Journal of Statistical Physics, Vol. 15, No. 5, 1976

Studies in Nonlinear Stochastic Processes. I. Approximate Solutions of Nonlinear Stochastic Differential Equations by the Method of Statistical Linearization

Aaron B. Budgor[1]

Received April 27, 1976

A method based on a variational procedure is presented which provides simple and useful approximate solutions to a wide variety of nonlinear stochastic differential equations. This method of "statistical linearization" is most successful when the stochasticity of the differential equation is due to excitations which are normally distributed or harmonic with random phase. Effects due to deviations from normality can be corrected for in a systematic fashion. Comments regarding existence and uniqueness are given and some error bounds arising from the use of statistical linearization are computed.

KEY WORDS: Stochastic processes; nonlinear stochastic equations; statistical linearization; autocorrelation functions; spectral densities.

1. INTRODUCTION

Nonlinear stochastic differential equations occur in diverse areas of physics and engineering, such as transport theory in nonequilibrium statistical mechanics,[1,2] mode–mode coupling in critical dynamics,[3] turbulent flow,[4] and the structural response to random excitations.[5] The stochastic nature of these problems either comes from our desire to characterize a system having many degrees of freedom in terms of a description which requires much less information, or from a strict lack of information about the system. In either case, one equation used in modeling such problems is the generalized Langevin equation

$$(\partial/\partial t)G_j(t) = H_j(t, \{G_l(s)\}) + F_j(t), \qquad j, l = 1,..., N, \quad t_0 \leqslant s \leqslant t \tag{1}$$

Supported by the National Science Foundation under Grants NSF GP 32031X and MPS72-04353 A03.

[1] Department of Chemistry, University of California—San Diego, La Jolla, California.

subject to the condition

$$\langle F_j(t)\rangle_0 = 0 \tag{2}$$

Here the $\{G_j(t)\}$ are a set of N time-dependent system variables, $H_j(t, \{G_l(s)\})$ is a function, linear or nonlinear, which contains the explicit dependence of the rate of change of $G_j(t)$ upon the values of the chosen variables $G_l(s)$ between some initial time t_0 and t, and $F_j(t)$ is a fluctuating external force whose mean value (averaged over some initial distribution of the F_j), as expressed by Eq. (2), is zero. If the functionals H_j satisfy the condition

$$H_j(t, \{G_l(s)\}) = H_j(t, \{G_l(t)\}), \qquad t_0 \leqslant s \leqslant t \tag{3}$$

the corresponding Langevin equation is Markovian; otherwise it is non-Markovian.

An essential difficulty with current methods used in solving the Langevin equation is that exact solutions can only readily be obtained for systems in which the random noise is Gaussian-delta-correlated,

$$\langle F_j(t_1)F_k(t_2)\rangle = d\delta(t_1 - t_2)\,\delta_{kj} \tag{4}$$

with d a constant expressing the strength of the delta function. In such a case the random process is Markovian and the Fokker–Planck equation can be used to calculate the various statistical quantities for the observables in question. Markovian, however, implies memoryless (instantaneous) and clearly many processes do involve memory and time lag.

This paper will be the first in a series of publications devoted to a method which has had some widespread publicity in Soviet and American engineering literature and has been used extensively by control and systems analysts to construct solutions for a variety of nonlinear stochastic equations. This approach, the method of statistical or equivalent linearization,[6] was originated by Krylov and Bogoliubov[7] for the treatment of nonlinear systems under deterministic excitations and was later extended by Booton,[8] Caughey,[9] and Crandall[10] to random processes. Statistical linearization is based on the replacement in Eq. (1) of all nonlinear functions by linear functions which in some sense are equivalent statistically. By requiring that the mean squared error due to this replacement be minimal, one obtains an exact reproduction of the mean and an approximate expression for the dispersion. Of course, the first two moments do not completely characterize a distribution, but for practical purposes, knowing these is sufficient for a determination of the measurable quantities of a system.

We begin in Section 2 by describing the general theory and procedure involved in the statistical linearization of nonlinear stochastic differential equations and its connection to the theory of optimization of nonlinear transformations. Mention is made of the fact that for Gaussian processes statistical linearization is indeed an exact optimization theory.

This is of paramount consequence when it is realized that even for linear stochastic differential equations simple solutions of practical import are rarely obtained except when excitations are normally distributed or harmonic with a random phase.

In Sections 3 and 4 we examine the conditions under which nonlinear systems subjected to Gaussian noise can be suitably approximated by a linear representation. In such a situation, subordinate to certain applicability and normalization criteria, statistical linearization may be used—even for non-Gaussian external fluctuations—as a valid procedure in computing the average properties of the response.

When the above-mentioned criteria are not completely obeyed, systematic power-series correction terms (which statistical linearization truncates at the linear term) are derived which take account of these deviations from normality. This, unfortunately, complicates the simple analytical results of statistical linearization and necessitates solving an integral equation of Hammerstein type. Comments regarding existence and uniqueness are made and the error bounds due to the truncated statistical linearization solution are computed.

In another publication the results of this work will be extended to the case when the external excitation is Gaussian noise coupled to a deterministic harmonic signal.

Application of the procedures derived here for Gaussian input noise are deferred to a companion paper[11] in which variances, autocorrelation functions, and spectral densities are computed for Langevin equations of the Duffing oscillator type

$$\ddot{x} + \alpha\dot{x} + (x + \beta f(x)) = F(t) \tag{5}$$

where $x(t)$ is the position of the oscillator at time t, $f(x)$ is some nonlinear function of x, $F(t)$ is an external fluctuating force, and α and β are constant coefficients of damping and stiffness, respectively. We consider this equation rather than the more usual form of the Duffing equation where $f(x) = x^3$, since the Duffing equation has more widespread physical application than Brownian motion of an anharmonic oscillator in a heat bath.[12] It can be shown that (5) is also the equation for the forced motion due to an external random force of a rectangular plate supported by immovable hinges and subject to linear stress–strain relations.[5] For nonlinear stress–strain relations a more general equation results[13]

$$\ddot{x} + \alpha\dot{x} + K\tan(\pi x/2L) = F(t), \qquad -L < x < L \tag{6}$$

which for small x/L deviates from the Duffing model by a term of the order $(x/L)^5$.

Particular emphasis is placed on the Duffing equation itself since comparison with previous work by Bixon and Zwanzig (BZ)[12] and Morton and Corrsin[14] is possible. The work of MC is especially noteworthy in that our general procedure enables us to obtain a number of their spectral density approximations which were derived by them via diagrammatic techniques.

