June 1966

USCEE Report 171

UNIVERSITY OF SOUTHERN CALIFORNIA

SCHOOL OF ENGINEERING

ESTIMATES OF THE STATISTICS OF RANDOMLY VARYING PARAMETERS OF LINEAR SYSTEMS

S. M. Brainin

G. A. Bekey

ELECTRONIC SCIENCES LABORATORY

Engineering

June 1966

USCEE Report 171

ESTIMATES OF THE STATISTICS OF RANDOMLY VARYING PARAMETERS OF LINEAR SYSTEMS

S. M. Brainin

G. A. Bekey

Department of Electrical Engineering
University of Southern California
Los Angeles, California

This research was supported in part by the National Aeronautics and Space Administration under Grant No. NGR-05-018-022.

TABLE OF CONTENTS

ABSTRACT

This paper describes an approach to the identification of linear dynamic systems with random parameters. A technique, based on the Fokker-Planck equation, is developed for estimating the statistics of randomly varying parameters in systems whose differential equation is known a priori. The resulting mathematical model then applies to a family of similar systems rather than to an individual.

The Fokker-Planck equation is applied to a first order linear system with a random (Gaussian) parameter. First order differential equations for the moments of the system output are generated and used to estimate the first and second moments of the random parameters. The identification procedure is based on regression techniques.

The results of an application of the technique to a specific first order system are presented with a theoretical discussion of the accuracy of the identification procedure. The concluding section of the paper discusses the extension of the techniques to higher order linear and non-linear systems.

1. INTRODUCTION

The identification problem is concerned with the determination of a mathematical relationship which describes the input-output behavior of an unknown system. A large class of identification procedures is concerned with systems where the form of the mathematical model is assumed to be known a priori and only parameters values are unknown. In such cases, the identification is reduced to finding the values of the parameter set which minimize an appropriate functional of the difference in performance of the model and system. Typical of such procedures are relaxation methods, gradient techniques and random search algorithms [1, 2, 3]. Techniques of this type have been applied to such diverse problems as identification of the response characteristics of a human operator in a manual control task or the identification of the parameters of a malfunctioning system for diagnostic purposes.

However, it must be noted that the class of identification techniques outlined above is primarily suited to systems with deterministic parameters. Uncertainties are normally assumed to be present only in the measurement of system outputs. When, as in the case of the human operator modeling problem cited above, it is suspected that the system has randomly varying parameters the identification concentrates on the mean values of the parameters and not on their distribution. The randomness is accounted for by resort to Monte Carlo techniques, involving a large number of repeated experiments. This paper presents an approach to the extension of the present theory to the identification of the statistics of random parameters in a single experiment.

The function of identification takes on a broader aspect if the end result represents a class of systems rather than an individual. Such is the case for random parameters. By identifying the statistics of the parameters, the resulting model applies to a family of similar systems. Alternately, many individual systems cannot be defined in terms of deterministic parameters. A case in point is the biological system (e.g., the human operator) whose parameters are not constant when observed over any given period of time, but can be represented to a closer degree of reality by means of statistical parameters. The research reported here is devoted to the problem of random parameter identification.

In recent years two approaches to noise theory have developed. The first method is heavily dependent on the concept of stationarity and the spectrum of the random process. This technique has been employed by Rice and is known as Rice's [4] method. The second method is the method of Fokker-Planck or the diffusion process method [6]. Historically, this latter method is based on the theory of Brownian motion where the path of a particle of colloidal size in a viscous fluid is perturbed by molecular collisions.

In applying this latter technique we seek the probability distributions for a random variable whose mean has a bandwidth very much smaller than the disturbing and superimposed noise. So, if a process $y(t)$ is defined as $y(t) = S(t) + n(t)$, the sum of signal and noise, then it can be shown for any small interval of time Δt, the first moment, $\overline{y}(t) = S(t)$, the second moment $y^2(t) = 0(\Delta t)$, and all higher moments vanish in the limit as the time increment Δt approaches zero. Simply this means that in the time Δt, the noise can fluctuate quite rapidly, but $y(t)$ changes in the mean only as $S(t)$, and only small perturbations of $y(t)$, about $\overline{y}(t)$ can occur in small intervals of time Δt. The actual course which a Brownian particle will take depends only on the instantaneous values of its physical parameters and is entirely independent of its whole previous history. Thus, the path of a Brownian particle falls in the classification of a Markhoff process, with independent increments.

There are many noise processes which occur in nature, such as thermal and shot noise, which when processed through a system, have the macroscopic properties of the diffusion process. Thus, the method of Fokker-Planck can be abstracted and applied in general to stochastic variables which fit this rather broad model. In this paper the Fokker-Planck equation is derived for a one dimensional random process. It is then applied to a system described by a linear differential equation containing a random parameter. As a result of this application, first order differential equations for the various moments of the system output are generated. These are then used to identify the first and second moments of the random parameter.

2. THEORY

Let $y(t)$ be a random Markhoff process and $W_2(y_o/y, t)$ its second order conditional probability density function, where y_o is the value of y at $t = 0$. Then, for an incremental change in time Δt this density function can be described as

$$W_2 (y_o/y, t + \Delta t) = \int_{-\infty}^{\infty} W_2 (y_o/z, t) W_2 (z/y, \Delta t) dz \qquad (1)$$

where $z = y - \Delta y$ and Δy is the increment in $y(t)$. If only small changes in y can occur in the time Δt this density function also satisfies the so-called Fokker-Planck differential equation

$$\frac{\partial W_2}{\partial t} = - \frac{\partial}{\partial y} [A_1 (y) W_2] + \frac{1}{2} \frac{\partial^2}{\partial y^2} [A_2(y) W_2] \qquad (2)$$

where the $A_n(y)$ are the conditional moments of the variations about the expected value of $y(t)$ at any instant t, as defined by

$$A_n(y) = \lim_{\Delta t \to 0} \frac{1}{\Delta t} \int_{-\infty}^{\infty} (\Delta y)^n W_2 (y/y + \Delta y, \Delta t) dy \qquad (3)$$

and $A_n(y) = 0$ for all $n > 2$.

