Processing of Random Signals

by

William K. George, Jr.
State University of New York at Buffalo

with Appendices by

Paul D. Beuther
State University of New York at Buffalo
and
John L. Lumley
Cornell University

Introduction

The experimentalist who is involved in dynamic flow measurement is almost always interested in some type of random signal processing. This interest can arise because the flow itself is random, as in turbulent flow. Or it can arise from concern about whether measured data represent real flow phenomena or merely statistical fluctuations in the data. These interests can occur simultaneously. For example, in the problem of determining periodicities in the turbulent flow behind a moving blade row the question of distinguishing between fluctuations in the data and the true spectral peaks is crucial to the experiment.

In this review article, attention will be confined to the problem of measuring and interpreting statistics (to the second moment) of a stationary random signal. After an introduction to the nature of stationary processes, we shall review the great strides of the last two decades in the processing of discretely sampled signals. The advances of the last decade in the understanding of randomly sampled data will be reviewed and these techniques will be compared to the even more recent advances in understanding the burst-processor laser-Doppler anemometer. Finally, some new ideas for measurement using both continuous and random sampling will be introduced.

No attempt will be made to be comprehensive, either in the sense of covering all topics of interest or in the sense of providing detailed references to the material. Rather, the approach adopted is tutorial and is intended to both in-

troduce the material and provide a framework for subsequent learning. This article should be viewed in the context of the other articles in this volume which also address topics relevant to signal processing and interpretation — especially the review articles by Buchhave, Kovasznay and Van Atta.

The historical roots of this presentation are many, but particular credit must be given to the excellent articles on the subject found in references (1) - (5). In no sense should this article be viewed as a substitute for careful reading of these and the other references.

The Ensemble Average

The concept of an *ensemble average* is familiar in some sense to every layman. To compute such an average we simply add the individual realizations of the process (supposed random) and divide by the number of realizations; that is,

$$X_N = \frac{1}{N} \sum_{i=1}^{N} x_i \qquad\qquad [1]$$

where X_N is the "average" computed on the basis of N realizations and the x_i denote the individual realizations.

By the very act of using such an algorithm, the layman is expressing a belief that a *true average* exists and that the computed number X_N is representative of it. He even has a primitive idea of convergence in that he believes that the more realizations he has, the better will be his estimate of the true value. If $\overline{x}$ denotes the true average, we express this convergence formally as

$$\lim_{N \to \infty} X_N = \overline{x} \qquad\qquad [2]$$

The fact that these concepts are widespread is probably an expression of a basic belief (conditioned or inherited) that there is some underlying order in the universe; in effect, we insist that even random events are orderly.

We can quantify the above ideas by noting that since our estimator X_N given by equation [1] is the sum of random variables, it is itself a random variable. Therefore, we can define the variance of X_N as

$$\text{var}\left\{ X_N \right\} = \overline{(X_N - \overline{x})^2} \qquad\qquad [3]$$

where the overbar denotes the true ensemble average and is assumed to exist. The question of convergency can now be expressed as: Does the variance of estimator, X_N, become vanishingly small as the number of realizations, N, becomes large?

$$\lim_{N \to \infty} \quad \text{var}\{X_N\} \overset{?}{\to} 0 \tag{4}$$

It is easy to show (and is well-known) that if the samples are identically distributed and are statistically independent, then

$$\text{var}\{X_N\} = \frac{1}{N} \text{var}\{x\} \tag{5}$$

where VAR$\{x\}$ is the mean square fluctuation in the ensemble given by

$$\text{var}\{x\} = \overline{(x-\overline{x})^2} = \sigma_x^2 \tag{6}$$

If an acceptable fluctuation in our estimate is given by ϵ where

$$\epsilon^2 \equiv \frac{\text{var}\{X_N\}}{\overline{x}^2} , \tag{7}$$

then for N realizations,

$$\epsilon = \frac{1}{\sqrt{N}} \cdot \left[\frac{\sigma_x}{\overline{x}}\right] \tag{8}$$

Thus the variability of our estimator is proportional to the relative fluctuation of the random variable itself $(\sigma_x/\overline{x})$ and is inversely proportional to the square root of the number of *independent* samples. Clearly the estimator given by equation [1] converges as N $\to\infty$ if the samples are identically distributed and independent.

We have assumed to this point that the estimator converges to the *true* mean. When this is not true the estimator is said to be *biased*. Examples of biased estimators will be introduced later.

Time Series, the Time Average, and the Ergodic Hypothesis

Let us now imagine a continuous signal u(t) which is random. By this we mean both that the time evolution of a single signal can not be predicted by a deterministic function of time nor can the value at a particular time be predicted for an ensemble of realizations. If we further imagine a number of different experiments all evolving under the same physical constraints, we can immediately apply our concept of an ensemble average and speak of the ensemble average at a particular instant in the evolution of the process. The conditional and periodic averages discussed elsewhere in this volume are simply variations on this ensemble average approach.

In many situations in which flow measurements are to be undertaken it is inconvenient, or impossible, to carry out a number of independent experiments. In such situations a time average defined by

$$U_T = \frac{1}{T} \int^T u(t)dt \qquad\qquad [9]$$

is often employed with the hope that U_T will approach the *true* average as the length of the integration interval increases. That is, we hope that

$$\lim_{T \to \infty} U_T = \bar{u} \qquad\qquad [10]$$

It is obvious that this operation of time averaging makes no sense unless the averages themselves are independent of time — at least in the measurement interval. When $\bar{u}$ is time-independent, we say that the process is *statistically stationary* and have some reason to believe that the operation of time averaging might make sense. Our hopes can be summarized in the *ergodic* hypothesis which for our purposes can be stated as:

The Ergodic Hypothesis

Time averages converge to a mean value as the averaging time becomes large. Moreover, the time average always converges to the same value regardless of when the averaging process is initiated. Volumes have been written about the ergodic hypothesis. For the physical scientist, however, it is simply a statement that the world is working as we think it should.

The *ergodic hypothesis* has a number of implications about the nature of random time series. Of particular interest to us are the following:

(i) Random variables must become uncorrelated (and, in fact, statistically independent) at large time delays.
     Symbolically,

$$\overline{u(t)\, u\, (t+\tau)} \to 0 \text{ as } \tau \to \infty \qquad\qquad [11]$$

(ii) An integral scale exists and is a measure of the memory of the process (time to lose correlation).

For our purposes, the integral scale can be defined from the autocorrelation as:

$$T_u = \int_0^\infty \rho_u (\tau)\, d\tau \qquad\qquad [12]$$

where $\rho_u(\tau)$ is the autocorrelation coefficient defined by

$$\sigma_u^2 \; \rho_u(\tau) = \overline{[u(t) - \overline{u}] \; [u(t') - \overline{u}]} = B_u(\tau) \qquad [13]$$

where

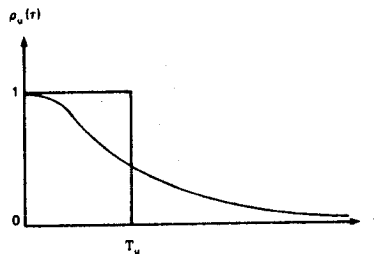$$\sigma_u^2 = \overline{[u(t) - \overline{u}]^2} = \mathrm{var}\{u\} \qquad [14]$$

Fig. 1. Typical autocorrelation showing integral scale from equation 12

From Fig. 1, T_u is seen to be the intersection with the time delay axis of a rectangle having unit height and area equal to that under autocorrelation coefficient curve. The figure makes it clear that T_u is a measure of the time over which the signal is correlated, as required.

We note that other integrals scales can be defined; all of these will be about the same. Also, if we generate a new process from u(t), say F[u(t)], the integral scale of F will be no larger than that of the original process.

The Convergence of Time Averages
We are now prepared to answer the two major questions of random signal processing:

1.   Do my time averages converge to the correct value?
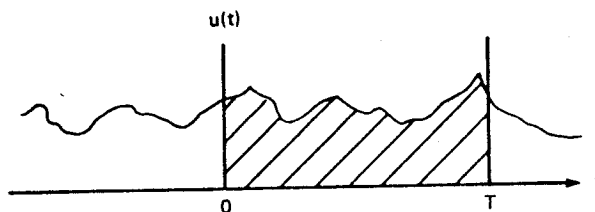2.   At what rate do my time averages converge?

Fig. 2. Illustration that the time average is a random variable (shaded area).

The operation of time averaging which we have defined is illustrated in Fig. 2 by the area bounded by the random signal and the averaging interval. It is clear that our estimator U_T is also a random quantity. The answer to the first question is easily obtained by ensemble averaging U_T. Since the signal is assumed stationary the operations of ensemble averaging and integration commute and we have

$$\overline{U}_T = \frac{1}{T} \int_0^T \overline{u(t)} \, dt = \overline{u}$$

Thus in our previously accepted terminology: U_T is an unbiased estimator of $\overline{u}$. We answer the question of convergence as before by examining the variance of U_T.

$$\text{var}\{U_T\} = \overline{[U_T - \overline{u}]}^2 = \frac{1}{T^2} \iint_0^T \overline{[u(t) - \overline{u}] [u(t') - \overline{u}]} \, dt dt' \qquad [15]$$

The integrand is just the autocorrelation of the signal $u(t)$ given by equation [13]. After a partial integration (see ref. 3) we have

$$\text{var} \{U_T\} = \frac{2 \, \text{var}\{u\}}{T} \int_0^T \rho(\tau) [1 - \frac{\tau}{T}] \, d\tau \qquad [16]$$

Since $\rho \rightarrow 0$ as T becomes large, this reduces to

$$\text{var}\{U_T\} \cong \frac{2 \, \text{var}\{u\}}{T} \int_0^T \rho_u(\tau) \, d\tau = \frac{2 \, T_u}{T} \, \text{var}\{u\} \qquad [17]$$

The relative error in the estimator is then given by

$$\epsilon^2 = \frac{\text{var}\{U_T\}}{\overline{u}^2} \cong \frac{2T_u}{T} \frac{\text{var}\{u\}}{\overline{u}^2} \quad [18] \quad \text{or} \quad \epsilon = \sqrt{\frac{2T_u}{T}} \cdot \left[\frac{\sigma_u}{\overline{u}} \right] \qquad [19]$$

The analogy between equations [19] and [8] is obvious. In effect, the number of independent samples in our average is

$$N = \frac{T}{2T_u} \qquad [20]$$

Thus segments of our time record of two integral scales in length contribute to the average as though they were statistically independent. This is illustrated in Fig. 3.
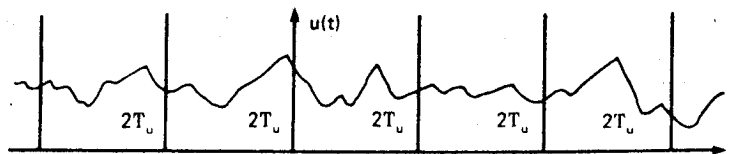

Fig. 3. Segments of stationary random process illustrating pieces of the record

The results given in equations [19] and [20] occur over and over again in the analysis of time series. We only assumed that the signal u(t) was stationary.

