Journal of Statistical Physics, Vol. 3, No. 2, 1971

The Closure Approximation in the Hierarchy Equations

G. Adomian[1]

Received December 4, 1970

The expectation of the solution process in a stochastic operator equation can be obtained from averaged equations only under very special circumstances. Conditions for validity are given and the significance and validity of the approximation in widely used hierarchy methods and the "self-consistent field" approximation in nonequilibrium statistical mechanics are clarified. The error at any level of the hierarchy can be given and can be avoided by the use of the iterative method.

KEY WORDS: Hierarchy equations; closure approximation; local independence; truncation; stochastic differential equation; perturbation; stochastic operator equation; random operator.

In the quantum mechanical many-body problem and in the theory of turbulence,[1] in the scattering or propagation of waves in random media,[2–5] and other applications,[4,6] stochastic equations[4,5,7–9] arise. The dynamical equations lead to an infinite hierarchy[10] of coupled equations (hierarchy equations) in which ensemble averages of interest are related to successively higher-order terms. A truncation procedure or closure approximation is made which terminates the hierarchy at some level to obtain a solution. Despite considerable and continuing use of this method in theoretical physics, its validity has not been adequately discussed. In general, it can be shown to be incorrect except where the randomness is relatively insignificant so that perturbation methods are satisfactory[2,5] (then, the methods are equivalent) and except also for a highly singular case (Dirac measure space) lacking in physical interest. To see this precisely and to inquire into the significance of the closure approximation, let us consider the relationship between the expected solution of a stochastic equation and the solution of the associated averaged (deterministic) equation.[7]

Supported by the National Aeronautics and Space Administration (Grant NGR 11-003-020) and partially supported by the Office of Naval Research (Contract N 00014-69-A-0423 Themis).

[1] Department of Mathematics, University of Georgia, Athens, Georgia.

822/3/2-3

If $(\Omega, \mathscr{F}, \mu)$ is a probability measure space with Ω a nonempty abstract set or sample space, $\mathscr{F}$ a Borel field of subsets of Ω, and μ a complete probability measure on $\mathscr{F}$, and if, further, $\mathscr{X}$ is a Banach space with $\mathscr{G}$ a σ-algebra of Borel subsets of $\mathscr{X}$, then a random function $\{x(t, \omega), \omega \in \Omega, t \in T\}$ can be considered a mapping from Ω into $\mathscr{X}$, where $\{\omega : x(\omega, t) \in \mathscr{G}\} \in \mathscr{F}$ for $t \in T$. Now, we consider a stochastic transformation of the random function x into another random function y, i.e., a linear stochastic operator[11,12] $H(\omega)$ mapping $\Omega \times \mathscr{X}$ into an undefined space $\mathscr{M}$ supposing $H(\omega)[x]$ is for every $x \in \mathscr{X}$ a random variable with values in $\mathscr{M}$. Thus, the output or transformed stochastic process $y(\omega) = H(\omega)[x(\omega)]$ represents the stochastic transformation or action of the random operator on the input stochastic process x.[2] Conversely, we may have a differential equation for the system—a stochastic differential equation[9,11] of the form $\mathscr{L}(\omega)[y(\omega)] = x(\omega)$, or simply $\mathscr{L}y = x$, where $x(\omega)$ is a random variable with values on $\mathscr{X}$: suppose, for example, $x(t)$ is a stationary random function representing an input voltage to an RC filter and $y(t)$ is a stationary random function representing the voltage across the capacitor. $\mathscr{L}$ is the operator $1 + RC\, d/dt$. If we suppose that C is randomly time varying, $\mathscr{L}$ is a random operator.

Let us therefore consider the expected solution of the stochastic differential equation $\mathscr{L}y = x$ where $x(t, \omega')$, $t \in T$, $\omega' \in \Omega'$ is a stochastic process and

$$\mathscr{L}(t, \omega) = \sum_{\nu=0}^{n} a_\nu(t, \omega)\, d^\nu/dt^\nu$$

is a stochastic (differential) operator by virtue of the stochastic coefficient processes $a_\nu(t, \omega)$, $t \in T$, $\omega \in \Omega$. If we have $\mathscr{L}(t, \omega)[y(t, \omega)] = x(t, \omega)$, appropriate decompositions on the probability spaces lead to our more useful form.

