A New Treatment of Rough Surface Scattering

C. Macaskill and P. Cao
School of Mathematics and Statistics
University of Sydney
N.S.W. 2006

Revised January 18, 1996

Abstract

This paper treats the problem of the scattering of a narrow beam at a randomly rough surface where the height distribution is Gaussian and the correlation of heights is Gaussian. The field at the surface is assumed to satisfy a Dirichlet boundary condition. The problem is formulated using the standard integral equation approach. In the case when the rms height is very large, numerical solution of this equation for a realization of the surface involves the solution of a very large matrix problem, particularly if the beam is scattered to other parts of the surface that are at a considerable distance from the initial illumination. Here we provide a straightforward technique, motivated by ray theory, that allows only those parts of the surface that are actually illuminated at any stage to be treated. Thus for the first reflection of the beam one solves an integral equation over the initially illuminated area. This solution is used to determine the parts of the surface which give rise to doubly scattered waves, and one then solves the integral equation over the newly illuminated area, and so on, until all reflections have been accounted for. This procedure can be carried out with no explicit reference to any ray model, working simply with the discrete matrix form of the integral equation. In particular, for 1-D surface variation, a strategy of successive sweeps from left to right and back again (for a beam incident from the left) is shown to be successful. Any standard iterative or direct solution technique can be used for the sequence of sub-problems so introduced. The general method is also applicable to scattering from surfaces varying in two dimensions.

Introduction

Over the last fifteen years in the optics, acoustics and radio-physics literature, there have been many papers discussing direct numerical methods for scattering of waves from rough surfaces, in particular for surfaces that may be defined as realizations of a random function. Examples for the 2-D problem where the surface varies in one dimension only are the papers by Fung and Chen (1985), Kachoyan and Macaskill (1987), Soto-Crespo and Nieto-Vesperinas (1989) and Thorsos (1988). For the 3-D scalar problem the work of Tran and Maradudin (1992), Macaskill and Kachoyan (1993) and Tsang et al (1994) is relevant, and the full electromagnetic scattering problem is treated by Tran and Maradudin (1994). Most of these papers consider the problem of uniform or close to uniform illumination by a broad incident beam, or illumination by some approximation to a plane wave over a contiguous region. In a series of papers, Saillard and Maystre (1988,1990) and Maystre et al (1991) have demonstrated that one can deal with extended illumination by looking at the scattering of overlapping narrow beams, and of course in many applications one is interested in narrow-beam illumination directly. As pointed out by Maystre and Saillard it is difficult to deal with the case where the surface is initially illuminated in one closely concentrated region, but the local surface geometry is such as to cause illumination of a distant but similarly concentrated part of the surface. Similarly, if the surface roughness is extreme, then the incident beam may be scattered several times before escaping from the surface, and these reflections are often at quite widely separated regions.

In this paper we discuss a method that allows a straightforward treatment of such problems. Basically the idea is to solve exactly for the initially illuminated part of the surface, and then to use this solution to find patches that are illuminated by the field reflected from this initial patch. One then solves for this second scattered field exactly, and so on, until all reflections are dealt with, and the wavefield is reflected away from the surface. This is easily visualized in a ray picture, and accordingly we find that for parameter regimes where the ray approximation is a good approximation, the present method can be used to find the full solution of the governing integral equation very effectively. This gives a method that can be used to break problems involving surfaces of very large extent into manageable and yet physically relevant sub-problems, by automatically identifying only those parts of the surface which are actually illuminated at any given order of reflection. This is demonstrated explicitly here only for surfaces varying in one dimension, but the formulation is also valid for scattering from two-dimensional surfaces, and offers a promising way forward for these numerically very taxing problems. As each sub-problem is equivalent to a standard scattering problem over a limited spatial domain, all the specialized numerical techniques that have been evolved to deal with such problems are still appropriate. Furthermore, these sub-problems turn out to be better conditioned in a numerical sense than is the case with other approaches that use the whole surface, and thus are more amenable to solution by iterative techniques.

In essence, we are considering integral equations that after discretization lead to matrix problems with *solutions* that are sparse. The difficulty is to identify the precise structure of the solution, given only the different (typically) sparse structure of the right hand side of the matrix system, i.e. the incoming wave. In this paper we use simple physical ideas, borrowed from ray theory, in order to identify this structure. We then show how essentially the same picture can be built up directly from the integral equation, without any consideration of rays, in a straightforward but rigorous manner that can be implemented easily in practice.

1. Mathematical formulation

The scattering geometry is shown in Figure 1, where $\mathbf{k}_i$ and $\mathbf{k}_s$ represent the incident wavevector (at incident angle θ) and the scattered wavevector (with scattering angle θ_s) respectively. We consider the 2-D case only here, where the surface varies in one dimension only, with the incident field and the scattered field in the $x-z$ plane.

We consider a surface $z=\zeta(x)$ with Gaussian statistics and a Gaussian correlation function, with rms height h and correlation length l. The scalar wavefield satisfies a Dirichlet boundary condition, $\Psi=0$, which corresponds to scattering at the rough surface of a perfect conductor in the electromagnetic case or to scattering at a pressure release surface in the acoustic case. The field away from the surface can be written as

$$\Psi(x,z)=\Psi_0(x,z)+\Psi_s(x,z), \tag{1}$$

where the incoming wavefield Ψ_0 corresponds to a finite-width beam, given by

$$\Psi_0(x,z)=\frac{1}{2}\sqrt{\frac{A}{\pi}}\int_0^{\pi} e^{-A(\frac{\theta'-\theta}{2})^2} e^{ik[(x-x_0)\cos\theta'-z\sin\theta']}d\theta' \tag{2}$$

with $A=(kg\sin\theta)^2$ (see Thorsos, 1988). Here k is the magnitude of the incident wavenumber and x_0 denotes the centre of the beam. This corresponds to an integral over plane waves, with mean incident angle θ, and a Gaussian weighting function, dependent on the parameter g, which controls the angular width of the incident field. Clearly the limit $g\to\infty$ corresponds to an incident plane wave, while when g is made smaller the beam width decreases (see appendix A for details of the numerical evaluation of this initial field). The relevant non-dimensional parameters in the problem are then h/l and kl.