Another example which leads to a similar result is the time average estimator for the autocorrelation defined by

$$B_T(\tau) = \frac{1}{T} \int_0^{T-|\tau|} u(t)u(t+\tau)\, dt \qquad [21]$$

Note at the outset that this estimator is biased since

$$\overline{B_T(\tau)} = \frac{1}{T} \int_0^{T-|\tau|} \overline{u(t)u/t+\tau)}\, dt = \frac{T-|\tau|}{T}\, B(\tau) \qquad [22]$$

Clearly this bias could have been avoided by dividing by $T-|\tau|$ instead of T. It will be seen in the following paragraphs that the unbiased estimator does not converge for large time lags and is therefore unsuitable; hence, the widespread use of the biased estimator.

The variance of our estimator(s) can be computed in the following manner:

$$\text{var}\{B_T\} = \overline{[B_T(\tau) - B(\tau)]^2} = \frac{1}{T^2} \iint_0^{T-|\tau|} \overline{u'(t)u'(t+\tau)u'(t_1)u'(t_1+\tau)}\, dt\, dt_1 \qquad [23]$$

where the primes denote only the fluctuating values. By assuming that the fourth order moments of u(t) are jointly Gaussian

$$\overline{uu'u''u'''} = \overline{uu'}\cdot\overline{u''u'''} + \overline{uu''}\cdot\overline{u'u'''} + \overline{uu'''}\cdot\overline{u'u''} \qquad [24]$$

and modelling the autocorrelation of u(t) as a simple exponential

$$\rho_u(\tau) = \exp\{-|\tau|/T_u\} \qquad [25]$$

it can be shown that (see ref. 1)

$$\text{var}\{B_T\} \cong \frac{2T_u}{T}\, \text{var}\{u\} \qquad [26]$$

The relative error is then

$$\epsilon^2 = \frac{\text{var}\{B_T\}}{B^2(\tau)} = \frac{2T_u}{T} \cdot \left[\frac{1}{\rho_u(\tau)}\right]^2 \qquad [27]$$

since var $\{u\}$ = B(o). Thus the relative accuracy to which we have determined the autocorrelation with this estimator decreases with increasing time lag since the autocorrelation goes to zero. Alternately, the larger the lag at which we want to compute the autocorrelation, the longer we must average.

A similar derivation reveals the reason for abandoning the unbiased estimator (see ref. 1). The occurrence of a term $T-|\tau|$ in the denominator of equation [27] increases the relative error with time lag. As long as the largest time lag of interest is much less than the averaging time, there is no significant difference between the biased and unbiased estimators.

Spectral Estimates From Finite Record Lengths

We first consider the problem of Fourier decomposing a stationary random signal (of infinite record length). Originally this subject was approached through the use of Fourier-Stieltjes integrals which, to the uninitiated, appeared strange and complicated. The subject has now been considerably simplified, thanks to the introduction of generalized functions into the Fourier analysis. For a complete review, the interested reader is referred to the comprehensive monograph by Lumley (ref. 4).

The Fourier transform of a signal u(t) in the usual sense is defined by

$$\hat{u}(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i\omega t} \, u(t)dt \qquad [28]$$

For this transform to exist, u(t) must satisfy a number of conditions including smoothness and vanishing at infinity (see any standard calculus text). The inverse transform is given by

$$u(t) = \int_{-\infty}^{\infty} e^{-i\omega t} \, \hat{u}(\omega) \, d\omega \qquad [29]$$

Stationary random functions of time do not, in general, satisfy the necessary conditions for their Fourier transform to exist in the above sense. However, if one agrees to work in the domain of generalized functions so that our signal of infinite length is approached as the limit of a sequence of functions whose Fourier transforms exist, then we can define the Fourier transform of the stationary random signal as the limit of the Fourier transforms of the members of the sequence. Under a set of conditions which need not concern us here, this limit can be assumed to exist. Hence, we can write the Fourier transform of u(t) as in equation [28] and denote it by $\hat{u}(\omega)$, if we agree to say that this is the Fourier transform of u(t) *in the sense of generalized functions.* This seems a small price to pay since once we have made this qualification we can treat $\hat{u}(\omega)$ as though it were an ordinary Fourier transform (for our purposes, at least).

It should be obvious that for any given realization of u(t), there will be a particular realization of $\hat{u}(\omega)$. In other words, if u(t) is random, so must be $\hat{u}(\omega)$

Since u(t) is assumed stationary (eg. $\overline{u(t)\ u(t+\tau)}$ is function of τ only), it is reasonable to expect some constraints on the expected values of the Fourier transforms. It can be shown that a consequence of stationarity is that the Fourier coefficients are uncorrelated at different frequencies; that is

$$\overline{\hat{u}(\omega)\hat{u}(\omega')} = 0 \; ; \; \omega \neq \omega' \tag{30}$$

In fact, we can write

$$\overline{\hat{u}(\omega)\hat{u}(\omega')} = S(\omega)\delta(\omega-\omega') \tag{31}$$

where $S(\omega)$ is the spectrum of the signal u(t) and $\delta(\omega-\omega')$ is the familiar Dirac delta-function.

It is well-known and follows immediately that the spectrum is the Fourier transform of the autocorrelation

$$S(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i\omega\tau}\ B(\tau)\ d\tau \tag{32}$$

and that the autocorrelation is the inverse Fourier transform of the spectrum

$$B(\tau) = \int_{-\infty}^{\infty} e^{-i\omega\tau}\ S(\omega)d\omega \tag{33}$$

By evaluating the last equation at $\tau=0$ we have

$$B(0) = \sigma^2 = \int_{-\infty}^{\infty} S(\omega)d\omega \tag{34}$$

Thus, as is well known, the integral of the spectrum over all frequencies yields the variance (or mean square) of the signal.

Now that we are assured (in principle, at least) that a spectrum exists, we turn to the problem of its estimation from a finite length of record. Equation [28] suggests that an appropriate estimator for the Fourier transform $\hat{u}(\omega)$ might be

$$\hat{u}_T(\omega) = \frac{1}{2\pi} \int_{-\frac{T}{2}}^{\frac{T}{2}} e^{i\omega\tau}u(t)dt \tag{35}$$

The interval $(-\frac{T}{2}, \frac{T}{2})$ has been chosen instead of (0,T) for symmetry and simplicity with no loss of generality.

That this is a good choice is confirmed by the fact that the average value of the spectral estimator defined by

$$S_T(\omega) = 2\pi \frac{|\hat{u}_T(\omega)|^2}{T} \tag{36}$$

does reproduce the correct value of spectrum as T becomes large. This is easily seen by carrying out the operations implied by equations [35] and [36] to obtain

$$\overline{S_T(\omega)} = \frac{1}{2\pi} \int_{-T}^{T} e^{i\omega\tau} B(\tau) \cdot \left[1 - \frac{|\tau|}{T} \right] d\tau \tag{37}$$

Clearly from equation [37],

$$\lim_{T\to\infty} \overline{S_T(\omega)} = S(\omega) \tag{38}$$
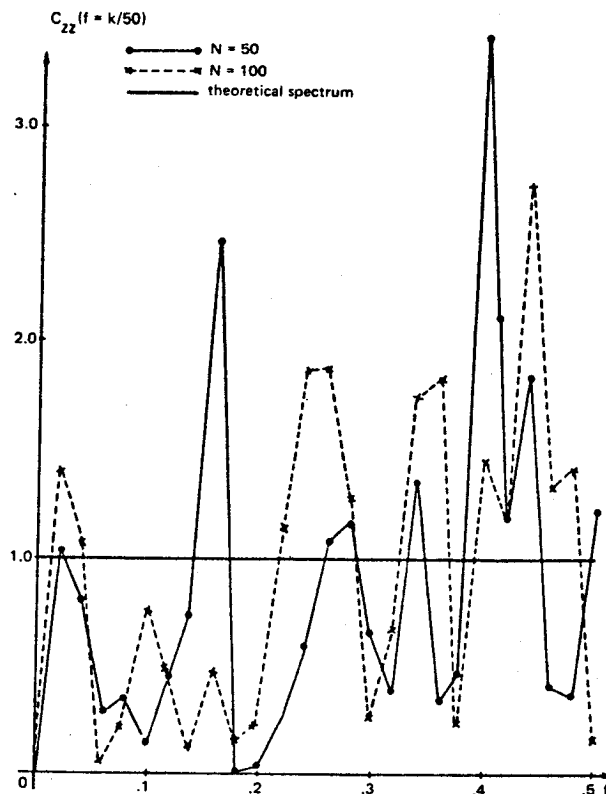
and the estimator is unbiased.

Fig. 4. Spectral estimates for two different record lengths of white noise signal illustrating 100 % relative error (adapted from ref. 1).

Unfortunately, unlike our previous estimators we have a problem with the variability of $S_T(\omega)$. This is illustrated in Figs. 4 and 5 taken from ref. (1) which show both the true spectrum and the estimated spectrum. It is clear

that the scatter is unacceptable and that things do not improve with increasing record length.
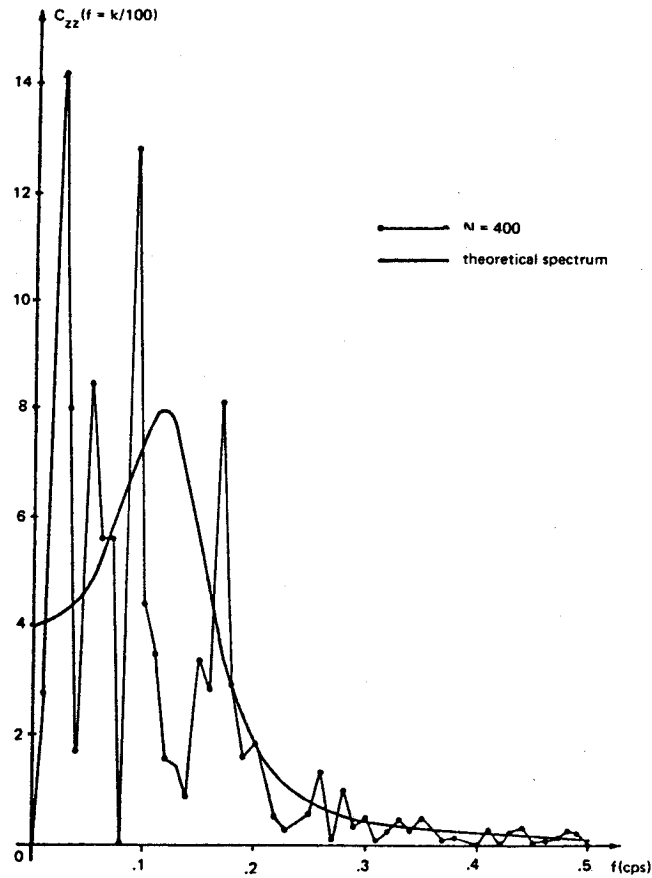

Fig. 5. Spectral estimate for filtered noise signal illustrating 100 % relative error (adapted from ref. 1)

To see why this is so we compute the variance of our estimator. By assuming that fourth-order moments of u(t) are jointly normal, we can obtain (see ref. 1)

$$\text{var}\left\{ S_T(\omega) \right\} = [S(\omega)]^2 \tag{39}$$

It immediately follows that the relative error is unity!

$$\epsilon^2 = \frac{\text{var}\left\{ S_T(\omega) \right\}}{[S(\omega)]^2} = 1 \tag{40}$$

Thus the rms fluctuations are as big as the spectrum!

