The Resolving Power of Gross Earth Data

George Backus and Freeman Gilbert

Summary

A gross Earth datum is a single measurable number describing some property of the whole Earth, such as mass, moment of interia, or the frequency of oscillation of some identified elastic-gravitational normal mode. We show how to determine whether a given finite set of gross Earth data can be used to specify an Earth structure uniquely except for fine-scale detail; and how to determine the shortest length scale which the given data can resolve at any particular depth. We apply the general theory to the linear problem of finding the depth-variation of a frequency-independent local Q from the observed quality factors Q of a finite number of normal modes. We also apply the theory to the non-linear problem of finding density vs depth from the total mass, moment, and normal-mode frequencies, in case the compressional and shear velocities are known.

1. Introduction

In the present paper we describe and apply a method for finding the resolving power of a finite set of gross Earth data when these data are used to reconstruct the internal structure of the Earth. The discussion is a continuation of that begun in Backus & Gilbert (1967, hereafter called Inverse I) but the present paper is self-contained. We introduce one refinement of terminology, a distinction between gross Earth functionals and gross Earth data, described immediately below.

2. The nature of the inverse problem

An Earth model is a mathematical abstraction which suffices for some geophysical discussion. For example, in much of observational seismology an Earth model need consist only of four functions of position, the density ρ, the bulk modulus κ, the shear modulus μ, and the local quality factor Q (assumed independent of frequency). In geochemistry an Earth model would have to include the concentrations of various elements or compounds as functions of position and time. In geomagnetism, an Earth model would include electrical conductivity as a function of position. We will consider only Earth models which consist of a finite number of functions of the position vector $\mathbf{r}$, say $m_1(\mathbf{r}), ..., m_n(\mathbf{r})$. Such an ordered n-tuple of functions can be thought of as a single n-dimensional vector-valued function $\mathfrak{m}(\mathbf{r})$; we will call it an n-dimensional Earth model.

All the Earth models under consideration will constitute a subset of some linear space $\mathfrak{M}$. Certain members of this linear space will be clearly without geophysical interest. For example, if the space $\mathfrak{M}$ consists of all seismological Earth models $(\rho, \kappa, \mu, Q^{-1})$, then those models for which ρ or κ or μ or Q^{-1} is negative at some

169

locations are of no geophysical interest and are included only to make $\mathfrak{M}$ a linear space. The members of $\mathfrak{M}$ which are of geophysical interest will be defined by a collection $\mathscr{R}$ of inequalities ($\rho > 0$, $\kappa > 0$, $\mu \geqslant 0$, $Q^{-1} > 0$ in the seismological example) and will constitute a subset of $\mathfrak{M}$ which we denote by $(\mathfrak{M}, \mathscr{R})$. We call the members of this subset 'reasonable' Earth models.

Any rule g which assigns to each reasonable Earth model $\mathfrak{m}$ a real number $g(\mathfrak{m})$ will be called a gross Earth functional. Thus a gross Earth functional is simply a real valued function on $(\mathfrak{M}, \mathscr{R})$. If $\mathfrak{M}$ consists of seismological Earth models, examples of gross Earth functionals are total mass, moments of inertia, frequencies and decay rates of elastic-gravitational normal modes, rotational splitting parameters, and seismic travel times between particular sources and receivers.

The value which a particular gross Earth functional is observed to have for the real Earth will be called a gross Earth datum. Suppose we have chosen for study some particular finite collection $\mathscr{G}$ of gross Earth functionals $g_1, ..., g_N$, which may or may not be linear. We assume that their values have all been measured accurately for the real Earth. Errors in the measurements are an interesting and essential complication which we plan to treat in a later paper. Let $\gamma_1, ..., \gamma_N$ be the measured values which the functionals $g_1, ..., g_N$ in $\mathscr{G}$ are observed to take for the real Earth. We define a '$\mathscr{G}$-acceptable' Earth model to be any reasonable model $\mathfrak{m}$ (i.e. any member of $(\mathfrak{M}, \mathscr{R})$) which exactly satisfies the N equations $g_1(\mathfrak{m}) = \gamma_1, ..., g_N(\mathfrak{m}) = \gamma_N$. The collection of all $\mathscr{G}$-acceptable Earth models we denote by $(\mathfrak{M}, \mathscr{R}, \mathscr{G})$. Since $\mathscr{G}$ is necessarily finite, it is intuitively reasonable that $(\mathfrak{M}, \mathscr{R}, \mathscr{G})$ will be 'very' infinite. In Inverse I we showed that if $\mathfrak{M}$ consists of all seismological Earth models ($\rho, \kappa, \mu, Q^{-1}$) and $\mathscr{G}$ is finite, then under very general assumptions $(\mathfrak{M}, \mathscr{R}, \mathscr{G})$ is either empty or an infinite-dimensional (usually curved) submanifold of $\mathfrak{M}$. The argument is the same for other types of Earth models.