The scattered field Ψ_s is given by the Helmholtz integral formula

$$\Psi_s(x,z) = \frac{k}{4}\int_S \psi(x')H_0^{(1)}(k|\mathbf{r}-\mathbf{r}'|)dx' \tag{3}$$

where ψ, the normal derivative of the wavefield, is given by

$$\psi(x) = \frac{i}{k}\left(-\frac{\partial\Psi}{\partial z} + \zeta'(x)\frac{\partial\Psi}{\partial x}\right)_{z=\zeta(x)}. \tag{4}$$

In equation(3), $\mathbf{r}$ is the position vector for a point away from the surface and $\mathbf{r}'$ is the position vector for a point on the surface, respectively, i.e., $\mathbf{r} = (x,z)$ and $\mathbf{r}' = (x', \zeta(x'))$. $H_0^{(1)}$ is the zero-order Hankel function of the first kind. As discussed by Meecham (1956), with both $\mathbf{r}$ and $\mathbf{r}'$ confined to the surface, the normal-derivative operator applied to equation (3) with use of the normal-derivative of (1) gives

$$\psi(x) + \int_S K(x,x')\psi(x')\,dx' = 2\psi_0(x). \tag{5}$$

which is a Fredholm integral equation of the second kind for ψ. The normal derivative of the incident field is given by

$$\psi_0(x) = -\frac{1}{2}\sqrt{\frac{A}{\pi}}\int_0^{\pi} e^{-A(\frac{\theta'-\theta}{2})^2} e^{ik[(x-x_0)\cos\theta' - \zeta(x)\sin\theta']}[\sin\theta' + \zeta'(x)\cos\theta']d\theta' \tag{6}$$

and the kernel function is then given by

$$K(x,x') = \frac{ik}{2}\frac{H_1^{(1)}(k\rho)}{\rho}[\zeta(x') - \zeta(x) - (x'-x)\zeta'(x)], \tag{7}$$

where $H_1^{(1)}(y) = -dH_0^{(1)}(y)/dy$, $\rho = \sqrt{(x-x')^2 + (\zeta(x)-\zeta(x'))^2}$.

In this paper we consider a narrow beam field of the form (6) incident on the rough surface. Often a truncated plane wave or broad beam field tapered at the edges is considered in the determination of the scattering cross section for random rough surfaces. The broad tapered beam in particular is used to reduce edge diffraction (Thorsos, 1988), which can be particularly significant for very rough surfaces. Saillard and Maystre (1988) have shown that an alternative method is to consider narrow beam illumination of an extended surface. One can then either take a superposition of many offset narrow beams, solving for each small problem separately and then using the principle of superposition, or equivalently illuminate the central part of a long surface and then repeat the problem over an ensemble of such surfaces. We note also that many practical experimental set-ups involve narrow beams which are then scanned over the surface, so it is important to understand the detailed mechanisms for scattering of such beams and to have a general method for dealing with the problem.

Consider Figure 2, where the incident beam illuminates 'patch 1', and is then reflected away from the surface. If we look for points on the surface where $|\psi_0(x)| > \epsilon$, where ϵ is some threshold value, we find five patches. Patches 2,3,4 and 5 are not in fact illuminated but can be thought of as due to the incident field 'passing through the surface'.

We now decompose $\psi = \psi_1 + \psi_2 + \cdots + \psi_N$, where, as will be demonstrated later, ψ_1 corresponds to the field reflected just once from the surface, $\psi_1 + \psi_2$ corresponds to the total field scattered once or twice, and $\psi_1 + \psi_2 + \cdots + \psi_i$ corresponds to the total field reflected anything up to i times. We assume that there is a maximum number of such reflections, say N.

Let S_i be the part of the surface illuminated at the ith scattering. The contribution to the total field from this patch S_i is then ψ_i, and we can then define the integral operator α_{ij} from the relation

$$\alpha_{ij}\psi_j = \int_{S_j} K(x,x')\psi_j(x')\,dx', \quad x \in S_i.$$

which describes the radiation on patch S_i due to scattering from patch S_j. Consider now all $x \in S_1$ such that $\psi_{0,1} \equiv 2\psi_0(S_1) \neq 0$. Then by equation (5)

$$\psi_1 + \alpha_{11}\psi_1 = \psi_{0,1}. \tag{8}$$

We define $\psi_1(x)$ to be the solution of (8) if $x \in S_1$ and to be zero otherwise. The integral equation (8) describes first reflections, or scatterings, and is *exact* if there are no second reflections, as we take the set S_1 to include patches one to five, in Figure 2 e.g. Equation (5) then reduces to (8) directly as $\psi = 0$ for any x that is not in S_1. However, equation (8) is not exact if there are any double reflections on the surface. (More precisely, any double reflections that involve points on the surface that are not in S_1: in our physical discussion we assume that the beam is narrow enough, and the wavenumber large enough that these are the only double reflections. The algorithm is not affected by these considerations, however.) How do we identify such second reflections? We determine the scattered field due to ψ_1 and evaluate it on the surface. Any patches which are not contained in S_1, but for which this scattered field is non-zero, are indicative of second reflections, which have not been taken care of in the solution of equation (8). An alternative but equivalent way of thinking about this is to substitute $\psi = \psi_1 + \phi_2$ into the full integral equation (5), where ϕ_2 is the correction that takes second and higher order scattering into account. We have for any $x \in S$:

$$\psi_1(x) + \phi_2(x) + \int_S K(x,x')\phi_2(x')\,dx' = 2\psi_0(x) - \int_{S_1} K(x,x')\psi_1(x')\,dx'$$