The reason for the above can be seen by examining equation [37]. We define a new function $h_T(\tau)$ by

$$h_T(\tau) = \begin{cases} 1 - \dfrac{|\tau|}{T} & ; |\tau| < T \\ 0 & ; |\tau| > T \end{cases}$$ [41]

and denote its transform by $H_T(\omega)$. Equation [37] can be rewritten as

$$\overline{S_T(\omega)} = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i\omega\tau}\ \overline{u(t)u(t+\tau)}\ h_T(\tau)\ d\tau$$ [42]

or using Parseval's theorem

$$\overline{S_T(\omega)} = \int_{-\infty}^{\infty} S(\omega_1)\ H_T(\omega-\omega_1)\ d\omega_1$$ [43]

Thus, our spectral estimator is viewing the spectrum through a window whose width is determined by the averaging time T ($\Delta\omega_T = 2\pi/T$). As T increases, $h_T(\tau)$ gets wider and $H_T(\omega)$ gets narrower. In the limit as $T \to \infty$, $H_T(\omega)$ simply selects a single realization of the spectrum. While this explains the lack of bias, it also accounts for the variance since it means that we have used only a single independent sample for our estimate!

There are two ways in which we can make our spectral estimator converge with increasing averaging time:

1. We can ensemble average independent estimates of the spectrum based on non-overlapping record segments. This technique can be viewed as either doing the experiment many times or as subdividing a long record.

2. We can use the fact that the estimates at different frequencies are un-correlated when separated by more than $\Delta\omega_T = 2\pi/T$, and average over different frequencies. This process is called smoothing. The fact that adjacent estimates are correlated only over $\Delta\omega_T$ follows immediately from the fact that the true Fourier coefficients (which are uncorrelated at different frequencies) are seen through the window $H_T(\omega)$ by virtue of the finite record length.

The rate of convergence due to method (1) is simply the same as that for any ensemble average ($\epsilon \sim 1/\sqrt{N}$). For method (2) it should be clear that the number of independent samples making the smoothed estimate depends on the relative value of the effective window width $\Delta\omega$ and the correlation width $\Delta\omega_1$.

Thus, $N \sim \Delta\omega/\Delta\omega_T$ and the relative spectral error is given by

$$\epsilon^2 = \frac{2\pi}{\Delta\omega T} \tag{44}$$

This does go to zero as $T \to \infty$ as required.

There are many ways in which the averaging (or smoothing) processes above can be implemented. All involve the choice of a window in either the time domain $D(\tau)$ or the frequency domain, $W(\omega)$. The smoothed spectrum is then given by

$$S_M(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i\omega\tau} \ B_q(\tau) \ D(\tau) d\tau \tag{45}$$

where

$$B_q(\tau) = B(\tau)h_T(\tau) \tag{46}$$

or

$$S_M(\omega) = \int_{-\infty}^{\infty} S_q(\omega_1) \ W(\omega-\omega_1) \ d\omega_1 \tag{47}$$

where $S_q(\omega)$ is the Fourier transform of $B_q(\tau)$.

TABLE I   Lag and Spectral Windows (from Figure 6)

Lag Window		Spectral Window	Effective Bandwidth ($\Delta\omega$)
Rectangular	$D_R(\tau) = \begin{cases} 1, & \lvert\tau\rvert \leqslant M \\ 0, & \lvert\tau\rvert > M \end{cases}$	$W_R(\omega) = \frac{M}{\pi}\left\lbrace\frac{\sin \omega M}{\omega M}\right\rbrace$	$\dfrac{\pi}{M}$
Bartlett	$D_B(\tau) = \begin{cases} 1 - \frac{\lvert\tau\rvert}{M}, & \lvert\tau\rvert \leqslant M \\ 0, & \lvert\tau\rvert > M \end{cases}$	$W_B(\omega) = \frac{M}{2\pi}\left\lbrace\frac{\sin \omega M/2}{\omega M/2}\right\rbrace^2$	$\dfrac{3\pi}{M}$
Tukey	$D_T(\tau) = \begin{cases} 1/2\left\lbrace 1+\cos\frac{\pi\tau}{M}\right\rbrace, & \lvert\tau\rvert \leqslant M \\ 0, & \lvert\tau\rvert > M \end{cases}$	$W_T(\omega) = \frac{M}{2\pi}\left\lbrace\frac{\sin \omega M}{\omega M}\right\rbrace\left[\frac{1}{1 - (\omega M/\pi)^2}\right]$	$\dfrac{8\pi}{3M}$
Parzen	$D_p(\tau) = \begin{cases} 1-6\left\lbrace\frac{\tau}{M}\right\rbrace^2 +6\left\lbrace\frac{\lvert\tau\rvert}{M}\right\rbrace^3, & \lvert\tau\rvert \leqslant \frac{M}{2} \\ 2\left\lbrace 1-\frac{\lvert\tau\rvert}{M}\right\rbrace^3, & \frac{M}{2} \leqslant \lvert\tau\rvert \leqslant M \\ 0, & \lvert\tau\rvert > M \end{cases}$	$W_p(\omega) = \frac{3M}{8\pi}\left\lbrace\frac{\sin \omega M/4}{\omega M/4}\right\rbrace^4$	$\dfrac{3.77\pi}{M}$

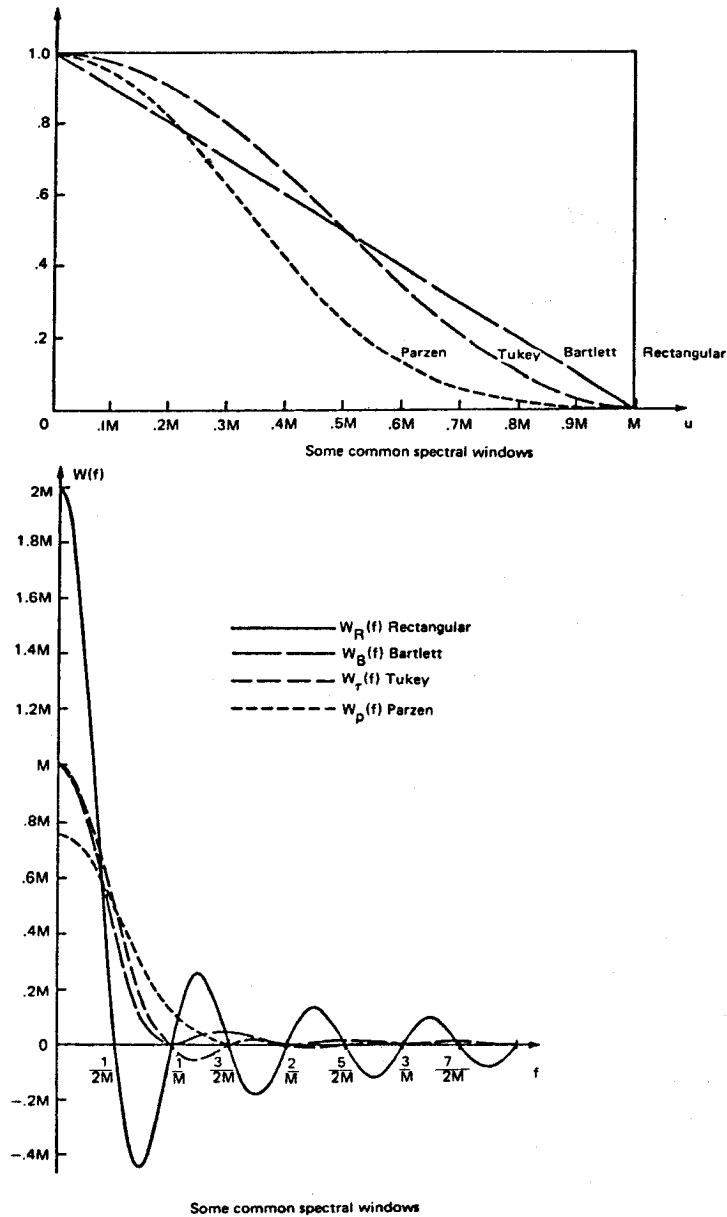Some common spectral windows

Some common spectral windows

Fig. 6. Lag and spectral windows commonly in use (adapted from ref. 1). For definitions and effective bandwidths see Table I.

Some sample windows are shown in Fig. 6 in the form in which they are usually applied to digital data. For analog estimation the window is determined by the characteristics of the band-pass filter used in the analysis.

That we have not solved all of our problems is illustrated by Fig. 7 from ref. 1 which shows the effect of various window widths on a spectrum estimated from a fixed length of record. For the narrowest window, the spectrum cannot be determined because of the variability of the estimator (which peaks are real?), while for the widest window the spectrum, although smooth, cannot be believed. Unfortunately, *the smoothing process introduces a bias.*
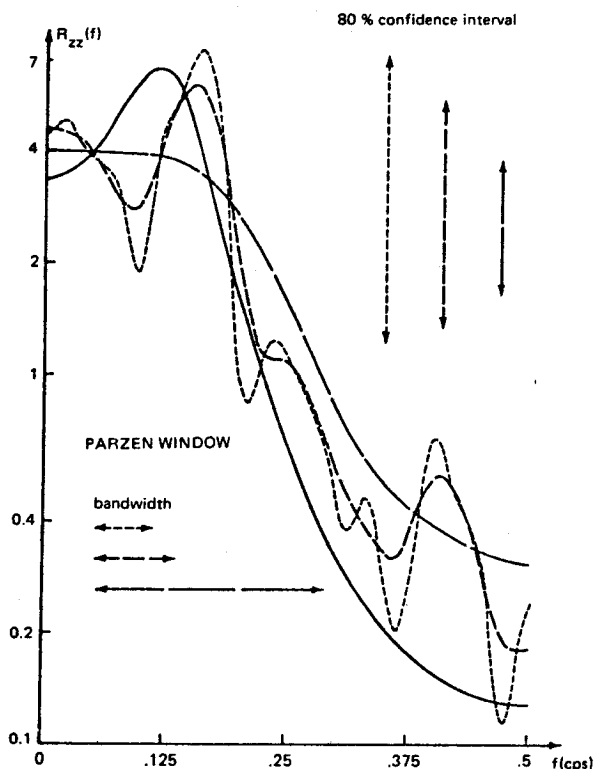

Fig. 7. Spectral estimates illustrating dependence of relative error and bias on window width (from ref. 1). Solid line is true spectrum.

It is not difficult to show that this bias can be related to the curvature of the spectrum (ref. 1) and simply places a constraint on the maximum window width (and minimum averaging time) which can be used. Approximately we must have

$$\Delta\omega \ll \left[S(\omega) \Big/ \frac{d^2}{d\omega^2} S(\omega) \right]^{1/2} \quad \text{to avoid bias.} \tag{48}$$

Discretely Sampled Signals

The emergence over the past 30 years of the digital computing machine as a major experimental tool and especially the rapid advances in mini- and micro-computer technology of the last decade have made this mode of signal processing the most common. Even on-line hard-wired correlators and spectrum analyzers are no longer analog devices but rather dedicated microprocessors operating on digitized inputs.

In the preceding section we have examined the basic character of random signals and the limitations imposed on our ability to measure them. In this section we shall look at the additional problems introduced by digitizing this signal with a sampling device which takes samples at a fixed rate. We should not expect that the process of digitization will eliminate any of the considerations of the preceding section. Rather, the important question is: how much information is lost in the conversion process?