A brief derivation of the Fokker-Planck equation is given in Appendix A.

Assume now that the system to be identified is defined by a first order differential equation:

$$\dot{y}(t) + B(t)y(t) = x(t) \qquad (4)$$

where $x(t)$ is the driving function, $B(t)$ is the random parameter, and $y(t)$ is the system output. It is assumed that $B(t) = b + n(t)$ where b is a constant and $n(t)$, is wide bandwidth Gaussian noise with zero mean and variance σ^2.

The coefficients of the Fokker-Planck equation $A_1(y)$, $A_2(y)$ defined in equation (3) are derivable from equation (4). Taking a small increment in time Δt, so that y(t) can be approximated by $\Delta y/\Delta t$, we obtain the approximate relations

$$\Delta y \cong x\,\Delta t - By\,\Delta t$$

$$\Delta y^2 \cong + B^2 y^2 \Delta t^2 - 2By\,x\,(t)\,\Delta t^2 + x^2\,\Delta t^2$$

Using the definition of the moments $A_1(y)$ and $A_2(y)$ from (3):

$$A_1(y) = \lim_{\Delta t \to 0} \left[\frac{\overline{x\,\Delta t - By\,\Delta t}}{\Delta t} \right] = -by + x \tag{5a}$$

$$A_2(y) = \lim_{\Delta t \to 0} \left[\frac{1}{\Delta t} \overline{(by\,\Delta t + ny\,\Delta t)^2} - \frac{1}{\Delta t} \overline{(2byx\,\Delta t^2 + 2nyx\,\Delta t^2)} + \frac{1}{\Delta t} \overline{(x\,\Delta t)^2} \right]$$

$$A_2(y) = \lim_{\Delta t \to 0} \left[x^2\,\Delta t - 2bxy\Delta t + b^2 y^2\,\Delta t + \iint_t^{t+\Delta t} \frac{\overline{n(t_1)\,n(t_2)}\;y(t_1)\;y(t_2)}{\Delta t} \right] dy_1 dy_2 = y^2 \sigma^2$$

since $\overline{n(t_1)\,n(t_2)} = \sigma^2\,\delta(t_2 - t_1)$. See Reference 4.

It can also be shown that $A_n = 0$ for all $n > 3$. See Appendix B.

The corresponding Fokker-Planck equation is therefore, from equation (2):

$$\frac{\partial W}{\partial t} = -\frac{\partial}{\partial y} \Big((- by + x)\,W \Big) + \frac{1}{2}\,\frac{\partial^2}{\partial y^2} \Big(\sigma^2\,y^2\,W \Big) \tag{6}$$

Where W is understood to mean $W_2\,(y_o/y,\,t)$.

The moments of y can now be defined in terms of W:

$$\overline{y^n} = \int_{-\infty}^{\infty} y^n \; W \; dy \tag{7}$$

This equation can be differentiated with respect to time to yield

$$\dot{\overline{y^n}} = \int_{-\infty}^{\infty} \frac{\partial y^n}{\partial t} W \, dy + \int_{-\infty}^{\infty} y^n \frac{\partial W}{\partial t} dy \tag{7a}$$

Now $\frac{\partial y^n}{\partial t} = n \, y^{n-1} \frac{\partial y}{\partial t} \cong \lim_{\Delta t \to 0} ny^{n-1} \frac{\Delta y}{\Delta t}$

and the first integral on the right of e.g. (7a) can be written as

$$\int_{-\infty}^{+\infty} \frac{\partial y^n}{\partial t} W \, dy = E \left\{ ny^{n-1} \lim \frac{\Delta y}{\Delta t} \right\} = n \lim_{\Delta t \longrightarrow 0} E \left\{ y^{n-1} \frac{\Delta y}{\Delta t} \right\}.$$

The response in y is of the Brownian motion type and is a Markhoff process with independent increments. Hence $E(y^k \Delta y) = 0$ for any finite k. Thus

$$\dot{\overline{y^n}} = \int_{-\infty}^{\infty} y^n \frac{\partial W}{\partial t} dy. \tag{8}$$

We can substitute for $\frac{\partial W}{\partial t}$ from the Fokker-Planck equation (6), yielding

$$\dot{\overline{y^n}} = \int_{-\infty}^{\infty} y^n \left\{ \left(-\frac{\partial}{\partial y} (-by + x) W \right) + \frac{\sigma^2}{2} \frac{\partial^2}{\partial y^2} (y^2 W) \right\} dy$$

$$\dot{\overline{y^n}} = \int_{-\infty}^{\infty} \frac{\partial}{\partial y} (by \; W - x \; W) \, dy + \frac{\sigma^2}{2} \int_{-\infty}^{\infty} y^n \frac{\partial^2}{\partial y^2} (y^2 \; W) \, dy$$

Let $I = \int_{-\infty}^{\infty} y^n \frac{\partial}{\partial y} (by\, W - x\, W)\, dy = y^n (by\, W - x\, W) \Big|_{-\infty}^{+\infty} - n \int_{-\infty}^{\infty} y^{n-1} (by\, W - xW)\, dy$

Now, by assumption $W(\pm \infty) = 0$ and $\frac{\partial W}{\partial y} (\pm \infty) = 0$, and

$$\therefore \; I = nb\, \overline{y^{n-1}}. \quad \text{Reference 5} \tag{9}$$

Similarly,

$$II = \int_{-\infty}^{\infty} y^n \frac{\partial^2}{\partial y^2} (y^2 W)\, dy = y^n \frac{\partial}{\partial y} (y^2 W) \Big|_{-\infty}^{\infty} - \int_{-\infty}^{\infty} n y^{n-1} \frac{\partial}{\partial y} (y^2 W)\, dy,$$

which reduces to

$$II = n(n-1)\, \overline{y^n} \tag{10}$$

Hence, the moments of the system output satisfy the differential equations

$$\dot{\overline{y^n}} = -nb\, \overline{y^n} + nx\, \overline{y^{n-1}} + \frac{\sigma^2}{2} n(n-1)\, \overline{y^n} \tag{11}$$

subject to the initial condition $y^n(0)$ for all n.