For points x not in S_1, ψ_1 and ψ_0 are both zero and we find

$$\phi_2(x) + \int_S K(x,x')\phi_2(x')\,dx' = -\int_{S_1} K(x,x')\psi_1(x')\,dx' \tag{9}$$

We now define the set S_2 to consist of all points that are not contained in S_1 for which the right hand side of (9) is non-zero. This right hand side corresponds to the field reflected from our initially illuminated patch S_1 and we see that we can regard this as analogous to equation (8), but with the singly scattered field taking the place of the incident beam. Then we let $\phi_2 = \psi_2 + \phi_3$ and *define* $\psi_2 = 0$ for x not in S_2. Then

$$\begin{aligned}\psi_2(x) + \int_{S_2} K(x,x')\psi_2(x')\,dx' = -\int_{S_1} K(x,x')\psi_1(x')\,dx' \\ - \phi_3(x) - \int_S K(x,x')\phi_3(x')\,dx'.\end{aligned} \tag{10}$$

If we consider $x \in S_2$ *only*, and set the sum of the last two terms in $\phi_3(x)$ to zero, we have the integral equation for ψ_2:

$$\psi_2 + \alpha_{22}\psi_2 = \psi_{0,2} = -\alpha_{21}\psi_1. \tag{11}$$

(We note that the regions $S_1, S_2, \ldots$ are not necessarily separate, i.e. they can overlap. Although successive regions S_1, S_2 or S_2, S_3 are always separate, in Figure 3, e.g., we see that in this triple reflection case S_1 and S_3 overlap. If no patches intersect with S_2, then in fact $\phi_3(x)$ is zero for $x \in S_2$.) We then solve equation (11) in order to determine ψ_2. Continuing in this way, we obtain the recurrence relation:

$$\begin{cases} \psi_1 + \alpha_{11}\psi_1 = \psi_{0,1}, \\ \psi_i + \alpha_{ii}\psi_i = \psi_{0,i} = -\alpha_{i\;i-1}\psi_{i-1}, \quad 2 \le i \le N. \end{cases} \tag{12}$$

The process is self-limiting, because at stage N, there is no scattered field on S from S_N, as we have assumed that there are no more than N reflections. We can show that the process is indeed self-limiting as follows. Consider the exact relation corresponding to (10) at stage N. We have, with $\psi = \psi_1 + \psi_2 + \cdots + \psi_{N-1} + \phi_N$, and for $x \in S_N$,

$$\begin{aligned}\psi_N(x) + \int_{S_N} K(x,x')\psi_N(x')\,dx' = -\int_{S_{N-1}} K(x,x')\psi_{N-1}(x')\,dx' \\ - \phi_{N+1}(x) - \int_S K(x,x')\phi_{N+1}(x')\,dx'\end{aligned} \tag{13}$$

where the sum of the last two terms in equation(13), which are dependent on ϕ_{N+1}, is zero for $x \in S_N$, as usual. For x not in S_N,

$$\phi_{N+1}(x) + \int_S K(x,x')\phi_{N+1}\,dx' = -\int_{S_N} K(x,x')\psi_N(x')\,dx', \tag{14}$$

where (12) is used to remove any contributions from $\psi_1, \psi_2, \ldots \psi_{N-1}$ if x is in one or more of $S_1, S_2, \ldots, S_{N-1}$.

However, by assumption there are no more than N reflections, so that the energy radiated from S_N leaves the surface and the integral on the right hand side of (14) vanishes for such x not in S_N. Therefore, $\phi_{N+1} + \int_S K(x,x')\psi_N(x')\,dx' = 0$ everywhere, and hence $\phi_{N+1} = 0$ everywhere (there are no eigensolutions as the integral operator is invertible). Thus the error in the solution of the integral equation is zero everywhere on S.

A discussion of the convergence of iterative methods, in particular the generalised conjugate gradient methods applied to this general class of scattering problem is given in Cao and Macaskill (1995). In that paper it was shown that the convergence depends, as would be expected, on the spectral radius of the governing matrix system and that the spectral radius scales with $h/l\sqrt{kl}$. It was found that about the best that can be achieved with current iterative techniques is to give convergence for the spectral radius less than about four or so. (This is to be contrasted with the simple Neumann iteration, where convergence is only achieved if the spectral radius is less than one.) The small sub-problems encountered in the procedure used in the present paper in general have been found to have smaller spectral radius than the full order system. Thus practical convergence for each sub-problem is good. This leaves aside the question of convergence to the true solution, however. One can address this as follows. In the limit as $kl \to \infty$, one can interpret the problem in ray terms. In the problems treated here all incident energy is eventually reradiated from the surface. In the ray limit there must be therefore be a finite number of reflections, as assumed above, so that the process is self-limiting and we have convergence to the true solution. When is this likely to break down? First, when $kl \to 0$, we expect that any narrow incident beam will undergo extreme angular broadening and so the narrow beam picture becomes inappropriate, and we are brought back to illumination over an extended surface, and the treatment of one full integral equation. For small to moderate h/l we then expect convergence, as described in Cao and Macaskill (1995). More generally, for large h/l and large, but finite, kl we can still expect to find a finite number of patches illuminated at the various orders of scattering, leading to a finite number of sub-problems, and we expect behaviour close to the ray-theoretic case. Clearly as kl is decreased, this correspondence becomes more tenuous; however, in the cases treated here, with moderate kl, there have been no convergence difficulties.

