Reprinted from JOURNAL OF PHYSICAL OCEANOGRAPHY, Vol. 7, No. 1, January 1977
American Meteorological Society
Printed in U. S. A.

On the Nonlinear Theory for Gravity Waves on the Ocean's Surface. Part I: Derivations

B. L. WEBER AND D. E. BARRICK

National Oceanic and Atmospheric Administration, Wave Propagation Laboratory, Sea State Studies Program, Boulder, Colo. 80302

(Manuscript received 17 February 1976, in revised form 10 September 1976)

ABSTRACT

A general hydrodynamic solution is derived for arbitrary gravity-wave fields on the ocean surface by extending Stokes' (1847) original perturbational analysis. The solution to the nonlinear equations of motion is made possible by assuming that the surface height is periodic in both space and time and thus can be described by a Fourier series. The assumption of periodicity does not limit the generality of the result because the series can be made to approach an integral representation by taking arbitrarily large fundamental periods with respect to periods of the dominant ocean waves actually present on the surface. The observation areas and times over which this analysis applies are assumed small, however, compared to the periods required for energy exchange processes; hence an "energy balance" (or steady-state) condition is assumed to exist within the observed space-time intervals. This in turn implies the condition of statistical stationarity of the Fourier height coefficients when one generalizes to a random surface. Part I confines itself to the formulation of a perturbation solution (valid to all orders) for the higher order terms resulting from a two-dimensional arbitrary periodic description of the surface height. The method is demonstrated by deriving (to second order) the height correction to the sea and (to third order) the first nonzero correction to the lowest order gravity-wave dispersion relation.

1. Introduction

In recent years, it has become evident that radio waves can be used to measure an appreciable portion of the directional ocean wave-height spectrum. Barrick (1972) has presented a theory that relates this spectrum to the radar Doppler spectrum, which is observed when radio waves are scattered from the sea surface.

The purpose of this paper is 1) to derive the hydrodynamic part[1] of Barrick's theoretical results and 2) to derive the general correction term for the deep-water gravity-wave dispersion relation.

The approach taken here to solve the nonlinear hydrodynamic equations for ocean waves is similar to the approach that was used by Stokes (1847). That is, the general form of the solution is first postulated and then the details of this solution are performed starting with the equations of motion. In his classic work, Stokes sought a solution for a single gravity wave that propagates with a rigid, periodic profile and a constant velocity. He found that the wave height contained higher spatial frequency components in addition to the fundamental sinusoid predicted by the linear solution and that the higher order correction to the wave velocity depended upon the wave height.

The present problem is to find a general periodic wave train (consisting of an arbitrary number of individually distinct gravity waves) whose profile need not be rigid and whose spectral components need not all have the same phase velocity. The condition that the wave train is periodic allows the wave height to be expanded in a spatial and temporal Fourier series, which greatly simplifies the solution. The assumption of periodicity is a mathematical device that does not limit the generality of the solution because the fundamental spatial and temporal periods of the series can be made large compared with the physical area and time interval over which observations are made. In fact when these periods become very large—approaching infinity—we intend to use the fact that the series converge to integrals in the Riemann sense. It is assumed that such a periodic wave train can be used to give an approximate description of real ocean waves. The apparent stochastic character of these waves can then be included, if desired, by allowing the wave heights to be random variables.

It is more common to use a spatial and temporal Fourier-Stieltjes representation for analyzing ocean

[1] The electromagnetic scattering part of the theory will be derived in another paper.

waves. For example, Tick (1959) and Huang (1971) obtained straightforward nonlinear solutions by using this approach, but they also neglected the wave-height dependence in the dispersion relation. Tick's first and second-order perturbation derivations are essentially the same as those presented here, although Tick confined himself to one-dimensional ocean waves. Huang obtained a Fourier-Stieltjes integral equation relating wave height to velocity potential (to all orders); this expression could have been expanded in a perturbation series to solve for wave height correctly to second order, but that result was not pursued in Huang's treatment. Huang and Tung (1976) derived a general dispersion relation which was a function of wave height but which was also a function of space and time. However, their derivation appears inconsistent because they did not treat the wave frequency as a function of space and time throughout the entire derivation.

Phillips (1957, 1960) and Hasselmann (1962, 1963a,b) used a Fourier-Stieltjes integral for the spatial coordinates alone, leaving the wave height a general function of time. In this way, the original equations of motion, which contain both space and time derivatives, are replaced by differential equations with time derivatives only. Longuet-Higgins and Phillips (1962) and Benney (1962) used this general formulation to study the wave height dependence of wave velocity for the simpler case of one-dimensional wave trains. They used formal expressions for wave height that are essentially Fourier series, where the Fourier coefficients are slowly varying functions of time. This approach leads to a number (equal to the number of terms in the Fourier series) of coupled differential equations. Therefore, it is understandable why these studies were restricted to cases with small numbers of waves. The approach used in the present paper allows for an arbitrary number of waves (of different spatial periods and directions) by requiring that the sea surface be periodic.

Although we have neglected the energy transfer due to wind-wave interactions, wave-wave interactions, viscous damping, etc., the present problem is well-defined and soluble. The perturbation approach used by Stokes is valid mathematically (see, Lamb, 1932, p. 420); thus, we believe that the present generalization of Stokes' solution is also valid. In addition, since the energy transfer rates are relatively low compared with the periods of gravity waves, we expect the present description of these waves to give an accurate physical picture of the sea surface just as Stokes' waves closely resemble simple wave trains such as swells. The various energy transfer mechanisms are discussed in detail elsewhere by Phillips (1966), Hasselmann (1966), Miles (1967), Willebrand (1975) and Whitham (1967), just to name a few. Our main concern here is the correct description of the sea surface over times and areas such that energy transfer is not a dominant feature in the propagation of the gravity waves on the surface.