Therefore, the first two moment equations become

$$\dot{\overline{y}} + b\, \overline{y} = x \tag{12}$$

$$\dot{\overline{y^2}} + (2b - \sigma^2)\, \overline{y^2} = 2\, x\, \overline{y} \tag{13}$$

These are the first order differential equations for the first and second moments of the system's output. In an actual application of the foregoing theory, a mathematical model will be assumed which has the same differential equation as the system (equation 4) but whose parametric values are unknown. In that case, b and σ^2, the mean and variance of the random parameter are unknown. The identification of these parameters will be demonstrated in the discussion to follow.

3. IDENTIFICATION

Since the model differential equation is assumed to have the same form as the system, the differential equations for the moments of the model output (derivable from the Fokker-Planck equation) must also have the same form as equations (12) and (13). Therefore, let the identification model consist of the following differential equations:

$$\dot{m}_1 + C_1 m_1 = x \tag{14}$$

$$\dot{m}_2 + (2C_1 - C_2) m_2 = 2x\, m_1 \tag{15}$$

where m_1 is the estimator of $\bar{y}$, with assumed initial value m_{10}

m_2 is the estimator of $\overline{y^2}$, with assumed initial value m_{20}

C_1 is the estimator of b

and C_2 is the estimator of σ^2

The driving function x will be common to both the system and model equations.

The procedure consists of first identifying C_1 in equation (14) by means of a discrete regression technique, and then computing C_2 from equation (15). Questions of convergence for the regression identification algorithm are discussed in Reference 7.

For the determination of C_1, both the system (Equation 4) and model (Equation 14) response functions were sampled with a sampling interval ΔT. This was established by the time increment employed for the integration algorithm for the differential equations.

The noise component of the parameter B was generated by a digital Gaussian random number generator, whose samples are mutually independent. However, an effective bandwidth (2W cps) was superimposed on the noise samples by the sampling interval. Since the total noise power is σ^2, the time duration of the sampled signal is T seconds, and the bandwidth is 2W, then from sampling theory

$$\sigma^2 = \frac{1}{2WT} \sum_{k=1}^{2WT} n_k^{\,2} \tag{16}$$

where the n_k are the noise samples. Thus $2WT = N$ where N is the number of samples, and the effective sampling rate is $2W = N/T = \frac{1}{\Delta T}$ samples/second. The total noise power is distributed over the bandwidth of 2W cps. Hence, the spectral density in this region is $\sigma^2/2W$.* Since 2W is

---

* The estimator for $\frac{\sigma^2}{2W}$ will be denoted by $C_2^{\,1}$ in the rest of the paper, to distinguish it from C_2 (the estimator of σ^2 alone).

much larger than $b/2\pi$, this noise looks like white noise* to the system with a correlation function.

$$R(\tau) \cong \frac{\sigma^2}{2W}\,\delta(\tau). \quad \text{See Reference 8.}$$

The identification was based on the following criterion function:

$$\varphi(\bar{y};\, C_1) = \sum_{i=1}^{N} (m_{1i} - \bar{y}_i)^2 = \sum_{i=1}^{N} e_i^{\,2}$$

where $\bar{y}_i$ and m_{1i} are the corresponding samples of the system's sliding average response and the first moment model equation response, equation (14).

---

* The white noise character of the multiplicative noise, as experienced by the system can be described in the following heuristic way:

The differential equation of the system can be written in the following form:

$$\dot{y} + (b + n)\, y = x$$

where n is the wide band additive noise. Now transposing the product ny to the right hand side of this equation

$$\dot{y} + by = x - ny$$

If for the moment x is considered zero, ny is the driving function to a first order linear differential equation. The bandwidth of ny is the convolution in frequency of the bandwidth of n, and the bandwidth of y.

The bandwidth of y is restricted to $b/2\pi$. The bandwidth of ny is thus slightly larger than the bandwidth of n. In passing through the first order differential equation, the bandwidth of ny is reduced to the bandwidth of y again. Thus, if the bandwidth of $n \gg b/2\pi$ the bandwidth of $ny \gg b/2\pi$ and ny is effectively white noise to the system.

The change of φ with respect to C_1, is given by

$$\frac{\partial \varphi}{\partial C_1} = \sum_{i=1}^{N} e_i \quad \frac{\partial e_i}{\partial C_1} = 2 \sum_{i=1}^{N} e_i \left(\frac{\partial m_1}{\partial C_1}\right)_i \tag{18}$$

since only the model equation is sensitive to changes in C_1. In order to find stationary point for φ, with respect to C_1, an increment ΔC_1 must be found to reduce $\frac{\partial \varphi}{\partial C_1}$ to zero. Let the incremented value $C_{11} = C_{10} + \Delta C_1$ and the corresponding model output be approximated as

$$m_{11} = m_{10} + \Delta C_1 \frac{\partial m_1}{\partial C_1}$$

where C_{10} and m_{10} are the initial value of C_1 and m_1 respectively. Now substituting these expressions into (18), and equating it zero, the parameter increment ΔC can be found:

$$0 = \sum_{i=1}^{N} \left(m_{1i} + \Delta C_1 \left(\frac{\partial m_1}{\partial C_1}\right)_i - \bar{y}_i\right) \left(\frac{\partial m_1}{\partial C_1}\right)_i \tag{19}$$

and

$$\Delta C_1 = - \frac{\sum_{i=1}^{N} e_i \left(\frac{\partial m_1}{\partial C_1}\right)_i}{\sum_{i=1}^{N} \left[\left(\frac{\partial m_1}{\partial C_1}\right)_i\right]^2}$$