We now derive the above equations directly from the integral equation (5) and show that the solution of the N smaller problems is the exact solution of the original problem, under the assumption that at any stage $\psi_{0,i}$ can be determined to be either zero or non-zero (corresponding in practice to $|\psi_{0,i}| < \epsilon$ or $|\psi_{0,i}| > \epsilon$ respectively). As before, let $\psi = \psi_1 + \psi_2 + \cdots + \psi_N$, where $\psi_k(x) \neq 0$ only for $x \in S_k$, and where we *define* the set S_k to be all x not contained in S_{k-1} such that

$$\int_{S_{k-1}} K(x, x')\psi_{k-1}(x')\, dx' \neq 0, \quad k > 1. \tag{15}$$

Using this definition, $S_{N+1} \equiv \emptyset$, as there are no more than N reflections. Then the integral equation (5) becomes

$$\sum_{k=1}^{N} \psi_k(x) + \int_S K(x, x') \sum_{k=1}^{N} \psi_k(x')\, dx' = 2\psi_0(x),$$

or, on interchanging the order of integration and summation

$$\sum_{k=1}^{N} \left[\psi_k(x) + \int_S K(x, x')\psi_k(x')\, dx' \right] = 2\psi_0(x). \tag{16}$$

If x is not in $\bigcup_{k=1}^{N} S_k$, then $\psi_0(x) = 0$ and $\int_S K(x, x')\psi_k(x')\, dx' = 0$, and we find $\sum_{k=1}^{N} \psi_k(x) = 0$, or $\psi = 0$. If, by contrast, x is in one or more S_k, then

$$\begin{aligned} &q(x, S_1) \left[\psi_1(x) + \int_{S_1} K(x, x')\psi_1(x')\, dx' - 2\psi_0(x) \right] \\ &+ \sum_{\substack{k>1, \\ x \in S_k}} \left[\psi_k(x) + \int_{S_k} K(x, x')\psi_k(x')\, dx' + \int_{S_{k-1}} K(x, x')\psi_{k-1}(x')\, dx' \right] = 0. \end{aligned} \tag{17}$$

This follows as the only non-zero contributions from $\int_S K(x, x')\psi(x')\, dx'$ at $x \in S_k$ come from either $x' \in S_k$ or $x' \in S_{k-1}$, by our definition of the set S_k in equation (15) above. In (17), we have used the fact that $\psi_k(x) = 0$ if x not in S_k. The expression $q(x, S_1)$ is given by

$$q(x, S_1) = \begin{cases} 1, & \text{if } x \in S_1; \\ 0, & \text{otherwise.} \end{cases}$$

Equation (17), which is equivalent to the original integral equation (5), is satisfied exactly for any x if we require

$$\begin{aligned} &\psi_1(x) + \int_{S_1} K(x, x')\psi_1(x')\, dx' = 2\psi_0(x), \quad x \in S_1 \\ &\psi_k(x) + \int_{S_k} K(x, x')\psi_k(x')\, dx' = -\int_{S_{k-1}} K(x, x')\psi_{k-1}(x')\, dx', \quad x \in S_k, \quad 1 < k \leq N. \end{aligned} \tag{18}$$

We know that for Fredholm equations of the second kind of the form (5) there is a unique solution and thus equation (18), which solves (5) exactly, is that solution. These equations are precisely those given above as equations (8) and (12). We note that the formulation extends immediately to the three-dimensional scattering problem of scattering from a two-dimensional rough surface.

2. Physical discussion

How does this new method compare with a straightforward Neumann iteration? The Neumann iteration forms the iterative sequence:

$$\psi^{n+1} = -\int_S K(x,x')\psi^n(x')\,dx' + \psi^0(x), \tag{19}$$

with $\psi^0 = 2\psi_0(x)$.

The Kirchhoff approximation corresponds to setting $\psi = \psi^0 = 2\psi_0$. For a narrow beam initial field, the iteration in (19) corresponds to applying the Kirchhoff approximation repeatedly, as the field is scattered to each new portion of the surface. Consider the iteration in (19) used for situation like that shown in Figure 2. On patch 1, at least for a small amplitude surface, the method will give a good approximation to $\psi(x)$ everywhere. However, on patches 2, 3, 4 and 5 (see Figure 2) it will give a completely incorrect solution, corresponding to the 'false illumination' shown in the figure. Moreover, these false illuminations may give rise to false second and higher order reflections, that cannot easily be removed at a later stage. Thus later iterations will mainly be used to remove the effects caused by the fact that shadowing is not correctly dealt with in the Kirchhoff approximation, or equivalently, in any one iteration as in (19).

For a plane wave initial field, the repeated Kirchhoff approximation, i.e. the Neumann iteration, in many cases converges to the correct solution after sufficient iterates have been used. As described in Thorsos and Jackson (1991) false illuminations as described above are certainly taking place, but if enough iterations are performed, these are eventually nullified and the correct solution is regained (if the spectral radius is small enough). In such a case, the iteration is sensible, as the whole surface must be treated at any iteration because of the uniform illumination. For a problem with narrow beam illumination, however, where only isolated patches of the surface are illuminated, it is clear the Neumman iteration will often involve large portions of the surface on which there is in reality never any illumination; this is particularly so if the initial beam is near to grazing incidence, or if the surface is extremely rough. The present method, by contrast, automatically limits the spread of spurious illumination; for example in Figure 2, patches 2,3,4 and 5 are exactly treated in the solution of equation (8), and thenceforth play no further part. Indeed, in the case of surfaces varying in one dimension, a variation on the general procedure, which is described in detail below, means that we never have to deal with *any* false illumination. Furthermore, the physical processes are clearly identified, so that the first iteration deals with all single reflections, the next iteration adds the result of second reflections and so on.