This paper confines itself to the derivational details and their justification. Fourier-series forms are used in this paper. A companion paper shows that this generalized two-dimentional solution checks, in the appropriate limiting cases, with the simpler but well-established results of Stokes (1847) for wave velocity and height corrections for a single wave; with Longuet-Higgins and Phillips' (1962) phase velocity correction for one wave due to the presence of another colinear wave; and with Tick's (1959) result for the second-order wave height of a one-dimensional wave-train profile. It is shown how the Fourier series approach can be converted to integral form suitable for statistical averaging processes. Finally, that paper gives several applications of these derivations to physical situations, which provide some appreciation for the utility of the results. Thus we believe that this work represents the first truly complete generalization of Stokes' technique which stands up to comparisons with all of the previously accepted specialized cases.

2. The generalization of Stokes' problem

The problem to be solved here is a generalization of the problem that was solved earlier by Stokes (1847). In that problem, Stokes sought a periodic wave train which propagates with a rigid profile and a constant velocity. The present problem is to find a general periodic wave train whose profile need not be rigid and whose spectral components need not all have the same phase velocity. The method of solution is based upon the perturbation technique used by Stokes. Also, the equations of motion are essentially the same simplified hydrodynamic equations that Stokes employed.

To begin with, the ocean is assumed to be infinitely deep and unbounded along its surface. Also, atmospheric effects are taken to be absent so that the interface is a free surface. Next, the water is treated as a homogeneous fluid that is incompressible, inviscid and without surface tension. All of these restrictions are generally accepted as being approximately valid for the description of the free propagation of gravity waves. Phillips (1966), Lamb (1932) and Batchelor (1970) are just a few who discussed these simplifying restrictions in detail. These assumptions can now be used to simplify the general hydrodynamic equations.

The first equation derives from the conservation of mass equation, which reduces to $\nabla \cdot \mathbf{v}=0$, where $\mathbf{v}$ equals the local velocity of the water. For an inviscid fluid, initially irrotational motion will remain irrotational (i.e., $\nabla \times \mathbf{v}=0$). In this case, a velocity potential ϕ can be defined such that $\mathbf{v}=\nabla\phi$, and the conservation of mass is then expressed by

$$\nabla^2\phi=0. \tag{1}$$

For these same conditions the Navier-Stokes equation (or conservation of linear momentum equation) at the surface can be used to obtain

$$\left[\frac{\partial\phi}{\partial t}+\frac{1}{2}\nabla\phi\cdot\nabla\phi\right]_{z=\eta}=-g\eta, \quad (2)$$

where η is the vertical displacement of the surface due to waves. The coordinate system for this equation was chosen so that the positive (vertical) z axis is up and the x and y axes are in the plane (flat-earth approximation) of the undisturbed surface at $z=0$.

The third and final equation comes from the requirement that the surface remain intact. Then the vertical velocity $v_z=\partial\phi/\partial z$ of the water at the surface must equal the vertical velocity $d\eta/dt$ of the surface. That is,

$$\left[\frac{\partial\phi}{\partial z}\right]_{z=\eta}=\frac{\partial\eta}{\partial t}+\nabla\eta\cdot[\nabla\phi]_{z=\eta}. \quad (3)$$

Now a periodic waveform for η is sought. Therefore, η is expanded in spatial and temporal Fourier series as follows:

$$\eta(\mathbf{r},t)=\sum_{\mathbf{k},\omega}\eta(\mathbf{k},\omega)\exp[i(\mathbf{k}\cdot\mathbf{r}-\omega t)], \quad (4)$$

where $\mathbf{r}=\hat{x}x+\hat{y}y$ gives the position in the x, y plane. All of the spatial and temporal frequencies are harmonics of the fundamental frequencies, which may be taken to be infinitesimal.

It is obvious from (4) that η is a function of only the x and y spatial coordinates and time t. However, the velocity potential ϕ depends upon the z coordinate also. The form of the z dependence in ϕ is determined by (1) and the condition that $\phi\rightarrow 0$ as $z\rightarrow-\infty$. Therefore,

$$\phi(\mathbf{r},z,t)=\sum_{\mathbf{k},\omega}\phi(\mathbf{k},\omega)\exp[kz+i(\mathbf{k}\cdot\mathbf{r}-\omega t)]. \quad (5)$$

These expressions (4) and (5) give a general form for a periodic wavetrain. The fact that $\eta(r,\mathbf{t})$ and $\phi(\mathbf{r},z,t)$ are real physical quantities requires that the conditions $\eta^*(\mathbf{k},\omega)=\eta(-\mathbf{k},-\omega)$ and $\phi^*(\mathbf{k},\omega)=\phi(-\mathbf{k},-\omega)$ be satisfied by the series (4) and (5), respectively.

The fact that (2) and (3) are nonlinear suggests a perturbation approach. Stoker (1957), Tick (1959) and others preferred to make the perturbation expansions on $\eta(\mathbf{r},t)$ and $\phi(\mathbf{r},z,t)$. However, Whitham (1974) and Cole (1968) pointed out that such an approach may omit the amplitude dependence in the dispersion relation. On the other hand, Stokes' perturbation method does produce an amplitude dependence for the dispersion relation, and thus it is believed to be more general.

