$$W(\omega) = \begin{cases} \dfrac{\gamma}{|S_{XY}(\omega)|} & \dfrac{SNR_X(\omega)\, SNR_Y(\omega)}{SNR_X(\omega) + SNR_Y(\omega) + 1} \leqq \gamma \\ \dfrac{K(\omega)}{|S_{XY}(\omega)|} & \dfrac{SNR_X(\omega)\, SNR_Y(\omega)}{SNR_X(\omega) + SNR_Y(\omega) + 1} > \gamma \end{cases}$$

This can also be written more compactly as

$$W(\omega) = \frac{1}{|S_{XY}(\omega)|} \left\{ \tfrac{1}{2}\, \gamma\, [Sgn(K(\omega) - \gamma)+1] + \tfrac{1}{2}\, K(\omega)|\, Sgn(\gamma - K(\omega)) + 1] \right\}$$

When this weighting is used there is a resulting increase in resolution without the spurious spikes in the time domain. This weighting replaces the portion of the spectrum having a high SNR with a constant, γ, and the low SNR portion with the same weighting as used by the ML processor. A typical example of the cross-spectral shaping provided by the ML filter and the filter discussed above are shown in Fig. 3.2 and Fig. 3.3.

3.7 Robust Deconvolution of Correlation Functions

A generalized cross-correlator is discussed in this section which can provide improved resolution in delay estimation measurements [52]. The prefilters for this GCC are robust (in the minimax sense) linear filters which have found use in linear equalization, [40]. The presence of the prefilters, in effect, deconvolves the cross-correlation function. This GCC is used primarily for locating acoustic emissions from partial discharge in high voltage, solid cast transformers. Both the electrical and acoustic emissions are used in the time delay estimation technique for locating the emission sites. Improved resolution is needed in power transformers because of the complicated multipath signals that are present in the transformer winding and this GCC is effective in combating the problems associated with the

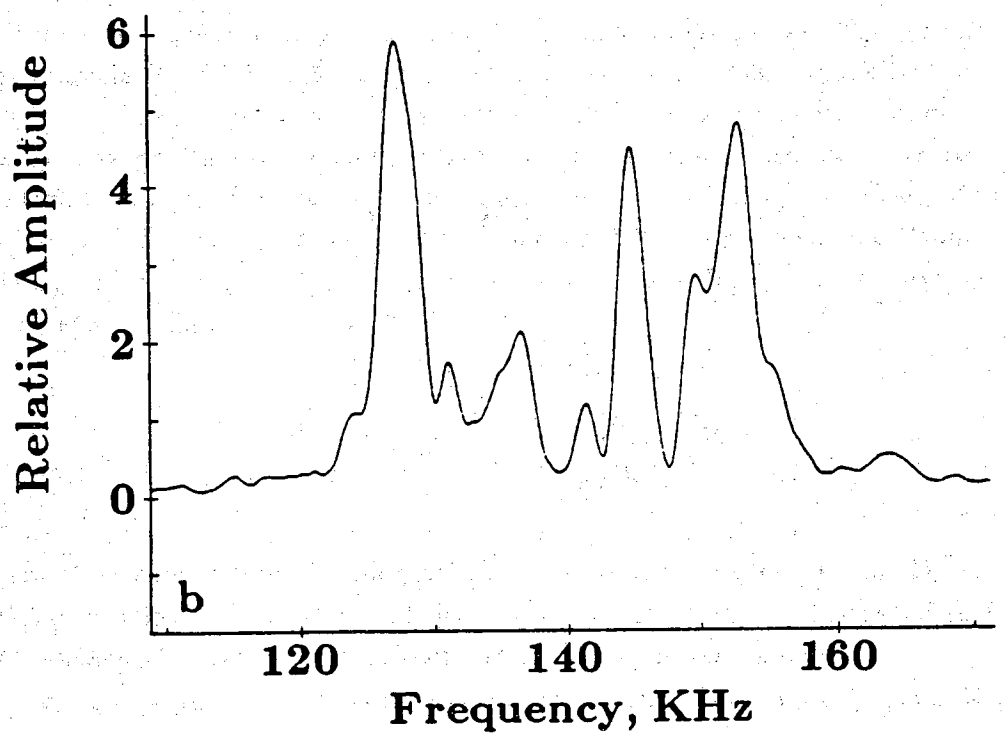


Figure 3.2. Output ESD using the ML GCC weighting.

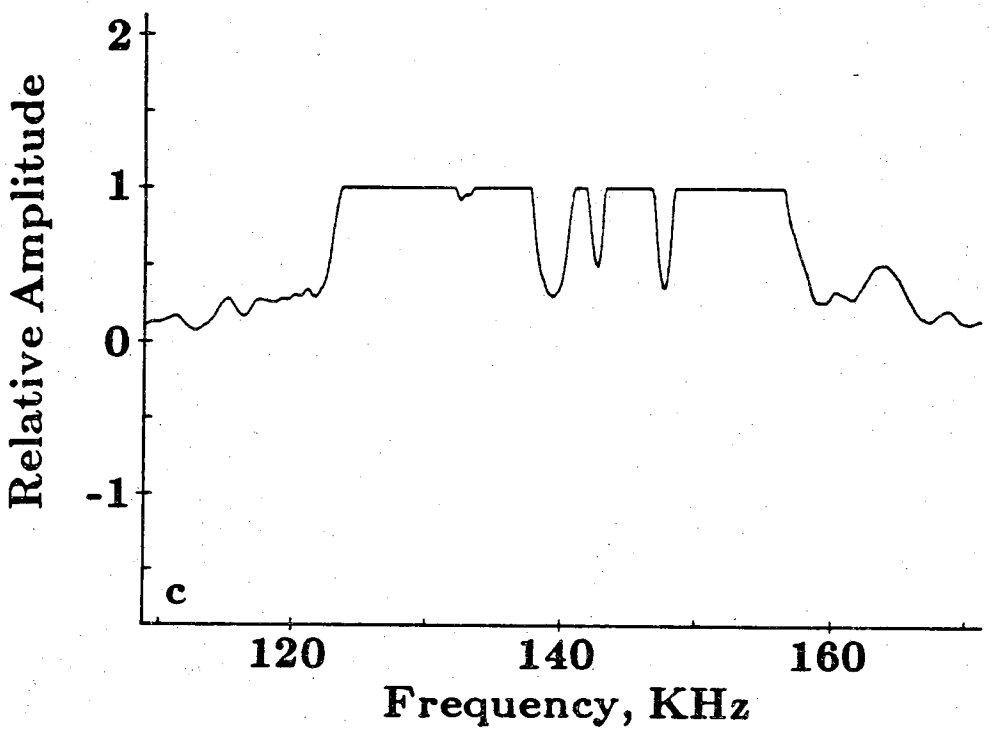


Figure 3.3. Output ESD using the modified PHAT GCC weighting.