JOURNAL OF COMPUTATIONAL PHYSICS 2, 288–305 (1968)

On the Realization of Random Surfaces

M. G. RUSBRIDGE[1]

Culham Laboratory, Abingdon, Berkshire, England

ABSTRACT

For certain numerical experiments on the behavior of fluctuations in a plasma it is useful to be able to construct examples of random surfaces. These are functions of two variables which are members of ensembles which can be characterized only by statistical properties. We consider only homogeneous surfaces in which these statistical properties are independent of position. The problem we consider is this: given a correlation function and sets of random numbers, to operate on these sets so as to produce successively members of our ensemble of functions with the specified correlation function. Two possible methods of doing this are described and contrasted, one based on Fourier synthesis and the other on the theory of optimum linear prediction.

1. INTRODUCTION

While the theory of random functions of one variable is well understood, random functions of two variables have received much less attention. Such functions can be described as *random surfaces*. From a survey of their properties [1], I find, as a rule of thumb, that only those properties which can be described in terms of the value of the function itself and its derivatives at a point are well understood. We are in almost complete ignorance of properties which cannot be described in this way and may be called nonlocal. For example, the statistical properties of contours of the surface are practically unknown. Such properties may be approached through Monte-Carlo-type calculations, whereby examples of random surfaces are set up and examined numerically. Such an approach may indeed be suggestive enough ultimately to generate approximate analytical solutions to these problems. In this report I describe and compare two methods for constructing numerical examples of random surfaces. This work was undertaken mainly to form the basis of a specific Monte-Carlo calculation of the scattering of microwave radiation by the turbulent plasma in the Zeta discharge.

In Section 2, I delineate the problem more precisely, indicate the general method of its solution, and briefly discuss the method of Fourier synthesis. Section 3

[1] Present address: University of Manchester Institute of Science and Technology, P.O. Box 88, Sackville St., Manchester 1, England.

contains a summary of the theory of optimum linear prediction and an application of it to the construction of examples of random functions of time, and in Section 4 its extension to two dimensions and application to the construction of random surfaces is discussed.

2. Random Surfaces

By analogy with the definition of a random process (see, for example, Bendat [2]), a random surface may be defined as a member of an ensemble of functions $\{f_k(x, y)\}$, $-\infty < x < \infty$, $-\infty < y < \infty$, which can be characterized only through its statistical properties. The index k separates different members of the ensemble. Averages over the ensemble are denoted by brackets $\langle\ \rangle$. We shall assume $\langle f_k(x, y)\rangle = 0$ for all x and y. The principal statistical properties are the variance V, defined by

$$V = \langle f_k^2(x, y)\rangle$$

and the correlation function defined by

$$R(\xi, \eta) = \frac{\langle f_k(x, y)\, f_k(x + \xi, y + \eta)\rangle}{V}.$$

We shall consider only homogeneous surfaces for which V and R are independent of x and y. We shall further assume that the probability distribution of the surface is Gaussian in a generalized sense, so that the specification of V and R is sufficient to specify all higher order moments and correlation functions (S. F. Edwards, private communication). This corresponds to the experimental situation: we have information about the form of R but usually no independent information about higher order correlations.

Finally we assume that the surface is isotropic, i.e.

$$R(\xi, \eta) \equiv R(\rho) \qquad \text{where} \quad \rho = (\xi^2 + \eta^2)^{1/2}.$$

This restriction could be removed in principle.

What we are after, then, is a method of generating a particular $f_k(x, y)$. Both the methods to be described are basically operations applied to sets of random numbers, chosen from a Gaussian distribution, to generate values of the function at a mesh of suitably spaced points in the (x, y) plane. Of course, if the points were spaced so far apart that the values were statistically independent, then the random numbers themselves would serve as values of the function, after a trivial normalization to give the correct value of the variance V. But in this case we should have lost the fine structure of the surface represented by the correlation function, and the aim of the present work is to keep in this structure and examine the effects of varying it. The spacing of the points, therefore, must certainly be

small compared with the integral scale L of the correlation function, which is defined by

$$L = \int_0^\infty R(\rho)\, d\rho \tag{1}$$

and corresponds to the usual intuitive notion of the correlation length.

More generally, we must note that physically occurring correlation functions contain (at least) two independent length scales, of which one is L and the other can be defined by the behavior of R near $\rho = 0$. Since $R''(0)$ is a negative-definite quantity ([2], p. 21), $R(\rho)$ must be of the form

$$R(\rho) \simeq 1 - \rho^2/\lambda^2 + \cdots \tag{2}$$

for small ρ, where λ may be called the *gradient scale*. We shall be interested in problems in which the effect of varying λ is to be examined; so, clearly, the spacing of the points should be small compared with a typical value of λ, and this, in turn, is usually smaller than L.

Of course, only a finite portion of each surface can be generated. It will be adequate, however, if the linear dimensions of the sample are large compared with L.