Therefore, the Fourier coefficients for wave height $\eta(\mathbf{k},\omega)$, velocity potential $\phi(\mathbf{k},\omega)$ and the frequency ω are expanded in perturbation series as follows:

$$\eta(\mathbf{k},\omega)=\eta_1(\mathbf{k},\omega)+\eta_2(\mathbf{k},\omega)+\cdots, \quad (6)$$

$$\phi(\mathbf{k},\omega)=\phi_1(\mathbf{k},\omega)+\phi_2(\mathbf{k},\omega)+\cdots, \quad (7)$$

$$\omega=\omega_0+\omega_1+\omega_2+\cdots, \quad (8)$$

where the subscripts give the perturbation order. For example, $\eta_2\sim\eta_1\eta_1$, $\phi_1\sim\eta_1$, $\phi_2\sim\eta_1\eta_1$, etc. Similarly, $\omega_1\sim\eta_1$, $\omega_2\sim\eta_1\eta_1$, but ω_0 is independent of wave height. Thus η_1 is considered an independent parameter and $\mathbf{k}$ is the independent variable of the present problem. The perturbation approach is valid if the wave heights are sufficiently small such that

$$\sum_{\mathbf{k},\omega}|\eta(\mathbf{k},\omega)|\times k<1. \quad (9)$$

This condition limits the slope of the surface to small values so that the perturbation expansion will converge.

Later, it will become evident that the wave heights of various orders in (6) do not all exist in the same domains of wave vector frequency space. In other words, the dispersion relation is in general different for each order of the ocean wave. For example, the first- and second-order wave-height spectra do not overlap in wave vector frequency space. Therefore, it will become convenient to use different notation for the wavevectors and frequencies of different orders of ocean waves.

3. The first-order solution

The perturbation expansions (6), (7) and (8) can be used along with the Fourier series (4) and (5) in order to solve the equations of motion for a periodic wave train. In this solution, the first-order wave heights $\eta_1(\mathbf{k},\omega)$ are arbitrary and all of the other variables are expressed in terms of them. The solution begins by substituting the Fourier series (4) and (5) into the equations of motion (2) and (3).

When the Fourier series are substituted into the equations of motion, the exponential $\exp(kz)$ in (5) becomes $\exp[k\eta(\mathbf{r},t)]$ because these equations are evaluated at the surface $z=\eta(\mathbf{r},t)$. Then the exponential is represented by its power series, and the wave height in each term is replaced by its Fourier series (4). Finally, the resulting equations are integrated over the spatial and temporal periods of the wave train. Because of the orthogonality of the Fourier components, (2) becomes

$$\begin{aligned}&-i\omega\phi(\mathbf{k},\omega)+\sum_{\mathbf{k}',\omega'}[-i\omega'k'\phi(\mathbf{k}',\omega')\eta(\mathbf{k}-\mathbf{k}',\omega-\omega')\\&+\tfrac{1}{2}[k'|\mathbf{k}-\mathbf{k}'|-\mathbf{k}'\cdot(\mathbf{k}-\mathbf{k}')]\phi(\mathbf{k}',\omega')\phi(\mathbf{k}-\mathbf{k}',\omega-\omega')]\\&+\sum_{\mathbf{k}',\omega'}\sum_{\mathbf{k}'',\omega''}\left[-i\omega'\frac{k'^2}{2}\phi(\mathbf{k}',\omega')\eta(\mathbf{k}'',\omega'')\right.\\&\times\eta(\mathbf{k}-\mathbf{k}'-\mathbf{k}'',\omega-\omega'-\omega'')+\tfrac{1}{2}(k'k''-\mathbf{k}'\cdot\mathbf{k}'')(k'+k'')\\&\left.\times\phi(\mathbf{k}',\omega')\phi(\mathbf{k}'',\omega'')\eta(\mathbf{k}-\mathbf{k}'-\mathbf{k}'',\omega-\omega'-\omega'')\right]\\&+O(4)=-g\eta(\mathbf{k},\omega). \quad (10)\end{aligned}$$

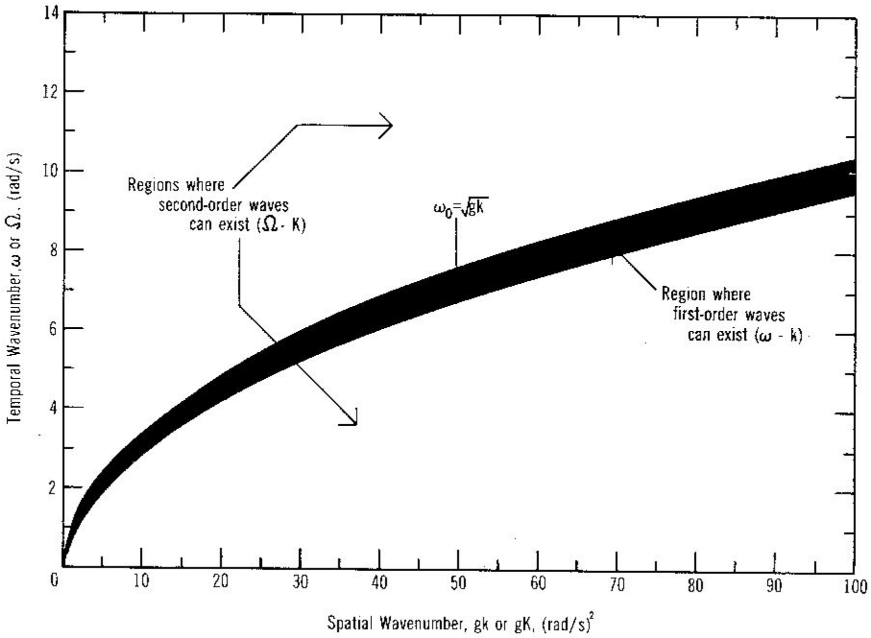

FIG. 1. Gravity-wave dispersion relationship diagram. First-order waves exist in heavily shaded region centered on $\omega_0=\sqrt{gk}$. Second-order waves exist in remainder of diagram.