The general method described in section 1 has the difficulty that one is continuously dealing with patches of false illumination, at considerable computational cost. It turns out that this is easily remedied in 2-D problems, where the surface varies in one dimension only. We now describe the modification to the algorithm that ensures that there is no need for direct treatment of patches where there is false illumination. Consider the case in Figure 2 again. The beam is incident from the left. In our idealized case above, we assumed one could determine a well-defined boundary between illuminated and non-illuminated parts of the surface. For an initial field satisfying the scalar wave equation this is only approximately true. Therefore, it is necessary to define some threshold value of $|\psi_0(x)|$, say ϵ, below which we approximate the value of the illumination as zero. Once this has been decided, one can sweep all x values from left to right, searching for the first patch on which $|\psi_0(x)| \geq \epsilon$, i.e. we define the set S_1 to consist of all points on patch 1 in Figure 2, i.e. we ignore false illumination in patches 2,3,4 and 5. For patch 1 defined in this way, we then find ψ_1 by inverting the discretized integral equation (8) on this domain; in all calculations used here we used a generalized conjugate gradient method to carry such inversions out (see, e.g. Cao and Macaskill, 1995) but in principle (usually inefficiently) one could just invert the discrete analogue of the system using gaussian elimination. Once this has been done, the scattered field from patch 1, given by $\alpha_{21}\psi_1$ is calculated and subtracted from $2\psi_0 = \psi_{0,1}$. Thus we have

$$\psi_{0,2} = \psi_{0,1} - \alpha_{21}\psi_1,$$

which can be thought of as a new incident field. We then continue our left to right scan, continuing from the right hand end of the previously determined patch 1, and now searching for the first patch where the new incident field $\psi_{0,2}$ is non-zero. In the case of Figure 2, there will be no such patches, because there are no second reflections. But what has happened to patches 2,3,4 and 5 where the false illumination is seen? We know, from physical considerations, that there is no illumination away from patch 1, so that at any of patches 2,3,4 and 5, the true solution is $\psi_1(x) = 0$ (indeed in this particular case $\psi(x) = 0$). Therefore, on these patches, the integral relation corresponding to equation (8) reduces to $\psi_{0,1} = \alpha_{21}\psi_1$, and hence $\psi_{0,2}$ will be identically zero there, and our algorithm will automatically ignore such false patches.

In a situation such as that shown in Figure 3, by contrast, we will still need to take account of second reflections. In this particular case, these will not be encountered until the second

sweep, which must run from right to left. When the second reflection has been picked up (i.e. when the patch where $|\psi_{0,2}| \geq \epsilon$ has been identified) we find we have equation (11) to solve. Once this has been done, the final solution is known with an error roughly of $O(\epsilon)$, as all reflections have been taken care of. In general as many as N sweeps (counting each left or right sweep as one) are required to deal with N reflections. However, the jth sweep may include some contributions from reflections of order $j+1$ or higher. In principle, further sweeps can now be used to refine the solution, using smaller values of ϵ. In fact, the values of ϵ used are obviously directly related to the size of matrix problem that must be solved at each iteration; for efficient implementation it is best to ensure that this size is chosen so that patches at least of the order of the beam width are treated. In appendix B, we give a concise description of this algorithm.
The problem of false illumination is more difficult to deal with for 2-D surfaces, but there are various sweeping strategies that could be tried. However, for the 2-D surface problem, it is possible that the simplicity of the basic algorithm described in section 1 may make it more attractive than methods that attempt to eliminate false illumination.

3. Numerical results for one-dimensional surfaces

In the first numerical example, we test the case in Figure 4, where the initial region illuminated by a narrow beam is small but the field is scattered to another distant part of the surface. The Gaussian surface has 4096 points and 128 points per correlation length; this surface was picked from a number of realizations to illustrate the basic mechanisms involved. In this case, $h/l = 1$ and the incident angle $\theta = 32°$. The ray traces, given in Figure 4(a), illustrate the basic scattering mechanism, and it clear that this is a case where most of the scattered energy is reflected once or twice only. (Note that this simple ray trace takes no account of the angular spread of the incident beam.) For the full calculations, we use $g/l = 0.3$ and $kl = 50$, and then the incident beam has beam width $W_x \approx 1.27$ correlation lengths at the mean surface level and angular width $W_\theta \approx \pm 9°$ where these quantities are given in the formulas (A3) and (A4). The finite angular width of the beam means that some of the smaller peaks will also be illuminated after the beam is first reflected from the surface, as shown in Figure 4(c), although this is not apparent in the ray trace in Figure 4(a). The situation shown here is not rare for Gaussian surfaces with large rms height even if the incident field is narrow and the incident angle is moderate. It is the advantage of the method developed in this paper that it is able to deal with the beam problem in which the relatively small illuminated regions are separated by extremely long distances.