Clearly, y is $y(t, \omega, \omega')$ and we must solve for it in a statistical sense. Assuming existence of the integrals, the expected solution $\langle y \rangle$, or $E\{y\}$, is defined by $\int_\Omega \int_{\Omega'} y(t, \omega, \omega')\, d\mu(\omega)\, d\mu'(\omega')$, where μ and μ' are the appropriate measures over Ω and Ω'. The expectation $\langle x(t, \omega') \rangle = \int_{\Omega'} x(t, \omega')\, d\mu'(\omega')$. The average operator L will be defined by $L[\cdot] = \langle \mathscr{L}(t, \omega)[\cdot] \rangle = \int_\Omega \mathscr{L}(t, \omega)[\cdot]\, d\mu(\omega)$. Writing $\langle \mathscr{L} \rangle$ without indicating the operand, we write $\mathscr{L} = L + \mathscr{R}$, where L is a deterministic operator and $\mathscr{R}(t, \omega)$ is a random operator whose expectation is zero. The average operator L is clearly obtained from the stochastic operator $\mathscr{L}$ by replacing the coefficient processes by their means.

If $\mathscr{L}$ is replaced by L and the input x by $\langle x \rangle$, the result is a simple deterministic equation, or so-called "averaged" equation, to solve. Then $L[y] = \langle x \rangle$. It has been pointed out that the solution $\xi = L^{-1}\langle x \rangle$ is not, in general, equal to the expectation $\langle y \rangle$ of the solution process y in the stochastic equation, but further examination clarifies closure approximations which are still receiving general use. For example, the "self-consistent field" approximation used in nonequilibrium statistical mechanics (Vlasov equation for a plasma) is of this type. The following will clarify the significance of the approximation, which we wish to discourage and replace with methods which are discussed in Refs. 12–16.

[2] A forthcoming paper will develop precisely the notion of a stochastic operator on random processes in a general framework.

Suppose we simply average the equation $\mathscr{L}y = x$, writing as a consequence $\langle \mathscr{L}y \rangle = \langle x \rangle$. The left side is identical to

$$\int_{\Omega} \int_{\Omega'} \mathscr{L}(t, \omega)[y(t, \omega, \omega')]\, d\mu(\omega)\, d\mu'(\omega')$$

Suppose further that $\mu(\omega)$ can be written as the symbolic Dirac delta function in the sense of distributions, i.e., $d\mu(\omega) = \delta(\omega - \omega_0)\, d\omega$, a very singular measure on the probability space of the operator which, in effect, makes a deterministic operator out of the stochastic operator by picking out one realization or making them all identical. Then, we have

$$\mathscr{L}(t, \omega_0) \int_{\Omega'} y(t, \omega_0, \omega')\, d\mu'(\omega')$$

which can only be $L[\langle y \rangle]$, or $L\langle y \rangle$ if we drop the unnecessary brackets. This is quickly verified, since the use of the Dirac measure $d\mu(\omega) = \delta(\omega - \omega_0)\, d\omega$ means $\int_{\Omega} \mathscr{L}(t, \omega)[\cdot]\, d\mu(\omega)$ is $\mathscr{L}(t, \omega_0)$ and L by definition. Hence, we have $L[\langle y \rangle] = \langle x \rangle$, which means the expected solution $\langle y \rangle$ of the stochastic equation is the solution ξ of the averaged equation if the δ-function measure is used on the space of $\mathscr{L}$, i.e., $\langle \mathscr{L}[y] \rangle_{\mathscr{L},y} = \langle L[y] \rangle = L[\langle y \rangle]$. (This was recognized by Bharucha-Reid.[7])

The above simply means that, if there is no randomness in the operator and the stochastic operator $\mathscr{L}$ can be replaced by the deterministic operator $\langle \mathscr{L} \rangle$ or, equivalently, L, we have thrown away the $\mathscr{R}$. Hence, $\langle \mathscr{L}y \rangle = \langle Ly \rangle = L\langle y \rangle$ or $\langle \mathscr{L} \rangle \langle y \rangle$.