As we accumulate measurements of more gross Earth data, the infinite-dimensional manifold of acceptable Earth models shrinks. If there are more data, the are fewer Earth models which can satisfy them all. Formally, if $\mathscr{G}_1 \subseteq \mathscr{G}_2$ (here $\subseteq$ means set-theoretic inclusion) then $(\mathfrak{M}, \mathscr{R}, \mathscr{G}_2) \subseteq (\mathfrak{M}, \mathscr{R}, \mathscr{G}_1)$. But since the larger collection $\mathscr{G}_2$ of measured gross Earth functionals is still finite, the smaller manifold $(\mathfrak{M}, \mathscr{R}, \mathscr{G}_2)$ of acceptable Earth models is still infinite-dimensional. We can never evade the issue of studying infinite dimensional manifolds of acceptable Earth models.

It follows that the 'inverse problem' for a given finite collection $\mathscr{G}$ of gross Earth functionals consists in describing the infinite-dimensional manifold $(\mathfrak{M}, \mathscr{R}, \mathscr{G})$ of all $\mathscr{G}$-acceptable Earth models, i.e. all models $\mathfrak{m}$ in $(\mathfrak{M}, \mathscr{R})$ to which the gross Earth functionals $g_1, ..., g_N$ in $\mathscr{G}$ assign the observed values $\gamma_1, ..., \gamma_N$. In Inverse I we showed how this formulation of the inverse problem makes that problem well-posed from the point of view of a digital computer, and enables the computer to generate particular $\mathscr{G}$-acceptable Earth models. We also discussed briefly the problem of exploring the whole manifold $(\mathfrak{M}, \mathscr{R}, \mathscr{G})$, and pointed out that one of the most interesting questions to a geophysicist is whether the different members of $(\mathfrak{M}, \mathscr{R}, \mathscr{G})$ differ only in their fine-scale detail, i.e. whether the infinite-dimensionality of $(\mathfrak{M}, \mathscr{R}, \mathscr{G})$ arises simply from the fact that a finite number of measurements gives us only a finite resolving power. Obviously the answer to this question will depend on $\mathscr{G}$. In Inverse II we gave some examples of qualitatively different Earth models which fit the normal mode data within the errors of observation. In the present paper we describe and apply a general method for discovering whether the observed values of a given finite set $\mathscr{G}$ of gross Earth functionals determine a 'reasonable' Earth model $\mathfrak{m}$ which is unique except for fine-scale resolution. When such a model is determined, we show how to estimate the resolving length of the data at any depth in the Earth.

As in Inverse I and II, we consider only gross Earth functionals which are Fréchet differentiable (Dunford & Schwartz 1958) with respect to some convex norm on $\mathfrak{M}$,

and we assume that the Earth models are spherical. That is, the functions of position which are to be determined from the data depend only on the distance r from the centre of the Earth. The modifications required to treat non-spherical Earth models will become obvious as the discussion proceeds. Usually it will suffice to 'correct' the observed data for the effects of small asphericities (including rotation), as discussed in Inverse I, Dahlen (1968) and Backus & Gilbert (1961).

One of the main applications of the present discussion is to the question of the variation of density ρ with depth in the Earth. If the gross Earth functionals in $\mathscr{G}$ are frequencies of normal modes, and if we know the compressional velocity $v_P(r)$ and the shear velocity $v_S(r)$, then we can use for the space of Earth models $\mathfrak{M}$ simply the space of functions $\rho(r)$, while the set $\mathscr{R}$ of inequalities which define 'reasonable' Earth models consists of the single inequality $\rho \geqslant 0$. Most geophysicists would accept as a second member of $\mathscr{R}$ the inequality $d\rho/dr \leqslant 0$, or even perhaps the adiabatic fluid stability condition

$$\frac{d\rho}{dr} + \frac{3g(r)}{3v_P{}^2 - 4v_S{}^2} \leqslant 0$$

where $g(r)$ is the local acceleration of gravity. These last two inequalities are by no means laws of nature. The mantle might permit $d\rho/dr > 0$ if it has finite strength or very high viscosity or a certain type of convective regime.

If we want to use the normal mode data to determine v_P and v_S as well as ρ, then we can take for $\mathfrak{M}$ the linear space of function triples (ρ, v_P, v_S). The set $\mathscr{R}$ becomes $\rho \geqslant 0$, $0 \leqslant 2v_S \leqslant 3^{\frac{1}{2}} v_P$, and perhaps one of the inequalities on $d\rho/dr$.