For the surface given in Figure 4, we compare the results of the new method with a full integral equation solution of the 4096 point problem, carried out on a CM5 parallel computer with 32 vector processors and 1 Gbyte of memory (see Cao, 1995, for further details). In Figure 5, we give the distribution of scattered energy. On comparison with Figure 4, we see that the scattered energy at 10° corresponds mainly to first order reflection, while the peak at 145° is due to doubly reflected waves. Due to the truncation error involved in the new method, some small difference between the full integral equation solution and the approximate solution occurs in the forward grazing direction. As described in Appendix B, in practice we employ a maximum array size M, which is roughly comparable with the patch size, which in turn is taken to be of the order of the beam width. In this case, all the forward reflections are second reflections, and so just two sweeps with a maximum array size $M = 166$ gives a very good result. The third sweep refines the accuracy of the solution, but at considerable cost in extra computation time. The corresponding values of ψ on the surface for two or three sweeps are given in Figure 6(a) and Figure 6(b) respectively. For a given set of physical parameters, some degree of experimentation is needed to determine suitable values for the maximum matrix size M and the initial value of the threshold ϵ. In this case, $\epsilon = .01$, $M = 166$ and only 3 sweeps were performed. Calculations using the new method were carried out on a scalar workstation (HP9000/755) which performs roughly 20-30 times fewer floating operations per second than the CM5 for the calculations treated here. Moreover, very little memory was required for the calculation on the workstation. Details are given in Table 1.

We now move to a consideration of the calculation of ensemble averaged results for the beam scattering problem using the new method. In Figure 7, the results obtained using the new method are compared with those from the full integral equation method. In each case the results are for the same 384 surface realizations where each surface has 4096 points and 128 points per correlation length. The scaled rms height, $h/l = 2$, which corresponds to a mean slope angle of 70°, and the scaled wavenumber $kl = 15$. For the incident beam, $\theta = 45°$, $W_x = 4.2$ correlation lengths and $W_\theta = \pm 6.7°$. Based on the statistics of a ray simulation (see Figure 8), we know that 99% of the rays undergo seven or fewer scatters before escaping from the surface, so the maximum number of sweeps N is taken to be seven and the maximum array size M is restricted to 512, since the intial beam width is about four correlation lengths (i.e. 512 points). It has been found that there are two strong peaks in the energy spectrum, corresponding to the direct backscatter direction and the specular direction. The solid line in Figure 7 is the result using the new method with seven sweeps, and the dashed line is the result of the full integral equation treatment for the same incident beam. Clearly the new method works well, even for this difficult problem. As can be seen, agreement is good over the whole range of scattering angles, and energy is well conserved in both cases. We note that the backscatter peak is not exactly at the incident angle in Figure 7. This is due to the inadequate number of realizations; each realization takes about 40 minutes of computer time on the workstation used. Although this is an improvement on the time taken to solve the full problem, it is still computer intensive. One might then expect that one could use the new algorithm on the CM5 parallel computer to tackle the problem more effectively. However, although the storage requirement is dramatically reduced, the computer time is increased on the parallel machine when the new algorithm is used, mainly due to the large number of small randomly sized matrix problems that has to be treated, which means that arrays are continually taken in and out of CM5 memory, by contrast with the full inversion problem. Thus at this stage the algorithm is not suited to a parallel environment, although no major efforts have been made to explore the possibilities.

Accordingly, the benchmark was calculated on the CM5 using the full 4096×4096 arrays, over 1024 surface realizations, and employing a generalized conjugate gradient method (see Cao and Macaskill, 1995, for a discussion of the application of this method to this class of problem). The results are shown in Figure 8, with a three point nearest neighbour smoothing applied to the spectral estimates. Also shown are the results of a ray simulation, similar to that carried out in Macaskill (1991), but using 10,240 surfaces (the ray simulation is a relatively quick calculation so that results can be calculated using enough surfaces so that there is essentially no error due to inadequate sampling). Here for each surface a ray trace was carried out for each of 1300 rays launched at equal spatial intervals over the middle portion of the surface. The final scattered angle for each ray was then determined and the results averaged over all the surfaces used. We note that this procedure applied to a single surface will give different results to the full beam solution, as the illuminated portion of the surface is different, and there is no tapering in the ray calculation. However, we expect the ensemble average results to be the same, as is borne out by the fact that the ray results show almost exact agreement with the full solution of the integral equation, indicating that the scattering processes are almost purely geometric for this class of surface with relatively high $kl = 15$. For completeness, we show the results of ray simulations at a range of rms scaled heights in Figure 9. The forward and backscattered peaks are very apparent, particularly for the larger scaled rms heights.

We now consider the convergence of the procedure for the extreme case $h/l = 2$, with the maximum matrix size $M = 512$. The measure of error used in the general procedure is given in Appendix B as $\left[\frac{\langle \psi_i^* \cdot \psi_i \rangle}{\max(|\chi|)}\right]^{1/2}$. In Figure 10 we see that this error, averaged over 50 realizations, approaches zero as the number of sweeps is increased, with the error approximately halving with each sweep in this case. Note however, that the time to perform each sweep increases as the sweep number increases, which is a weakness of this method, i.e. good solutions are obtained quickly, but ultimate refinement is numerically expensive. We see also that the automatic procedure for the variation of the threshold ϵ given in Appendix B does indeed lead to an accompanying reduction in this threshold as the error decreases. When the threshold approaches zero, then we approach the case treated theoretically in the body of the paper, so that in the limit the exact solution is obtained. If a fixed value of ϵ were to be used, by contrast, there would be an asymptotic limit beyond which the error could not be reduced, which would depend directly on the value of epsilon chosen. A typical realization is considered in Figure 11, where we show the rms error for this approximate solution procedure, relative to the solution obtained from an exact inversion of the full matrix, again for the case $h/l = 2$, and for three values of the maximum patch size $M = 256, 512$ and 1024. We see that after an intial increase in error with sweep number, which is more apparent for the smaller values of M, the error decreases uniformly. Although the convergence rate is greater with larger M, this is compensated for by the fact that each iteration is quicker with smaller M, so that in fact the dependence of the total time taken is only weakly dependent on the precise choice of this maximum matrix size.