We shall calculate the values of the function at a square mesh of points x_i, y_j, $i, j = 1...N$, with mesh spacing δ and N points on a side. The above requirements amount to

$$\delta \ll \lambda \ll L \ll N\delta. \tag{3}$$

But evidently if "$\ll$" is interpreted to imply an order-of-magnitude difference, then $N \approx 10^3$, which would in turn require the generation of $N^2 \approx 10^6$ numbers. This is quite impracticable. The largest surface I have computed has $N = 25$, and this takes about 2-minutes' computing time. It may be shown that to satisfy (3) strictly would require of the order of 10^5 times as long. Accordingly, half an order-of-magnitude difference is the most that can be obtained; the following set of values is typical:

$$\delta = 1, \qquad \lambda = 2.42,$$
$$L = 6.33, \qquad N = 25.$$

An actual surface corresponding to these values is shown in Fig. 4b. Whether the surfaces are still useful under this restriction can only be decided in connection with a specific problem. With these values $N\delta/L$ is only about 4, but the whole surface contains about $(N\delta/L)^2 \approx 15$ independent regions.

It should be clear from all this that the essential difficulty of the problem is that neighboring points are not independent. This suggests the first possible approach to a solution, which is to seek a set of elements that are independent, from which the surface can be constructed, and to assign to each a value randomly chosen from a suitable distribution. For a homogeneous surface these elements are the

Fourier components, and accordingly we describe this as the method of Fourier synthesis. The other procedure is to take explicit account of the correlations between neighboring points; this can be done by an extension of the theory of optimum linear prediction.

These methods are really suitable only when the distribution of the f_k is Gaussian. In each case the value of f_k is given ultimately as a linear combination of random numbers, and although the central limit theorem will not strictly apply because the coefficients may be of very different sizes, nevertheless we should expect the distribution to tend towards the Gaussian form, irrespective of the distributions of the random numbers themselves. I have used Gaussian distributions for both the f_k and the random numbers.

The method of Fourier synthesis will be described briefly. The two-dimensional power spectrum $G(k_x, k_y)$ is defined by

$$G(k_x, k_y) = \frac{4}{\pi^2} \int_0^\infty \int_0^\infty R(x, y) \cos k_x x \cos k_y y \, dx \, dy$$

where k_x and k_y are the x and y components of the wave vector $\mathbf{k}$. We set up a mesh of points in k-space with spacing δk, say. Each point on the mesh represents a Fourier mode, and to each such mode we assign a random amplitude $A(k_x, k_y)$, chosen from a Gaussian distribution with variance $G(k_x, k_y)\,\delta k^2$, and a random phase $\phi(k_x, k_y)$ chosen from a distribution uniform in the interval $(-\frac{1}{2}\pi, \frac{1}{2}\pi)$ (the rest of the full interval $(-\pi, \pi)$ is represented by negative values of A). Then

$$f_k(x, y) = \sum_{k_x, k_y} A(k_x, k_y) \cos(k_x x + k_y y + \phi(k_x, k_y)) \tag{4}$$

is the function we require.

This procedure was programmed and carried out on the Culham Laboratory KDF-9 computer, but it proved to be very slow, with each surface taking of the order of 15 minutes computing time. Since any problem we may wish to investigate using these surfaces will be statistical, requiring the computation of many surfaces, this is depressingly long, and since the other method was found to require only two minutes per surface, I did not pursue this method further.

A partial explanation for the slowness of the Fourier synthesis method is as follows. The number of modes must be at least of the same order as the number of spatial points, i.e., N^2; each mode contributes to each spatial point in Eq. (4), giving a computing time proportional to N^4. In the other method there are still N^2 spatial mesh points, but for each of these we need take account of correlations with only N_0^2 other points, where $N_0 \sim L/\delta$. The total computing time therefore is proportional to $N_0^2 N^2$, and, other things being equal we would achieve an increase in speed of $\sim N^2/N_0^2 = N^2\delta^2/L^2 \approx 15$ in our example above. We have taken advantage of the independence of different regions of the surface.

3. Optimum Linear Prediction

Suppose for the moment that we are considering a random process, i.e., a random function of time belonging to an ensemble $\{f_k(t)\}$ which we shall assume to be stationary i.e. with variance independent of the time. Suppose we know the values of a particular function $f(t)$ at discrete times in the past, given by

$$t = t_{-n} = t_0 - n\Delta t, \qquad 0 \leqslant n,$$

and wish to construct the function at times in the future given by

$$t = t_n = t_0 + n\Delta t, \qquad 0 < n.$$

It is adequate to construct the value at $t = t_1$ only, after which further values can be constructed by iteration. Then $f(t_1)$ is a random variable which is to some extent correlated with $f(t_0)$, $f(t_{-1})$, etc. We may consider the subensemble of functions $\{f_s(t)\}$ with the following properties: every $f_s(t)$ is a member of $\{f_k(t)\}$, and for every $f_s(t)$,

$$f_s(t_{-n}) = f(t_{-n}), \qquad 0 \leqslant n.$$

Thus the subensemble includes all those functions whose values in the past fit the known set of values, and by "constructing" the value of $f(t_1)$ we mean choosing at random a particular f_s and evaluating $f_s(t_1)$. We shall denote subensemble averages by $\langle\ \rangle_s$. Evidently the (subensemble) mean and variance of $f_s(t_1)$ are in general functions of $f(t_0)$, $f(t_{-1})$, etc., and this expresses the correlation between them. Let us consider the simplest case, when the mean and variance are functions only of $f(t_0)$, i.e., the "present", rather than the past. Such a process is by definition a *Markov process*, and if we write $q = R(\Delta t)$, where R is the correlation function, the following results may be established:

$$\begin{aligned} \langle f_s(t_1)\rangle_s &= qf(t_0), \\ V_s(t_1) &= 1 - q^2, \end{aligned} \tag{5}$$

where $V_s(t_1)$ is the variance of $f_s(t_1)$, and we assume the ensemble variance to be unity. Moreover the correlation function must be of the form

$$R(\tau) = \exp\left(-\frac{|\tau|}{\Delta t}\ln\frac{1}{q}\right);$$

see, for example, Bendat ([2] pp. 215 and 178). The converse is true: if $R(\tau)$ has this form then $f(t_1)$ can be chosen from a Gaussian distribution with parameters given by Eq. (5). Thus for this case our problem is solved.