We begin by defining a sampling function $g(t)$ which simply selects values of the process being sampled at programmed times. If we agree that $g(t)$ can be represented as a generalized function and that we can work with such functions, then the sampled signal [denoted by $u_0(t)$] is most easily represented as

$$u_0(t) = u(t)\, g(t) \tag{49}$$

It is clear that $u_0(t)$ only has meaning in the sense of generalized functions.

The question we can now ask is: How much information about $u(t)$ is left in $u_0(t)$? To answer this we need to select a form for $g(t)$. An appropriate choice is easily shown to be

$$g(t) = \Delta t \sum_{n=-\infty}^{\infty} \delta(t - n\Delta t) \tag{50}$$

Whether we sum or simply regard $g(t)$ as a sequence is irrelevant since the delta functions never overlap. The factor Δt is introduced to keep $g(t)$ dimensionless.

We first compute the time averaged mean value of $u_0(t)$.

$$U_{0_T} = \frac{1}{T}\int_0^T u_0(t)\,dt = \frac{1}{T}\int_0^T u(t)g(t)\,dt = \frac{\Delta t}{T}\sum_{n=1}^{N} u(n\Delta t) \tag{51}$$

But $T/\Delta t$ is simply the number of samples and we have

$$U_{o_T} = \bar{u} = \frac{1}{N} \sum_{n=1}^{N} u(n\Delta t) \qquad [52]$$

which is reminiscent of our ensemble average.

It is immediately obvious that our sampled function indeed retains all the information on the moments of $u(t)$. It is also obvious from our previous considerations that we can simultaneously minimize the number of data to be handled while maximizing the convergence rate of the estimator by insuring that the samples are, in effect, statistically independent by choosing

$$\Delta t = 2T_u \qquad [53]$$

Thus the optimal sampling rate is one sample for every two integral scales in time.

If we analyze our estimator for the autocorrelation we find that its digital counterpart is

$$B_{o_T}(n\Delta t) = \frac{1}{N^2} \sum_{i=1}^{N-n} \sum_{j=1}^{N-n} u(i\Delta t) u([j+n]\Delta t) \qquad [54]$$

where the time lag is given by $n\Delta t$.

A detailed analysis of the convergence of this estimator will reveal what we could have guessed; namely, that all of these time lag products do not contribute to the convergence of the estimator, only those separated by two integral scales. Thus a much more efficient estimator which would converge at the same rate would be

$$\tilde{B}_{o_T}(n\Delta t) = \frac{1}{N'} \left\{ u(o)u(n\Delta t) + u(2T_u)u(2T_u + n\Delta t) + \right. \qquad [55]$$

$$\left. \ldots + u(2[N'-1]T_u) u(2[N'-1]T_u + n\Delta t) \right\}$$

where $N' = T/2T_u$ and is, in effect, the number of independent samples. The algorithm computes a time-lag pair for the appropriate time delay, then leaps down the record $2T_u$ before computing another. This is illustrated in Fig. 8.

This method of computing the autocorrelation was suggested to this author by Lumley in 1972 who also showed that a very efficient spectral estimator could be derived from it. Since this suggestion was unpublished, it has been included here as Appendix II with the appropriate credit given.
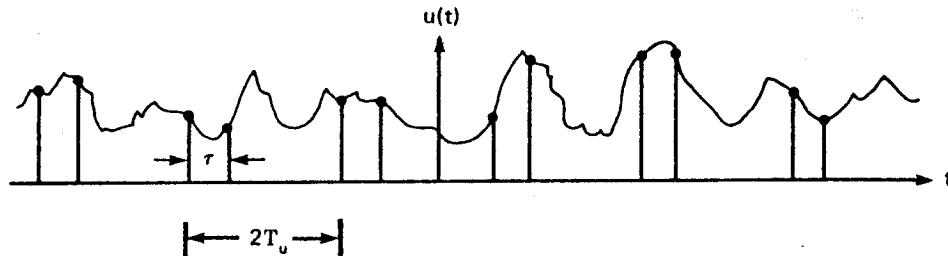
u(t)

|◄—— 2T_u ——►|

Fig. 8. Illustration of skip method for computing autocorrelations in which pairs of samples are taken only every two integral scales.

Spectral Analysis of Discretely Sampled Signals

It is straightforward to compute the Fourier transform (in the sense of generalized functions) of the signal $u_0(t)$.

$$\hat{u}_0(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i\omega\tau} u(t)g(t) \, dt \qquad [56]$$

Substitution for $g(t)$ and some manipulation yields the spectrum of $u_0(t)$ as

$$S_0(\omega) = \sum_{n=-\infty}^{\infty} S(\omega - 2\omega_n) \qquad [57]$$

where

$$2\omega_n = \frac{2\pi}{\Delta t} \qquad [58]$$

Thus our discretely sampled signal produces not only the desired result, $S(\omega)$, but an infinity of similar spectra all shifted by frequency ω_n. The phenomenon is illustrated in Fig. 9 from ref. (1).

It is clear that if the signal $u(t)$ has spectral components above the frequency ω_n, these components are aliased to lower frequencies with the result that the desired spectrum can no longer be distinguished. The phenomenon is called aliasing and ω_n, the folding frequency, is called the *Nyquist frequency*.

The *Nyquist criterion* can be summarized as follows: The spectral information in a discretely sampled time signal can be retained only if the sampling rate is *at least* twice the highest frequency present in the original signal.

Thus for digital processing of random signals we have conflicting requirements. We must sample at small time intervals to avoid aliasing, yet the additional data contribute little to the convergence of our estimators. The "waste factor" is

$2T_u/\Delta t$ which for wide-band signals can be very large indeed.

The problem of processing the large data blocks necessitated by the Nyquist and convergence criteria has received a large amount of attention since the infancy of digital processing techniques. As we shall see, there are several clever algorithms and techniques that have been developed to reduce the computational effort. A number of the commonly used and new approaches are summarized below.
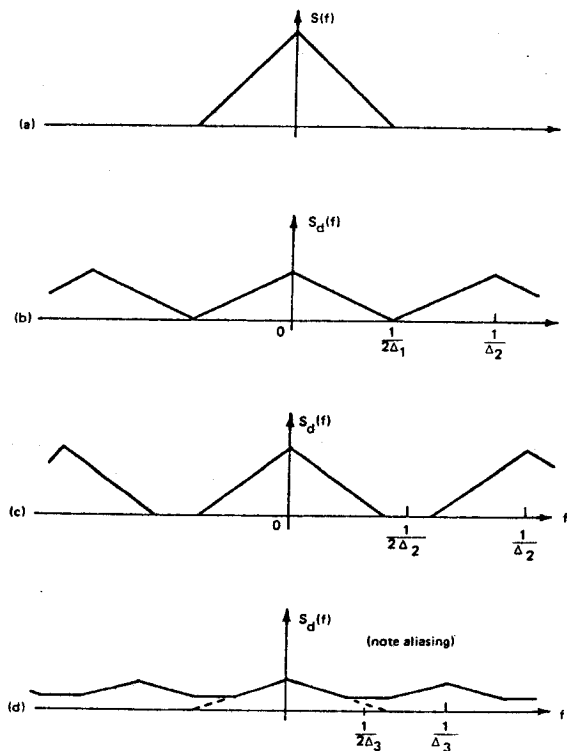
Fig. 9. Effect of sampling rate on measured spectra. Original spectrum (uppermost); aliased spectrum (lowermost) (adapted from ref. 1)

1. The time-lag product approach

This is a direct approach based on a discrete version of the time average estimator for the autocorrelation, Equation [54]. The spectrum is then computed by multiplying this autocorrelation by the appropriate lag window and taking its discrete transform. Reference (6) provides a complete discussion of the application of this approach.

2. **The Fast Fourier Transform**

This is a very fast and efficient algorithm for computing the Fourier coefficients of the signal by utilizing a bit-reversal of the binary address of the time-series data. The smoothed spectrum is then computed by squaring the coefficients and convolving them with the appropriate window. The savings over the direct approach is approximately $\ell n\, N/N$ where N is the number of data points (see ref. (7) for complete discussion).

3. **The Low-Pass FFT**

This is a variation on the FFT in which the incoming signal is divided by frequency into band-limited blocks. The first block might be low-pass filtered at $0.1\,\omega_n$, the second at $0.2\,\omega_n$, the third at $0.4\,\omega_n$, etc.

Each block has a length inversely proportional to the cutoff frequency of the preceding block; that is, the higher blocks have shorter records. Each block is processed by the F.F.T. and smoothed by filters whose width increases with block width. Not only does this technique allow substantially larger spectral samples to be computed on a fixed core machine, but a substantial savings in computational costs is achieved, at some increases in experimental complexity. For a complete discussion see reference (8).

4. **The Auto-Regressive Spectral Technique**

Although not in common use in the fluid dynamics community, these techniques are evolving rapidly elsewhere. Because of this a detailed summary by Beuther of how these techniques work is included as Appendix I. In brief, an autoregressive series is fitted to the incoming time-data which optimizes the ability to predict the next data point. The spectrum is then simply determined by the coefficients of the autoregression and the mean-square error in the fit. A bonus is that the spectrum is already smoothed in a manner which minimizes both bias and relative error and is thus optimally filtered in some sense.

5. **The Skip Technique**

We have already mentioned this efficient estimator due to Lumley which utilizes only *independent samples of the time-lag products*. This method is described in detail in Appendix II. For most applications to turbulence measurement it appears to be the most efficient and is particularly well-suited to small computers.

To conclude this section we note that all spectral estimates must be smoothed in the same manner as for the continuous signal. The rate of convergence of the smoothed estimator is again proportional to inverse product of the window width in frequency space and the averaging time. As for the continuous signal, bias can result from this smoothing operation.

Randomly Sampled Data

There is no reason that the process of converting a continuous signal to a digital one has to be carried out at uniformly spaced time intervals. One of the most significant advances in signal processing over the past ten years has been the understanding and appreciation of the merits of random sampling. That is, instead of sampling at uniformly spaced intervals in time, the intervals are themselves random and are governed by some statistical process.

The development of the theory of random sampling is due to a number of independent investigators; among them are Gaster and Roberts, Mayo, Shapiro and Silverman, and others. For a comprehensive review the reader is referred to the article in this proceedings by Mayo. Our purpose here is to briefly review the basic principles and place the theory into the context of the overall problem of signal processing.

We begin by insisting that our random sampling process be *statistically independent of the process being sampled.* As will be seen later this *excludes* the burst-mode LDA from consideration. We choose a sampling function g(t) so that it selects values of u(t) at random instants in time. The sampled signal is then given by

$$u_o(t) = u(t)\, g(t) \qquad\qquad [59]$$

as before.

A suitable choice for g(t) is

$$g(t) = \delta(t - t_i), \quad i = 1, \ldots, \infty \qquad\qquad [60]$$

where the t_i are uncorrelated. If the average sample rate is denoted by ν, it is easy to show that

$$\bar{g} = \nu \qquad\qquad [61]$$

and

$$\overline{g(t)\, g(t')} = \nu^2 + \nu\delta(t' - t) \qquad\qquad [62]$$

and

$$\overline{(g - \bar{g})(g' - \bar{g})} = \nu\delta(t' - t) \qquad\qquad [63]$$

Since, by hypothesis, u(t) and g(t) are statistically independent, it follows immediately that

$$\overline{u_o(t)} = \nu \overline{u} \tag{64}$$

$$\overline{u_o'(t)\, u_o'(t')} = \nu^2\, \overline{u'(t)u'(t')} + \nu\, \overline{u^2}\, \delta(t'-t) \tag{65}$$

An immediate consequence is that the autocorrelation and spectrum of $u_o(t)$ are given by:

$$B_o(\tau) = \nu^2\, B(\tau) + \nu\, \overline{u^2}\, \delta(\tau) \tag{66}$$

and

$$S_o(\omega) = \nu^2\, S(\omega) + \frac{\nu}{2\pi}\, \overline{u^2} \tag{67}$$

Thus with the exception of the spike at the origin in the autocorrelation and the corresponding white noise added to the spectrum, all of the information contained in the original signal has been retained.