If $\mathscr{R}$ is not ignored, $\langle \mathscr{L}y \rangle \neq \langle \mathscr{L} \rangle \langle y \rangle$, since $\langle \mathscr{R}y \rangle$ will not separate except within the confines of a perturbation approximation. The Dirac measure on the space of $\mathscr{L}$ means $\mathscr{L}$ is no longer stochastic. Then the expectation *is* obtained by averaging everything and the averaging method is justified only in this highly singular case (and in some pathological cases[3]).

If the measure spaces Ω' and Ω are identical, we have (now further omitting the t for simplicity in notation)

$$\mathscr{L}(\omega)\, y(\omega) = x(\omega)$$

$$\langle x(\omega) \rangle = \int x(\omega)\, d\mu(\omega) = \langle \mathscr{L}y \rangle = \int \mathscr{L}(\omega)\, y(\omega)\, d\mu(\omega)$$

which, in the case of the δ-function measure, is

$$\langle x \rangle = \mathscr{L}(\omega_0)\, y(\omega_0) = L\langle y \rangle$$

since

$$\langle \mathscr{L} \rangle = \int \mathscr{L}(\omega)\, \delta(\omega - \omega_0)\, d\omega = \mathscr{L}(\omega_0) = L$$

and

$$\langle y \rangle = \int y(\omega)\, \delta(\omega - \omega_0)\, d\omega = y(\omega_0)$$

[3] Where the rates of variation are very different, a fact not known before the solution is made.

To verify that no other possibilities exist, write $\langle \mathscr{L}y\rangle = L\langle y\rangle$ by assumption without the Dirac measure. Then,

$$\begin{aligned}\langle y\rangle &= L^{-1}\langle x\rangle \\ &= L^{-1}\iint \mathscr{L}(t,\omega)[y,\omega,\omega')]\,d\mu(\omega)\,d\mu'(\omega') \\ &= L^{-1}\iint (L+\mathscr{R})[y]\,d\mu\,d\mu' \\ &= \langle y\rangle + L^{-1}\iint \mathscr{R}(t,\omega)[y(t,\omega,\omega')]\,d\mu(\omega)\,d\mu'(\omega') \\ &= \langle y\rangle + L^{-1}\langle \mathscr{R}y\rangle\end{aligned}$$

i.e., $\langle \mathscr{R}y\rangle$ must be zero; unless the second term is zero for all t, the assumed separability leads to a contradiction. We can see that this should be the case by writing $\mathscr{L}y = x$ as $(L+\mathscr{R})\,y = x$, or $y = L^{-1}x - L^{-1}\mathscr{R}y$, and then averaging to get

$$\langle y\rangle = L^{-1}\langle x\rangle - L^{-1}\langle \mathscr{R}y\rangle$$

If we want $\langle y\rangle$ to equal $L^{-1}\langle x\rangle$, then $\langle \mathscr{R}y\rangle$ must be zero. The δ-function simply picks out a nonrandom operator from the $\mathscr{R}$ i.e., its average, which is absorbed into L. For other measures, we can now assume $\mathscr{R}$ is "small" i.e., $\mathscr{R} = \epsilon\mathscr{L}_1$ where $\mathscr{L}_1$ is a small stochastic perturbation on L, with $\langle \mathscr{L}_1\rangle = 0$. Then,

$$\langle \mathscr{R}y\rangle = \langle \epsilon\mathscr{L}_1(y_0 + \epsilon y_1 + O(\epsilon^2))\rangle,$$

where $y_0 = \langle y\rangle$. Consequently, $\langle \mathscr{R}y\rangle = \langle \mathscr{R}\langle y\rangle\rangle + O(\epsilon^2) = \langle \mathscr{R}\rangle\langle y\rangle + O(\epsilon^2)$.[4] The first term is zero when $\mathscr{R}$ is zero mean. Within the confines of perturbation theory, the separation is valid for any measure, or it is valid for the singular measure without assuming $\mathscr{R}$ is small. Thus, $\langle \mathscr{R}y\rangle = 0$ can be taken as a necessary and sufficient condition for the validity of the separation. It can be zero because $\mathscr{R}$ is zero or, using a perturbation series, because $\langle \mathscr{R}\rangle$ is zero.