Similarly, the other equation of motion [(3)] becomes

$$k\phi(\mathbf{k},\omega)+\sum_{\mathbf{k}',\omega'} k'^2\phi(\mathbf{k}',\omega')\eta(\mathbf{k}-\mathbf{k}',\omega-\omega')$$

$$+\sum_{\mathbf{k}',\omega'}\sum_{\mathbf{k}'',\omega''}\frac{k'^3}{2}\phi(\mathbf{k}',\omega')\eta(\mathbf{k}'',\omega'')$$

$$\times\eta(\mathbf{k}-\mathbf{k}'-\mathbf{k}'',\omega-\omega'-\omega'')+O(4)=-i\omega\eta(\mathbf{k},\omega)$$

$$+\sum_{\mathbf{k}',\omega'}-\mathbf{k}'\cdot(\mathbf{k}-\mathbf{k}')\phi(\mathbf{k}',\omega')\eta(\mathbf{k}-\mathbf{k}',\omega-\omega')$$

$$+\sum_{\mathbf{k}',\omega'}\sum_{\mathbf{k}'',\omega''}-\mathbf{k}'\cdot\mathbf{k}''k'\phi(\mathbf{k}',\omega')\eta(\mathbf{k}'',\omega'')$$

$$\times\eta(\mathbf{k}-\mathbf{k}'-\mathbf{k}'',\omega-\omega'-\omega'')+O(4). \quad (11)$$

Next, the perturbation expansions will be used in (10) and (11) in order to separate the various orders of perturbation solution. The first-order terms in (10) are equated to give

$$-i\omega_0\phi_1(\mathbf{k},\omega)=-g\eta_1(\mathbf{k},\omega), \quad (12)$$

while the corresponding terms in (11) are

$$k\phi_1(\mathbf{k},\omega)=-i\omega_0\eta_1(\mathbf{k},\omega). \quad (13)$$

Hence

$$(\omega_0^2-gk)\eta_1(\mathbf{k},\omega)=0. \quad (14)$$

This equation can be satisfied if either the wave height $\eta_1(\mathbf{k},\omega)$ vanishes or

$$\omega_0^2=gk. \quad (15)$$

Therefore, for finite wave heights (15) becomes a dispersion relation which must be satisfied. It is important to note, however, that this condition (15) puts a restriction only upon the value of the first term ω_0 of the perturbation expansion of the frequency ω. The higher order terms in (8) are determined by the higher order terms in (10) and (11). Also, it is the total frequency ω, and not just ω_0, which must be a harmonic frequency of the fundamental frequency. In Stokes' case of a rigid profile, the frequency of the first-order wave is the fundamental frequency. In the general case with many first-order waves present, the fundamental frequency need not be equal to the frequency of any of these waves.

4. The second-order solution

The second-order solution to (10) is

$$-i\omega_0\phi_2(\mathbf{k},\omega)-i\omega_1\phi_1(\mathbf{k},\omega)+\sum_{\mathbf{k}',\omega'}\{-i\omega_0'k'\phi_1(\mathbf{k}',\omega')$$

$$\times\eta_1(\mathbf{k}-\mathbf{k}',\omega-\omega')+\tfrac{1}{2}[k'|\mathbf{k}-\mathbf{k}'|-\mathbf{k}'\cdot(\mathbf{k}-\mathbf{k}')]$$

$$\times\phi_1(\mathbf{k}',\omega')\phi_1(\mathbf{k}-\mathbf{k}',\omega-\omega')\}=-g\eta_2(\mathbf{k},\omega). \quad (16)$$

Likewise, the second-order solution to (11) can be written

$$k\phi_2(\mathbf{k},\omega)+\sum_{\mathbf{k}',\omega'}\mathbf{k}'\cdot\mathbf{k}\phi_1(\mathbf{k}',\omega')\eta_1(\mathbf{k}-\mathbf{k}',\omega-\omega')$$

$$=-i\omega_0\eta_2(\mathbf{k},\omega)-i\omega_1\eta_1(\mathbf{k},\omega), \quad (17)$$

where terms on both sides of (11) have been combined.

There are two ways of attempting to satisfy (16) and (17) and solving for the desired second-order wave height η_2 and/or velocity potential ϕ_2. These can best be explained by dividing the frequency wavenumber $(\omega-k)$ domain into two distinct regions as illustrated in Fig. 1. The heavily shaded ridge is the region, for a given k, where the lowest order solution (zero-order) for ω must satisfy the dispersion equation (15) given in the preceding section, viz., $\omega_0=\pm\sqrt{gk}$. Since we have expanded ω in a perturbation series about ω_0, then $\omega_1, \omega_2, \ldots$ should be increasingly small corrections to ω_0 by the very nature of the perturbation expansions. The center of the deep-shaded ridge is therefore defined by $\omega_0=\sqrt{gk}$, and the width of the ridge (within which $\omega_1, \omega_2, \ldots$ must lie), defining the complete ω, is kept small. We call the region within this ridge the region of first-order waves; we continue to denote the temporal frequencies in this region by ω, ω_0, ω_1, etc., i.e., lower case. Over the remainder of the diagram, no such restriction is required on the lowest order temporal frequency; to avoid confusion, we henceforth redefine the frequencies and wavenumbers corresponding to this region as $\mathbf{k}\equiv\mathbf{K}$ and $\omega\equiv\Omega$, with $\mathbf{K}$ as before being the independent variable. Just as with ω, we expand Ω in a perturbation series, $\Omega=\Omega_0+\Omega_1+\Omega_2+\cdots$; but here we do *not* require that $\Omega_0=\sqrt{gK}$. It will be shown that second-order ocean waves lie in this lightly shaded region exclusive of the ridge around $\omega_0=\sqrt{gk}$ and that the two regions are in fact nonoverlapping. In other words, second-order waves cannot even approximately satisfy the first-order dispersion relationship normally identified with freely propagating ocean waves.