In practice, the nuisance due to the spike at the origin in the autocorrelation can be eliminated by simply agreeing that we will never include self-products when we multiply signals together. It is easy to show that this corresponds to writing

$$\overline{g(t)\, g(t')} = \nu^2 \tag{68}$$

from which it follows that

$$B_o(\tau) = \nu^2\, B(\tau) \tag{69}$$

$$S_o(\omega) = \nu^2\, S(\omega) \tag{70}$$

Thus, by agreeing to remove self-products from any algorithms, we need only the sample rate to recover the desired information.

To derive practical estimators we turn to time-averaged value of $u_o(t)$ and its moments. For the mean value we have simply

$$U_{o_T} = \frac{1}{T}\int_o^T u_o(t)dt = \frac{\nu}{T}\, \sum_{i=1}^{N} u_i \tag{71}$$

where N is the number of samples which arrive in the interval (o, T) and the u_i

are the randomly sampled realizations of the signal. For T large enough, $\nu T \cong N$ and an unbiased estimator for the mean can be written as

$$U_{TR} = \frac{1}{N} \sum_{i=1}^{N} u_i \qquad [72]$$

An immediate question is: How does this estimator converge? It is easy to show in the same manner as before that

$$\text{var} \{U_{TR}\} = \frac{2T_u}{T} \left[1 + \frac{1}{2\nu T_u}\right] \cdot \text{var} \{u\} \qquad [73]$$

from which it follows that the relative error is

$$\epsilon^2 = \frac{\text{var} \{U_{TR}\}}{\bar{u}^2} = \frac{2T_u}{T}\left[1 + \frac{1}{2\nu T_u}\right] \cdot \left[\frac{\sigma_u}{\bar{u}}\right]^2 \qquad [74]$$

Note that as $\nu T \to \infty$, this reduces to our previous result for continuous signals; while for $\nu T \to 0$ we obtain

$$\lim_{\nu T \to 0} \epsilon \cong \sqrt{\frac{1}{\nu T}} \cdot \frac{\sigma_u}{\bar{u}} \cong \frac{1}{N} \frac{\sigma_u}{\bar{u}} \qquad [75]$$

These limits are reasonable since as the sampling rate becomes very high relative to the time scale of the signal, we reproduce the continuous signal. As the sampling rate becomes very low, all the samples are, in effect, statistically independent.

A simple estimator for the mean square value of the original signal is given by

$$\sigma^2_{TR} = \frac{1}{N} \sum_{i=1}^{N} \left[u_i - U_{TR}\right]^2 \qquad [76]$$

It is straightforward to show that the relative error is analogous to that in equation [74].

The best estimator for estimating the autocorrelation and spectrum from the randomly arriving samples is still a matter of debate (see the article by Mayo in this volume). The approach adopted here will be to first obtain the spectrum, and then compute the inverse transform to obtain the autocorrelation.

The finite time estimate for the Fourier coefficients of $u_o(t)$ is given by

$$\hat{u}_{o_T}(\omega) \equiv \frac{1}{2\pi} \int_0^T e^{i\omega\tau} u'_o(t)dt = \frac{1}{2\pi} \sum_{i=1}^{N} e^{i\omega\tau} u'(t_i) \qquad [77]$$

It follows that a spectral estimator for $S_o(\omega)$ is

$$S_{o_T}(\omega) = 2\pi \frac{|\hat{u}_{o_T}(\omega)|^2}{T} = \frac{1}{2\pi T} \sum_{\substack{i=1 \\ i \neq j}}^{N} \sum_{j=1}^{N} e^{i\omega(t_i - t_j)} u'(t_i)u'(t_j) \qquad [78]$$

where we have employed our prohibition on self-products.

From equation [78] we can immediately obtain an estimator for $S(\omega)$, the desired spectrum, as

$$S_{TR}(\omega) = \frac{1}{2\pi\nu^2 T} \sum_{\substack{i=1 \\ i \neq j}}^{N} \sum_{j=1}^{N} e^{i\omega(t_i - t_j)} u'(t_i)u'(t_j) \qquad [79]$$

This estimator is unbiased *and unaliased*. The lack of aliasing is a consequence of the random sampling and is independent of the mean sample rate. Thus random sampling would appear to have significant advantages over the traditional discrete sampling technique.

The absence of aliasing is not without price, however. The price is the increased relative error due to the random sampling. It can be shown by assuming that fourth order moments are jointly Gaussian and by noting that   [80]

$$\overline{gg'g''g'''} = \nu^4 + \nu^3 \left[\delta(t''-t) + \delta(t'''-t) + \delta(t''-t') + \delta(t'''-t')\right] + \nu^2 \left[\delta(t''-t)\delta(t'''-t') + \delta(t''-t')\delta(t'''-t)\right]$$

that the relative error of the estimator $S_{TR}(\omega)$ is given by

$$\epsilon^2 = \frac{\text{var}\{S_{TR}(\omega)\}}{[S(\omega)]^2} = \left\{1 + \frac{S(o)}{S(\omega)} \cdot \left[\frac{1}{2\nu T_u}\right]\right\}^2 \qquad [81]$$

Since the spectrum is assumed to fall as frequency increases, the effect of the random sampling is to increase the relative error with frequency. Note that as the sample rate becomes large, the relative error reduces to the result for continuous signals. Clearly the relative error is never less than unity and is therefore unacceptable. Hence we must smooth the estimator as before.

There are several techniques by which the estimator above can be smoothed to make it converge. These are:

(i) **Block Averaging**
    This approach is certainly the most straightforward to implement and simply involves averaging the spectral estimators obtained from independent record lengths of the process. As in the continuous case, the relative error is decreased by the inverse square root of the number of independent samples.

(ii) Spectral Window

As for the continuous case this can be done either in the time or frequency domain. The simplest approach is to insert a lag-window into the estimator of equation [79] to obtain

$$S_{TRM}(\omega) = \frac{1}{2\pi\nu^2 T} \sum_{\substack{i=1 \\ i \neq j}}^{N} \sum_{j=1}^{N} e^{i\omega(t_i - t_j)} \, u'(t_i)u'(t_j)D(t_i - t_j) \qquad [82]$$

where $D(\tau)$ can be any of the windows already introduced. It is tedious but straightforward (see ref. 9 or proceed as above) to show that the relative error is the error given by equation [81] reduced by a factor of $2\pi/\Delta\omega T$ where $\Delta\omega$ is the effective window width in the frequency domain. We have

$$\epsilon^2 = \frac{2\pi}{\Delta\omega T} \left\{ 1 + \frac{S(o)}{S(\omega)} \cdot \frac{1}{2\nu T_u} \right\}^2 \qquad [83]$$

A similar result is obtained if spectrum estimated by equation [79] is convolved with a filter of width $\Delta\omega$. In view of the increased error as frequency increases, this approach can have significant advantages if the window varies with the frequency at which the spectrum is estimated. Gaster and Roberts (10) suggest that for power law spectrum a filter whose width increases logarithmically with frequency would be desirable. Clearly combinations of all the above approaches could be used to advantage.

(iii) The Time-Slot Approximation

This is best illustrated by first taking the inverse Fourier transform of $S_{TR}(\omega)$ to obtain

$$B_{TR}(\tau) = \frac{1}{\nu^2 T} \sum_{\substack{i=1 \\ i \neq j}}^{N} \sum_{j=1}^{N} u'(t_i)u'(t_j)\delta(t_i - t_j - \tau) \qquad [84]$$

This represents a collection of random realizations of the autocorrelation. We now group these realizations into slots of width $\Delta\tau$ by defining the time-slot approximation to the autocorrelation as

$$B_s(n\Delta\tau) = \frac{1}{\Delta\tau} \int_{(n-\frac{1}{2})\Delta\tau}^{(n+\frac{1}{2})\Delta\tau} B_{TR}(\tau)d\tau \qquad [85]$$

which yields the "smoothed" autocorrelation as

$$B_s(n\Delta\tau) = \frac{1}{\Delta\tau} \left(\frac{1}{\nu^2 T} \right) \sum_{\substack{i=1 \\ i \neq j}}^{N} \sum_{j=1}^{N} u(t_i)u(t_j) \qquad [86]$$

where

$$(n-\tfrac{1}{2})\Delta\tau < |t_i-t_j| < (n+\tfrac{1}{2})\Delta\tau$$

This algorithm can be directly applied to the incoming data to compute the autocorrelation which can, in turn, be Fourier transformed to yield a discrete spectrum. This transform can even incorporate a lag window to improve convergence in the usual manner.

There is some confusion in the literature as to what constitutes the proper relative error for spectra computed from the time-slot approximation. The operation defined by equation [85] is, in fact, a convolution of the autocorrelation with a time window. In frequency space this means that the spectrum is simply filtered by the window function which is assumed broadband since the time window $\Delta\tau$ is narrow. Since the variance is proportional to the spectrum squared, it is reduced by this low-pass filtering, but only at the expense of bias. Thus the primary cause for convergence is the window used in computing the smoothed spectrum from the time-slot correlation, and the convergence is identical to equation [83]. This is, in fact, the conclusion of Mayo from empirical evidence.
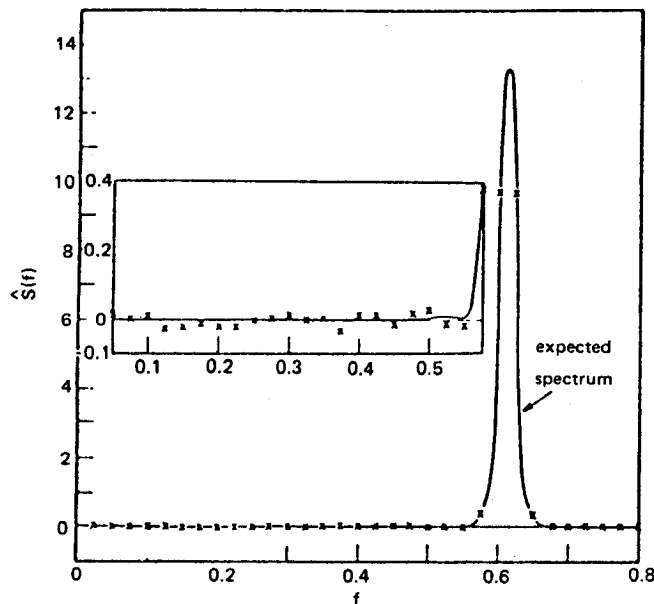

Fig. 10. Spectrum of sine wave from random samples $(2\pi\nu/\omega = 1.7)$ (from ref. 10).