4. Conclusion

In this paper we have introduced a new technique for dealing with the scattering of narrow beams from rough surfaces when the rms height of the surface is large. It has been shown that for surfaces with well-defined features, such as are evident if the surface has a Gaussian correlation function, the governing integral equation can be broken up into a sequence of smaller problems, corresponding to the different numbers of reflections of the beam at the surface. The advantage of the technique is that it mirrors the physical process, and hence identifies the appropriate sub-problems that *must* be solved if the scattering problem is to be treated correctly, without dealing with parts of the surface that are never illuminated. Moreover, very long (in principle arbitrarily long) surfaces can be treated without the need for large amounts of computer memory, and any standard technique already used for such problems can incorporate this new approach quite simply. On scalar computers, the method also provides accurate approximations for the scattering problem more quickly than other techniques.

Further work using this approach can proceed in several directions. First, the method has not been tested with surfaces with a wide range of scales, where the geometric analogy is not so obviously applicable. Secondly, the general formulation also applies to scattering from 2-D surfaces, but an important part of the application of the technique to such problems will be the determination of an effective search procedure for the successively illuminated patches that correspond to the various levels of reflection or scattering. This has not yet been done.

Acknowledgement

The comments of Dr Eric Thorsos, Dr Bernard Kachoyan and an anonymous referee on an earlier version of this paper were very valuable. This work was supported in part by the Australian Research Council.

Appendix A: the incident beam

The form (2) satisfies the Helmholtz wave equation exactly. However, there is no general analytic expression for this integral and it is difficult to evaluate numerically. The same is true for the normal derivative, given in (6), which is needed here.

An approximate form can be obtained by expanding $(\theta' - \theta)$ in a Taylor series at $\theta' = \theta$ and retaining terms up to $(\theta' - \theta)$ or $(\theta' - \theta)^2$, as in Thorsos (1988). This approximate form satisfies the wave equation to $O(1/\sqrt{A})$ or $O(1/A)$ respectively, and is thus appropriate if the rms height of the surface is not too large. Equivalent forms of Ψ_0 are used in Saillard and Maystre (1988) and Tran and Maradudin (1992). One can then determine the corresponding approximate form of the normal derivative.

This approximate solution can be used to give a general indication as to how the beam width and angular width depend upon the parameter g and the wavenumber k. If we make the substitution $\delta = (\theta' - \theta)$, expand $ik[\cos(\delta + \theta)(x - x_0) - \sin(\delta + \theta)z]$ in δ and retain terms to $O(\delta)$, the integral in (2) is then given by

$$\Psi_0(x, z) = \frac{1}{2}\sqrt{\frac{A}{\pi}} e^{ik[(x-x_0)\cos\theta - z\sin\theta]} \int_{-\theta}^{\pi-\theta} e^{-\alpha\delta^2 + 2\beta\delta} d\delta, \qquad (A1)$$

where $\alpha = A/4$ and $\beta = -ik[(x - x_0)\sin\theta + z\cos\theta]/2$.

If we extend the limits of integration to $\pm\infty$ by assuming that the integrand in (A1) is negligible outside the interval $[-\theta, \pi - \theta]$, then (A1) becomes

$$\Psi_0(x, z) = e^{ik[(x-x_0)\cos\theta - z\sin\theta]} e^{-[(x-x_0)+z\cot\theta]^2/g^2}. \qquad (A2)$$

If the mean surface at $z = 0$ is used as a reference level for measuring the Gaussian beam width, the magnitude of Ψ_0 is given by

$$|\Psi_0| = e^{-\frac{(x-x_0)^2}{2\sigma^2}}, \quad \sigma = g/\sqrt{2}.$$

If we then use $|x - x_0| < 3\sigma$ as an approximation to the full width of the Gaussian beam, then the beam width W_x is given by

$$W_x = 6\frac{g}{\sqrt{2}} \approx 4.24g. \qquad (A3)$$

The angular width of the incident beam given in (2) can thus be expressed approximately by

$$W_\theta = 2\sqrt{\frac{\pi}{2A}} \approx \frac{2.5}{kg\sin\theta}. \qquad (A4)$$

Since $W_x W_\theta \approx \frac{10}{k \sin\theta}$, a narrow beam with small angular spread can only be achieved if the wavenumber k is large, and if θ is not too small, i.e. if the incident angle is not too close to grazing. In the work here, typically an angular spread of approximately $\pm 10°$ was used.

In practice, for very large rms heights, it was found that this approximate solution, or high order refinements, were not accurate enough, and so a full numerical evaluation of the incident beam field was required. To evaluate the normal derivative ψ_0, as given in (6), the interval of integration $[0, \pi]$ is divided into n subintervals by setting

$$n = 2^w$$

where

$$w = \frac{\log[(kl)^2 + 10m]}{(\log 2 + 0.5)}.$$

Here [ ] denotes taking the integer part and m is the number of points per correlation length. It was found that the composite trapezoidal rule was the most efficient for this problem, probably because of the oscillatory nature of the integrand. (If the integrand were periodic on $[0, \pi]$, then the trapezoidal rule would obviously be appropriate; however this is not the case, because of the Gaussian taper function).