Due to the first order approximations employed, this process does not lead to a minimum φ in a single trial and must be repeated several times, so that after the j-th iteration

$$C_1^{(j+1)} = C_1^{(j)} + \Delta C_1^{(j)}$$

The partial derivative of m_1 with respect to C_1 is a sensitivity coefficient or influence coefficient (9) derived in the following manner. Differentiate equation (14) with respect to C_1

$$\frac{\partial}{\partial C_1}(\dot{m}_1) + \frac{\partial}{\partial C_1}(C_1 m_1) = \frac{\partial}{\partial C_1} x$$

and let the sensitivity function X_1 be defined by

$$X_1 \overset{\Delta}{=} \frac{\partial m_1}{\partial C_1} .$$

Then, assuming that m_1 is continuous in both C_1 and t_1 the order of differentiation can be interchanged and

$$\dot{X}_1 + bX_1 = -m_1, \quad X_1(0) = 0 \tag{20}$$

This equation was programmed with the system equation (4) and regression equations (17), (19) in order to compute the increments $\Delta C^{(j)}$. The computation procedure consisted of the following steps:

1. For a given pair of values of b and σ^2 a set of k system response samples are derived from the system equation. In addition the mean and variance of this data is computed, i.e., $\bar{y}$, σ_y^2.
2. For an estimated value of C_1, the first moment m_1 is generated by means of the first model equation, equation (14).
3. Simultaneously, the sensitivity factor X_1 is generated by means of equation (20).
4. The increment $\Delta C_1^{(j)}$ is then computed so as to reduce the criterion function, equation (19).
5. This process is repeated until the criterion function is minimized.
6. The second moment C_2 is then computed from equation (15), in steady state knowing, C_1, σ_y^2, $\bar{y}^2$ and 2W, as follows: In steady state, for a step function input

$$m_2 \cong \frac{2\ A\ m_1}{2C_1\ (1 - C_2'/2C_1')}$$

and

$$m_1 \cong A/C_1, \text{ where } x(t) = A$$

Further, $C_2/2C_1 >> 1$, so that the expression for m_2 can be written as

$$m_2 = m_1^{\,2}\,(1 + C_2'/2C_1).$$

However, when φ is minimum $m_1 \cong \bar{y}$, and the condition required to yield C_2' is $m_2 \cong \overline{y^2}$. Hence

$$\overline{y^2} = \bar{y}^2 + \bar{y}^2 \; \frac{C_2'}{2C_1}$$

yielding

$$C_2' = 2C_1 \frac{\sigma_y^{\,2}}{\bar{y}^{\,2}} = \frac{C_2}{2W}$$

or

$$C_2 = 4\,W\,C_1 \frac{\sigma_y^{\,2}}{\bar{y}^2} \tag{21}$$

where C_2 estimates the noise variance of the parameter B, i.e. σ^2.

Ordinarily, x(t) need not be a step but can be a stationary random process. In this case it can be shown (10) that the equivalent driving function is $\overline{x(t)}$, the mean value of the sample function of the ensemble exciting the model. $\overline{x(t)}$ is a step function and the preceding analysis applies.

4. EXPERIMENTAL RESULTS AND ERROR ANALYSIS

4.1 ESTIMATION OF THE MEAN

The error in identifying b, the average value of the unknown parameter, consists of a precision error that is due to the experiment and a bias error, due to the noise. The latter is due to the statistical fluctuation of the disturbing noise, resulting in a non-zero mean for a finite sample space.

After the transient has become negligible the criterion function employed for the regression can be approximated in the steady state by

$$\Phi_{ss} = \sum_{j=1}^{k} \left(\frac{A}{C_1} - \overline{\left(\frac{A}{B}\right)} \right)^2 \tag{22}$$

where the term under the summation sign is independent of the index j and

A is the magnitude of the step input

C_1 is the estimated value of b

$B = b + n$

and n is the additive Gaussian noise with zero mean and variance σ^2 and k noise samples were employed.

$\overline{(A/B)}$ can be considered as $\frac{A}{b + \overline{n}}$. For an infinite sample space $E(n) = 0$, but for a finite sample population $\overline{n}$ has some non-zero value. Hence, the effective b is the algebraic sum of the actual parameter value and the sample mean of the noise.

Let $b' = b + \overline{n}$

For any one run, all the terms under the summation are constant; hence,

$$\Phi_{ss} = \sum_{j=1}^{k} \left(\frac{A}{b'}\right)^2 \frac{e_1'^2}{C_1^2} = \left(\frac{A}{b'}\right)^2 \frac{e_1'^2}{C_1^2} \tag{23}$$

where $e_1' = b' - C_1$

k = total number of error samples

$\overline{\Phi}_{ss}$ = measured value of the criterion function

The value of $e_1 = C_1 - b = -e_1' + \overline{n}$ can be found from equation (23) where $\overline{y} = A/b'$.

$$e_1 = \overline{n} - \frac{C_1}{\overline{y}} \sqrt{\frac{\overline{\Phi}_{ss}}{k}} \tag{24}$$

and $$\frac{C_1}{b} - 1 = \frac{\overline{n}}{b} - \frac{C_1}{b} \frac{1}{\overline{y}} \sqrt{\frac{\overline{\Phi}_{ss}}{k}}$$

Thus

$$\frac{C_1}{b} = 1 + \frac{\overline{n}}{b} - \frac{1}{\overline{y}} \sqrt{\frac{\overline{\Phi}_{ss}}{k}} + \frac{\overline{n}}{b\overline{y}} \sqrt{\frac{\overline{\Phi}_{ss}}{k}} \tag{25}$$

Experimental results of C_1/b are plotted in Figure 1, as a function of σ^2, the actual noise variance. For small values of k, the precision error

$$\frac{1}{\overline{y}} \sqrt{\frac{\overline{\Phi}_{ss}}{k}}$$

is predominant. This can be observed from the negative slope of the curves for $k = 89$ points, since $\overline{\Phi}_{ss}$ will increase with increasing noise. However, as k increases the error terms are reduced, as is apparent from Figure 1. The cross-coupling term is small relative to the other terms and can be neglected. The variance of the error between the actual and estimated values of the first moment can be found from equation (24). Since these errors can be considered independent, the variance of e_1 is expressed by:

$$\text{Var } e_1 = \text{Var } \overline{n} + \text{Var}\left(\frac{C_1}{\overline{y}} \sqrt{\frac{\overline{\Phi}_{ss}}{k}}\right)$$

where

$$\text{Var}\ \bar{n} = \sigma^2/k$$

$$\text{Var}\left(\frac{C_1}{\bar{y}}\sqrt{\frac{\varphi_{ss}}{k}}\right) = \frac{C^2{}_1\ \varphi_{ss}}{k\ (\bar{y})^4}\ \frac{\sigma_y{}^2}{k} + \frac{C_1{}^2}{4k\varphi}\ \sigma_\varphi{}^2$$

and $\sigma_{\bar{\Phi}}{}^2$ = the variance of $\bar{\Phi}$ and $\sigma_y{}^2$ = the variance of y.

$$\text{Var}\ e_1 = \frac{\sigma^2}{k} + \frac{C_1}{k}\left(\frac{\varphi_{ss}\sigma^2{}_y}{k\bar{y}^4} + \frac{\sigma_\varphi{}^2}{4\varphi_{ss}}\right) \tag{27}$$

From equation (27) it is seen that this variation becomes small as k increases.