The inequalities $\mathscr{R}$ that define reasonable Earth models play no role in the discussion until after all the gross Earth data have been used. Then those inequalities serve to eliminate certain otherwise $\mathscr{G}$-acceptable Earth models. Therefore in what follows we will usually make no explicit mention of $\mathscr{R}$, and will write the manifold of $\mathscr{G}$-acceptable Earth models as $(\mathfrak{M}, \mathscr{G})$.

3. Linear gross Earth functionals

The simplest inverse problems are those for which all the gross Earth functionals in $\mathscr{G}$ are *linear* functionals on $\mathfrak{M}$. We may assume without loss of generality that the members of $\mathscr{G}$ are linearly independent, for if one of them is a linear combination of the others, say

$$g_{N+1} = \sum_{i=1}^{N} a_i g_i,$$

then the observations must satisfy

$$\gamma_{N+1} = \sum_{i=1}^{N} a_i \gamma_i.$$

Otherwise there are no $\mathscr{G}$-acceptable Earth models. But then obviously γ_{N+1} gives no information about the Earth which is not already contained in $\gamma_1, ..., \gamma_N$, so we may delete g_{N+1} from $\mathscr{G}$ with no loss of information. When the set $\mathscr{G}$ consists of N linearly independent linear functionals $g_1, ..., g_N$ on $\mathfrak{M}$, then the manifold $(\mathfrak{M}, \mathscr{G})$ of $\mathscr{G}$-acceptable Earth models is an affine hyperplane in $\mathfrak{M}$ with codimension N (Kato 1966).

To simplify this first discussion, we assume that the Earth models in $\mathfrak{M}$ are single functions m of radial distance r. In Section 4 we discuss the modifications required to handle n-dimensional Earth models $\mathbf{m}(r) = (m_1(r), ..., m_n(r))$.

Fréchet differentiability and linearity imply that each of the linear functionals $g_1, ..., g_N$ in $\mathcal{G}$ has the form

$$g_i(m) = \int_0^1 G_i(r)\, m(r)\, dr \qquad (3.1)$$

where $G_i(r)$ is a known function of r because g_i is a known linear functional. (Here and throughout this paper we take the radius of the Earth to be 1.) We denote the measured values of $g_1, ..., g_N$ for the real Earth by $\gamma_1, ..., \gamma_N$. Thus the N equations which determine $(\mathfrak{M}, \mathcal{G})$, the set of $\mathcal{G}$-acceptable Earth models, are simply

$$\gamma_i = \int_0^1 G_i(r)\, m(r)\, dr, \quad i = 1, ..., N. \qquad (3.2)$$

The numbers $g_i(m)$ can be thought of as generalized moments of the function m. Our problem is, what can we say about a function m in $\mathfrak{M}$ when all we know about it is that its N moments (3.1) have the values (3.2) (and that it satisfies the 'reasonable' inequalities $\mathcal{R}$).

A partial answer to this question was given in Inverse I. Let $\mathcal{G}_\perp$ be the set functions $\tilde{m}$ in $\mathfrak{M}$ such that $g_1(\tilde{m}) = ... = g_N(\tilde{m}) = 0$. Then $\mathcal{G}_\perp$ is an infinite-dimensional linear subspace of $\mathfrak{M}$. If m_0 is any $\mathcal{G}$-acceptable Earth model, then the manifold $(\mathfrak{M}, \mathcal{G})$ of $\mathcal{G}$-acceptable Earth models consists precisely of the functions $m_0(r) + \tilde{m}(r)$ where $\tilde{m}$ is any member of $\mathcal{G}_\perp$. Certain of these models will be eliminated by the inequalities $\mathcal{R}$.

It is difficult to visualize an infinite-dimensional function space, like $\mathcal{G}_\perp$, so we seek some way of describing all the functions in $(\mathfrak{M}, \mathcal{G})$ in terms of their appearance as functions of r. Given any particular point r_0 between 0 and 1, we try to compute from the observed data $\gamma_1, ..., \gamma_N$ a number $\langle m \rangle_{r_0}$ which is in some sense an average of the values of m in a short interval containing r_0. If such a local average can be obtained from $\gamma_1, ..., \gamma_N$ alone, it will be the same for all $\mathcal{G}$-acceptable Earth models m, and its value is a positive statement about the Earth to which we can attach as much confidence as we attach to the observed data. From a finite number of moments of m the best average we can hope to obtain near r_0 will be a weighted average which (we hope) heavily emphasizes points close to r_0 and gives very little weight to distant points. In principle we cannot hope to do better. The length of the interval around r_0 that contains the heavily weighted values of m can be thought of as the resolving length of the data near r_0, or the resolving length of the average $\langle m \rangle_{r_0}$. Clearly we seek that average, calculable from the moments $g_1(m), ..., g_N(m)$, which at r_0 has the shortest possible resolving length compatible with good suppression of contributions from distant r. Our problem is to quantify this qualitative description of a desirable local averaging process.