Examples of spectra computed by Gaster and Roberts from randomly sampled data are shown in Figs. 10 and 11. In Fig. 10, a sine wave is sampled so that the mean sample rate is less than the Nyquist frequency. Clearly the expected peak is reproduced. In Fig. 11, a low-passed noise is sampled so that the mean sample rate corresponds to the breakpoint in the spectrum. Although the increased variability at the high frequencies is obvious, the spectrum is accurately reproduced.
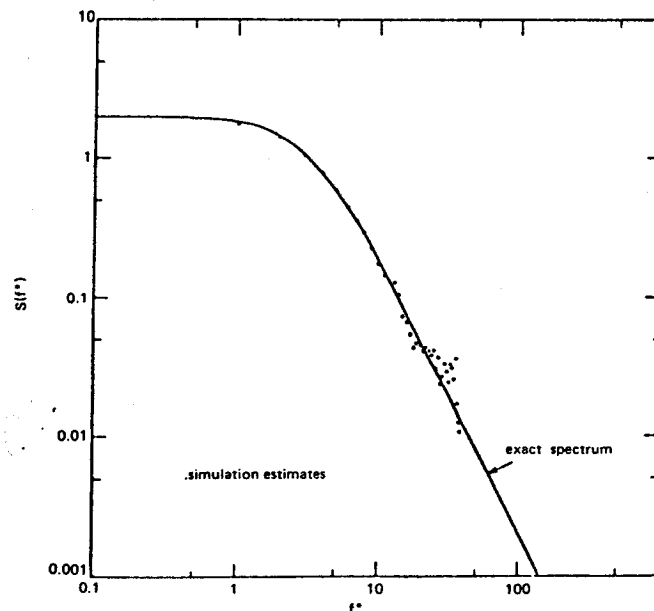

Fig. 11. Spectrum of low-passed noise $(v/f_o = 1)$(from ref. 10).

The Burst-Mode Laser Doppler Anemometer

The purpose of this section is not to discuss the burst-mode LDA as a tool for flow measurement but rather to concentrate on the interesting signal analysis problems it presents. The phrase "burst-mode LDA" is used to refer to an LDA operating in such a way that for most of the time there are no scattering particles in the measuring volume and there is never more than one particle at a time. As the particles arrive randomly in time, the processor measures their velocity and makes this data available as a digital word for further processing. As we shall see, the processor must also provide other information if the signal is to be interpreted correctly. For a complete description of this instrument the reader is referred to the review article by Buchhave in this proceedings and to references (11) and (12).

The randomly arriving particles do, in fact, randomly sample the flow velocity. Unlike the sampling scheme just discussed, however, the sampling process is *not* independent of the process being sampled. This is easily seen if one considers that the particles are assumed randomly distributed at statistically independent locations *in space* and carried to the measuring volume by the flow. Thus, arrival time and flow are correlated.

We can correctly analyze this instrument if we design a sampling function which samples the velocity at the spatial location of the particle. This was done in references (12) and (13) with the following result: the sampled velocity is given by

$$u_o(t) = \int_{\text{all space}} U(\underline{a},t)g_1(\underline{a})w[\underline{x}(\underline{a},t)]\,d^3\underline{a} \qquad [87]$$

where $U(\underline{a},t)$ is the Lagrangian velocity of the flow which the particles are assumed to follow, the function $w(\underline{x})$ accounts for the fact that the measuring volume is of finite extent, and $g_1(\underline{a})$ is a random function which accounts for the presence or absence of a particle at a particular Lagrangian coordinate.

Because of the motion the particle is moved by the flow to a location given by

$$x = x(\underline{a},t) = \int_o^T U(\underline{a},t_1)\,dt_1 \qquad [88]$$

Thus, when the particle wanders into the measuring volume, the function $w(\underline{x})$ "turns on" the signal $u_o(t)$. A typical scattering volume is shown in Fig. 12 and the signal $u_o(t)$ resulting from the randomly arriving particles is shown in Fig. 13.
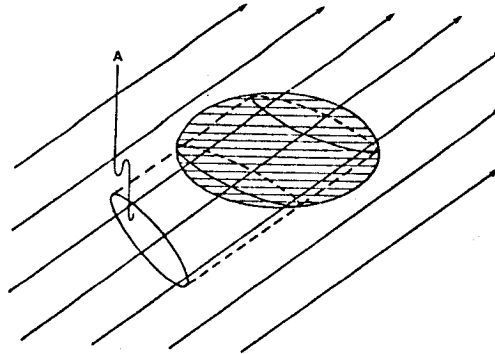

Fig. 12. Schematic showing measuring volume defined by laser Doppler anemometer.
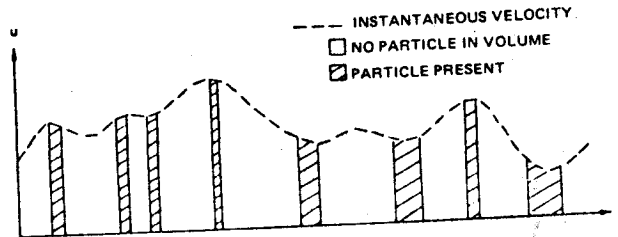
Fig. 13. Typical velocity signal as sampled by individual and randomly arriving particles.

By a series of arguments (see reference (13)), we can transform the dependence on Lagrangian coordinates to Eulerian (or spatial) ones and write

$$U_0(t) = \int\limits_{\text{all space}} u(\underline{x},t)\, w(\underline{x})\, g(\underline{x},t) d^3\underline{x} \qquad [89]$$

where $u(\underline{x},t)$ is the velocity at the point $\underline{x}$ and $g(\underline{x},t)$ has the following statistical properties:

$$\overline{g(\underline{x},t)} = \mu \qquad [90]$$

$$\overline{g(\underline{x},t)\, g(\underline{x}',t')} = \mu\delta(\underline{x}-\underline{x}' - \underline{U}\cdot[t'-t]) + \mu^2 \qquad [91]$$

and μ is the average number of particles per unit volume.

We compute the moments of $u_0(t)$ directly as

$$u_0 = \int\limits_{\text{all space}} \overline{u(\underline{x},t)}\;\overline{g(\underline{x},t)}\; w(\underline{x}) d^3\underline{x} \qquad [92]$$

$$B_0(\tau) = \overline{u_0(t)u_0(t')} = \iint\limits_{\text{all space}} \overline{u(x,t)u(\underline{x}',t')}\;\; \overline{g(\underline{x},t)g(\underline{x}',t')} \cdot w(\underline{x})w(\underline{x}')d^3\underline{x}d^3\underline{x}' \qquad [93]$$

where we have used the fact that velocity and the occurrence of a particle in space are independent. It follows immediately that (see ref. (11))

$$\overline{u_0(t)} = (\mu V)\,\overline{u} \qquad [94]$$

and

$$[95]$$

where we have assumed that the effect of averaging the velocity and its space-time correlation over the scattering volume is negligible. The function $\rho_1(\tau)$ is easily shown to be a "spike-like" function of width proportional to the mean transit time of particles crossing the volume.

Since μV corresponds to the expected number of particles in the volume and is assumed small, the "spike-like" term will dominate the spectrum in the same manner the self-product terms dominated the random time sampling process discussed earlier. The "cure" here is the same as before: we agree to never include self-products in any computation. This corresponds to eliminating the delta function in equation [91] yielding

$$\overline{g(\underline{x},t)\, g(\underline{x}',t')} = \mu^2 \tag{96}$$

from which it follows that

$$B_o(\tau) = (\mu V)^2 B(\tau) \tag{97}$$

and

$$S_o(\omega) = (\mu V)^2 S(\omega) \tag{98}$$

Thus the moments and spectra are completely recoverable, in principle, from the information available from the particles. That complete flow statistics could be obtained from the randomly arriving particles was suspected long before it was proven in 1975 (see for example ref. (14)). Unfortunately it was not until after this (and from this) that correct estimators were formulated (see ref. (11), (12) and (13)).

The mean value can be calculated from the time integral as before:

$$
\begin{aligned}
U_{o_T} &= \frac{1}{T} \int_0^T u_o(t)\,dt \\
&= \frac{1}{T} \left\{ \int_{t_1}^{t_1+\Delta t_1} u_o(t)\,dt + \int_{t_2}^{t_2+\Delta t_2} u_o(t)\,dt + \ldots + \int_{t_N}^{t_N+\Delta t_N} u_o(t)\,dt \right\}
\end{aligned}
\tag{99}
$$

where the t_i are the arrival times and the Δt_i are the *residence times* or times that the signal is on.

If we assume that the velocity is relatively constant while the particle is traversing the volume, this immediately yields a practical estimator for the mean velocity as

$$U_{TBP} = \frac{\sum\limits_{i=1}^{N} u_i \, \Delta t_i}{\sum\limits_{i=1}^{N} \Delta t_i} \qquad\qquad [100]$$

where in the denominator we have used the fact that as the number of realizations becomes large, the "on-time" is given by

$$\mu VT = \sum\limits_{i=1}^{N} \Delta t_i \qquad\qquad [101]$$

A similar analysis for the mean square fluctuating velocity yields

$$\sigma^2_{TBP} = \frac{\sum\limits_{i=1}^{N} (u_i - U)^2 \Delta t_i}{\sum\limits_{i=1}^{N} \Delta t_i} \qquad\qquad [102]$$

It is easy to show that the relative error in the estimator is the same as for the randomly sampled case discussed earlier if the mean sampling rate ν is interpreted as the mean data rate given by

$$\nu_{eff} = (\mu V)/T_p \qquad\qquad [103]$$

where T_p is the average particle residence time and is approximately given by $d/\bar{u}$ where d is the effective width of the measuring volume.
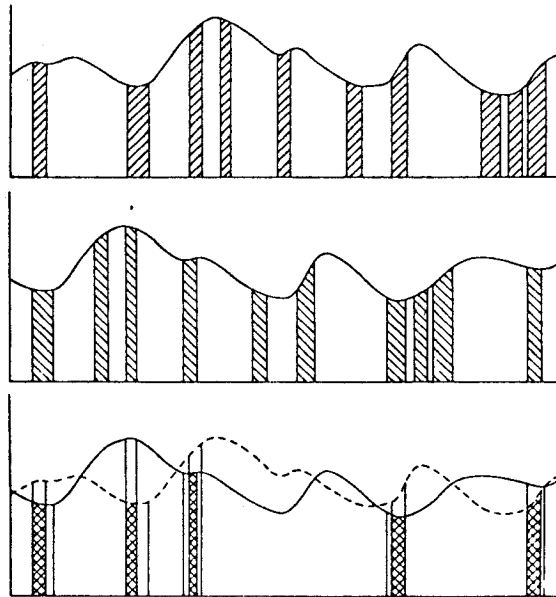

Fig. 14. Computation of the autocorrelation from burst processors using only the overlap times. The middle trace is the upper trace displaced by amount τ as shown