2. OPTIMAL TRANSFORMATIONS AND STATISTICAL LINEARIZATION

Consider the dynamical system

$$Z: F(t) \to X(t) \tag{7}$$

in which the transformation Z maps some initial random process F at time t_0 into some final random process X at time t. The general theory of optimization of nonlinear transformations[15] tells us that for given random processes X and F which are expressible either analytically or in terms of experimental data, it is possible to determine some optimal transformation Z^*, $Z^*(F) \neq Z(F)$, relating F to X such that the mean squared error upon replacement of Z by Z^* is minimized,

$$\lim_{T\to\infty} (1/2T) \int_{-T}^{T} [Z(F) - Z^*(F)]^2 \, dt = \eta = \min \tag{8}$$

A necessary and sufficient condition for this to occur is that Z^* exactly reproduce the mean value of Z. The map Z^* is a characteristic of the dynamical system and may be linear or nonlinear, time dependent or time independent. In particular, Z^* for the nonlinear integral transformation

$$X(t) = \int_0^\infty K[F(t - \tau), \tau] \, d\tau \tag{9}$$

or alternatively, in terms of its nonlinear differential equation analog,

$$Q(d/dt)Y(t) + R(d/dt)X(t) = U(d/dt)F(t) \tag{10}$$

can be shown by a variational procedure to be a nonlinear Wiener–Hopf integral equation.[15] In (10) we define $Q(d/dt)$, $R(d/dt)$, and $U(d/dt)$ to be linear differential operators of time and $Y(t) = f(x)$ to be a nonlinear function of the random process $X(t)$.[2]

Under the stringent conditions that X and F be Gaussian, though not necessarily delta-correlated, Z^* reduces to a linear transformation. Statistical linearization is an approximation to the *general theory of optimal transformations in the mean squared sense* in that it always assumes Z^* to be linear,

[2] We limit ourselves in this paper to odd nonlinearities in the independent variables. For even nonlinearities a somewhat different minimization procedure needs to be applied.

regardless of the distributions of the random processes X and F, but becomes an exact optimization theory when the random processes are Gaussian-distributed. Thus, for Gaussian random processes, statistical linearization results in an exact reproduction of mean values and a minimization in the deviations of the second moment statistics.

Specializing this statement to (10), we note that in the integral representation F is nonlinear due to a term containing $Y(t)$. However, Z^* will be linear if $Y(t)$ is replaced by $h_1X(t)$. Applying a variational procedure to (8) by minimizing η with respect to h_1,

$$\partial\eta/\partial h_1 = 2(h_1\langle X^2\rangle_T - \langle f(X)X\rangle_T) = 0 \tag{11a}$$

we obtain an optimal Z^* on setting[3]

$$h_1 = \langle f(X)X\rangle_T/\langle X^2\rangle_T \tag{11b}$$

The brackets with the subscript T containing the random processes denote time averages. For ergodic dynamical systems this may be replaced by ensemble averages. In such cases the subscript T will be deleted.

3. STATISTICAL LINEARIZATION AND STATIONARY AUTOCORRELATION FUNCTIONS

Prior to our derivation of the bounds of applicability of statistical linearization we shall need the following definitions and expansions.

For a random process $A(t)$ we define the stationary autocorrelation function $R_{AA}(\tau)$ and its spectral density $S_{AA}(\omega)$ by

$$R_{AA}(\tau) = \langle [A(t) - \langle A(t)\rangle][A(t+\tau) - \langle A(t+\tau)\rangle]\rangle \tag{12a}$$

$$S_{AA}(\omega) = \int_{-\infty}^{+\infty} e^{-i\omega\tau} R_{AA}(\tau)\, d\tau \tag{12b}$$

By extension, the autocorrelation function of some function of $A(t)$, $f(A(t))$, can be written as

$$R_{f(A)f(A)}(\tau) = \iint_{-\infty}^{+\infty} [f(A_1) - \langle f(A_1)\rangle][f(A_2) - \langle f(A_2)\rangle]p(A_1, A_2)\, dA_1\, dA_2 \tag{13}$$

where the independent variables A_1 and A_2 are notational shorthand for the random processes $A(t)$ and $A(t+\tau)$.

[3] This variational procedure may be generalized to include nonlinearities whose structure is dependent upon a number of internal degrees of freedom. Thus, a linear approximation to the nonlinearity $Y(t) = f(X_1, X_2, \ldots, X_n)$ in which the $X_i(t)$ are independent and $\langle X_i(t)\rangle = 0$ is $Y(t) = \sum_{i=1}^{n} h_i X_i(t)$, where the h_i are defined by the relations $h_i = \langle f(X_1, \ldots, X_n)X_i\rangle_T/\langle X_i^2\rangle_T$.

It therefore follows, due to a result of Barret and Lampard,[16] that for a two-dimensional probability density belonging to a particular class of functions Λ such that

$$p(A_1, A_2) = p(A_1)p(A_2) \sum_{n=0}^{\infty} b_n \theta_n^{(1)}(A_1)\theta_n^{(2)}(A_2) \tag{14a}$$

$R_{f(A)f(A)}(\tau)$ has a series representation

$$R_{f(A)f(A)}(\tau) = \sum_{n=1}^{\infty} a_n{}^2 \rho_{AA}^n(\tau) \tag{14b}$$

The functions $\theta_n^{(i)}$ are normalized orthogonal polynomials with respect to weighting functions $p(A_i)$ (one-dimensional probability densities) obeying the property

$$\int_{-\infty}^{+\infty} p(A_i)\theta_m^{(i)}(A_i)\theta_n^{(i)}(A_i)\, dA_i = \delta_{mn} \tag{14c}$$

with the coefficients b_n and a_n expressed as

$$b_n = \iint_{-\infty}^{+\infty} p(A_1, A_2)\theta_n^{(1)}(A_1)\theta_n^{(2)}(A_2)\, dA_1\, dA_2 \tag{14d}$$

$$a_n = \int_{-\infty}^{+\infty} f(A)p(A)\theta_n(A)\, dA \tag{14e}$$

$\rho_{AA}(\tau)$ is the normalized autocorrelation function defined in terms of the variance $\sigma_A{}^2$,

$$\rho_{AA}(\tau) = R_{AA}(\tau)/\sigma_A{}^2 \tag{15a}$$

$$\sigma_A{}^2 = \langle [A(t) - \langle A(t)\rangle]^2\rangle = R_{AA}(0) \tag{15b}$$

Forms of specific utility to this and to future work occur when the random process $A(t)$ is Gaussian or harmonic of amplitude a and with uniformly distributed phase. Then relation (14a) becomes, for the Gaussian case,

$$\begin{aligned} p(A_1, A_2) &= \frac{1}{2\pi\sigma_{A_1}\sigma_{A_2}[1 - \rho_{A_1A_2}^2(\tau)]^{1/2}} \\ &\quad \times \exp\left[- \frac{\sigma_{A_1}^2 A_1{}^2 - 2\sigma_{A_1}\sigma_{A_2}\rho_{A_1A_2}(\tau) + \sigma_{A_1}^2 A_2{}^2}{2\sigma_{A_1}^2\sigma_{A_2}^2[1 - \rho_{A_1A_2}^2(\tau)]} \right] \\ &= \frac{1}{2\pi\sigma_{A_1}\sigma_{A_2}} \left\{ \exp\left[-\frac{1}{2}\left(\frac{A_1{}^2}{\sigma_{A_1}^2} + \frac{A_2{}^2}{\sigma_{A_2}^2} \right) \right] \right\} \\ &\quad \times \sum_{n=0}^{\infty} \frac{1}{n!} \rho_{A_1A_2}^n(\tau) \mathrm{He}_n\left(\frac{A_1}{\sigma_{A_1}} \right) \mathrm{He}_n\left(\frac{A_2}{\sigma_{A_2}} \right) \end{aligned} \tag{16a}$$