We consider only averages of m which depend linearly on m, so that the root-mean-square, for example, is excluded. A linear average of m can always be written in the form

$$\langle m \rangle_{r_0} = \int_0^1 A(r_0, r)\, m(r)\, dr \qquad (3.3)$$

where the averaging kernel A is 'unimodular'; that is,

$$\int_0^1 A(r_0, r)\, dr = 1. \qquad (3.4)$$

The averaging kernel A may be a distribution in r rather than an ordinary function. Ideally, we would like to evaluate equation (3.3) with $A(r_0, r) = \delta(r - r_0)$ where δ is the Dirac delta distribution. With only a finite number of moments $g_1(m), ..., g_N(m)$ available for computing the local average of m near r_0, we cannot hope to obtain weighting functions A which are so localized.

We are assuming that our average $\langle m \rangle_{r_0}$, whatever it turns out to be, depends only on $g_1(m), ..., g_N(m)$. Since $\langle m \rangle_{r_0}$ and $g_1(m), ..., g_N(m)$ all depend linearly on the function m, it follows rigorously (as proved in Appendix A) that $\langle m \rangle_{r_0}$ must depend linearly on $g_1(m), ..., g_N(m)$. Therefore there exist constants $a_1(r_0), ..., a_N(r_0)$, depending on the fixed point r_0, such that

$$\langle m \rangle_{r_0} = \sum_{i=1}^{N} a_i(r_0) g_i(m) \tag{3.5}$$

for any Earth model m, while for any $\mathcal{G}$-acceptable Earth model, including the real Earth,

$$\langle m \rangle_{r_0} = \sum_{i=1}^{N} a_i(r_0) \gamma_i. \tag{3.6}$$

From equation (3.1) we see that $\langle m \rangle_{r_0}$ has the form (3.3) with

$$A(r_0, r) = \sum_{i=1}^{N} a_i(r_0) G_i(r). \tag{3.7}$$

In short, the weighted averages of m which we can compute from the values of the moments $g_1(m), ..., g_N(m)$ consist precisely of the linear combinations (3.5); and the weight function or averaging kernel for any such average is equation (3.7). Then our problem is to choose the constants $a_1(r_0), ..., a_N(r_0)$ in equation (3.7) so that $A(r_0, r)$ satisfies equation (3.4) and resembles $\delta(r - r_0)$ as closely as possible. Among all unimodular linear combinations of the known kernels $G_1(r), ..., G_N(r)$, which one is most nearly $\delta(r - r_0)$?

If the kernels $G_1, ..., G_N$ are unfortunately chosen, none of their unimodular linear combinations will resemble $\delta(r - r_0)$. An example is $G_i(r) = \sin \pi i r$, $i = 3, 4, ..., N+2$. No linear combination of these kernels is a unimodular function $A(r_0, r)$ which, when used in equation (3.3), gives a useful estimate of a local average value of $m(r) = \sin \pi r$. However, if $G_i(r) = \sin \pi i r$ with $i = 1, 2, ..., N$, then the choice

$$a_i(r_0) = \frac{2}{\pi} \sin \pi i r_0$$

gives for $A(r_0, r)$ in equation (3.7) simply the Dirichlet kernel for the Fourier sine series:

$$A_N(r_0, r) = S_N(r - r_0) + S_N(r + r_0)$$

where

$$S_N(2x) = \frac{\sin \pi (2N + 1) x}{2 \pi \sin \pi x}.$$

Either simple inspection or the theory of Fourier sine series assures us that for large N the foregoing Dirichlet kernel is a good approximation to $\delta(r - r_0)$. The width of the peak in $A_N(r_0, r)$ for r near r_0 is approximately $2N^{-1}$.

In a real geophysical problem the kernels $G_i(r)$ in equation (3.1) usually will be generated numerically. There will be no analogue of the theory of Fourier series available to suggest how we ought to pick the coefficients $a_i(r_0)$ in equation (3.7) in order to make $A(r_0, r)$ a good unimodular approximation to $\delta(r - r_0)$, that is to give $A(r_0, r)$ a high, narrow peak near r_0 and very small values elsewhere (small

'sidebands' in the terminology of filter theory). We need a real-valued functional which measures the 'δ-ness' of arbitrary unimodular functions $A(r)$. Many such 'δ-ness criteria' are available, so we are free to choose one which facilitates numerical computation.