Appendix B: the algorithm for 1-D surfaces

Let n be the number of points on the surface and S_i be a patch on the surface for which $|\psi_0(x,\zeta(x))| > \epsilon$, where S_i is a subset of S and we define ϵ as a threshold value. The parameter δ is the required relative error in the unknown ψ. The maximum size of patch allowed is M. Then the algorithm for the method can be briefly described as follows:

ALGORITHM

start: Set $\chi := 2\psi_0$ and $\psi := 0$, given ϵ, δ and M
iterate: **For** $i = 1$ **step** 1 **until** convergence **do**
(1) $k := 0$, $p_s := 1$ (if i odd) or $p_s := n$ (if i even), $p_n = n + 1 - p_s$;
(2) **iterate**: **For** $p = p_s$ **step** 1 **until** $p = p_n$ **do**
(2a) $k := k + 1$;
(2b) determine patch $S_i^{(k)}$, on which $|\chi| > \epsilon$;
(2c) solve $\psi_i^{(k)}(x) + \alpha_{ii}^{(k)}(x,x')\psi_i(x)^{(k)} = \chi(x), \quad x, x' \in S_i^{(k)}$;
(2d) $\psi_{0,i+1}^{(k)} := A_i\psi_i^{(k)}$, here A_i has dimension $(n \times m)$,
m is the size of $S_i^{(k)}$;
(2e) $\chi := \chi - \psi_{0,i+1}^{(k)}$;
(2f) $\psi := \psi + \psi_i^{(k)}, \quad x \in S$.

CONVERGENCE CRITERION

There are two options for the stopping criterion:

$$1. \qquad \left[\frac{\langle \psi_i^* \cdot \psi_i \rangle}{\max(|\chi|)}\right]^{1/2} < \delta;$$

$$2. \qquad i > N,$$

where ψ_i^* is the complex conjugate of ψ_i. If high accuracy is required, then the first option with a small value of δ is required. However, a large number of iterations, usually more than N, is then needed. The second option is appropriate if the maximum number of reflections N is known for the scattering problem given.

The values of ϵ and M, whose choice depends on the initial beam width on the surface, will have some influence on the speed of convergence. In order to run the program efficiently, ϵ is readjusted repeatedly in step (2b). If the size of $S_i^{(k)} > M$, we then enlarge ϵ by multiplying by 1.5 (say) and perform the sweep again. At the end of each stage i, ϵ is reduced to ϵ/i^2. Step (2c) is where the major computation takes place. Both direct methods and iterative methods can be considered for the solution of the linear equation. The condition of these small linear systems turns out to be much better than that for the full system of order n, i.e. the spectral radius is smaller. This means that iterative methods can be used with more success. For very rough surfaces, however, the reduced problem (2c) is still strongly non-symmetric, so nonsymmetrizable iterative methods, such as the generalized congugate gradient method, are still needed (see the discussion in Cao and Macaskill, 1995). The fact that the spectral radius is reduced, however, means that more extreme problems may be treated with these iterative techniques if the new algorithm is used.

References

Cao, P. (1995). "Wave scattering from randomly rough surfaces", PhD thesis, School of Mathematics and Statistics, University of Sydney, Sydney, 187pp.

Cao, P. and Macaskill, C. (1995). "Iterative techniques for rough surface scattering problems", *Wave Motion* **21**, 209–229.

Fung, A. K and Chen, M. F. (1985). "Numerical simulation of scattering from simple and composite random surfaces", *J. Opt. Soc. Am.*, **A2**, 2274–2284.

Kachoyan, B. J. and Macaskill, C. (1987). "Acoustic scattering from an arbitrarily rough surface", *J. Acoust. Soc. Am.*, **82**, 1720–1726.

Macaskill, C. (1991). "Geometric optics and enhanced backscattering from very rough surfaces", *J. Opt. Soc. Am. A* **8**, 88–96.

Macaskill, C. and Kachoyan, B. J. (1993). "Iterative approach for the numerical simulation of scattering from one- and two-dimensional rough surfaces", *Appl. Opt.* **32**, 2839–2847.

Maystre, D., Saillard, M. and Ingers, J. (1991). "Scattering by one- or two-dimensional randomly rough surfaces", *Waves in Random Media*, **3**, 143–155.

Meecham, W.C. (1956). "On the use of the Kirchhoff approximation for the solution of reflection problems", *J. Rat. Mech. Anal.*, **5**, 323–334.

Saillard, M. and Maystre, D. (1988). "Scattering from random rough surfaces: A beam simulation method", *J. Opt. (Paris)*, **19**, 173–176.

Saillard, M. and Maystre, D. (1990). "Scattering from metallic and dielectric rough surfaces", *J. Opt. Soc. Am. A*, **7**, 982–990.

Soto-Crespo, J.M. and Nieto-Vesperinas, M. (1989) "Electromagnetic scattering from very rough random surfaces and deep reflection gratings", *J. Opt. Soc. Am. A*, **6**, 367–384.

Thorsos, E. I. (1988). "The validity of the Kirchhoff approximation for rough surface scattering using a Gaussian spectrum", *J. Acoust. Soc. Am.*, **83**, 78–92.

Thorsos, E. I. and Jackson, D. R. (1991). "Studies of scattering theory using numerical methods", *Waves in Random Media*, **3** S165-S190.

Tran, P. and Maradudin, A. A. (1992). "Scattering of a scalar beam from a two-dimensional randomly rough hard wall: Enhanced backscattering", *Phys. Rev. B*, **45**, 3936–3939.

Tran, P. and Maradudin, A. A. (1994). "Scattering of electromagnetic waves from a randomly rough 2D metallic surface", *Opt. Comm.*, **110**, 269–273.

Tsang, L., Chan, C. H. and Pak, K. (1994). "Backcattering enhancement of a two-dimensional random rough surface (three-dimensional scattering) based on Monto Carlo simulations", *J. Opt. Soc. Am. A*, **11**, 711–715.