The largest experimental confidence interval for C_1 thus occurred for σ^2 = 20, b = 50 and k = 89. This represents an extreme case for the observed data.

Where C_1 = 49.33

$\sigma_y{}^2$ = .00352

estimated $(\sigma_{\bar{\Phi}}{}^2/\bar{\Phi})$ = .5

$\bar{\Phi}_{ss}$ = .05

$\bar{y}$ = 1.0116

The Var $e_1 \simeq .294$, its standard deviation is .543 and the 95 percent confidence interval for C_1/b is therefore $(C_1/b \pm .022)$. Thus the number of required observations lies between 89 and 489.

4.2 ESTIMATION OF THE PARAMETER VARIANCE

The estimation error for σ^2 can be determined with the aid of equation (21), which can be rewritten as

$$C_2 = \frac{2C_1}{\Delta T} \frac{\sigma_y^2}{\bar{y}^2} \tag{28}$$

where ΔT is the sampling interval, $\Delta T = \frac{1}{2W}$.

From the differential calculus the experimental fluctuation of C_2 can be expressed in terms of the fluctuations of C_1, σ_y^2 and $\bar{y}^2$ to a first order of approximation:

$$\Delta C_2 = \frac{\partial C_2}{\partial C_1} \Delta C_1 + \frac{\partial C_2}{\partial \sigma_y^2} \Delta \sigma_y^2 + \frac{\partial C_2}{\partial \bar{y}} \Delta \bar{y}$$

Then

$$e_2 = \Delta C_2 = \frac{2\sigma_y^2}{(\Delta T)\ \bar{y}^2} \Delta C_1 + \frac{2C_1}{\Delta T\ \bar{y}^2} \Delta\sigma_y^2 - \frac{4C_1\ \sigma_y^2}{\Delta T\ \bar{y}^3} \Delta\bar{y} \tag{29}$$

The increments in $\Delta\bar{y}$ and $\Delta\sigma_y^2$ will be in error due to statistical fluctuation, with equal likelihood of either sign. Furthermore, the errors due to $\Delta\sigma_y^2$ and $\Delta\bar{y}$ are due to the same source. This is also true for ΔC_1, as k gets large. Thus, a conservative estimate of the error C_2 is made by assuming the three contributing errors to be dependent. Where

$$\Delta C_1 = e_1 = \bar{n} - \frac{C_1}{\bar{y}} \sqrt{\frac{\bar{\Phi}_{ss}}{k}}$$

$$\Delta\bar{y} = \sigma_y / \sqrt{k}$$

Thus

$$\Delta\sigma_y^2 = \sqrt{\frac{3}{k}}\ \sigma_y^2$$

Proof: That $\Delta\sigma_y^{\ 2} = \sqrt{\frac{3}{k}}\ \sigma_y^{\ 2}$

$$\sigma_y^{\ 2} = \frac{1}{k} \sum_{j=1}^{k} (y_j - \bar{y})^2$$

$$\text{Var}\ \sigma_y^{\ 2} = \text{Var}\ \frac{1}{k} \sum_{j=1}^{k} (y_j - \bar{y})^2 = \frac{1}{k^2} \sum_{j=1}^{k} \text{Var}\ (y - \bar{y})^2$$

$\text{Var}\ (y_i - \bar{y})^2 = E\ (y_j - \bar{y})^4 = 3\ [E(y_j - \bar{y})^2]^2 = 3\sigma_y^{\ 4}$, by the gaussian assumption.

$$\text{Var}\ \sigma_y^{\ 2} = \frac{3}{k}\ \sigma_y^{\ 4} \quad \text{Q.E.D.}$$

Now

$$\frac{e_2}{C_2} = \frac{\sigma^2 - C_2}{C_2} = \frac{\sigma^2}{C_2} - 1 \tag{30}$$

or

$$\frac{\sigma^2}{C_2} = 1 + \frac{e_2}{C_2} = 1 + \left[\frac{\bar{n}}{C_1} - \frac{1}{y}\sqrt{\frac{\Phi_{ss}}{k}} + \sqrt{\frac{3}{k}} + \frac{2}{y}\ \frac{\sigma_y}{\sqrt{k}}\right]$$

Since the bracketed term is small compared to 1.

$$\frac{C_2}{\sigma^2} = 1 - \frac{\bar{n}}{C_1} + \frac{1}{y}\sqrt{\frac{\varphi_{ss}}{k}} - \sqrt{\frac{3}{k}} - \frac{2}{y}\ \frac{\sigma_y}{\sqrt{k}} \tag{31}$$

Equation (32) gives the expression for C_2/σ^2_{max}, the estimated variance normalized with the maximum σ^2.

$$\frac{C_2}{\sigma^2_{max}} = \left(1 - \frac{\bar{n}}{C_1} + \frac{1}{y}\sqrt{\frac{\Phi_{ss}}{k}} - \sqrt{\frac{3}{k}} - \frac{2\sigma_y}{\bar{y}\sqrt{k}}\right) \frac{\sigma^2}{\sigma^2_{max}} \tag{32}$$

Figure 2 illustrates some of the experimental results for C_2. The improvement with an increased number of sample points is again apparent. For large k,

$$\frac{C_2}{20} \simeq 1 - \frac{\bar{n}}{C_1} \frac{\sigma^2}{20}$$

Thus, the error in C_2 varies directly as $\bar{n}$ and inversely as C_1.