We pick a function $J(r_0, r)$ which vanishes when $r = r_0$ and increases monotonically as r increases or decreases away from r_0, but otherwise is arbitrary. Examples are

$$J(r_0, r) = (r - r_0)^2$$

and

$$J_\sigma(r_0, r) = 2\sigma^2\{1 - \exp[-(r - r_0)^2/2\sigma^2]\} \qquad (3.8)$$

where 2σ is a positive number which measures the width of the trough in J_σ when r is near r_0. Then for any function $A(r)$ we can define

$$\Delta_J(A) = \int\limits_0^1 J(r_0, r) A(r)^2 \, dr. \qquad (3.9)$$

Among all unimodular functions A, those which have smaller values of $\Delta_J(A)$ will be more concentrated toward r_0, and will tend to have higher, narrower peaks at r_0 and smaller magnitudes away from r_0. For the linear combinations of the form equation (3.7), that is, of the form

$$A(r) = \sum_{i=1}^N a_i \, G_i(r),$$

$\Delta_J(A)$ is a positive-definite quadratic function of $a_1, ..., a_N$. If we choose the constants a_i so as to minimize $\Delta_J(A)$ subject to the linear constraint (3.4), we obtain that unimodular linear combination of the known kernels $G_1(r), ..., G_N(r)$ which most nearly resembles $\delta(r - r_0)$ when resemblance is measured by the δ-ness criterion (3.9) defined by our particular choice of the function J.

The coefficients $a_1, ..., a_N$ in equation (3.7) which minimize equation (3.9) subject to equation (3.4) clearly satisfy the $N + 1$ equations

$$\left.\begin{aligned}
\sum_{j=1}^N \left[\int\limits_0^1 J(r_0, r) G_i(r) G_j(r) \, dr\right] a_j + \left[\int\limits_0^1 G_i(r) \, dr\right] \lambda = 0 \\
\sum_{j=1}^N \left[\int\limits_0^1 G_j(r) \, dr\right] a_j = 1
\end{aligned}\right\} \qquad (3.10)$$

where the $(N+1)$'st unknown, λ, is a Lagrange multiplier.

Our numerical procedure is to evaluate the coefficients in (3.10) by numerical integration and to solve equation (3.10) for $a_1, ..., a_N$ and λ. Then we substitute $a_1, ..., a_N$ into equation (3.7) and inspect the $A(r_0, r)$ which results. If this $A(r, r)$ resembles $\delta(r - r_0)$ in having a high, narrow peak near r_0 and being of very small magnitude elsewhere, then the integral (3.3) can be thought of as an average of the values of m near r_0; and the value of that average for any $\mathcal{G}$-acceptable Earth model is equation (3.6) with the same coefficients $a_i(r_0)$ just used to compute $A(r_0, r)$. The width of the peak in $A(r_0, r)$ measures the resolving power of the data, $\gamma_1, ..., \gamma_N$, near r_0.

If the solution $A(r_0, r)$ obtained from the foregoing calculation does not resemble a blurred delta function, and if $J(r_0, r)$ does not become nearly constant as $|r - r_0|$ increases, then we can conclude that none of the linear combinations of $G_1, ..., G_N$ resembles $\delta(r - r_0)$ and that from the given data it is not possible to obtain an

approximation to an average of m near r_0. In the case of a pathological J, one which is nearly constant except for a very narrow trough of width $2\sigma \ll 1$ at r_0, this negative conclusion may not be valid. If $G_1, ..., G_N$ have unimodular linear combinations which resemble $\delta(r-r_0)$ but if all such linear combinations have a peak width considerably greater than 2σ, then *vis a vis* the linear combinations of $G_1, ..., G_N$ the pathological function J behaves almost as if it were constant. It does a very poor job of selecting δ-like unimodular linear combinations of $G_1, ..., G_N$. For such pathological J's, failure of the calculation (3.10) to produce a good approximation to $\delta(r-r_0)$ from among the linear combinations of $G_1, ..., G_N$ does not mean that no such linear combination exists.

4. Other criteria of δ-ness

Many variations of the δ-ness criterion suggest themselves. Subject to equations (3.4) and (3.7) we might try to minimize

$$\int_0^1 J(r_0, r) \, |A(r)| \, dr,$$

or we might try to minimize

$$\Delta_L(A) = \max \{|A(r)|: 0 \leqslant r \leqslant r_0 - L \text{ or } r_0 + L \leqslant r \leqslant 1\}$$

and then vary L so as to produce the narrowest peak in A compatible with an acceptable sideband level. Or, subject to equations (3.4), (3.7) and the constraint that $A(r) \geqslant 0$ for all r, we might try to minimize

$$\int_0^1 J(r_0, r) \, A(r) \, dr.$$

In short, the whole technology of filter design is at our disposal, but the filter windows are unimodular linear combinations of the given kernels $G_1, ..., G_N$ rather than Nth degree trigonometric polynomials.