and for the harmonic case

$$p(A_1, A_2) = \frac{1}{4\pi^2} \int_{-\infty}^{+\infty}\int J_0(a(u^2 + v^2 + 2uv \cos \omega\tau)^{1/2}) e^{-i(uA_1 + vA_2)}\, du\, dv$$

$$= \frac{1}{\pi^2} (a^2 - A_1{}^2)^{-1/2}(a^2 - A_2{}^2)^{-1/2}$$

$$\times \sum_{n=0}^{\infty} \epsilon_n T_n\left(\frac{A_1}{a}\right) T_n\left(\frac{A_2}{a}\right) \cos(n\omega\tau) \tag{16b}$$

when

$$A_i{}^2/a^2 < 0, \qquad \epsilon_n = \begin{cases} 1, & n = 0 \\ 2, & n \geqslant 1 \end{cases}$$

and otherwise

$$p(A_1, A_2) \equiv 0$$

In (16a), $p(A)$ is given by

$$p(A) = (2\pi\sigma_A{}^2)^{-1/2} \exp(-A^2/2\sigma_A{}^2) \tag{17a}$$

and $\mathrm{He}_n(x)$ is the Hermite polynomial with orthogonality condition

$$\int_{-\infty}^{+\infty} \mathrm{He}_n(x)\mathrm{He}_m(x) \exp(-x^2/2)\, dx = (2\pi)^{1/2} n!\, \delta_{mn} \tag{17b}$$

Thus, $\theta_n(x)$ and b_n are

$$\theta_n(x) = (n!)^{-1/2}\mathrm{He}_n(x) \tag{17c}$$

$$b_n = \rho^n_{A_1A_2}(\tau) \tag{17d}$$

In (16b), $p(A)$ is given by

$$p(A) = (1/\pi)(a^2 - A^2)^{-1/2} \tag{18a}$$

and $T_n(x)$ is the Tchebycheff polynomial of the first kind with orthogonality condition

$$(1/\pi) \int_{-1}^{+1} \epsilon_n T_m(x) T_n(x)(1 - x^2)^{-1/2}\, dx = \delta_{mn} \tag{18b}$$

Thus, $\theta_n(x)$ and b_n are

$$\theta_n(x) = \epsilon_n^{1/2} T_n(x/a) \tag{18c}$$

$$b_n = \cos(n\omega T) \tag{18d}$$

We are now in a position to derive two procedures for calculating $R_{xx}(t)$, the autocorrelation function for the solution of (10) when $F(t)$ is Gaussian.

These will respectively be referred to, in conjunction with the parlance of MC,[14] who obtained identical equations via diagrammatic techniques, as the cumulant discard and quasilinear Green's function methods. The effects of combined Gaussian and harmonic excitations on $R_{xx}(t)$ will be discussed elsewhere.

3.1. Cumulant Discard Method

By Fourier-transforming (10) and multiplying by its complex conjugate, we obtain the spectral density equation for Gaussian unit-delta-correlated noise,

$$|\hat{R}(i\omega)|^2 S_{xx}(\omega) + |\hat{Q}(i\omega)|^2 S_{YY}(\omega) + \hat{R}(i\omega)\hat{Q}^*(i\omega)S_{xY}(\omega) + \hat{R}^*(i\omega)\hat{Q}(i\omega)S_{Yx}(\omega) = |\hat{U}(i\omega)|^2 S_{FF}(\omega) \quad (19)$$

This may be confirmed by noting that since

$$\hat{X}^*(\omega')\hat{X}(\omega) = \int_{-\infty}^{+\infty}\int e^{-i\omega t}e^{i\omega' t'} R\left(\frac{d}{dt}\right) X(t) R\left(\frac{d}{dt'}\right) X(t')\, dt\, dt'$$

$$= \hat{R}(i\omega)\hat{R}^*(i\omega') \int_{-\infty}^{+\infty}\int e^{-i\omega t}e^{i\omega' t'} X(t)X(t')\, dt\, dt' \quad (20a)$$

on changing variables to $t - t' = \tau$ and $\kappa = (t + t')/2$ and averaging over the product of the stationary random processes, we find that (20a) reduces to

$$\hat{X}(\omega')\hat{X}(\omega) = \hat{R}(i\omega)\hat{R}^*(i\omega') \int_{-\infty}^{+\infty} d\tau \int_{-\infty}^{+\infty} d\kappa\, e^{-i\omega(\kappa + \tau/2)} e^{i\omega'(\kappa - \tau/2)} R_{xx}(\tau)$$

$$= 2\pi\delta(\omega - \omega')\hat{R}(i\omega)\hat{R}^*(i\omega') \int_{-\infty}^{+\infty} d\tau\, e^{-i\omega\tau} R_{xx}(\tau)$$

$$= 2\pi\delta(\omega - \omega')\hat{R}(i\omega)\hat{R}^*(i\omega')S_{xx}(\omega) \quad (20b)$$

This produces the first term in (19). The other terms are obtained in a similar fashion. This result, although exact, is not very useful since the explicit forms of $S_{xY}(\omega)$, $S_{Yx}(\omega)$, and $S_{YY}(\omega)$ are unknown due to the fact that their corresponding correlation functions depend on the higher order correlations between $X(t)$ and $Y(t')$ and between $Y(t)$ and $Y(t')$.

As an approximation scheme we can utilize the quasinormal assumption, which asserts that the higher order correlations between $X(t)$ and $Y(t')$ and $Y(t)$ and $Y(t')$ are decomposable as joint Gaussian processes and therefore

expandable in terms of Hermite polynomials. Thus, by virtue of relation (14b),

$$R_{xY}(\tau) = a_1\sigma_x\rho_{xx}(\tau) = (a_1/\sigma_x)R_{xx}(\tau) = R_{Yx}(\tau) \tag{21a}$$

$$R_{YY}(\tau) = \sum_{n=1}^{\infty} a_n{}^2\rho_{xx}^n(\tau)$$

and therefore

$$S_{xY}(\omega) = S_{Yx}(\omega) = (a_1/\sigma_x)S_{xx}(\omega), \qquad S_{YY}(\omega) = \sum_{n=1}^{\infty} a_n{}^2 S_n(\omega) \tag{21b}$$

where $S_\nu(\omega)$ is the Fourier transform of the νth-degree normalized autocorrelation function $\rho_{xx}(\tau)$,

$$\begin{aligned} S_\nu(\omega) &= \int_{-\infty}^{+\infty} e^{-i\omega\tau}\rho_{xx}^\nu(\tau)\,d\tau \\ &= \int_{-\infty}^{+\infty} e^{-i\omega\tau}\left[(1/2\pi\sigma_x{}^2)\int_{-\infty}^{+\infty} e^{i\omega\tau}S_{xx}(\omega)\,d\omega\right]^\nu d\tau \end{aligned} \tag{22a}$$