The variance of the estimated error for σ^2 can be found from equation (31).

$$\text{Var } e_2 = \text{Var} \left(- \frac{\bar{n}}{C_1} + \frac{1}{\bar{y}} \sqrt{\frac{\varphi_{ss}}{k}} - \sqrt{\frac{3}{k}} - \frac{2\sigma_y}{\bar{y}\sqrt{k}} \right) \sigma^2$$

$$\text{Var } (e_2/\sigma^2) = \left(\frac{\sigma^2}{C_1^{\,2} k} + \frac{\left(\sqrt{\varphi_{ss}} + 2\sigma_y\right)^2}{\bar{y}^4\, k^2} + \frac{\sigma_\varphi^{\,2}}{4k\, \bar{y}^2 \varphi_{ss}} + \frac{3\sigma_y^{\,2}}{\bar{y}^2\, k^2} \right) \tag{33}$$

For the extreme case used to illustrate the confidence interval for e_1, we find the standard deviation from equation 30, for $e_2/20$ to be .04. Hence the 95 percent confidence limits for $e_2/20$ are $(e_2/20 \pm .08)$ for $\sigma^2 = 20$, and the number of observations should be slightly greater than 489.

Conclusion

By means of a model derived from the Fokker-Planck differential equation, it was possible to set up a computation algorithm for the identification of the second order statistics of a random parameter. This was done for a first order linear system. However, the method is applicable in principle to any order linear system containing a finite number of finite parameters. It may also be possible to extend the method to certain classes of nonlinear systems by using approximation techniques to represent probability density and correlation functions. More work will be required to examine this hypothesis.

Although the identification technique employed herein was by means of regression, this is arbitrary. The method can be employed with any identification procedure, provided only that the model moment equations exist and are known.

Once the first moment or the average parameter value has been established, the method can be used to estimate any required number of the moments of the system's output, assuming only that the input process is stationary.

REFERENCES

1. Beckenbach, E. F., "Modern Mathematics for the Engineer," McGraw Hill, New York 1956, Chapter 17.

2. Eykhoff, Pieter, "Some Fundamental Aspects of Process Parameter Estimation," IEEE Transactions on Automatic Control, October 1963.

3. J. J. Munson and A. I. Rubin, "Optimization by Random Search on the Analog Computer," National Simulation Conference, Dallas Conference, Dallas, Texas, October 25, 1958.

4. Rice, S. O., Mathematical Analysis of Random Noise, Bell System Technical Journal, Volume 54, No. 7. Also reprinted in "Selected Papers on Noise and Stochastic Processes," edited by N. Wax, Dover Press, 1954.

5. Middleton, David, An Introduction to "Statistical Communication Theory",

6. Wang, M. C. and Uhlenbeck, G. E., "On the Theory of Brownian Motion II," Reviews of Modern Physics, Vol. 17 Nos. 2 and 3. Also reprinted in "Selected Papers on Noise and Stochastic Processes," edited by N. Wax, Dover Press, 1954.

7. Identification of Non-Linear Dynamic Systems by Regression Analysis Methods, Dissertation by R. B. McGhee, University of S. California, June 1963.

8. Blackman, R. B. and Tukey, J. W. "The Measurement of Power Spectra," Dover Publications, 1958.

9. Meissinger, H. F., "The Use of Parameter Influence Coefficients in Computer Analysis of Dynamic Systems." "Proceedings of the Western Joint Computer Conference," May 1960.

10. Estimation of the Statistics of Randomly Varying Parameters of Linear Systems. Dissertation by S. M. Brainin, University of Southern California, June 1966.