Let us now take a simple case of $\mathscr{R}$ involving no derivatives, i.e., $\mathscr{L} = L + \alpha(t,\omega)$, where the stochastic equation is $[L + \alpha(t,\omega)]\,y(t,\omega,\omega') = x(t,\omega')$. The term $\langle \mathscr{R}y\rangle$ is now

$$\langle \alpha(t,\omega)\,y(t,\omega,\omega')\rangle = \iint \alpha(t,\omega)\,y(t,\omega,\omega')\,d\mu(\omega)\,d\mu'(\omega')$$

Clearly, the δ-function measure gives the desired result since $\langle \mathscr{L}\rangle = 0$. If $\langle \mathscr{R}\rangle$ is nonzero, it would be absorbed into L in any event; we chose $L = \langle \mathscr{L}\rangle$. If we use other measures over Ω, we get $\int \alpha(t)\,y(t,\omega')\,d\mu'(\omega')$. Since $\langle \alpha(t,\omega)\rangle = \int \alpha(t,\omega)\,d\mu(\omega)$

[4] Let $\Omega = (-1/2,\ 1/2)$ with Lebesgue measure, $\omega \in \Omega$. Let $\mathscr{L} = d/dt + \epsilon\omega$, with $0 < \epsilon \ll 1$, and $x(t,\omega) = t + \omega$. Hence, $L = d/dt$ and $R = \epsilon\omega$. $\langle R\rangle = \epsilon/2$. Now, $\langle \mathscr{R}y\rangle = \langle \epsilon\omega y\rangle = \langle \epsilon\omega y_0 + \epsilon y_1 + O(\epsilon^2) = \langle \epsilon\omega\langle y\rangle\rangle + O(\epsilon^2) = (\epsilon/2)\langle y\rangle + O(\epsilon^2) = \langle R\rangle\langle y\rangle + O(\epsilon^2)$.

is assumed to be zero, but α, or $\mathscr{R}$, is assumed to be nontrivial,[5] we cannot have the double integral equal to zero *for all t* and cannot therefore have $\langle \mathscr{L} y \rangle = \langle \mathscr{L} \rangle \langle y \rangle$.

We have seen that, except within the confines of a perturbation theory, the averaging is justified only for this highly singular measure to make both the operator and the solution process in $\langle \mathscr{L} y \rangle$ be deterministic and separable into $\langle \mathscr{L} \rangle \langle y \rangle = L \langle y \rangle$. Otherwise, the solution ξ of $L[y] = \langle x \rangle$ is not $\langle y \rangle$.[6]

In the commonly used hierarchy equation method of solving stochastic equations for various moments, an essential truncation or closure approximation is used to terminate an otherwise infinite system of equations to get a solution. (This truncation assumption has been discussed by Kraichnan, Keller, Bharucha-Reid, Richardson, and others; however, its exact validity has remained a matter of doubt.) It is, however, precisely equivalent to the separation $\langle \mathscr{L} \rangle \langle y \rangle$ if done at the first stage of the hierarchy. At the second level (supposing again $\mathscr{L} = L + \mathscr{R}$, where $\mathscr{R}$ is the random part of the operator and $\langle \mathscr{L} \rangle = L$), we obtain $\langle \mathscr{R} L^{-1} \mathscr{R} y \rangle$, which is separated by handwaving arguments into $\langle \mathscr{R} L^{-1} \mathscr{R} \rangle \langle y \rangle$ so the resulting equation can be solved for $\langle y \rangle$. The procedure is valid within the framework of perturbation theory, i.e., if $\mathscr{R} = \epsilon \mathscr{L}_1 + \epsilon^2 \mathscr{L}_2 + O(\epsilon^3)$, where $\mathscr{L}_1$, $\mathscr{L}_2$ are small (random) perturbations of L.