From the latter equality in (22a) it is a simple matter to derive the convolution recursion relation for $S_\nu(\omega)$:

$$S_\nu(\omega) = (1/2\pi)\int_{-\infty}^{+\infty} S_1(\omega - \omega')S_{\nu-1}(\omega')\,d\omega', \qquad \nu = 2, 3,\ldots \tag{22b}$$

$$S_1(\omega) = S_{xx}(\omega)/\sigma_x{}^2$$

Substitution of relations (21b) into (19), dividing through by $|\hat{R}(i\omega)|^2$, and utilizing the identities (22b) yields the cumulant discard spectral density equation

$$S_{xx}(\omega) = |\psi(i\omega)|^2 S_{FF}(\omega) - |\psi_1(i\omega)|^2 \sum_{n=2}^{\infty} a_n{}^2 S_n(\omega) \tag{23}$$

with $\psi(i\omega)$ and $\psi_1(i\omega)$ given by

$$\psi(i\omega) = \frac{\hat{U}(i\omega)/\hat{R}(i\omega)}{1 + h_1[\hat{Q}(i\omega)/\hat{R}(i\omega)]} \tag{24a}$$

$$\psi_1(i\omega) = \frac{\hat{Q}(i\omega)/\hat{R}(i\omega)}{1 + h_1[\hat{Q}(i\omega)/\hat{R}(i\omega)]} \tag{24b}$$

The coefficient h_1, $h_1 = h_1(\sigma_x)$, is the coefficient of statistical linearization as defined by (11b) and can be identified via (14e) as

$$h_1 = a_1/\sigma_x \tag{25}$$

By Fourier-transforming (23) and utilizing the convolution property of

Fourier integrals, one obtains the cumulant discard correlation function equation (Appendix A),

$$R_{xx}(t) = R_0(t) - \sum_{n=2}^{\infty} \frac{a_n^2}{\sigma_x^{2n}} \int_{-\infty}^{+\infty} dt'\, R_1(t') R_{xx}^n(t - t') \tag{26}$$

where $R_0(\tau)$ and $R_1(\tau)$ are the Fourier transforms of $\psi(i\omega)$ and $\psi_1(i\omega)$, respectively.

3.2. Quasilinear Green's Function Method

By rewriting (10) as a linear inhomogeneous differential equation in $X(t)$

$$R\left(\frac{d}{dt}\right)X(t) + h_1 Q\left(\frac{d}{dt}\right)X(t) = U\left(\frac{d}{dt}\right)F(t) + \phi[X(t)] \tag{27}$$

it can be transformed into a Green's function convolution equation

$$X(t) = \int_0^{\infty} G(\tau)\left\{U\left(\frac{d}{d(t-\tau)}\right)F(t-\tau) + \phi[X(t-\tau)]\right\} d\tau \tag{28a}$$

with

$$\phi[X(t)] = Q(d/dt)[h_1 X(t) - Y(t)] \tag{28b}$$

In obtaining the result in (28a), we set, without loss of generality, all initial conditions equal to zero since they can be absorbed into the noise term $U(d/dt)F(t)$.

The autocorrelation function $R_{xx}(\tau)$ is then, upon application of (12a),

$$\begin{aligned} R_{xx}(\tau) &= \int_0^{\infty}\int d\tau_1\, d\tau_2\, G(\tau_1)G(\tau_2)\left\{U\left(\frac{d}{d(t-\tau_1)}\right)F(t-\tau_1) + \phi[X(t-\tau_1)]\right\} \\ &\quad \times \left\{U\left(\frac{d}{d(t+\tau-\tau_2)}\right)F(t+\tau-\tau_2) + \phi[X(t+\tau-\tau_2)]\right\} \\ &= \int_0^{\infty}\int d\tau_1\, d\tau_2\, G(\tau_1)G(\tau_2) \\ &\quad \times [R_{U(\cdot)FU(\cdot)F}(\alpha) + R_{\phi U(\cdot)F}(\alpha) + R_{U(\cdot)F\phi} + R_{\phi\phi}(\alpha)] \end{aligned} \tag{29}$$

where $\alpha = \tau + \tau_2 - \tau_1$.

For mathematical tractability, as was done for the cumulant discard approximation, we employ the quasinormal assumption for $R_{\phi U(\cdot)F}(\alpha)$ and $R_{\phi\phi}(\alpha)$, which, together with definition (25), yields

$$\begin{aligned} R_{\phi U(\cdot)F}(\alpha) &= h_1\sigma_x\sigma_F\rho_{Q(\cdot)XU(\cdot)F}(\alpha) - \sigma_F a_1 \rho_{Q(\cdot)XU(\cdot)F}(\alpha) = R_{U(\cdot)F\phi}(\alpha) = 0 \\ R_{\phi\phi}(\alpha) &= (h_1\sigma_x)^2\rho_{Q(\cdot)XQ(\cdot)X}(\alpha) - 2h_1\sigma_x a_1 \rho_{Q(\cdot)XQ(\cdot)X}(\alpha) \\ &\quad + \sum_1^{\infty} a_n^2 \rho^n_{Q(\cdot)XQ(\cdot)X}(\alpha) \end{aligned} \tag{30}$$

Introduction of the equalities in (30) into (29) results, for known $R_{U(\cdot)FU(\cdot)F}(\alpha)$, in a nonlinear integral equation for $R_{xx}(\tau)$ and its higher order correlates

$$R_{xx}(\tau) = \int_0^\infty\int d\tau_1\, d\tau_2\, G(\tau_1)G(\tau_2)\left[R_{U(\cdot)FU(\cdot)F}(\alpha) + \sum_{n=2}^{\infty} \frac{a_n^2}{\sigma_x^{2n}} R^n_{Q(\cdot)XQ(\cdot)X}(\alpha)\right] \quad (31)$$

which, after changing to "center-of-mass" coordinates $t = \tau_1 - \tau_2$ and $T = \tau_1 + \tau_2$, reduces to a single integral convolution equation[4] (quasilinear Green's function equation)

$$R_{xx}(\tau) = R_0(\tau) + \sum_{n=2}^{\infty} \frac{a_n^2}{\sigma_x^{2n}} \int_{-\infty}^{+\infty} d\tau'\, R_1(\tau') R^n_{xx}(\tau - \tau') \quad (32)$$

where the functions $R_0(t)$ and $R_1(t)$ are identical to those functions entering in (26).

By comparing (32) with (26), one sees that the cumulant discard and quasilinear Green's function methods differ only with respect to a change in sign in the higher autocorrelation function terms. Furthermore, to first order in $R_0(t)$ both solutions are identical to the one obtained by the method of statistical linearization; that is, by equating $Y(t) = h_1 X(t)$. Thus, $\phi[X(t)]$ in (28b) is identically zero and the autocorrelation function $R_{xx}(t)$ of the solution to the linear inhomogeneous differential equation (27), just as for Eq. (10), reduces to $R_0(t)$.

Comments about the error bounds due to statistical linearization and the quasinormal assumption will be deferred to the next two sections.