In order to eliminate the first possible way of solving (16) and (17), let us attempt to solve for η_2 and ϕ_2 at frequencies ω which could lie in the first-order zone. Since ω, ω' and ω'' all appear as arguments of first-order wave heights and velocity potentials, to lowest order we saw that we must require that $\omega_0=\sqrt{gk}$, $\omega_0'=\sqrt{gk'}$ and $\omega_0''(=\omega_0-\omega_0')=\sqrt{g|\mathbf{k}-\mathbf{k}'|}$. If these three equations cannot be satisfied, then in general *no* solution can exist within the ridge for ω, since if the lowest perturbation order ω_0 fails to meet the requirement, then the overall frequency ω also fails (i.e., each perturbation order must be satisfied separately).

To show this zero-order failure, the three separate dispersion equations $\omega_0^2=gk$, $\omega_0'^2=gk'$ and $\omega_0''^2=gk''$ can be combined to give

$$\omega_0^2+\frac{(3-\hat{k}\cdot\hat{k}'')}{2}\omega_0\omega_0''+\omega_0''^2=0,$$

which has real roots only if $\hat{k}\cdot\hat{k}''=-1$. The solution in this event is $\omega_0=-\omega_0''$ (requiring also $k=k''$), which leads to $\omega_0'=0$ and $\mathbf{k}'=0$. Thus of all the terms in the series of (16) and (17), only one is permissible: that with $\mathbf{k}'$ and $\omega_0'=0$. There cannot be any waves with $k'=0$, however, because this would produce a change in the mean level of the ocean, requiring the creation or destruction of water. We already defined the $x-y$ plane to lie in the mean plane of the ocean, implying that $\eta_1(\mathbf{0},\omega)=0$. Hence, the summations in (16) and (17) must vanish, leaving

$$-i\omega_0\phi_2(\mathbf{k},\omega)-i\omega_1\phi_1(\mathbf{k},\omega)+g\eta_2(\mathbf{k},\omega)=0, \quad (18)$$

$$k\phi_2(\mathbf{k},\omega)+i\omega_0\eta_2(\mathbf{k},\omega)+i\omega_1\eta_1(\mathbf{k},\omega)=0. \quad (19)$$

Now if we multiply the first equation by k, multiply the second by $i\omega_0$, and add the two, then using the fact that $\omega_0^2=gk$ we obtain $i\omega_1k\phi_1(\mathbf{k},\omega)+\omega_0\omega_1\eta_1(\mathbf{k},\omega)=0$. Upon substitution of (13) into this, we have $2\omega_0\omega_1\eta_1(\mathbf{k},\omega)=0$. Since in general $\eta_1(\mathbf{k},\omega)$ is not equal to zero (i.e., η_1, $\mathbf{k}$, and hence ω_0, are the independent variables of the problem), then $\omega_1\equiv0$. This leaves (18) and (19) for the second-order wave height and velocity potential identical identical to (12) and (13) for the equivalent first-order quantities. Hence, there is no unique second-order solution for η_2 and ϕ_2 (within the "ridge" of Fig. 1 where first-order waves can exist) which is dependent upon and expressible in terms of η_1 and/or ϕ_1. Since there is no difference between the ϕ_2 and η_2 remaining in (18) and (19) and ϕ_1 and η_1 in (12) and (13), we can define all quantities which lie in the ridge and satisfy these required first-order equations as first order.

Therefore we have established the following facts thus far in this section:

1) First-order waves—by definition—exist within the narrow region in ω centered about ω_0, with $\omega_1, \omega_2, \ldots$ being higher order (perturbation) corrections which are small compared to ω_0; η_1, ϕ_1, and ω_0 satisfy the first-order relation given by (12), (13) and (15).

2) The first-order correction to ω_0 (*viz.*, ω_1) is identically zero. Therefore, the first possible correction to the lowest-order dispersion relation is second-order; this fact might have been suspected from Stokes' original derivations as well as later works (e.g., Lamb, 1932).

3) Second-order waves η_2 (and velocity potentials ϕ_2) which are dependent upon and expressible in terms of double products of η_1 and/or ϕ_1 cannot exist within the narrow region in ω near ω_0 which contain the first-order waves (such that the lowest-order dispersion equation $\omega_0^2=gk$ is satisfied). They can exist and will be derived subsequently over the remainder of the temporal frequency region.

We now go back to (16) and (17) and solve for η_2 and ϕ_2 in the region where they can in fact exist: that region of Fig. 1 considerably away from the first-order ridge. As mentioned before, we change wavenumber notation to upper case (i.e., $\mathbf{K},\Omega$) to indicate that in this region the first-order dispersion relation does not apply, i.e., $\Omega_0^2\neq gK$. Thus whenever $\mathbf{k}$ and ω appear, we change to $\mathbf{K}$ and Ω. However, $\mathbf{k}'$ and ω' remain, and all arguments of η_1 and ϕ_1 are still required to satisfy the first-order dispersion equation; i.e., $\omega_0'^2=gk'$ and $(\Omega_0-\omega_0')^2=g|\mathbf{K}-\mathbf{k}'|$. Also, the second term on the left side of (16) and the last term on the right side of (17) are zero, because, as shown above, ϕ_1 and η_1 cannot exist in the second-order region (where $\mathbf{k}=\mathbf{K}$ and $\omega=\Omega$). Thus (16) can be rewritten as follows:

$$-i\Omega_0\phi_2(\mathbf{K},\Omega)+\sum_{\mathbf{k},\omega}\left[-\omega_0^2-\tfrac{1}{2}\omega_0(\Omega_0-\omega_0)\left(1-\frac{\hat{k}\cdot(\mathbf{K}-\mathbf{k})}{|\mathbf{K}-\mathbf{k}|}\right)\right]$$
$$\times\eta_1(\mathbf{k},\omega)\eta_1(\mathbf{K}-\mathbf{k},\Omega-\omega)=-g\eta_2(\mathbf{K},\Omega), \quad (20)$$

where (13) was used to replace the first-order potentials inside the summation in (16). Similarly, (17) becomes

$$K\phi_2(\mathbf{K},\Omega)+\sum_{\mathbf{k},\omega}-i\omega_0\hat{k}\cdot\mathbf{K}\eta_1(\mathbf{k},\omega)\eta_1(\mathbf{K}-\mathbf{k},\,\Omega-\omega)=-i\Omega_0\eta_2(\mathbf{K},\Omega). \quad (21)$$

These last two equations can be combined to obtain an expression for $\eta_2(\mathbf{K},\Omega)$ and one for $\phi_2(\mathbf{K},\Omega)$.