b ESTIMATE

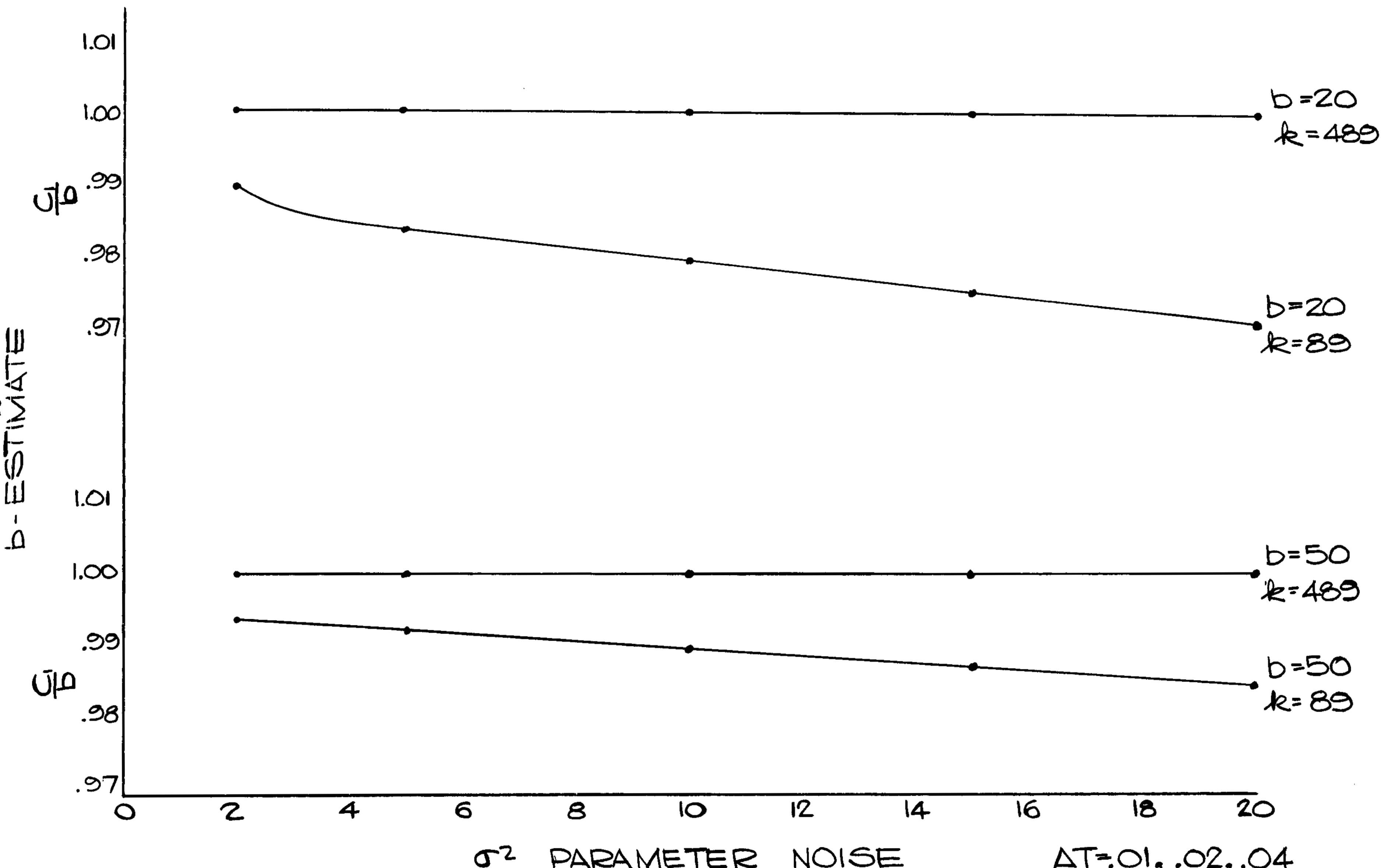

FIG 1

σ^2 ESTIMATE

$C_1 = 20$
$\Delta T = .02$

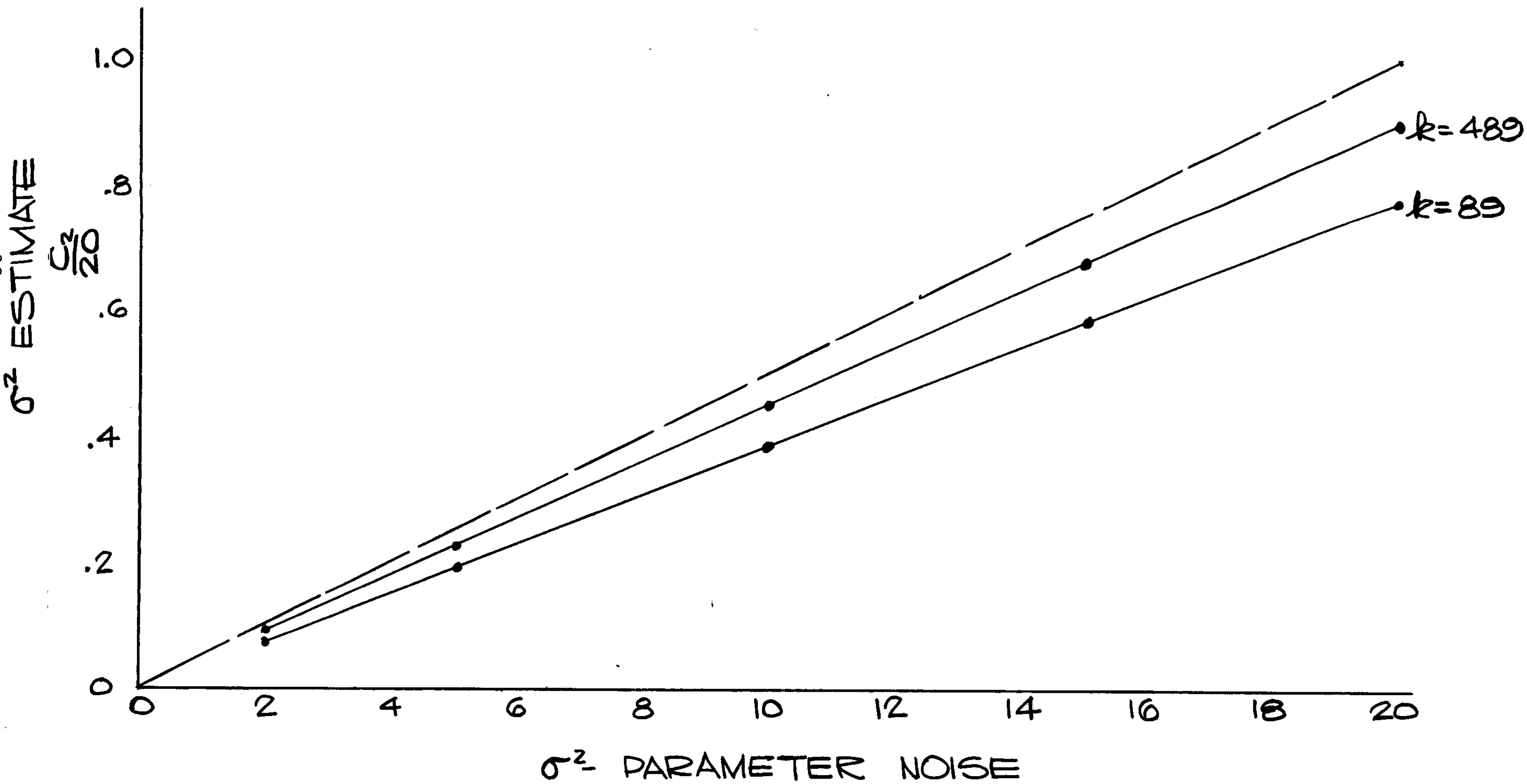

FIG 2

APPENDIX A

The development of the Fokker-Planck follows essentially the development of Reference 5.