An estimator for the autocorrelation is more difficult to interpret since both $u_o(t)$ and $u_o(t+\tau)$ must on "on" to contribute to the time integral. This is illustrated in Fig. 14. A time-slot approximation to the autocorrelation is easily shown to be given by

$$B_T(n\Delta\tau) = \frac{1}{N_{\Delta\tau}(\sum_{i,j} \Delta t_{ij})} \sum_{\substack{i=1 \\ i \neq j}}^{N} \sum_{j=1}^{N} u_i u_j \, \Delta t_{ij} \qquad [104]$$

where Δt_{ij} is the overlap time and only those $N_{\Delta\tau}$ realizations satisfying

$$(n-\frac{1}{2})\Delta\tau < |t_i - t_j| < (n+\frac{1}{2})\Delta\tau$$

can contribute. It can be shown (most directly by transforming the spectrum derived below) that an estimator which is easier to implement is given by

$$B_{TS}(n\Delta\tau) = \frac{1}{\Delta\tau}\left[\frac{T}{\sum_{i=1}^{N}(\Delta t_i)^2}\right] \sum_{\substack{i=1 \\ i \neq j}}^{N} \sum_{j=1}^{N} u_i u_j \, \Delta t_i \Delta t_j \qquad [105]$$

A direct Fourier transform can be defined from equation [35] as

$$\hat{u}_{o_T}(\omega) = \frac{1}{2\pi} \int_{-T/2}^{T/2} e^{i\omega\tau} u_o(t)dt \qquad [106]$$

which is approximately given by

$$\hat{u}_{o_T}(\omega) = \frac{1}{2\pi} \sum_{i=1}^{N} e^{i\omega t_i} u_i \Delta t_i \qquad [107]$$

(The neglected variation of t in the interval $(t_i, t_i + \Delta t_i)$ can be shown to be equivalent to low-pass filtering at the inverse residence time, which introduces no further approximations than have already been made).
From this a spectral estimator is readily calculated to be

$$S(\omega) = \frac{T}{2\pi(\sum_{i=1}^{N} \Delta t_i)^2} \sum_{\substack{i=1 \\ i \neq j}}^{N} \sum_{j=1}^{N} e^{i\omega(t_i - t_j)} u_i u_j \Delta t_i \Delta t_j \qquad [108]$$

It is tedious but straightforward to show that the relative error of this estimator is also given by equation [81] with ν given by equation [103]. As might be expected the convergence of the smoothed estimator is given by equation [83].

The preceding analysis of the LDA along with experimental evidence is given in a series of recent papers (ref. (12), (15), (11)). Also included in the latter is a detailed description of the hardware and software for implementing these

algorithms and an assessment of the errors resulting from ignoring the need for residence time information.

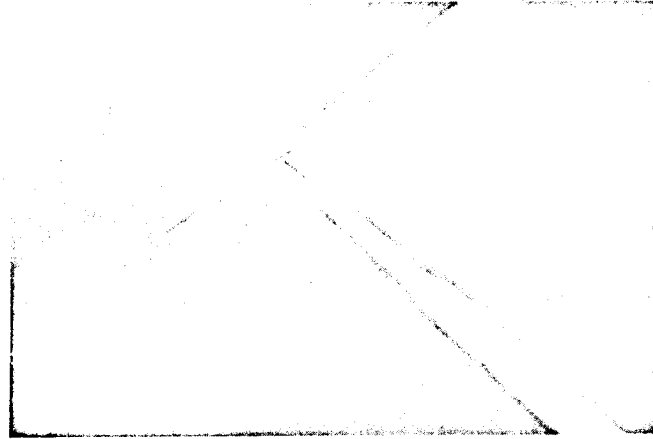**Mixed Mode Data Processing**

Fig. 15. Photograph showing simultaneous use of LDA and hot-wire anemometer

Our success with randomly sampled data leads as to inquire whether a mixed-mode signal processing system might be possible. An example of where such a mode might be desirable is illustrated in Fig. 15 which shows the simultaneous use of a laser Doppler system and a hot-wire anemometer to measure cross-correlations. The primary advantage of such a system over conventional hot-wire techniques is that there is no wake generated by the upstream probe. Moreover, the use of a downstream hot-wire instead of another LDA considerably reduces experimental complexity and cost.

To analyze this problem we assume the upstream LDA to be operating in the burst mode analyzed earlier. We note that any signal generated at the hot-wire probe by the scattering particles will be uncorrelated with the velocities and can be treated as noise (for example, water drops will be rapidly evaporated causing random spikes in the signal).

The mode of data analysis is the following: We continuously monitor the hot-wire signal. When the LDA receives a burst, we measure the particle's velocity and residence time, and send a delayed "freeze" command to the hot-wire monitor. We then compute an instantaneous realization of the space-time

cross-correlation at *all* time lags of interest by multiplying the single LDA rea-
lization by the appropriately delayed hot-wire signal. Each succeeding particle
arrival is treated in the same manner and the accumulated realizations of the
correlation are averaged to yield the time-averaged cross-correlation.

For simplicity, we consider only the case where the LDA and hot-wire are mea-
suring the same velocity. Extension of the analysis to spatially separated sig-
nals is straightforward. The analytical statement of the procedure described
above is then

$$C_o(\tau) = \frac{1}{T} \int_0^T u_o(t)\, u(t+\tau)\, dt \tag{109}$$

where $u_o(t)$ is the LDA burst signal given by equation [89], $u(t)$ is the hot-wire
signal, and the interval (o,T) is determined by the first and last burst, (Note:
$t+\tau > T$ and $t+\tau < o$ are allowed).
It is easy to show that

$$C_T(\tau) = \frac{1}{\mu VT}\; C_o(\tau) \tag{110}$$

and that this is an unbiased estimator of the cross-correlation. Hence an algo-
rithm consistent with the approximations in $u_o(t)$ of the preceding section is

$$C_T(n\Delta\tau) = \frac{\sum\limits_{i=1}^{N} u_o(t_i)u(t_i + n\Delta\tau)\, \Delta t_i}{\sum\limits_{i=1}^{N} \Delta t_i} \tag{111}$$

It remains to examine the convergence of our estimator C_T. Proceeding as be-
fore we write

$$\mathrm{var}\,\{C_T(\tau)\} = \frac{1}{(\mu VT)^2} \iint_0^T \overline{u_o(t)u_o(t_1)u(t+\tau)u(t_1+\tau)}\; dt\,dt_1 - [C(\tau)]^2 \tag{112}$$

By substituting for $u_o(t)$, the integrand is readily obtained as

$$\overline{u_o(t)u_o(t_1)u(t+\tau)u(t_1+\tau)} \tag{113}$$

$$= \iint\limits_{\text{all space}} \overline{u(\underline{x},t)u(\underline{x}',t_1)u(\underline{x}_o,t+\tau)u(\underline{x}_o,t_1+\tau)}\;\; \overline{w(\underline{x})w(\underline{x}'),g(\underline{x},t)g(x_1,t_1)} \cdot d^3\underline{x}\, d^3x'$$

where $\overline{g(\underline{x},t)\, g(\underline{x}',t')}$ is given by equation [91].

After considerable manipulation, it can be shown that the relative error is given
by

$$\epsilon^2 = \frac{\text{var}\left\{C_T(\tau)\right\}}{[C(\tau)]^2} \cong \frac{2T}{T}\left[\frac{1}{\rho_c^2(\tau)} + \frac{1}{\nu_{eff}T}\right] \qquad [114]$$

where ν is given by equation [103]. Comparison of this with equation [27] shows that the added variability is measured by the second term which is independent of the lag. Thus for modest sampling rates relative to the integral scale, the estimate converges nearly as well as for the continuous case.

A similar convergence criterion can be shown to exist if the randomly sampled signal is randomly sampled in time instead of a burst LDA signal. Moreover, it is straightforward to derive convergence criteria for smoothed spectral estimators using the previously illustrated techniques.

A Final Word
In this paper we have reviewed the development of our current understanding of random signal analysis. We have concentrated on the estimation of mean quantities by time average with particular attention to the problems of bias and convergence. Of particular interest because of its recent development, has been the randomly sampled process which holds considerable promise for application as digital transducer components become more common.

We have entirely avoided the practical and theoretical problems which arise from questions relating to frequency response, dynamic range, quantization error and noise. All of these play a significant role in real processing. For more information on this subject the interested reader is referred to refs. (5) and (7).

Bibliography

1. Jenkins, G.M. and D.G. Watts (1968) *Spectral Analysis and Its Applications.* Holden-Day, San Francisco.

2. Lumley, J.L. and H.A. Panofsky (1964) *The Structure of Atmospheric Turbulence,* Interscience, New York.

3. Tennekes, H. and J.L. Lumley (1972) *A First Course in Turbulence,* MIT Press, Cambridge.

4. Lumley, J.L. (1970) *Stochastic Tools in Turbulence,* Academic Press, New York.

5. Van Atta, C.W. (1974) Sampling Techniques in Turbulence Measurements, in *Annual Review of Fluid Mechanics,* Vol. 6, Academic Press, Palo Alto (Van Dyke, Vincenti, and Wehausen, coeditors).

6. Blackman, R.B. and J.W. Tukey (1959) *The Measurement of Power Spectra,* Dover Publications, Inc., New York.

7. Enochson, L.D. and R.K. Otnes (1968) *Programming and Analysis for Digital Time Series Data,* The Shock and Vibration Monograph Series, SVM-3, Naval Publication and Printing Office, Washington.

8. Pierce, R.E. (1972) Spectral Analysis by Large Block Transforms, Ph.D. Dissertation, Department of Aerospace Engineering, The Pennsylvania State University.

9. Gaster, M. and J.B. Roberts (1975) Spectral Analysis of Randomly Sampled Signals, *J. Inst. Maths. Appl.,* 15, 195-216.

10. Gaster, M. and J.B. Roberts (1977) The spectral analysis of randomly sampled records by a direct transform, *Proc. R. Soc. Lond A.,* 354, 27-58.

11. Buchhave, P., W.K. George and J.L. Lumley (1979). The measurement of turbulence with the laser Doppler anemometer, *Annual Reviews in Fluid Mechanics,* vol. 11 (Van Dyke and Wehausen, editors).

12. Buchhave, P. (1979) Errors and Correction Methods in Turbulence Measurements with the LDA, Ph.D. Dissertation, Department of Mechanical Engineering, State University of New York at Buffalo.

13. George, W.K. (1976) Limitations to measurement accuracy inherent in the laser Doppler signal, in *The Accuracy of Flow Measurements by Laser Doppler Methods,* Copenhagen, (P. Buchhave et al., editors).

14. Stevenson, W. (ed) (1974) *Proceedings of the Second International Workshop on Laser Velocimetry,* Purdue University, W. Lafayette, Ind.

15. Buchhave, P. and W.K. George (1978) Bias corrections in turbulence measurements by the laser Doppler Anemometer, *Proceedings of the Third Workshop on Laser Doppler Anemometry,* Purdue University, W. Lafayette, Ind. (W. Stevenson, ed.).

16. Akaike, H. (1972) "Use of an Information Theoretic Quantity for Statistical Model Identification. *Proceedings of the Fifth Hawaii International Conference on System Sciences,* Western Periodicals, Inc. pp.249-250.

17. Parzen, E. (1974) "Some Recent Advances in Time Series Analysis: *IEEE Trans, Auto. Control,* Vol. AC-19 pp. 723-730.

18. Parzen, E. (1974) "Some Recent Advances in Time Series Analysis: SUNY/B Department of Statistical Science Technical Report No. 10.

19. Parzen, E. (1975) "Multiple Time Series: Determining the Order of Approximating Autoregressive Schemes" SUNY/B Department of Statistical Science Technical Report No. 23.