If we have used a non-pathological $J(r_0, r)$, such as

$$J(r_0, r) = (r-r_0)^2, \tag{4.1}$$

in our δ-ness criterion (3.9), and if the $A(r_0, r)$ we obtain resembles $\delta(r-r_0)$, then we can take $\langle m \rangle_{r_0}$, calculated from equation (3.5), as an estimate of $m(r_0)$, with the understanding that $\langle m \rangle_{r_0}$ is really an average, with resolving power given by the width of the peak in $A(r_0, r)$. Then as a function of r, $\langle m \rangle_r$ is an estimate of $m(r)$. It may turn out that $\langle m \rangle_r$ is very large for some r and very small for others. In that case, sideband contamination can very seriously affect the usefulness of $\langle m \rangle_{r_0}$ as an estimate of $m(r_0)$ when the latter is small. To get a local average of m near a point r_0 where m is small we would prefer an averaging kernel $A(r_0, r)$ which was very small at those r where $m(r)$ is large, even at the expense of accepting a wider peak near r_0.

To deal with this question, suppose that $m_1(r)$ is an estimate of $m(r)$, which may or may not be $\langle m \rangle_r$ obtained from the δ-ness criterion (3.9) with a non-pathological J. If $|m_1(r)|$ has very large variations, we can use

$$J(r_0, r) = (r-r_0)^2 \int_0^1 m_1(r)^2 \, dr + [m_1(r) - m_1(r_0)]^2 \tag{4.2}$$

or some similar function to generate a new δ-ness criterion (3.9). This new criterion penalizes $A(r)$ for being large where $m_1(r)$ is large; it will give local averages $\langle m \rangle_r$

which may be regarded as a second estimate $m_2(r)$ for $m(r)$. If $m_2(r)$ is very different from $m_1(r)$ we can repeat the whole process with m_2 instead of m_1 in equation (4.2).

Different δ-ness criteria will produce different optimal unimodular linear combinations (3.7), but if somehow we manage to generate coefficients $a_1, \ldots, a_N$ such that

$$\sum_{i=1}^{N} a_i\, G_i(r)$$

has a high, narrow peak at r_0 and very small magnitude elsewhere, then the procedure by which we obtained $a_1, \ldots, a_N$ is irrelevant to the fact that now we have a linear combination

$$\sum_{i=1}^{N} a_i\, \gamma_i$$

of the observations $\gamma_1, \ldots, \gamma_N$ which is approximately an average of m in an interval of known length about r_0. And inspection of the averaging kernel

$$\sum_{i=1}^{N} a_i\, G_i(r),$$

without reference to any theory, will give us this resolution length or resolving power, as well as the height and weight of the sidebands.

So far we have considered only one-dimensional Earth models. We are forced to modify our δ-ness criterion in discussing the resolving power of gross Earth functionals when the Earth models are n-dimensional. Let $\mathfrak{M}$ now be a linear space of Earth models which are ordered n-tuples of real functions of the radius r, say $\mathbf{m}(r) = \big(m_1(r), \ldots, m_n(r)\big)$. An example with $n = 3$ could be the seismological models $\big(\rho(r), v_P(r), v_S(r)\big)$. In such spaces, a Fréchet differentiable, linear gross Earth functional g_i has the form

$$g_i(\mathbf{m}) = \sum_{\nu=1}^{n} \int_0^1 G_{i\nu}(r)\, m_\nu(r)\, dr, \qquad (4.3)$$

where the kernels $G_{i\nu}(r)$ are known. Now the problem is to determine, for specified μ and r_0, an estimate $\langle \mathbf{m} \rangle_{\mu r_0}$ of the average value of $m_\mu(r)$ when r is near r_0; and to use in this determination only the observed values $\gamma_1, \ldots, \gamma_N$ of the linear gross Earth functionals $g_1, \ldots, g_N$ in some finite set $\mathscr{G}$.