This form of correlation function is, however, quite restrictive, and indeed nonphysical [3]. To remove the restriction we might consider what happens as we include a dependence on successively more and more past values of $f(t)$. Happily

this is unnecessary, for the problem has been solved in general by Wiener [4]. In fact we may identify $\langle f_s(t_1)\rangle_s$ with Wiener's optimum linear prediction of $f(t_1)$, and $V_s(t_1)$ with his mean-square error in this prediction. Wiener's solution is given for functions whose whole past is known, but it is simple to adapt it to the case of a discrete set of values. We assume that $\langle f_s(t_1)\rangle_s$ can be written as

$$\langle f_s(t_1)\rangle_s = \int_0^\infty h(\Delta t, \tau)\, f(t_0 - \tau)\, d\tau, \tag{6}$$

where $\Delta t = t_1 - t_0$; then Wiener shows that

$$V_s(t_1) = 1 - \int_0^\infty h(\Delta t, \tau)\, R(\tau + \Delta t)\, d\tau \tag{7}$$

and that the function h is the solution of the integral equation

$$R(\nu + \Delta t) = \int_0^\infty h(\Delta t, \tau)\, R(\nu - \tau)\, d\tau. \tag{8}$$

Wiener shows that this method is optimum in the sense that no other linear solution with smaller $V_s(t_1)$ can be found, i.e. that it represents the most information we can possibly have about $f(t_1)$ knowing the past of $f(t)$. Further, for a large class of functions $f(t)$ no nonlinear method of prediction can be any better. Wiener gives a formal solution of Eq. (8) which for our purposes is only useful as a proof of existence. However, it does disclose two interesting results, first that in the continuous case $h(\Delta t, \tau)$ is a function of $d/d\tau$ as well as τ in general, i.e., it is an operator, and second, that there exist forms for $R(\tau)$ for which the future of $f(t)$ is entirely determined by its known past, i.e., the mean-square error $V_s(t_1)$ reduces to zero. A sufficient condition for this is that the corresponding power spectrum $G(\omega)$ should decrease as fast as $e^{-\omega}$ for large ω. This is a bit alarming since physically occurring power spectra probably do just this, and this would imply that the whole function $f(t)$ depends on its values at $t \to -\infty$, quite at variance with our intuitive motion of a random function. However, fortunately, in the discrete case when the function is known only at times separated by Δt, the condition is much less strong, since our knowledge of the past of $f(t)$ is much less complete. For $V_s(t_1)$ to be finite, it is then sufficient that $G(\omega)$, which is now a periodic function with period $2\pi/\Delta t$, should have no more than a finite number of separated zeros in one period.

As a simple example of the solution of Eq. (8), Wiener shows that for $R = e^{-\alpha|\tau|}$ we have

$$h(\Delta t, \tau) = 2\delta(\tau)\, e^{-\alpha \Delta t},$$

where $\delta(\tau)$ is the delta-function. Thus, for example,

$$\begin{aligned}\langle f_s(t_1)\rangle_s &= \int_0^\infty h(\Delta t, \tau)\, f(t_0 - \tau)\, d\tau \\ &= f(t_0)\, e^{-\alpha \Delta t} \\ &= q f(t_0),\end{aligned}$$

in agreement with Eq. (5).

In the discrete case, Eqs. (6)–(8) are replaced by

$$\langle f_s(1)\rangle_s = \sum_0^J h(j)\, f(-j) \tag{9}$$

where

$$\begin{aligned} h(j) &\equiv h(\Delta t, j\Delta t), \\ f(-j) &\equiv f(t_0 - j\Delta t), \\ f_s(1) &\equiv f_s(t_0 + \Delta t) \equiv f_s(t_1), \end{aligned}$$

and

$$V_s(1) = 1 - \sum_0^J h(j)\, R(1 + j) \tag{10}$$

where

$$R(j) \equiv R(j\Delta t) \qquad \text{and} \qquad V_s(1) \equiv V_s(t_1),$$

and J is a number so great that $R(j)$ is negligible for all $j > J$ (such a number must always exist). The integral equation becomes

$$R(1 + n) = \sum_0^J h(j)\, R(n - j), \qquad n = 0, 1 \ldots J, \tag{11}$$

a set of simultaneous linear equations which can be solved numerically by the usual matrix-inversion methods. I have written a computer program to solve these equations and to construct examples of random processes. In Table I, I give the values of $h(j)$ obtained for a particular form for R given by

$$R(\tau) = \exp[ab - a(b^2 + \tau^2)^{1/2}] \tag{12}$$

for $a = 0 \cdot 2$, $b = 5 \cdot 0$, and $J = 24$. This form for $R(\tau)$, which I have used in all the remaining part of this work, is a convenient and physically realistic replacement for the simple form $e^{-\alpha|\tau|}$, to which it tends in the limit $b \to 0$. For convenience the properties of this function are summarized in Appendix I. It satisfies the requirement that the associated power spectrum be positive definite, and is chosen as an approximate representation of measured correlation functions in an experimental situation which is approximately two-dimensional and isotropic.