In closing, a few comments should be made about the existence and uniqueness of (32) and (26)—written $R_\pm(t)$, respectively—which upon comparison with (B.1b) of Appendix B are integral equations of Hammerstein type,

$$R_\pm(t) = \int_G R_1(t, \tau) f[\tau, R_\pm(\tau)]\, d\tau + R_0(t) \quad (33a)$$

or in operator form

$$R_\pm = \mathscr{K}\mathscr{F} R_\pm + R_0 \quad (33b)$$

Here G is a domain, bounded or infinite in E^n. Elegant operator-theoretic contraction mapping techniques have been applied by Kolodner[17] to illustrate existence and uniqueness of a solution to Hammerstein's equation in Hilbert space $\mathscr{H}$. The important theorems and accompanying definitions are presented in Appendix B.

One can summarize those results by the following: Given reasonable

[4] The proof of this statement may alternatively be seen by Fourier-transforming (31), changing to "center-of-mass" coordinates, and realizing that the result is of the same form as (23).

boundedness conditions on $\mathscr{K}_\lambda = (I - \lambda\mathscr{K})^{-1}\mathscr{K}$, λ any scalar, and monotonicity requirements on both $\mathscr{F}$ and $\mathscr{K}_\lambda$, then the existence and uniqueness of the quasilinear Green's function and cumulant discard equations, $R_\pm(t)$, depend on the bijectivity of $\mathscr{K}$ and on the boundedness, continuity, range, and domain of $\mathscr{F}$.

It appears that since $\rho(t)$ is bounded and has a Taylor series expansion, $f(t, \rho)$, being some polynomial of $\rho(t)$, is Lipschitzian on bounded sets. That $\mathscr{K}$ is bounded (for an example, see Ref. 11) therefore implies uniqueness via Theorem B.1, with the question of existence determined by the criterion of bijectivity, Theorems B.2 and B.4.

4. ERROR BOUNDS AND STATISTICAL LINEARIZATION

In the introduction it was mentioned that statistical linearization is an outgrowth of the Krylov–Bogoliubov (KB) asymptotic analysis to deterministic differential equations which are perturbed by periodic forcing functions. To motivate the subsequent discussion on statistical linearization it will be of interest to give some background to the KB method.

In short, for a given forcing function $F(t)$

$$F(t) = A_F \sin(\omega_F t - \nu) \tag{34}$$

the KB method tries to determine the conditions under which

$$R(d/dt)X + Q(d/dt)f(X, \dot{X}) = U(d/dt)F(t) \tag{35}$$

has a periodic solution sufficiently close to $X = A \sin \Omega t$, where Ω, depending on whether the frequency of the forcing function $F(t)$ is higher or lower than the equilibrium frequency ω_0 of the homogeneous part of (35), is either ω_F or ω_0.

That a system has a solution of the aforementioned type cannot be established a priori, but there are both qualitative and quantitative measures as to the validity of such an approximation. We shall not go into the details of this procedure since an exhaustive discussion is given by Šiljak.[16] Basically, however, one assumes a solution of the form

$$X(t) = x_1(t) + \epsilon\chi(t) \tag{36a}$$

where $x_1(t)$ and $\chi(t)$, respectively, denote the contributions of the first and higher harmonics,

$$x_1(t) = A_1 \sin \Omega_1 t \tag{36b}$$

$$\chi(t) = \sum_{k=2}^{\infty} A_k \sin(k\Omega_1 t + \theta_k) \tag{36c}$$

The small parameter ϵ ($\epsilon \ll 1$) associated with the function $\chi(t)$ in (36a) ensures that the first harmonic dominates in the solution of (35).

By substituting (36a) into (35), Fourier-expanding $f(X, \dot{X})$, and equating to zero all harmonics of the same order, we obtain an equation for the first harmonic having accuracy of degree ϵ,

$$R\left(\frac{d}{dt}\right)X_1 + Q\left(\frac{d}{dt}\right)\left(q + \frac{q'}{\Omega_1}\frac{d}{dt}\right)X_1 - U\left(\frac{d}{dt}\right)\frac{A_F}{A_1}\left(\cos\nu - \frac{\sin\nu}{\Omega_1}\frac{d}{dt}\right)X_1 = 0 \tag{37}$$

or

$$R'\left(\frac{d}{dt}\right)X_1 + Q\left(\frac{d}{dt}\right)\left(q + \frac{q'}{\Omega_1}\frac{d}{dt}\right)X_1 = 0$$

The coefficients $q = q(A_1, \Omega_1)$ and $q' = q'(A_1, \Omega_1)$ are the Fourier sine and cosine coefficients corresponding to the first harmonic in the expansion of $f(X, \dot{X})$,

$$q = (1/\pi)\int_0^{2\pi} f(A_1 \sin\xi, A_1\Omega_1 \cos\xi)\sin\xi\, d\xi \tag{38a}$$

$$q' = (1/\pi)\int_0^{2\pi} f(A_1 \sin\xi, A_1\Omega_1 \cos\xi)\cos\xi\, d\xi \tag{38b}$$

$$\xi = \Omega_1 t \tag{38c}$$

Similar equations to (37) can be found for the higher harmonics. Qualitative criteria which permit neglect of the higher harmonics, thus resulting in an accurate characterization of (35) by (37), are:

1. The polynomials $\hat{R}'(i\Omega)$ and $\hat{Q}(i\Omega)$ satisfy the following relations:

(a) $\deg \hat{Q}(i\Omega) < \deg \hat{R}'(i\Omega)$

(b) $|\hat{Q}(ik\Omega)/\hat{R}'(ik\Omega)|C_k \ll |\hat{Q}(i\Omega)/\hat{R}'(i\Omega)|C_1, \qquad k \geqslant 2$

(c) $\lim_{k\to\infty} |\hat{Q}(ik\Omega)/\hat{R}'(ik\Omega)| \to 0 \qquad \forall k$

Conditions (a)–(c) imply that the homogeneous equivalent linear differential equation (37) effectively damps out or attenuates the higher harmonics so that their contribution to the solution becomes negligible. The coefficients C_l, $l \geqslant 1$, in (b) are the amplitudes of the harmonics in the Fourier expansion of $f(X, \dot{X})$ with $C_1 = (q^2 + q'^2)^{1/2}$. In situations where the C_k cannot be considered small in comparison to C_1, (b) may be replaced by

(b') $\qquad |\hat{Q}(ik\Omega)/\hat{R}'(ik\Omega)| \ll |\hat{Q}(i\Omega)/\hat{R}'(i\Omega)|$

2. The polynomials $\hat{R}'(ik\Omega)$ cannot have purely imaginary zeros, $k = 1, 2, 3, \ldots$. This criterion guarantees stability of the solution.

3. The function $f(X, \dot{X})$ should have finite partial derivatives with respect to its independent variables X and $\dot{X}$, and should not be an explicit function of time. Thus $f(X, \dot{X})$ may belong to both the class of piecewise continuous and that of discontinuous functions of Heaviside type.

In summary, then, the KB method assumes that in order to investigate the periodic solutions of the nonlinear differential equation (35), one may consider an associated linear differential equation (37), subject to the proviso that the aforementioned applicability conditions are met, in which the nonlinearity $f(X, \dot{X})$ has been expanded in a Fourier series but truncated at the first harmonic.