As with one-dimensional Earth models, we restrict attention to local averages $\langle \mathbf{m} \rangle_{\mu r_0}$ which depend linearly on $\mathbf{m}$. Since $g_1(\mathbf{m}), \ldots, g_N(\mathbf{m})$ also depend linearly on $\mathbf{m}$ and since $\langle \mathbf{m} \rangle_{\mu r_0}$ is supposed to be calculated from $g_1(\mathbf{m}), \ldots, g_N(\mathbf{m})$ alone, it follows from Appendix A that the local average must depend linearly on the values of the gross Earth functionals. That is, there must be constants $a_{1\mu}(r_0), \ldots, a_{N\mu}(r_0)$ such that

$$\langle \mathbf{m} \rangle_{\mu r_0} = \sum_{i=1}^{N} a_{i\mu}(r_0)\, g_i(\mathbf{m}) \qquad (4.4)$$

for any Earth model $\mathbf{m}$ in $\mathfrak{M}$. But then equations (4.3) and (4.4) imply that

$$\langle \mathbf{m} \rangle_{\mu r_0} = \sum_{\nu=1}^{n} \int_0^1 A_{\mu\nu}(r_0, r)\, m_\nu(r)\, dr \qquad (4.5)$$

where

$$A_{\mu\nu}(r_0, r) = \sum_{i=1}^{N} a_{i\mu}(r_0)\, G_{i\nu}(r). \qquad (4.6)$$

In equation (4.6) μ and r_0 are fixed, so we have n functions of r, namely $A_{\mu 1}(r_0, r), \ldots, A_{\mu n}(r_0, r)$, which we abbreviate as $A_1(r), \ldots, A_n(r)$. For the given fixed μ and r_0 the abbreviated form of equation (4.5) is

$$\langle \mathfrak{m} \rangle_{\mu r_0} = \sum_{\nu=1}^{n} \int_0^1 A_\nu(r) \, m_\nu(r) \, dr, \tag{4.7}$$

and the abbreviated form of equation (4.6) is

$$A_\nu(r) = \sum_{i=1}^{N} a_i \, G_{i\nu}(r), \tag{4.8}$$

where a_i stands for $a_{i\mu}(r_0)$. In this abbreviated notation, if $\mathfrak{m}$ is a $\mathscr{G}$-acceptable Earth model, we have

$$\langle \mathfrak{m} \rangle_{\mu r_0} = \sum_{i=1}^{N} a_i \, \gamma_i. \tag{4.9}$$

Equations (4.7), (4.8) and (4.9) show us that from the data, $\gamma_1, \ldots, \gamma_N$, we can compute any weighted average of a $\mathscr{G}$-acceptable $\mathfrak{m}$ in which the averaging kernals have the form (4.8), and Appendix A assures us that these are the only weighted averages of $\mathfrak{m}$ which we can compute from $\gamma_1, \ldots, \gamma_N$ and which depend linearly on $\mathfrak{m}$. The problem is to find coefficients $a_1, \ldots, a_N$ which produce from the known kernels $G_{i\nu}(r)$ and equation (4.8) a set of averaging kernels $A_1(r), \ldots, A_n(r)$ such that equation (4.7) really does give a good estimate of the average value of $m_\mu(r)$ near r_0. We see that the estimate (4.7) will be contaminated not only by sideband contributions from $m_\mu(r)$ with r distant from r_0, but also by contributions from the other components of $\mathfrak{m}$. Ideally, we would like to be able to choose the constants $a_1, \ldots, a_N$ so that the $A_\nu(r)$ in equation (4.8) have the form

$$A_\nu(r) = \delta_{\mu\nu} \delta(r - r_0),$$

where $\delta_{\mu\nu}$ is the Kronecker delta. With only finitely many data about $\mathfrak{m}$ we cannot expect such good resolution.

In order to generate a good n-tuple of averaging kernels $\mathbf{A}(r) = (A_1(r), \ldots, A_n(r))$ in the form (4.8) we need a δ-ness criterion which tells us how closely an imperfect n-tuple of averaging kernels resembles the ideal n-tuple $(0, \ldots, 0, \delta(r - r_0), 0, \ldots, 0)$ which has the Dirac delta distribution in the μth position and zeros in the other $n-1$ positions.

One example of such a criterion is

$$\Delta_{J\mu}(\mathbf{A}) = \int_0^1 J(r_0, r)[A_\mu(r)]^2 \, dr + \sum_{\nu=1}^{n} (1 - \delta_{\nu\mu}) \int_0^1 [A_\nu(r)]^2 \, dr, \tag{4.10}$$

where $J(r_0, r)$ is a function of the sort used in equation (3.9). Then our numerical procedure for optimizing the coefficients $a_1, \ldots, a_N$ for a fixed μ and r_0 is to substitute equation (4.8) into equation (4.10) and to minimize $\Delta_{J\mu}(\mathbf{A})$ as a positive-definite quadratic function of $a_1, \ldots, a_N$, subject to the constraint

$$\int_0^1 A_\mu(r) \, dr = 1. \tag{4.11}$$

All of the remarks about δ-ness criteria for averaging of one-dimensional Earth models, including equation (4.2), then generalize to the n-dimensional models. In particular, if the different components $m_1(r), \ldots, m_n(r)$ have different physical

dimensions, we want to make them dimensionless and of order unity so that all the terms in the sum on the right in equation (4.10) have roughly the same order of magnitude.