Now we can follow a particular $f(t)$ through time by calculating $\langle f_s(1)\rangle_s$ and $V_s(1)$ [the latter is in fact independent of the particular $f(t)$ and can be computed once and for all when $h(j)$ is calculated], choosing $f(1)$ randomly from a Gaussian distribution with these parameters, and repeating the process iteratively to give successively $f(2)$, $f(3)$, etc. However, we must have some initial values to start

TABLE I

THE FUNCTION $h(j)$

j	$h(j)$
0	3.8672
1	−7.2708
2	9.2983
3	−9.4957
4	8.5006
5	−7.0519
6	5.6006
7	−4.3393
8	3.3151
9	−2.5126
10	1.8955
11	−1.4260
12	1.0707
13	−0.8024
14	0.6001
15	−0.4472
16	0.3314
17	−0.2431
18	0.1752
19	−0.1224
20	0.0809
21	−0.0485
22	0.0244
23	−0.0089
24	0.0017

such a process; and this is a difficulty because, since the process is supposed to have been going on since $t = -\infty$, such initial values are not to be obtained. There are two methods of starting such a process, one rigorous, the other practicable.

(1) In 1950, Zadeh and Ragazzini [5] extended Wiener's analysis to the case where information about the part of the function $f(t)$ is available only over a limited time T. The result is simple enough: in Eqs. (6), (7), (8) the upper limit of integration is T instead of ∞. Analogously, if we know $f(t)$ at N discrete times in the past, Eq. (11) becomes

$$R(1 + n) = \sum_{0}^{N-1} h(j, N)\, R(n - j) \tag{12}$$

assuming $N < J$ (otherwise there is no problem). Notice that h now depends on N as well as j. To start the process we choose a random value for $f(1)$ with mean zero and variance unity; this is consistent with the absence of any information about the past of $f(t)$. Now we can proceed iteratively, so that to obtain $f(N + 1)$ we solve Eq. (12) to obtain $h(j, N)$, and use this in Eqs. (9) and (10) where J is to be replaced by $N - 1$. This is continued until $N - 1 > J$ when Eqs. (9)–(11) can be used without change. The disadvantage of this is that it requires the solution of several sets of simultaneous equations, which is a time consuming process if J is large.

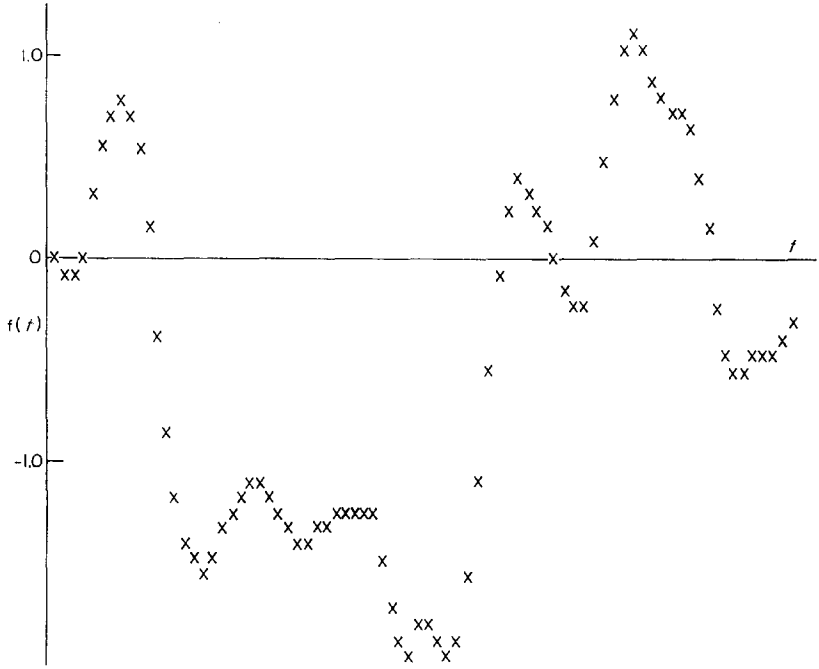

FIG. 1. An example of a random process, starting from a fixed origin.

(2) We can choose a set of $J + 1$ arbitrary values of f, follow the process as described above using a single h-function, and discard the first J' values of f, where $J' \gtrsim J$. The arbitrary set of values can be regarded as a particular sort of statistical fluctuation of $f(t)$, and if we follow the process long enough all "memory" of the initial state will have worn off. This happens within a time rather larger than the integral scale defined above.

This second method, which only requires the computation of one h-function, was the one actually used. As a test of the whole procedure, and to assist in choosing a value of J', I have computed 200 independent functions $f(t)$, following each for 80 time steps, with the initial values all set to zero. The correlation function used was given by Eq. (12) with $a = 0 \cdot 2$, $b = 5 \cdot 0$, and $J = 24$, so that the h-function is that shown in Table I. For this case the integral scale is $8 \cdot 3$. Figure 1 shows a particular $f(t)$, and Fig. 2 the r.m.s. amplitude (averaged over the ensemble of 200 functions) as a function of time. After 12 time steps this quantity shows no more than the expected statistical fluctuations, and the ensemble of functions can be regarded as quasi-stationary from then on. This is important for two reasons.