The second-order wave height becomes

$$\eta_2(\mathbf{K},\Omega)=\sum_{\mathbf{k},\omega}\sum_{\mathbf{k}',\omega'}A(\mathbf{k},\omega,\mathbf{k}',\omega')\eta_1(\mathbf{k},\omega)\eta_1(\mathbf{k}',\omega')\times\delta_{\mathbf{K}}^{\mathbf{k}+\mathbf{k}'}\delta_{\Omega}^{\omega+\omega'}, \quad (22)$$

where this expression has been written in a symmetrical form with the help of the Kronecker delta functions, and where

$$A(\mathbf{k},\omega,\mathbf{k}',\omega')=\begin{cases}\dfrac{1}{2}\left[k+k'+\dfrac{\omega_0\omega_0'}{g}(1-\hat{k}\cdot\hat{k}')\left(\dfrac{gK+\Omega_0^2}{gK-\Omega_0^2}\right)\right]\\ 0 \quad \text{if} \quad \mathbf{k}'=-\mathbf{k} \quad \text{and} \quad \omega'=-\omega.\end{cases} \quad (23)$$

This last expression was simplified by taking advantage of the lowest order dispersion relation.

Since it has been proven earlier that $\Omega_0^2\neq gK$, the denominator in (23) cannot vanish for any real values of $\mathbf{k}$, $\mathbf{k}'$, ω, ω' satisfying the Kronecker deltas. It can be shown that the two terms on the left side of (21) vanish for $K=0$. Because the velocity $\mathbf{v}=\nabla\phi$, it is obvious that $\mathbf{v}(\mathbf{k},\omega)=i\mathbf{k}\phi(\mathbf{k},\omega)$ for all perturbation orders. Now, it is always possible to choose our coordinate system so that the undisturbed ocean is stationary (i.e., there is no current). Hence, $\mathbf{v}(\mathbf{k},\omega)=0$ for $k=0$ to all perturbation orders and the first term on the left side of (21) vanishes. Similarly, the second term vanishes because K is a multiplying factor. As a result, $A(\mathbf{k},\omega,\mathbf{k}',\omega')$ in (23) is defined to be identically zero when $\mathbf{k}+\mathbf{k}'=0$.

Since second-order waves cannot satisfy the first-order dispersion equation, these waves are not "free" in that they do not remove energy from the first-order waves, and hence cannot propagate freely without the two first-order waves with wavenumbers $\mathbf{k}$, ω and $\mathbf{k}'$, ω'; they are said to be "trapped" or "evanescent" ocean waves. By the same token, whenever the first-order waves are present, the second-order waves will always accompany them.

The second-order wave height (22) was used by Barrick (1972) in order to calculate the theoretical radar Doppler spectrum that is continuously distributed about the first-order solution. The coefficient $A(\mathbf{k},\omega,\mathbf{k}',\omega')$, which appears in this wave-height expression and is given by (23), is equal to $i\Gamma_H$ in Barrick's integral expression for the Doppler spectrum. There is also another quantity in this integral which accounts for second-order electromagnetic contributions.

In a similar manner, (20) and (21) can be combined to obtain an expression for the second-order velocity potential. Hence,

$$\phi_2(\mathbf{K},\Omega)=\sum_{\mathbf{k},\omega}\sum_{\mathbf{k}',\omega'}B(\mathbf{k},\omega,\mathbf{k}',\omega')\eta_1(\mathbf{k},\omega)\eta_1(\mathbf{k}',\omega')\times\delta_{\mathbf{K}}^{\mathbf{k}+\mathbf{k}'}\delta_{\Omega}^{\omega+\omega'}, \quad (24)$$

where

$$B(\mathbf{k},\omega,\mathbf{k}',\omega')=\frac{-i\Omega_0\omega_0\omega_0'(1-\hat{k}\cdot\hat{k}')}{(gK-\Omega_0^2)}. \quad (25)$$

By using the same line of reasoning that followed (23), it is clear that $KB(\mathbf{k},\omega,\mathbf{k}',\omega')\rightarrow 0$ for $K\rightarrow 0$ because the current velocity vanishes by definition.[2]

5. The third-order solution

The higher order equations from (10) and (11) can in principle be solved to all perturbation orders. However, the complexity of these equations becomes prohibitive for large orders. In general, the nth order equations will have solutions for all orders of waves from first order up to and including nth order. That is, there are n different solutions for the nth order equations. Each of these solutions uses all of the previous solutions to all of the lower order equations.

The discussion here will be limited to the third-order solution for first-order waves. In other words, just as we did for second-order waves, we initially assume that these third-order waves can exist over all $\mathbf{k}-\omega$ space, both in the "ridge" region of Fig. 1 near which $\omega_0^2=gk$ and over the remaining region where $\Omega_0^2\neq gK$. We proved in the preceding section that second-order waves *cannot* exist in the first-order region near the ridge. A similar argument can be presented for third- and higher order waves. We will concern ourselves here with only those solutions which exist within the first-order wave region.