The spirit of this procedure follows through for the stochastic differential equation (10), but now, depending on the statistics of $F(t)$, the nonlinear function $Y(t)$ is expanded with the orthogonal polynomial basis expressing these statistics [see Eqs. (14)–(18)]. The method of statistical linearization truncates this expansion at the linear term since for $t > 0$, $|\rho(t)| < 1$ and the expansion coefficients a_n^2 of $R_{YY}(t)$ in (21a) are asymptotically at least $O[(n!)^{-1}]$. However, when t is small (ω large), $|\rho(t)| \to 1$ and the nonlinear distortions of $R_{YY}(t)$ will be important. If the equivalent linear equation to (10) obeys the applicability criteria of the KB method,[5] the nonlinear terms in $R_{YY}(t)$ become attenuated and the statistical linearization approximation will still hold true even for small t. This statement can be made rigorous by examining the spectral density equation (23) and its quasilinear Green's function analog. Writing the error in statistical linearization as $\mathscr{E}_{\pm}$,

$$\begin{aligned} \mathscr{E}_{\pm} &= \left|\frac{\hat{U}(i\omega)}{\hat{R}(i\omega)}\right|^2 S_{FF}(\omega) - \left|1 + h_1 \frac{\hat{Q}(i\omega)}{\hat{R}(i\omega)}\right|^2 S_{xx}(\omega) \\ &= \pm \left|\frac{\hat{Q}(i\omega)}{\hat{R}(i\omega)}\right|^2 \sum_{n=2}^{\infty} a_n^2 S_n(\omega) \end{aligned} \tag{39}$$

then, upon Fourier-transforming,

$$|\mathscr{F}\mathscr{E}_{\pm}| \leqslant \sup \sum_{n=2}^{\infty} a_n^2 \int_{-\infty}^{+\infty} |p(\tau)|\, |\rho^n(t-\tau)|\, d\tau = (e-2) \int_{-\infty}^{+\infty} |p(\tau)|\, d\tau \tag{40a}$$

where

$$p(\tau) = \mathscr{F}\left\{\left|\frac{\hat{Q}(i\omega)}{\hat{R}(i\omega)}\right|^2\right\} \tag{40b}$$

Equation (40a) states that the smaller the value of the integral of $p(\tau)$, the smaller the error made when using statistical linearization.

[5] Unless the external excitation is some combination of noise and a deterministic sinusoidal or impulse signal, $\hat{R}'(i\omega)$ is just $\hat{R}(i\omega)$.

When $F(t)$ is some combination of noise and a deterministic sinusoidal or impulse signal, the KB and statistical linearization methods can be used in conjunction with one another to determine the solution statistics. Discussion of this problem will be deferred to a subsequent paper.

5. DISCUSSION

For the case of normally distributed or harmonic excitations, linear stochastic differential equations have solutions which are, respectively, normally or harmonically distributed. We are not aware whether such simple linear transformations exist for other types of excitations. Since the method of statistical linearization reduces nonlinear stochastic differential equations to linear ones, within the approximations indicated in the body of the paper, these desirable properties of Gaussian and harmonic excitations will also hold true for those transformed nonlinear equations.

It is for these reasons that statistical linearization, perhaps coupled with the KB method, is such a powerful technique, with the strengths of this approach particularly evident when the stochasticity of the differential equation is due to Gaussian excitation. Then, if the modulus of the ratio $|\hat{Q}(i\omega)/\hat{R}(i\omega)|$ is such that the passband[6] of the transfer function is less than the frequency band of $S_{YY}(\omega)$, $Y(t)$ becomes "normalized" and $S_{YY}(\omega)$ may be restricted to the linear term (first term) in its Hermite polynomial expansion. Since the dependent variables are now all normal, the Green's function solution is therefore also normal. Deviations from this "normalization" condition lead to the requirement of additional terms in the expansion of $S_{YY}(\omega)$. If the excitation is some combination of a Gaussian and a harmonic signal, then after linearization the solution response, although Gaussian, will have shifted amplitudes and effective frequencies.

This approach can possibly be extended to non-Gaussian excitations, but now the nonlinear stochastic differential equation is appropriately linearized by virtue of the particular orthogonal polynomial expansion expressing the distribution characteristics of the excitation. At the time of writing, although bounds analogous to (40a) for the linearized error should be easily derivable, it is unknown to this author whether the previous applicability criteria (a)–(c) still hold or even if rule of thumb conditions similar to the bandpass arguments for normal signals (necessary to minimize distortion effects due to the nonlinearity) exist or can be found.

[6] For the linear transformation Z^*: $F(t) \to X(t)$ and its corresponding spectral density relation $S_{xx}(\omega) = |\psi(i\omega)|^2 S_{FF}(\omega)$, we determine sets $\{\omega_{j1}\}$, $\{\omega_j\}$, and $\{\omega_{j2}\} \ni$: for each j, $\omega_{j1} < \omega_j < \omega_{j2}$. The ω_j and ω_{jl} are found, respectively, by computing $\max_{\{\omega j\}}|\psi(i\omega)|^2$ and by the relation $|\psi(i\omega_{j1})|^2 = |\psi(i\omega_{j2})|^2 = \frac{1}{2}|\psi(i\omega_j)|^2$. Then for each j, the passband of the transfer function $|\psi(i\omega)|^2$ is $|\omega_{j1} - \omega_{j2}|$.

APPENDIX A. DERIVATION OF THE CUMULANT DISCARD AND QUASILINEAR GREEN'S FUNCTION AUTOCORRELATION EQUATIONS

Upon Fourier-transforming the cumulant discard or quasilinear Green's function equations

$$S_{xx}(\omega) = |\psi(i\omega)|^2 S_{FF}(\omega) \mp |\psi_1(i\omega)|^2 \sum_{n=2}^{\infty} a_n{}^2 S_n(\omega) \tag{A.1}$$

one obtains their respective autocorrelation function equations (26) and (32),

$$R_{xx}(\tau) = R_0(\tau) \mp \sum_{n=2}^{\infty} \frac{a_n{}^2}{\sigma_x^{2n}} \int_{-\infty}^{+\infty} d\tau' \, R_1(\tau') R_{xx}^n(\tau - \tau') \tag{A.2}$$

This is readily apparent on defining

$$\begin{aligned} R_0(t) &= (1/2\pi) \int_{-\infty}^{+\infty} e^{i\omega t} |\psi(i\omega)|^2 \, d\omega \\ R_1(t) &= (1/2\pi) \int_{-\infty}^{+\infty} e^{i\omega t} |\psi_1(i\omega)|^2 \, d\omega \\ R_{xx}(t) &= (1/2\pi) \int_{-\infty}^{+\infty} e^{i\omega t} S_{xx}(\omega) \, d\omega \end{aligned} \tag{A.3}$$

and utilizing (22a) and the convolution property of Fourier integrals, so that

$$|\psi_1(i\omega)|^2 S_v(\omega) = \int_{-\infty}^{+\infty} dt' \, e^{-i\omega t'} \int_{-\infty}^{+\infty} R_1(\tau) \rho_{xx}^v(t' - \tau) \, d\tau \tag{A.4}$$

Insertion of these definitions into the Fourier transform of (A.1) results in, for d-delta-correlated noise $F(t)$,

$$\begin{aligned} R_{xx}(t) &= dR_0(t) \mp \sum_{n=2}^{\infty} \frac{a_n{}^2}{2\pi} \int_{-\infty}^{+\infty} d\omega \, e^{i\omega t} \int_{-\infty}^{+\infty} dt \, e^{-i\omega t} \int_{-\infty}^{+\infty} R_1(\tau) \rho_{xx}^n(t' - \tau) \, d\tau \\ &= dR_0(t) \mp \sum_{n=2}^{\infty} a_n{}^2 \int_{-\infty}^{+\infty} \delta(t - t') \, dt' \, d\tau \, R_1(\tau) \rho^n(t' - \tau) \\ &= dR_0(t) \mp \sum_{n=2}^{\infty} \frac{a_n{}^2}{\sigma_x^{2n}} \int_{-\infty}^{+\infty} d\tau \, R_1(\tau) R_{xx}^n(t - \tau) \end{aligned} \tag{A.5}$$

By setting $d = 1$, (A.2) follows.