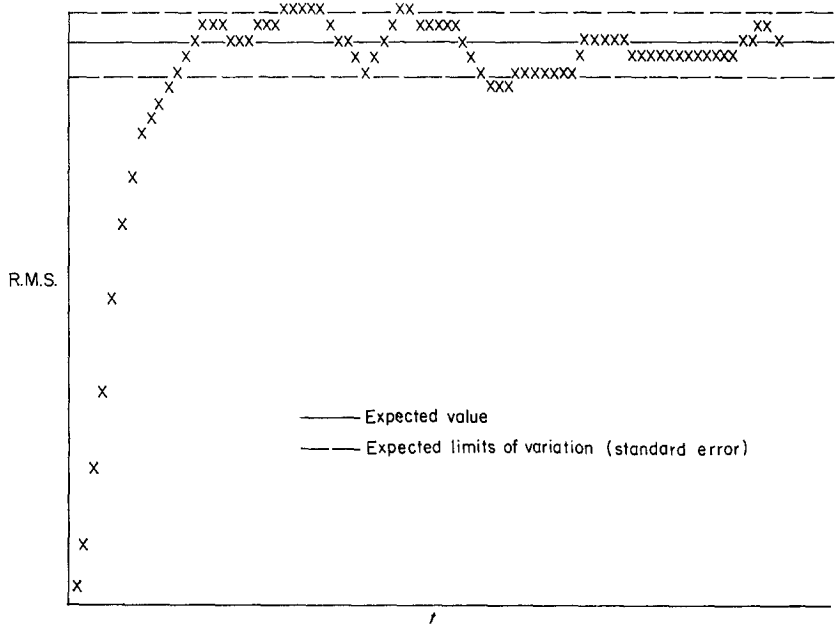

FIG. 2. The r.m.s. amplitude of 200 examples of random processes similar to that shown in Fig. 1. After 12 time steps, the value is constant and closely equal to the expected value shown, with fluctuations no larger than would be expected on the basis of the expected standard error limits.

First, it suggests that a reasonable value for J' is 12, or $\sim 1 \cdot 5$ times the integral scale, and second, it shows that the process is statistically stable. Some thought about the method suggests that numerical errors in the computation might act to increase the variance continuously by a random-walk process. Evidently this does not occur, or at least the effect is negligible over 80 time steps.

4. Two-Dimensional Prediction and Random Surfaces

Consider a mesh of points (x_i, y_j) with $x_i = i\delta$, $y_i = j\delta$, and suppose we know the value of a random function $f(x, y)$ at the mesh points $i \leqslant 0$, all j, and $i = 1$, $j < j'$. By an obvious notation we write

$$f(i, j) \equiv f(x_i, y_i).$$

As before, we introduce the subensemble $\{f_s(i, j)\}$ defined by $f_s(i, j) = f(i, j)$ at all the mesh points at which $f(i, j)$ is known. We wish to find means of calculating $\langle f_s(1, j')\rangle_s$. Figure 3 makes the situation clear. As before we write $\langle f_s(1, j')\rangle_s$ as a linear combination of all previously known values; dropping the dash on j we have

$$\langle f_s(1,j)\rangle_s = \sum_{i'=0}^{\infty} \sum_{j'=-\infty}^{\infty} h_2(i', j') f(-i', j - j')$$

$$+ \sum_{j'=1}^{\infty} h_1(j')\, f(1, j - j'). \tag{13}$$

We now wish to express the condition that $\langle f_s(1, j)\rangle_s$ be the best possible approximation to $f(1, j)$ in a least-square sense; i.e., the ensemble mean square error $V_s(1, j)$ should be a minimum, where

$$V_s(1, j) = \langle (f(1, j) - \langle f_s(1, j)\rangle_s)^2\rangle. \tag{14}$$

Before performing this minimization, however, it is convenient to rearrange Eq. (13). As it stands, there is no reason to believe $h_2(i', j')$ to be symmetrical in j'. We shall aim to replace it by a function which has this symmetry. Consider the subensemble $\{f_t(i, j)\}$ consisting of all those functions which are equal to $f(i, j)$ for $i \leqslant 0$. If we had not yet constructed any values of $f(i, j)$ for $i = 1$ it would be this subensemble which we should use to obtain the first value. Accordingly, we should write

$$\langle f_t(1,j)\rangle_t = \sum_{i'=0}^{\infty} \sum_{j'=-\infty}^{\infty} \bar{h}(i, j)\, f(-i', j - j'). \tag{15}$$

Now we can write

$$f(1, j) = \langle f_t(1, j)\rangle_t + \delta f(j) \tag{16}$$

and $\delta f(j)$ is now to be considered independent of all $f(i', j')$ for $i' \leqslant 0$. Substituting Eqs. (16) and (15) in (13), we can rewrite the result as

$$\langle f_s(1,j)\rangle_s = \sum_{i'=0}^{\infty} \sum_{j'=-\infty}^{\infty} h(i',j')\, f(-i', j-j') + \sum_{j'=1}^{\infty} h_1(j')\, \delta f(j-j'). \tag{17}$$