By following the pattern of the second-order solution and by using the first- and second-order results, it is possible to obtain the following simplified equations. The third-order expression for (10) becomes

$$-i\omega_0\phi_3(\mathbf{k},\omega)-i\omega_2\phi_1(\mathbf{k},\omega)+\sum_1=-g\eta_3(\mathbf{k},\omega). \quad (26)$$

Likewise, (11) leads to

$$k\phi_3(\mathbf{k},\omega)+\sum_2=-i\omega_0\eta_3(\mathbf{k},\omega)-i\omega_2\eta_1(\mathbf{k},\omega), \quad (27)$$

where the terms $\sum_1$ and $\sum_2$ are abbreviations which represent expressions that have the form of

$$\sum_{\mathbf{k}',\omega'}\sum_{\mathbf{k}'',\omega''}(\ldots)\eta_1(\mathbf{k}',\omega')\eta_1(\mathbf{k}'',\omega'')\times\eta_1(\mathbf{k}-\mathbf{k}'-\mathbf{k}'',\,\omega-\omega'-\omega'').$$

[2] Expressions obtained by Hasselmann (1962) for the second-order wave height and velocity potential coefficients can be shown to reduce exactly to our (23) and (25). Because that approach ignores corrections to the dispersion relation, however, wave height and velocity potential solutions of order higher than third will differ from those obtained with our formulation.

Now, both $\eta_3(\mathbf{k},\omega)$ and $\phi_3(\mathbf{k},\omega)$ can be simultaneously eliminated from (26) and (27) by using the first-order dispersion relation (15); thus, they are indeterminant and can be taken to be identically zero because they are physically indistinguishable from first-order waves.[2] One then obtains an expression for the second-order frequency term ω_2:

$$\omega_2\eta_1(\mathbf{k},\omega)=\sum_{\mathbf{k}',\omega'}\sum_{\mathbf{k}'',\omega''}(\ldots)\eta_1(\mathbf{k}',\omega')\eta_1(\mathbf{k}'',\omega'')\times\eta_1(\mathbf{k}-\mathbf{k}'-\mathbf{k}'',\omega-\omega'-\omega''). \quad (28)$$

At first, this equation seems to contain an inconsistency because ω_2 is real, while the first-order wave heights $\eta_1(\mathbf{k},\omega)$ are arbitrary complex parameters. However, this equation also implies that there must be four waves such that $\omega'''=\omega-\omega'-\omega''$ and $\mathbf{k}'''=\mathbf{k}-\mathbf{k}'-\mathbf{k}''$. As was discussed in detail earlier with respect to the second-order solution, this equality between the temporal frequencies must and does hold for all perturbation orders.

All of these equations of different perturbation order (i.e., $\omega_n'''=\omega_n-\omega_n'-\omega_n''$) cannot be simultaneously satisfied unless pairs of the frequencies are identically equal. That is, $\omega=\omega'$ and $\omega''=-\omega'''$, or $\omega=\omega''$ and $\omega'=-\omega'''$, or $\omega=\omega'''$ and $\omega'=-\omega''$. At the same time, there must be the analogous equalities among the spectral wavenumber vectors. In the case that $\omega=\omega'$ and $\omega''=-\omega'''$, for example, it is also true that $\mathbf{k}=\mathbf{k}'$ and $\mathbf{k}''=-\mathbf{k}'''$. Consequently, $\eta_1(\mathbf{k},\omega)$ is a factor in every term on both sides of (28). Thus, (28) reduces to

$$\omega_2=\omega_0\sum_{\mathbf{k}',\omega'}C(\mathbf{k},\omega,\mathbf{k}',\omega')|\eta_1(\mathbf{k}',\omega')|^2, \quad (29)$$

where

$$C(\mathbf{k},\omega,\mathbf{k}',\omega')=\frac{1}{2}\left[k'^2+\frac{\omega_0'}{\omega_0}\mathbf{k}\cdot\mathbf{k}'\left(2+\frac{k}{k'}\right)\right]\times[1-\tfrac{1}{2}\delta_{\omega'}^{\omega}\delta_{\mathbf{k}'}^{\mathbf{k}}-\tfrac{1}{2}\delta_{\omega'}^{-\omega}\delta_{\mathbf{k}'}^{-\mathbf{k}}]+A(\mathbf{k},\omega,\mathbf{k}',\omega')\times\left[-k'+\frac{\omega_0'}{\omega_0}\frac{\mathbf{k}\cdot\mathbf{k}'}{k'}\right]-\frac{B(\mathbf{k},\omega,\mathbf{k}',\omega')}{i\omega_0}\left[\mathbf{k}\cdot(\mathbf{k}+\mathbf{k}')-k|\mathbf{k}+\mathbf{k}'|+\frac{\omega_0'}{\omega_0}\frac{k}{k'}\mathbf{k}'\cdot(\mathbf{k}+\mathbf{k}')\right]. \quad (30)$$

The expressions for $A(\mathbf{k},\omega,\mathbf{k}',\omega')$ and $B(\mathbf{k},\omega,\mathbf{k}',\omega')$ are given by (23) and (25), respectively. Since $B(\mathbf{k},\omega,\mathbf{k}',\omega')$ contains a factor of i, the i in (30) is cancelled so that $C(\mathbf{k},\omega,\mathbf{k}',\omega')$ is clearly real. Here we see one advantage in retaining series formulations rather than integral notation, at least to this point. The step from (28) to (29) would be more difficult mathematically had we been using integrals.