This arises as follows. We have $\mathscr{L} y = (L + \mathscr{R}) y = x$. Averaging, we get $L \langle y \rangle + \langle \mathscr{R} y \rangle = \langle x \rangle$. Since $\langle \mathscr{R} y \rangle$ causes difficulties, we return to the original equation and write $Ly = x - \mathscr{R} y$, or $y = L^{-1} x - L^{-1} \mathscr{R} y$. Operating from the left by $\mathscr{R}$, we have $\mathscr{R} y = \mathscr{R} L^{-1} x - \mathscr{R} L^{-1} \mathscr{R} y$. Assuming statistical independence of $\mathscr{L}$ and x and averaging,

$$\langle \mathscr{R} y \rangle = \langle \mathscr{R} \rangle \, L^{-1} \langle x \rangle - \langle \mathscr{R} L^{-1} \mathscr{R} y \rangle$$

The first term on the right is zero if $\mathscr{R}$ has zero mean. (Of course,

$$\langle \mathscr{R} L^{-1} x \rangle \neq \langle \mathscr{R} \rangle \, L^{-1} \langle x \rangle$$

in general but only if statistical independence of $\mathscr{L}$ and x is assumed.[7]) The second term is generally separated with little or no justification into $\langle \mathscr{R} L^{-1} \mathscr{R} \rangle \langle y \rangle$. If we grant this for the moment, then $L \langle y \rangle - \langle \mathscr{R} L^{-1} \mathscr{R} \rangle \langle y \rangle = \langle x \rangle$ is easily solved for $\langle y \rangle$. The separation $\langle \mathscr{R} L^{-1} \mathscr{R} y \rangle \simeq \langle \mathscr{R} L^{-1} \mathscr{R} \rangle \langle y \rangle$ is similar to the separation $\langle \mathscr{L} y \rangle \simeq \langle \mathscr{L} \rangle \langle y \rangle$ except it is done at the next higher level of the hierarchy. It means exactly the replacement of the random operator $\mathscr{R} L^{-1} \mathscr{R}$ by its deterministic equivalent; then the left expectation bracket can be moved to the right until only y is enclosed by brackets. Just as we replaced $\mathscr{L}$ by $\langle \mathscr{L} \rangle$, i.e., by L, before, we have now replaced $\mathscr{R} L^{-1} \mathscr{R}$ by the correlation term $\langle \mathscr{R} L^{-1} \mathscr{R} \rangle$. At whatever level of the hierarchy we

---

[5] We are not interested in either y identically zero or $\langle y \rangle$ equal to zero [and further, we do not allow $y(t, \omega')$ to be written as $y_0 + \epsilon y_1 + \epsilon^2 y_2 + \ldots$ with y_0 equal to $\langle y \rangle$ in order to have a first term independent of ω'].

[6] One could conceivably write the exact equation for $\langle y \rangle$ in the form $\langle \mathscr{L}^{-1} \rangle^{-1} \langle y \rangle = \langle x \rangle$ and expand $\langle \mathscr{L}^{-1} \rangle^{-1}$ in a power series and truncate the series to get the perturbation result, but the inverse of a stochastic matrix is to be avoided.

[7] E.g., let $\Omega = (0, 1)$ with Lebesgue measure, $\omega \in \Omega$. Let $\mathscr{L} = (1 + \omega) \, d/dt$. Then, $L = d/dt$ and $\mathscr{R} = \omega \, d/dt$. Let $x(t, \omega) = t + \omega$. Now, $\langle \mathscr{R} L^{-1} x \rangle = (t/2) + (1/3)$, $\langle \mathscr{R} \rangle = (1/2) \, d/dt$, $\langle x \rangle = (1/2) + t$, and $L^{-1} \langle x \rangle = (t/2) + (t^2/2)$. Hence, $\langle \mathscr{R} \rangle \, L^{-1} \langle x \rangle = (1/4) + (t/2) \neq \langle \mathscr{R} L^{-1} x \rangle$.

make such a "closure approximation" or "local independence assumption," we are replacing a more and more complicated operator, depending on the level, of the form $\mathscr{R}L^{-1}\mathscr{R}L^{-1}\mathscr{R}\cdots\mathscr{R}L^{-1}\mathscr{R}$, by its expectation. Herein lies the significance of the closure approximation. Separation of $\langle\mathscr{L}y\rangle$ at the first stage, $\langle\mathscr{R}L^{-1}\mathscr{R}y\rangle$ at the second, or $\langle\mathscr{R}L^{-1}\mathscr{R}\cdots\mathscr{R}L^{-1}\mathscr{R}y\rangle$ at higher stages, from the y requires ignoring the randomness of the appropriate operator part, which is what we do by letting $d\mu(\omega) = \delta(\omega - \omega_0)\,d\omega$ on the measure space Ω. The error, or, at least, bounds on the error can now be determined by comparing the result for each stage of approximation with the results given in Refs. 12 and 13 and we propose to do this in a following paper.