In deriving this, note that *all* points $f(i, j)$ for $i \leqslant 0$ are included in the double summations in both Eqs. (13) and (15), and any linear combination of sums over these points can always be put in the form of the first term of Eq. (17). Now we substitute Eq. (17) in Eq. (14) and expand. We assume $\langle (f(1, j))^2\rangle = 1$, and obtain

$$\begin{aligned} V_s(1,j) &= 1 - 2\langle f(1,j)\langle f_s(1,j)\rangle_s\rangle + \langle(\langle f_s(1,j)\rangle_s)^2\rangle \\ &= 1 - 2\sum_{i'=0}^{\infty}\sum_{j'=-\infty}^{\infty} h(i',j')\langle f(i,j)\, f(-i', j-j')\rangle \\ &\quad + \sum_{i'=0}^{\infty}\sum_{j'=-\infty}^{\infty}\sum_{i''=0}^{\infty}\sum_{j''=-\infty}^{\infty} h(i',j')\, h(i'',j'')\langle f(-i', j-j')\, f(-i'', j-j'')\rangle \\ &\quad - 2\sum_{j'=1}^{\infty} h_1(j')\langle \delta f(j)\, \delta f(j-j')\rangle \\ &\quad + \sum_{j'=1}^{\infty}\sum_{j''=1}^{\infty} h_1(j')\, h_1(j'')\langle \delta f(j-j')\, \delta f(j-j'')\rangle. \end{aligned} \tag{18}$$

Here we have used the independence of $\delta f(j)$ and $f(i, j)$ for $i \leqslant 0$ to eliminate the cross-terms arising in the expansion of $(\langle f_s(1, j)\rangle_s)^2$, and to write

$$\langle f(1, j)\, \delta f(j - j')\rangle = \langle \delta f(j)\, \delta f(j - j')\rangle.$$

Inspection of Eq. (18) shows that we now have two completely independent sets of terms, each of which must be minimized separately, to obtain equations for $h(i', j')$ and $h_1(j')$. Now let us return to the subensemble $\{f_t(i, j)\}$ and consider the case where we have no information about $f(i, j)$ for $i = 1$. We now form the ensemble mean-square error as before, obtaining from Eq. (15),

$$\begin{aligned} V_t(1,j) &= 1 - 2\sum_{i'=0}^{\infty}\sum_{j'=-\infty}^{\infty} \bar{h}(i',j')\langle f(1,j)\, f(-i', j-j')\rangle \\ &\quad + \sum_{i'=0}^{\infty}\sum_{j'=-\infty}^{\infty}\sum_{i''=0}^{\infty}\sum_{j''=-\infty}^{\infty} \bar{h}(i', j')\, \bar{h}(i'', j'')\langle f(-i', j-j')\, f(-i'', j-j'')\rangle. \end{aligned} \tag{19}$$

Comparison of Eqs. (18) and (19) shows that when V_t and V_s are minimized we must have

$$\bar{h}(i, j) = h(i, j).$$

It follows that we may construct $f(1, j)$ by the following prescription: first construct $\langle f_t(1, j)\rangle_t$, then add to it a random function $\delta f(j)$, which is to be generated by an h-function determined by minimizing the last two terms in Eq. (18). The actual procedure for minimizing this equation is very closely parallel to that given by Bendat ([2], p. 181), and need not be reproduced here. The result is that $h(i', j')$ is the solution of the set of equations

$$R(1 + i, j) = \sum_{i'=0}^{J} \sum_{j'=-J}^{J} h(i', j')\, R(i - i', j - j') \tag{20}$$

for $i = 0, 1...J$, $j = -J...J$, where

$$R(i, j) \equiv R[\delta(i^2 + j^2)^{1/2}].$$

Substitution in Eq. (19) yields the following expression for the mean-square error, V_t,

$$V_t(1) = 1 - \sum_{i'=0}^{J} \sum_{j'=-J}^{J} h(i', j')\, R(1 + i', j'). \tag{21}$$

These equations can be simplified somewhat by noting that since $R(i, j)$ is symmetrical in j, from Eq. (20) $h(i', j')$ must be symmetrical in j'. This allows an appreciable reduction in the computing time and storage space required to solve Eq. (20).

The correlation function $R_\delta(j)$ of the random function $\delta f(j)$ is determined by the requirement that $f(1, j)$ given by Eq. (16) should have the correct correlation as a function of j. Once again, using the independence of $f(i, j)$ for $i \leqslant 0$ and $\delta f(j)$, we have

$$\langle f(1, j)\, f(1, j')\rangle = \langle\langle f_t(1, j)\rangle_t \langle f_t(1, j')\rangle_t\rangle + \langle \delta f(j)\, \delta f(j')\rangle$$

or

$$R(o, j) = V_t(1)\, R_\delta(j) + (1 - V_t(1))\, R_\phi(j), \tag{22}$$

where $R_\phi(j)$ is the correlation function of $\langle f_t(1, j)\rangle_t$ given by

$$(1 - V_t(1))\, R_\phi(j) = \sum_{i'=0}^{J} \sum_{j'=-J}^{J} h(i', j'')\, R(1 + i', j - j'). \tag{23}$$

We can now determine $\delta f(j)$ by the method of §3 as a random function having a variance given by $V_t(1)$ and a correlation function R_δ. Then Eq. (16) represents the solution to our problem, and we need only the starting procedure for a finite sample to complete it. Here again two starting procedures are possible, and it is

easily seen that what I have called in Section 3 the rigorous procedure requires here the calculation of $\sim J^2$ h-functions, which is quite out of the question. I have adopted the second alternative and set arbitrary values (actually zeros) of $f(i, j)$ in a band of width $J + 1$ around three sides of the mesh, with a band immediately inside it, again of width $J + 1$, within which the computed values of f are discarded. Figure (3) shows the situation more clearly.