One Dimensional Case

The basic starting point is the Smolukowski equation. See Reference 5. Let $y(t)$ be a random Markhoff process, and $W_2(y_o/y, t)$ its second order conditional density function. Where y_o is the value of y at time 0. Then for an incremental change in time this density function can be described as

$$W_2(y_o/y, t + \Delta t) = \int_{-\infty}^{\infty} W_2(y_o/z, t)\, W_2(z/y, \Delta t)\, dz \qquad \text{(A-1)}$$

$$z - y - \Delta y$$

Consider

$$I = \int_{-\infty}^{\infty} Q(y) \frac{\partial W_2(y_o/y, t)}{\partial t}\, dy \qquad \text{(A-2)}$$

where $Q(y)$ and all its derivatives $Q^{(n)}(y)$ exist, and vanish as $y \longrightarrow \pm \infty$ sufficiently rapidly for the convergence of all integrals in the following derivation, but is otherwise an arbitrary function of y. Replace $\partial W_2/\partial t$ by its limit form, use equation (A-1) and substitute in equation (A-2).

$$I = \lim_{\Delta t \to 0} \frac{1}{\Delta t} \left[\int_{-\infty}^{\infty} dy\, Q(y) \int_{-\infty}^{\infty} dz\, W_2(y_o/z, t)\, W_2(z/y, \Delta t) \right. \qquad \text{(A-3)}$$

$$\left. - \int_{-\infty}^{\infty} dz\, Q(x)\, W_2(y_o/z, t) \right]$$

All integrals in this development are to be considered as Stieljis integrals. Interchanging the order of integration and expanding Q(y) in a Taylor series about the point z then gives for the double integral in equation (A-3).

$$\int_{-\infty}^{\infty} dz\, W_2(y_o/z,\ t) \sum_{n=1}^{\infty} \frac{Q^{(n)}(z)}{n!} \int_{-\infty}^{\infty} (y-z)^n W_2(z/y,\ \Delta t)\ dy.$$

For n = 0, this term cancels the second term in equation (A-3) and

$$I = \sum_{n=1}^{\infty} \frac{1}{n!} \int_{-\infty}^{\infty} A_n(z)\, Q^{(n)}(z)\, W_2(y_o/z,\ t)\ dz \tag{A-3}$$

where by definition

$$A_n(z) = \lim_{\Delta t \to 0} \frac{1}{\Delta t} \int_{-\infty}^{\infty} (y-z)\ ^n W_2(z/y,\ \Delta t)\ dy \tag{A-4}$$

The coefficients $A_n(z)$ are the limits of the spatial moments of the increments in z, occurring during the interval Δt, given the present value z(t), (i.e. the value of $y(t-\Delta t)$). In other words the $A_n(z)$ are conditional moments of the variations about the sliding average of z(t)*, at any instant t.

It is assumed that all the coefficients $A_n(z)$ exist and are at most finite so that the series of equation (A-3) coverages. If equation (A-3) is now integrated by parts, and we employ the assumption that $Q^{(n)}(z) \to 0$, as $z \to \pm\infty$ sufficiently rapidly such that

$$Q^{(n)}\ A_n(z)\ W_2 \Big|_{-\infty}^{\infty} = 0,$$ for all n, we find that in changing the variable of integration from z to y that

$$\int_{-\infty}^{\infty} Q(y) \left[\frac{\partial W_2}{\partial t} - \sum_{n=1}^{\infty} \frac{(-1)^n}{n!} \frac{\partial^n}{\partial y^n} \left(A_n(y)\, W_2\right)\right] dy = 0 \tag{A-6}$$

---

* The sliding average is defined as the expectation of z(t), which for the non-stationary case, can be a function of time.

Since Q(y) is otherwise an arbitrary function

$$\frac{\partial W_2(y_o/y,\ t)}{\partial t} = \sum_{n=1}^{\infty} \frac{(-1)^n}{n!} \frac{\partial^n}{\partial y^n} \left(A_n(y)\ W_2(y_o/y,\ t\right) \qquad \text{(A-7)}$$

where $W_2(y_o/y,\ 0) = \delta(y - y_o)$. This is the series expression for the Smolukowski equation.

When the first and second moments $A_n(y)$ are non-zero and finite and all higher moments go to zero as $\Delta t \rightarrow 0$, this equation reduces to the Fokker-Planck partial differential equation, i.e., if only small changes in y can occur in small changes of time.

$$\frac{\partial W_2}{\partial t} = - \frac{\partial}{\partial y} (A_1(y)\ W_2) + \frac{1}{2} \frac{\partial^2}{\partial y^2} (A_2(y)\ W_2)$$

with the initial condition

$$W_2\ (y_o/y,\ 0) = \delta\ (y - y_o)$$

and the boundary conditions

$$W_2\ (\pm\ \infty) = 0, \quad \frac{\partial W_2}{\partial y} (\pm\ \infty) = 0$$

APPENDIX B

The purpose of this appendix is to prove that

$$A_n = 0 \text{ for all } n \geq 3, \text{ in equation (3)}$$

$$A_n = \lim_{\Delta t \to 0} \frac{1}{\Delta t} \overline{(\Delta y)^n}$$

Let $$\overline{\Delta y^n} = \int_t^{t+\Delta t} \cdots \int_t^{t+\Delta t} \overline{F(t_1)\, F(t_2) \ldots F(t_n)}\, dt_1\, dt_2 \ldots dt_n$$

where $$F(t) = x(t) - (b + n(t))\, y(t)$$

Since the incremental statistics are conditional on the present value of y, F(t) is conditionally a linear function of n(t). Thus, F(t) is a Gaussian random variable, and therefore

for $n \geq 3$ and odd, $\overline{\Delta y^n} = 0$.

For $n \geq 4$ and even, $\overline{F(t_1)\, F(t_2) \ldots F(t_n)}$ is equal to the sum of

$\dfrac{n!}{(2)^{n/2} \frac{n}{2}!}$ terms each of which is a product of n/2 pairs. Reference 4.

$$A_n = \lim_{\Delta t \to 0} \left[\frac{n!\, (\Delta t)^{-1}}{\frac{n}{2}!\; 2^{(n/2)}} \left(\iint_t^{t+\Delta t} \overline{F(t_1)\, F(t_2)}\, dt_1\, dt_2 \right)^{n/2} \right]$$

$$A_n = \lim_{\Delta t \to 0} \frac{n!}{\left(\frac{n}{2}\right)!\; 2^{n/2}} \left[\frac{(x^2\, \Delta t^2 + \sigma^2 y^2\, \Delta t)^{n/2}}{(\Delta t)} \right] = 0 \qquad \text{Q.E.D.}$$

USC Engineering