Higher order correction terms to the frequency ω in (8) can be computed in principle, but they are small compared to ω_2. Thus, the dispersion relation for first-order ocean waves is given (to second order) by $\omega=\omega_0+\omega_2$, where ω_0 is defined by (15) and ω_2 is defined by (29). It is recalled that the first-order correction term ω_1 was shown to be zero for these waves. Once the dispersion relation is known for first-order waves it is also known for second-order waves because the Kronecker-delta functions in (22) imply that $\Omega=\omega+\omega'$, where Ω is the frequency of a second-order wave and ω and ω' are the frequencies of first-order waves. Hence to second-order $\Omega=\omega_0+\omega_0'+\omega_2+\omega_2'$. Some interesting numerical examples will be presented in a companion paper.

6. Discussion and conclusions

The nonlinear solution presented here for a two-dimensional deep-water surface of arbitrary profile—periodic in space and time—can be evaluated to any perturbation order following the technique presented here. To second order, we determined expressions for the wave height and velocity potential and showed that those do not represent free waves, i.e., waves which follow (approximately) the first-order dispersion equation $\omega_0^2=gk$. We then carried the solution to third order to solve for the first nonzero correction to the first-order dispersion equation.

This solution gives a complete mathematical description (but only an approximate physical description) of the sea surface; it must be restricted in area and time. The sizes of the observed area and time intervals over which the solutions are valid are such that they are large compared to the spatial periods $2\pi/|\mathbf{k}|$ and temporal periods $2\pi/|\omega|$ of the dominant waves present, but small in terms of the areas and times over which energy transfer takes place (i.e., both nonlinear wave-wave energy transfer, energy transfer between the atmosphere and ocean and viscous effects). The waveheights $\eta_1(\mathbf{k},\omega)$ are in general complex random variables whose statistics change over areas and times larger than those required for energy transfer.

So long as the above area/time restrictions are understood, the Fourier series solution can be generalized to allow one to perform statistical averaging, and sums are readily converted to integrals, with average waveheight spectra evolving from the height coefficients. This process will be illustrated in a companion paper. In certain (but not all) averaging processes, one must define a length or time period associated with the series-integral conversion, and this quantity remains in the final result. For the radar problem, this length (or area) period must logically be taken as the actual areal resolu-

[2] These waves are interesting in another context because, as was shown in the works of Hasselmann (1963a,b) and Phillips (1966), since these waves do satisfy the first-order dispersion equation—and hence are indistinguishable physically from first-order waves—they can and do carry energy (in contrast with the evanescent or trapped second-order waves $\eta_2(\mathbf{K},\Omega)$). Therefore, nonlinear wave-wave energy redistribution of the original spectrum can and does occur via third-order waves.

tion cell observed by the radar. Likewise, the temporal period (if needed explicitly in the final result) would be the coherent observation time associated with the buoy or radar experiment. An example will be given in the companion paper.

REFERENCES

Barrick, D. E., 1972: Remote sensing of sea state by radar. *Remote Sensing of the Troposphere*, V. E. Derr, Ed., U. S. Govt. Printing Office, Washington, D. C., Chap. 12.

Batchelor, G. K., 1970: *An Introduction to Fluid Dynamics*. Cambridge University Press, 615 pp.

Benney, D. J., 1962: Nonlinear gravity wave interactions. *J. Fluid Mech.*, **14**, 577–584.

Cole, J. D., 1968: *Perturbation Methods in Applied Mathematics*. Blaisdell Publishing Co., p. 84.

Hasselmann, K., 1962: On the nonlinear energy transfer in a gravity-wave spectrum; Part 1. General theory. *J. Fluid Mech.*, **12**, 481–500.

——, 1963a: On the non-linear energy transfer in a gravity-wave spectrum; Part 2. Conservation theorems, wave-particle correspondence; Irreversibility, *J. Fluid Mech.*, **15**, 273–281.

——, 1963b: On the non-linear energy transfer in a gravity-wave spectrum, Part 3. Evaluation of the energy flux and swell-sea interaction for a Neumann spectrum. *J. Fluid Mech.*, **15**, 385–398.

——, 1966: Feynman diagrams and interaction rules of wave-wave scattering processes. *Rev. Geophys.*, **4**, 1–32.

Huang, N. E., 1971: Derivation of stokes drift for a deep water random gravity wave field. *Deep-Sea Res.*, **18**, 255–259.

——, and C. Tung, 1976: The dispersion relation for a nonlinear random gravity wave field. *J. Fluid Mech.*, **75**, 337–345.

Lamb, H., 1932 *Hydrodynamics*, 6th ed. Dover Publications, 738 pp.

Longuet-Higgins, M. S., and O. M. Phillips, 1962: Phase velocity effects in tertiary wave interactions. *J. Fluid Mech.*, **12**, 333–336.

Miles, J. W., 1967: On the generation of surface waves by shear flows, Part 5. *J. Fluid Mech.*, **30**, 163–175.

Phillips, O. M., 1957: On the generation of waves by turbulent wind. *J. Fluid Mech.*, **2**, 417–445.

——, 1960: On the dynamics of unsteady gravity waves of finite amplitude; Part 1. The elementary interactions. *J. Fluid Mech.*, **9**, 193–217.

——, 1969: *The Dynamics of the Upper Ocean*. Cambridge University Press, 261 pp.

Stokes, G. G., 1847: On the theory of oscillatory waves. *Trans. Cambridge Phil. Soc.*, **8**, 441–445.

Stoker, J. J., 1957: *Water Waves*. Interscience, 567 pp.

Tick, L. J., 1959: A nonlinear random model of gravity waves, I. *J. Math. Mech.*, **8**, 643–657.

Willebrand, J., 1975: Energy transport in a nonlinear and inhomogeneous random gravity wave field. *J. Fluid Mech.*, **70**, 113–126.

Whitham, G. B., 1967: Variational methods and applications to water waves, *Proc. Roy. Soc. London*, **A299**, 6–25.

——, 1974: Dispersive waves and variational principles. *Nonlinear Waves*, S. Leibovich and A. R. Seebass, Eds., Cornell University Press, Chap. V.