I have assumed that the result obtained in Section 3—that it is adequate to discard enough points to cover a distance $1 \cdot 5L$, applies here also—and in all the surfaces I have computed $J + 1 > 1 \cdot 5L/\delta$. It was not possible to make a rigorous statistical check of the whole procedure, as in the one-dimensional case, because it would have required to much computer time, but the surfaces I have computed at least show no obvious evidence of inhomogeneity near the edges.

The value of J is restricted by the method of solution. Equation (20) represents $(J + 1)(2J + 1)$ simultaneous equations. However, the symmetry of $h(i, j)$ in j reduces the number to $(J + 1)^2$. The method of solution then required at least $(J + 1)^4$ words of storage space in the computer, which means that J cannot be

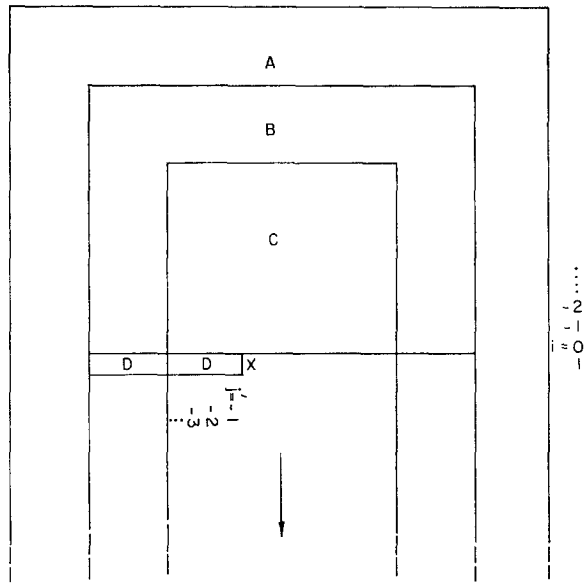

FIG. 3. The layout of the computation of a random surface. A is the region containing arbitrary values of $f(i, j)$ at the mesh points. B is the region in which the computed values are discarded. The regions B, C, and D together contain all the points that have been computed, and are used in forming the sub-ensemble $\{f_s(i, j)\}$. Only regions B and C are used in forming the sub-ensemble $\{f_t(i, j)\}$. X is the next point to be computed. The computation proceeds in the direction of the arrow.

much more than 10. This again limits L/δ to ~ 5. Table II shows the matrix $h(i, j)$ for a correlation function given by Eq. (12) with $a = 0.1840$, $b = 1 \cdot 0$, and $J = 9$. Evidently the components of h are very small for $i, j \sim J$ which gives us some reason to believe that the approximation introduced by the small value of J is not a bad one.

Figure (4) shows contour diagrams of two surfaces constructed by this method. To simplify the diagrams and bring out their main features more clearly only the positive- and zero-height contours are shown. The two surfaces have the same value of the integral scale $L = 6.33$ and variance unity, but differ in the other scale. Comparison shows the influence of the smaller scale on the roughness of the surface.

A pair of surfaces similar to these shown in Fig. (4) was computed for different values of a and b but using the same set of random numbers. Rather unexpectedly, it was found that the surfaces had the same general appearance. This suggested the conjecture that the smoother surface can in fact be derived from the rougher

TABLE II

THE MATRIX $h(i, j)$

i/j	0	1	2	3	4	5	6	7	8	9
0	.863	.304	.073	.042	.018	.012	.006	.005	.002	.004
1	−.416	−.215	−.074	−.038	−.020	−.011	−.007	−.004	−.003	−.003
2	.154	.090	.032	.015	.008	.004	.002	.001	.001	.000
3	−.057	−.036	−.014	−.006	−.003	−.002	−.001	−.000	−.000	−.000
4	.021	.014	.005	.002	.001	.000	.000	.000	.000	.000
5	−.008	−.005	−.002	−.001	−.000	−.000	−.000	−.000	−.000	−.000
6	.002	.002	.001	.000	.000	.000	.000	.000	.000	−.000
7	−.001	−.000	−.000	−.000	−.000	−.000	−.000	−.000	.000	−.000
8	.000	.000	.000	.000	.000	.000	.000	.000	.000	−.000
9	−.000	−.000	−.000	−.000	−.000	−.000	−.000	−.000	.000	−.000

by a filtering process whereby small wavelength modes are attenuated. This is far from obvious *a priori*, and I have not been able to prove or disprove it. It is however of some importance, since the one advantage that might be claimed for the Fourier synthesis method was that it would allow one to carry out just such a filtering operation on individual surfaces and examine the effects directly.

The random numbers used in these calculations were derived from a pseudo-random number generating routine in the computer. This yields numbers with a uniform distribution in the range $(-1, 1)$, but by adding five of them together a set of numbers could be obtained with a good approximation to a Gaussian distribution.

The construction of a surface on a mesh of side 25 by this method takes just over two minutes on the Culham Laboratory KDF-9 computer, if the h-functions are already known; to compute these also takes about 2 minutes. Of course this need be done only once for each ensemble of surfaces.