APPENDIX B. EXISTENCE AND UNIQUENESS OF THE HAMMERSTEIN INTEGRAL EQUATION[17]

The Hammerstein integral equation

$$y(s) = \int_G k(s, t) f[t, y(t)]\, dt \tag{B.1a}$$

or, in general,

$$y(s) = \int_G k(s, t) f[t, y(t)]\, dt + u(s) \tag{B.1b}$$

defined over the bounded or infinite domain $G \subseteq E^n$ ($s, t \in G$), can be written in operator-theoretic terms

$$y = \mathscr{K}\mathscr{F}y + u \tag{B.2a}$$

by defining $\mathscr{F}$ as the operator (on the space of all real-valued functions on G) $[\mathscr{F}(y)](t) = f(t, y(t))$ and $\mathscr{K}$ as the linear operator $[\mathscr{K}(y)](s) = \int_G k(s, t) y(t)\, dt$.

Thus, on a Hilbert space $\mathscr{H}$ with real or complex scalars, $\mathscr{K}$ and $\mathscr{F}$ define the mappings

$$\mathscr{K}: \mathscr{H} \to \mathscr{H}, \qquad \mathscr{F}: \mathscr{H} \to \mathscr{H} \tag{B.3}$$

An operator $\mathscr{T}$ defined on $\mathscr{H}$ will be called *monotonic*, designated $\mathscr{T} \geqslant 0$, provided that $\forall x_1, x_2 \in \mathscr{H}$ the real part of the inner product $(\cdot, \cdot)$ satisfies the inequality

$$\mathrm{Re}(x_1 - x_2, \mathscr{T}x_1 - \mathscr{T}x_2) \geqslant 0 \tag{B.4a}$$

If the operator satisfies the condition

$$\mathrm{Re}(x_1 - x_2, \mathscr{T}x_1 - \mathscr{T}x_2) \geqslant (\mathrm{Re}\,\mu)\|x_1 - x_2\|^2 \tag{B.4b}$$

it will be called μ *monotonic*, written $\mathscr{T} \geqslant \mu I$, with I being the identity. For $\mu > 0$, $\mathscr{T}$ is *strongly monotonic*.

A function $\mathscr{T}: \mathscr{H} \to \mathscr{H}$ is called *Lipschitzian* on $\mathscr{D} \subset \mathscr{H}$ if $\exists$ a number k, the Lipschitz constant, such that $\forall x_1, x_2 \in \mathscr{D}$,

$$\|\mathscr{T}x_1 - \mathscr{T}x_2\| \leqslant k\|x_1 - x_2\| \tag{B.5a}$$

The infimum of the set of Lipschitz constants is called the *Lipschitz norm* $\|\mathscr{T}\|_{\mathscr{D}}$,

$$\|\mathscr{T}\|_{\mathscr{D}} = \max\{\|\mathscr{T}x_1 - \mathscr{T}x_2\|/\|x_1 - x_2\|\}, \qquad x_1 \neq x_2 \tag{B.5b}$$

$\mathscr{T}$ is *Lipschitzian on bounded sets* if, for every bounded set $\mathscr{D} \subset \mathscr{H}$, $\|\mathscr{T}\|_{\mathscr{D}} < \infty$, and is called *bounded* if it maps bounded sets into bounded sets.

Lemma B.1. If $\mathscr{T}$ is a linear μ-monotonic function defined on $\mathscr{H}$, then $\mathscr{T}$ is bounded.

We now state, without proof, the necessary assumptions and theorems demonstrating the existence and uniqueness of $R_{\pm}(t)$. Details are given in Ref. 17. Convenient formulation of these proofs, however, will necessitate recasting (B.2a) into the following equivalent equations:

$$\mathscr{G}y = u \tag{B.2b}$$

$$y = \mathscr{W}_\lambda = \mathscr{K}_\lambda \mathscr{F}_\lambda + u_\lambda, \qquad \lambda \in \tau(\mathscr{K}) \tag{B.2c}$$

For the former, $\mathscr{G}$ ($\mathscr{G}$: $\mathscr{H} \to \mathscr{H}$) is defined by $\mathscr{G} = I - \mathscr{K}\mathscr{F}$, with $\mathscr{D}(\mathscr{G}) = \mathscr{D}(\mathscr{F})$.[7] The inverse of $\mathscr{G}$ (if it exists) will be denoted by ϕ.

For the latter equation, (B.2c), given any scalar λ and a set of scalars $\tau(\mathscr{K})$ for which $(I - \lambda\mathscr{K})$ is bijective,[8]

$$\mathscr{K}_\lambda = (I - \lambda\mathscr{K})^{-1}, \qquad \mathscr{F}_\lambda = \mathscr{F} - \lambda I, \qquad u_\lambda = L_\lambda u, \qquad L_\lambda = (I - \lambda\mathscr{K})^{-1} \tag{B.6}$$

Then, for specified minimal assumptions $\mathbf{H}_1$ and $\mathbf{H}_2$, and $\mu > \nu$,

$\mathbf{H}_1$: Suppose $\mathscr{K}$ is bounded and $\mathscr{D}(\mathscr{K}) = \mathscr{H}$; then $\exists$ a number ν, $(\nu, \infty) \subset \tau(\mathscr{K})$, such that $\forall \mu > \nu \|\mathscr{K}_\lambda\| \leqslant (\lambda - \nu)^{-1}$

$\mathbf{H}_2$: For some μ, $\mathscr{F} \geqslant \mu I$

the uniqueness of solutions to (B.2a) and consequently the existence of ϕ with domain $\mathscr{R}(\mathscr{G})$ follows with the theorem:

Theorem B.1. [Uniqueness]. Suppose $\mathbf{H}_1$ and $\mathbf{H}_2$ pertain. If $\mu > \nu$, then (B.2a) has at most one solution, and as a consequence $\mathscr{G}$ has an inverse ϕ: $\mathscr{R}(\mathscr{G}) \to \mathscr{D}(\mathscr{F})$.

The important implication of this theorem is that the existence of ϕ is completely independent of the continuity properties of $\mathscr{F}$.