20. Pagano, M. (1976) Time Series Analysis Course Notes, Department of Statistical Science, SUNY/B.

21. Parzen, E. (1976) Multivariate Time Series Analysis Course Notes, Department of Statistical Science, SUNY/B.

Appendix I: Autoregressive Spectral Estimation

by

Paul D. Beuther
State University of New York at Buffalo

Introduction
One of the major difficulties in estimating the spectrum of a turbulent process is reducing the variance in the spectral level. This variance is a statistical consequence of the Fourier transformation from time domain to frequency domain. Consider a stationary process, $Y(t)$, which has a "true" spectrum, $F(\omega)$. It can be shown that for a finite time segment, $Y(t), t=0, T$, the sample spectral density $F_T(\omega)$ is unbiased ($\overline{F_T(\omega)} \to F(\omega)$), but it is not consistent (VAR $F_T(\omega) \neq 0$). In fact, $VAR(F_T(\omega)) \sim [F(\omega)]^2$! Without some type of smoothing, a typical spectrum appears to be mostly noise.

Since the expected value for $F_T(\omega) \to F(\omega)$, one could certainly average many independent spectral realizations to obtain a better estimate for $F(\omega)$. However, in many cases it is impractical to obtain and analyze the necessary amount of data. In some experiments many realizations of the same event are not possible. In these situations other techniques are needed. Due to recent advances, autoregressive techniques are some of the most powerful methods available.

Theory of Autoregressive Spectra
A time series, $Y(t)$, can be represented as a sum of its weighted past plus some random "shock", $\epsilon(t)$. This can be represented mathematically as

$$Y(t) = a^*(1) \, Y(t-1) + a^*(2) \, Y(t-2) + \ldots + \epsilon(t). \qquad [1]$$

However, it is more common to define the process in the following manner:

$$Y(t) + \sum_{j=1}^{\infty} Y(t-j) \, a(j) = \epsilon(t) \text{ where } a(j) = -a^*(j). \qquad [2]$$

In practice there is an order p such that, for $j>p$, $a(j) \cong 0$. This is called a pth order autoregression and is written as:

$$Y(t) + \sum_{j=1}^{p} Y(t-j) \, a(j) = \epsilon(t). \qquad [3]$$

Techniques for determining p will be discussed in detail later. Since $a(j) = 0$ for $j<0$ and $j>p$, equation 3 can be rewritten as:

$$\sum_{j=-\infty}^{+\infty} a(j) \, Y(t-j) = \epsilon(t) \text{ where } a(0) \equiv 1.$$

[4]

The LHS is merely a convolution of a and Y which can be transformed to a frequency domain to read

$$F_a(\omega) \cdot F_Y(\omega) = F_\epsilon(\omega).$$

[5]

$F_a(\omega)$, $F_Y(\omega)$, and $F_\epsilon(\omega)$ are the spectral density functions for $a(j)$, $Y(t)$ and $\epsilon(t)$ respectively. From this it is easy to obtain an expression for the spectrum of Y:

$$F_Y(\omega) = \frac{F_\epsilon(\omega)}{F_a(\omega)}.$$

[6]

$\epsilon(t)$ is a random process, so $F_\epsilon(\omega)$ is a constant equal to the residual variance $\hat{\sigma}^2$. By substituting for $F_a(\omega)$ and $F_\epsilon(\omega)$ equation 6 becomes:

$$F_y(\omega) = \frac{\hat{\sigma}^2}{\left| \sum_{j=0}^{p} a(j)e^{ij\omega} \right|^2}$$

[7]

Calculating the a's for a pth Order Autoregression

The set of coefficients, $a(j)$, can be calculated from the covariance function $R(\tau)$. Transforming equation 4 into time-lag products results in the set of equations known as the Yule-Walker equations.

[8]

$$\sum_{j=0}^{p} R(\tau-j) \, a(j) = \hat{\sigma}^2 \delta(\tau) \qquad \tau = 0, 1, 2, \ldots p \text{ where } R(\tau) \equiv \frac{1}{T} \sum_{t=0}^{T-\tau} Y(t) \, Y(t-\tau).$$

This is a set of p linear equations with p unknowns, and can be solved by one of several means in a straightforward manner.

Deciding the Order p

There are two methods for determining the order p. The first was developed by Akaike (16), the second by Parzen (17). Both yield approximately the same result, but the theory behind each is quite different in approach. The Akaike method assumes that a given time series is a product of an autoregressive process with a "true" order p. Choosing p reduces to minimizing the function AIC(m), defined by

$$\text{AIC}(m) = \text{Ln}(\hat{\sigma}^2_m) + \frac{2m}{T}.$$

[9]

m is order, T the total number of points, and $\hat{\sigma}^2_m$ the residual variance for order m. The Parzen approach assumes that the given time series is actually the product of an infinite autoregressive process. The goal is to find the best estimate of this infinite series, which can be reduced to minimizing the function CAT (m), defined by

$$CAT(m) = 1 - \frac{\hat{\sigma}^2_\infty}{\sigma^2_m} + \frac{m}{T}. \tag{10}$$

m is the order, $\sigma^2_m = \frac{T}{T-m} \hat{\sigma}^2_m$ is the "unbiased" estimator of the residual variance and $\hat{\sigma}^2_\infty$ is the residual variance of an infinite order autoregression. Substituting for $\hat{\sigma}^2_\infty$ equation [10] can be rewritten as:

$$CAT(m) = \frac{1}{T} \sum_{j=1}^{m} \left(\frac{T-j}{T} \frac{1}{\hat{\sigma}^2_j} \right) - \frac{T-m}{T} \frac{1}{\hat{\sigma}^2_m}. \tag{11}$$

In practice the two methods give approximately the same order. It can be shown that the Akaike method is an upper bound of the Parzen criterion.

Extension of autoregressive techniques to multidimension time series is straightforward (see refs. (16) - (21)). However, the solution to the regression equations is more involved and may require more computer memory than some laboratory mini- and micro-computers have.

Appendix II:
Computation of Spectra and Cross Spectra by a Skipping Technique

by

John L. Lumley
Cornell University

It is straightforward to show (Lumley (4), for example) that if the mean of a quantity is computed by integrating in time over an interval of length T, the mean square relative error is given by (asymptotically)

$$\epsilon^2 \cong \frac{2T}{T} a^2 \tag{1}$$

where a^2 is the relative fluctuation level of the quantity being measured, and T is the integral scale. On the other hand, if N *independent* estimates of the same quantity are made, the mean square relative error is given by

$$\epsilon^2 \cong \frac{a^2}{N} \tag{2}$$

Comparison of [1] and [2] indicates that the interval T may be considered to contain $T/2T = N$ independent estimates. Evidently one point in each interval of length 2T will produce just as fast a convergence of the statistics. The extra points included in the integral over 2T do not improve the convergence, since they are not independent — they do not give new information.

This concept may be immediately applied to the measurement of spectra. Suppose it is desired to measure the spectrum of a process having no energy above a frequency $1/\eta$. The lowest frequency desired is $1/\tau$, where it is presumed that $\tau \geqslant T$. A tape of length T is taken, sufficient to provide the required accuracy by [1] and [2], and digitized at the Nyquist frequency $2/\eta$, providing $2T/\eta$ points. At each lag, $T/2T$ multiplications are required, followed by $T/2T$ additions, or T/T operations; there are $2\tau/\eta$ lags, for a total of $2\tau T/T\eta$ operations to produce the correlation. This must now be transformed by a fast Fourier transform of $2\tau/\eta$ points. If we indicate the number of operations required for this by $FFT(2\tau/\eta)$, the total number of operations is

$$2\tau T/T\eta + FFT(2\tau/\eta) \tag{3}$$

This may be compared with other ways of obtaining the spectrum. For example, each block of $2\tau/\eta$ points could be transformed by FFT, requiring

$(T/\tau)FFT(2\tau/\eta)$ operations; then each Fourier coefficient must be squared, re-quiring $2T/\eta$ operations; then the squared coefficients at each frequency must be added together, requiring $(2\tau/\eta)(T/\tau) = 2T/\eta$ operations. This gives a total of

$$4T/\eta + (T/\tau)FFT(2\tau/\eta) \qquad [4]$$

operations.

Still another way would involve taking the FFT of the entire tape, $FFT(2T/\eta)$, squaring the coefficients obtained, for $2T/\eta$ operations, and adding each group of T/τ together to form $2\tau/\eta$ block averages, requiring $2T/\eta$ operations, for a total of

$$4T/\eta + FFT(2T/\eta)$$

We may compare these by writing directly

$$\frac{\tau}{2T} \cdot \frac{4T}{\eta} + FFT(2\tau/\eta) \qquad [3]$$

$$\frac{4T}{\eta} + \frac{T}{\tau} FFT(2\tau/\eta) \qquad [4]$$

$$\frac{4T}{\eta} + FFT((T/\tau)(2\tau/\eta)) \qquad [5]$$

Now, ordinarily $\tau/2T \sim 1$, while $T/\tau \sim 10^2$, so that it is evident that method [3] is the cheapest. The ratio can be estimated from the expression $FFT(n) \sim n \log_2 n$, from which approximate expressions for [3], [4] and [5] may be derived; setting $\tau/2T = C$, we have

$$C \frac{4T}{\eta} + \frac{2\tau}{\eta} \log_2 \frac{2\tau}{\eta} \sim C \frac{4T}{\eta} \qquad [3]'$$

$$\frac{4T}{\eta} + \frac{T}{\tau} \cdot \frac{2\tau}{\eta} \log_2 \frac{2\tau}{\eta} \qquad [4]'$$

$$\frac{4T}{\eta} + \frac{T}{\tau} \frac{2\tau}{\eta} \log_2 \frac{2\tau}{\eta} + \frac{2T}{\eta} \log_2 \frac{T}{\tau} \qquad [5]'$$

If C = 1, the relative cost differences are

$$\frac{[4]' - [3]'}{[3]'} = 1/2 \log_2 \frac{2\tau}{\eta} \qquad [6]$$

$$\frac{[5]' - [3]'}{[3]'} = 1/2 \log_2 \frac{2T}{\eta} \qquad [7]$$

In turbulent flows τ/η ranges from 10 to 10^3; T/τ will ordinarily be at least 10^2 (for 10 % accuracy); thus $2 \times 10^3 \geqslant 2\tau/\eta > 20$, $2 \times 10^5 \geqslant 2T/\eta \geqslant 2 \times 10^3$. Thus roughly,

$$2 \leqslant \frac{[4]' - [3]'}{[3]'} \leqslant 5.5 \tag{8}$$

$$5.5 \leqslant \frac{[5]' - [3]'}{[3]'} \leqslant 9 \tag{9}$$

and the right-hand side of [9] may be even higher if greater accuracy is desired.

Hence, under the worst circumstances, the proposed method, leading to [3], is 1/3 the cost of the next best method, and may be far cheaper. It should be noted that tape reading times and storage requirements are the same for all methods; the proposed method gives a bonus, in the form of the correlation. This would only be obtained at extra cost in the other methods.

It must be noted that there may be other reasons for computing the Fourier coefficients than obtaining the spectrum, such as determining higher order statistics of the Fourier coefficients. In such a case, there is, of course, no substitute for the FFT. In addition, if it is desired to compute correlations to very long lags ($C \gg 1$), corresponding to very low frequencies★, the relative costs change, and [4]' may be less expensive than [3]' (although it is clear that [5]' will always be more expensive than [4]').

★ a typical need in calculating derivative spectra.

Acknowledgements