There is one final point which may bother some readers. We have treated a random function in space essentially as a process in time for computing purposes. In a temporal process it is, of course, obvious that the value of a function $f(t)$ can depend only on previous values of the function $f(t - \tau)$, $\tau > 0$. But surely, when we consider points along a line in space, say, the function $f(x)$ may depend on values both to right and to left of x, $f(x + \xi)$ for all ξ? The answer is that we have chosen to construct $f(x)$ in such a way that the points at which we know the values of the function are all to the left (say) of x, so that as far as our knowledge goes $f(x)$ resembles a time function. We could choose different methods in which points both to the right and to the left of x were known; but I believe all such methods involve the determination of many h-functions, and that the method I have used, involving only one h-function, is the simplest.

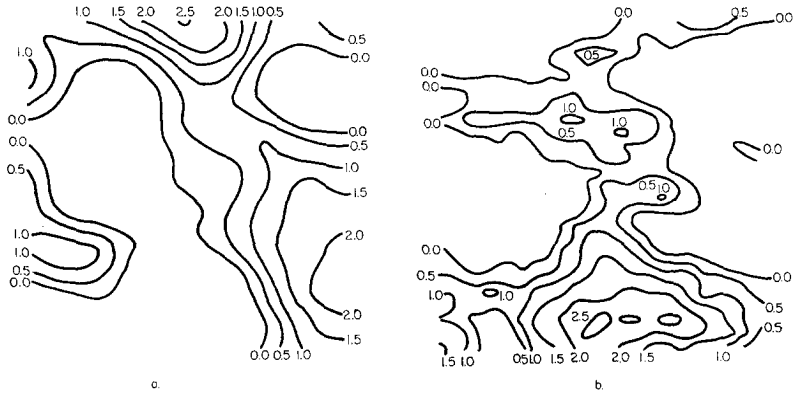

FIG. 4. Examples of random surfaces with correlation function given by Eq. (12), with the same integral scale $L = 6 \cdot 33$. The regions shown are squares of side 25. Only positive and zero-height contours are shown so as to bring out the structure more clearly. (a) $\lambda = 5.19$; (b) $\lambda = 2.42$.

5. CONCLUDING REMARKS

I have shown how specimens of random surfaces may be generated in a computer. While I was engaged on this work I received several ingenious suggestions for methods of producing surfaces by analogue means (e.g., throwing handfuls of wet sand onto a board) but while such methods may be very useful for particular problems they do not fulfill the basic condition of this work, namely that the surface produced must be a member of an ensemble with a specified correlation function.

APPENDIX I. THE CORRELATION FUNCTION

We have

$$R(\tau) = e^{ab}e^{-a(\tau^2+b^2)^{1/2}}.$$

The integral scale can be found from the transform, which is given by Erdelyi *et al* [6]: if

$$G(\omega) = \int_0^\infty R(\tau) \cos \omega\tau \, d\tau,$$

then

$$G(0) = \int_0^\infty R(\tau) \, d\tau \equiv L.$$

Thus the integral scale

$$L = be^{ab}K_1(ab).$$

In the limit $b \to 0$ for finite a, $K_1(ab) \to 1/ab$ and $L \to 1/a$, as expected since $R \to e^{-a|\tau|}$.

The gradient scale λ is given by

$$R(\tau) = 1 - \tau^2/\lambda^2 + \cdots \qquad \text{or} \qquad \left[\frac{d^2R}{d\tau^2}\right]_{\tau=0} = -\frac{2}{\lambda^2}.$$

This gives

$$\lambda^2 = 2b/a.$$

It can be shown that $1/b$ plays the part of a Kolmogorov microscale; the largest value of ω, for example, for which $G(\omega)$ is appreciable is $\approx 1/b$. This explains the association with surface "roughness" in Section 4, for example.

In Table III we give a series of values of a_1, b_1, and λ for constant $L = 6 \cdot 33$.

TABLE III

b	a	$1/a$	λ
3.84	0.2610	3.84	5.44
3.0	0.2222	4.5	5.19
2.0	0.2075	4.82	4.39
1.0	0.1840	5.44	3.30
0.5	0.1703	5.86	2.42

ACKNOWLEDGMENTS

It is a pleasure to acknowledge numberous entertaining discussions with D. J. H. Wort on this subject. I am grateful to F. M. Larkin who suggested the method of obtaining random numbers with a Gaussian distribution, and to G. O. Lester who constructed models of some of the surfaces.

REFERENCES

1. M. S. LONGUET-HIGGINS, "The Statistical Geometry of Random Surfaces" (Symposia in Applied Mathematics, Vol. XIII, p. 105). American Mathematical Society, Providence, Rhode Island (1962).
2. J. S. BENDAT, "Principles and Applications of Random Noise Theory." Wiley, New York (1958).
3. M. G. RUSBRIDGE, Culham Laboratory Report, CLM-R22 (1962).
4. N. WIENER, "Extrapolation, Interpolation, and Smoothing of Stationary Time Series." Technology Press, M.I.T., Cambridge, Massachusetts (1949).
5. L. ZADEH and J. RAGAZZINI, *J. Appl. Phys.* **21**, 645 (1950).
6. A. ERDÉLYI, *et al.* (Eds.), "Tables of Integral Transforms," Vol. I. McGraw-Hill, New York (1954).