The existence of solutions to (B.2a) is determined by imposing, in addition to the above minimal hypothesis, specific restrictions on $\mathscr{K}$ and $\mathscr{F}$. In particular, the restrictions on $\mathscr{F}$ are concerned with its domain, boundedness, continuity, and the range of $(\mathscr{F} + (\epsilon - \mu)I)$, $\epsilon > 0$. For $\mathscr{K}$ the restrictions are primarily directed toward its being bijective or not. The following theorems will consider situations where (i) $\mathscr{K}$ is bijective and $\mathscr{F}$ is Lipschitzian on bounded sets, (ii) $\mathscr{K}$ is bijective and $\mathscr{F}$ is bounded but not Lipschitzian, and (iii) $\mathscr{K}$ is not bijective.

[7] For any $\mathscr{T}$: $\mathscr{H} \to \mathscr{H}$, $\mathscr{D}(\mathscr{T})$ denotes its domain and $\mathscr{R}(\mathscr{T})$ its range.

[8] A function $\mathscr{K}$ is bijective on $\mathscr{H}$ if $\mathscr{K}^{-1}$ exists and $\mathscr{R}(\mathscr{K}) = \mathscr{H}$.

Theorem B.2. Assume that $\mathbf{H}_1$ and $\mathbf{H}_2$ hold true. If $\mathscr{D}(\mathscr{F}) = \mathscr{H}$ and $\mathscr{F}$ is Lipschitzian on bounded sets, then for $\mu > \nu$, ϕu exists and is the unique solution to (B.2a).

If for this theorem $\|\mathscr{K}\mathscr{F}\| < 1$, the contractive mapping theorem yields $\mathscr{G}$ by successive approximations. However, if $\mathscr{K}\mathscr{F}$ is neither Lipschitzian on $\mathscr{H}$ nor contractive on any subset of $\mathscr{H}$, one can still vary λ in (B.2c) in a way which places the norms of $\|\mathscr{K}_\lambda\|$ and $\|\mathscr{F}_\lambda\|$ into competition with each other. Thus, for some $\lambda > \nu$ and for a closed ball centered at u, $\|\mathscr{K}_\lambda\mathscr{F}_\lambda\| < 1$ and the contractive mapping theorem yields a $\mathscr{G}_\lambda = (I - \mathscr{K}_\lambda\mathscr{F}_\lambda)^{-1}$. In general, therefore, ϕu may be obtained by successive iterations on $\mathscr{W}_\lambda$.

Theorem B.3. Assume that $\mathbf{H}_1$ and $\mathbf{H}_2$ hold true and that $\mu' = \mu - \nu > 0$. If $\mathscr{K}$ is bijective, then ϕ is Lipschitzian on $\mathscr{H}$ with norm $\|\phi\| \leqslant \|\mathscr{K}^{-1}\|/\mu'$; furthermore, a solution to (B.2a) exists if for some $\sigma < \mu$, $\mathscr{R}(\mathscr{F}_0) = \mathscr{H}$.

Theorem B.4. Assume that $\mathbf{H}_1$ and $\mathbf{H}_2$ hold true and that $\mu > \nu$. If $\mathscr{K}$ is not bijective, then the existence of a solution to (B.2a) requires the boundedness of $\mathscr{F}$.

There are a number of ramifications to this theorem which affect not only the existence of (B.2a) but even the continuity of ϕ. If $\mathscr{K}$ is not bijective, then the smoothness properties of ϕ are reflected by those of $\mathscr{F}$. While ϕ is still defined on $\mathscr{R}(\mathscr{G})$, now it need not be continuous. If $\mathscr{F}$ is merely bounded, then ϕ is Hölder continuous[9] with exponent 1/2 on $\mathscr{D}(\mathscr{F}) \cap \mathscr{R}(\mathscr{G})$, with the modulus of continuity increasingly improving with continuity assumptions on $\mathscr{F}$. Thus, if $\mathscr{F}$ is Lipschitzian on bounded sets, so is ϕ.

ACKNOWLEDGMENTS

The subject of this research was suggested by Prof. K. E. Shuler and was carried out by the author during his stay as a Postdoctoral Fellow at UCSD. The author wishes to acknowledge many helpful discussions with Profs. Shuler and Lindenberg at UCSD and Prof. T. K. Caughey of the California Institute of Technology during the course of this work.

REFERENCES

1. R. Zwanzig, in *Proceedings of the Sixth IUPAP Conference on Statistical Mechanics*, The University of Chicago Press (1972), p. 241.
2. S. Nordholm and R. Zwanzig, *J. Stat. Phys.* **13**:347 (1975).

[9] A function $\mathscr{T}$ is called α-*Hölderian on* $\mathscr{D}$ ($0 \leqslant \alpha \leqslant 1$) if $\exists$ a number $k \ni$: $\forall x_1, x_2 \in \mathscr{D}$, $\|\mathscr{T}x_1 - \mathscr{T}x_2\| \leqslant k\|x_1 - x_2\|^\alpha$. It is called *Hölderian on bounded sets* if for each bounded set $\mathscr{D} \subset \mathscr{H}$ one can find a Hölder constant $k(\mathscr{D})$ and exponent $\alpha(\mathscr{D})$.

3. K. Kawasaki, *Phys. Rev.* **150**:291 (1966); *Ann. Phys.* **61**:1 (1970).
4. R. H. Kraichnan, in *Proceedings of the Sixth IUPAP Conference on Statistical Mechanics*, The University of Chicago Press (1972), p. 201.
5. Y. K. Lin, *Probabilistic Theory of Structural Dynamics*, McGraw-Hill, New York (1967), p. 258.
6. A. A. Pervozvanskii, Random Processes in Nonlinear Control Systems, in *Mathematics in Science and Engineering, Vol. 15*, Academic Press, New York (1965).
7. N. Krylov and N. Bogoliubov, Introduction á la mécanique nonlineaire: lés méthodes approchées et asymptotiques, *Ukr. Akad. Nauk Inst. Mécanique, Chaire de Phys. Math. Ann.* **1–2** (1937); transl. by S. Lefschetz, *Annals of Math. Studies*, Vol. 11, Princeton Univ. Press, Princeton, New Jersey (1947).
8. R. C. Booton, *Trans. IRE* **CT-1**:9 (1954).
9. T. K. Caughey, *J. Acoust. Soc. Am.* **35**:1706 (1963).
10. S. H. Crandall, *Zagadnienia Drgan Nielinowych* **14**:39 (1973).
11. A. B. Budgor, K. Lindenberg, and K. E. Shuler, *J. Stat. Phys.*, this issue, following paper.
12. M. Bixon and R. Zwanzig, *J. Stat. Phys.* **3**:245 (1971).
13. G. H. Klein, *J. Acoust. Soc. Am.* **36**:2095 (1964).
14. J. B. Morton and S. Corrsin, *J. Stat. Phys.* **2**:153 (1970).
15. J. K. Lubbock, *Proc. IEE* **107C**:60 (1959).
16. D. D. Šiljak, *Nonlinear Systems: The Parameter Analysis and Design*, Wiley, New York (1969).
17. I. I. Kolodner, *J. Math. Mech.* **13**:701 (1964).