Some additional comments may be useful. Since perturbation methods are limited to small randomness and solving "averaged" equations is only valid within the same framework or in the singular case (where the physical interest also vanishes) where the operator is no longer stochastic, the most satisfactory method so far available for the solution of a stochastic differential equation without highly restrictive assumptions is the iterative method of Adomian,[13] although a more general method may be possible[13,14] and is receiving further study. Applications are numerous, e.g., the propagation of electromagnetic waves in a stochastic medium.[4,5,12,16] Sibul[18] and Adomian[12] (see Part III) have shown that the iterative method is also useful in the wave propagation problem, extending without substantial difficulty to the case of integral equations and partial differential equations. In the case of integral equations, the relationship to the resolvent kernel[15] clarifies the solution. The method is not limited to expectation of the solution process. Covariance functions (or covariance matrices or mutual coherence functions) and other statistics can be found in terms of stochastic Green's functions.[12,14]

REFERENCES

1. R. H. Kraichnan, Dynamics of nonlinear stochastic systems, *J. Math. Phys.* **2**:124–148 (1961).
2. Y. M. Chen, On the scattering of waves by random distributions, *J. Math. Phys.* **5** (11): 1541–1546 (Nov. 1964).
3. V. Twersky, On the scattering of waves by random distributions, *J. Math. Phys.* **3**:700–715, 724–734 (1962).
4. G. Adomian, Linear stochastic operators, Ph.D. Dissertation (Physics), UCLA (1961), Chapter 4.
5. J. B. Keller, Stochastic equations and wave propagation in random media, in *Annual Symposia in Applied Mathematics*, Vol. XVI (Amer. Math. Soc., 1964).
6. J. C. Samuels and A. C. Eringen, On stochastic linear systems, *J. Math. Phys.* **38**:83–103 (1959).
7. A. T. Bharucha-Reid, On the theory of random equations, in *Annual Symposia in Applied Mathematics*, Vol. XVI (Amer. Math. Soc., 1964), pp. 40–69.
8. A. T. Bharucha-Reid, *Random Equations* (Academic Press, New York, in publication).
9. R. Syski, Stochastic differential equations, Chapter 8 in *Modern Nonlinear Equations*, T. L. Saaty, Ed. (McGraw-Hill, New York, 1967).
10. J. M. Richardson, The application of truncated hierarchy techniques in the solution of a stochastic linear differential equation, *Annual Symposia in Applied Mathematics*, Vol. XVI (Amer. Math. Soc., 1964).
11. G. Adomian, Linear stochastic operators, *Rev. Mod. Phys.* **35** (1): 185–207 (Jan. 1963).
12. G. Adomian, Linear random operator equations in mathematical physics, Part I, *J. Math. Phys.* **11** (3):1069–1084 (March 1970); Parts II and III to appear in Sept. 1971 and Oct. 1971.

13. G. Adomian, Theory of random systems, in *Proc. Fourth Internat. Conf. on Information Theory, Statistical Decision Functions, and Stochastic Processes* (Czechoslovak Academy of Sciences, Prague, 1967).
14. G. Adomian, Stochastic Green's functions, in *Proc. of Annual Symposia in Applied Mathematics*, Vol. XVI, Invited Paper (Amer. Math. Soc., 1964).
15. L. H. Sibul, Stochastic Green's functions and their relation to the resolvent kernels of integral equations, in *Proc. of the Fifth Allerton Conference on Circuit and Systems Theory* (University of Illinois, 1967).
16. G. Adomian and L. H. Sibul, Propagation in a stochastic medium, forthcoming paper.
17. Klimontovich, *The Statistical Theory of Nonequilibrium Processes in a Plasma* (MIT Press, 1967).
18. L. H. Sibul, Application of linear stochastic operator theory, Ph.D. Dissertation, Pennsylvania State University (1968).