5. Mantle dissipation

As an example of the analysis of linear gross Earth data for one-dimensional Earth models $m(r)$, we consider the problem of determining the dissipation at various depths in the mantle when the normal modes of elastic-gravitational oscillation of the Earth are excited. We take as gross Earth data the reciprocals Q^{-1} of the observed quality factors of some finite collection of normal modes. We assume that the Earth is isotropic and spherical and that its density ρ, bulk modulus κ, and shear modulus μ are known functions of the radial distance r. We also assume that the dissipation is a small perturbation ($Q \gg 1$) confined to the mantle, and that at each depth r the dissipation is described by a frequency-independent local Q, $Q(r)$. This means that the energy per gram dissipated in any local cycle of strain is independent of the speed with which the cycle is executed and is proportional to the square of the local strain amplitude. Then an Earth model is $m(r) = [Q(r)]^{-1}$ and a gross Earth datum is $\gamma_i = Q_i^{-1}$ for some particular normal mode, which we will call the ith mode. When $Q^{-1} = 0$, let the computed vector displacement field in a normal mode be $s(r, \theta, \lambda) e^{i\omega t}$, where r is radius, θ is colatitude, and λ is longitude. Let $2\sigma = \nabla s + (\nabla s)^T$ where T means transpose. Let $\mathbf{D}(r, \theta, \lambda) = \sigma - (\text{tr } \sigma)\mathbf{I}/3$ where $\mathbf{I}$ is the three-dimensional identity tensor and tr means trace. Define

$$\mathscr{D}(r)^2 = \frac{1}{4\pi} \int\limits_0^\pi d\theta \sin\theta \int\limits_0^{2\pi} d\lambda \, \text{tr}\,(D \cdot D^*)$$

where * means complex conjugate and the dot product of two second-order tensors is defined as in Gibbs & Wilson (1901). Define

$$\mathscr{S}(r)^2 = \frac{1}{4\pi} \int\limits_0^\pi d\theta \sin\theta \int\limits_0^{2\pi} d\lambda \, s \cdot s^*.$$

Then for the ith normal mode the gross Earth functional Q_i^{-1} has the form

$$g_i(m) = \int\limits_0^1 m(r) \, G_i(r) \, dr \tag{5.1}$$

where (Anderson & Archambeau 1964)

$$G_i(r) = \frac{r^2 \, \mu(r) \, \mathscr{D}_i(r)^2}{\omega_i^2 \int\limits_0^1 r^2 \rho(r) \, \mathscr{S}_i(r)^2 \, dr} . \tag{5.2}$$

In our numerical calculation of the kernels (5.2) we used the shear modulus and normal-mode wave functions obtained from model 5821 described in Appendix C. We considered sets $\mathscr{G}$ of 15 to 25 gross Earth functionals (values of Q^{-1} for normal modes). Our δ-ness criterion was equation (3.9) with J given by equation (3.8). We tried various trough-widths 2σ in equation (3.8), ranging from $2\sigma^2 = 1$ to $2\sigma^2 = 10^{-4}$. As expected from the number of functionals available, the optimal averaging kernel $A(r_0, r)$ depended only slightly on σ as long as $10^{-2} < 2\sigma^2 < 1$, but when $2\sigma^2$ became less than 10^{-2} the resemblance of $A(r_0, r)$ to $\delta(r - r_0)$ began to disappear.

From our kernels (5.2) we tried to obtain linear combinations $A(r_0, r)$ which were good approximations to $\delta(r-r_0)$ for $r_0 = 0.55$, 0.65. 0.75, 0.85, 0.94 and 0.98. We used five different sets of gross Earth functionals, described and labelled in Table 1. In Table 1, $_nT_l$ denotes a toroidal mode with angular order l and radial

Table 1

Name of set	Modes whose Q^{-1} constitute $\mathscr{G}$
$_n\mathscr{G}_l^S$	$_0S_0$, $_1S_0$, $_2S_0$, $_3S_0$, $_2S_1$, $_0S_2$, $_1S_2$, $_2S_2$, $_0S_3$, $_1S_3$, $_2S_3$, $_0S_4$, $_1S_4$, $_2S_4$, $_4S_4$, $_0S_5$, $_1S_5$, $_2S_5$, $_4S_5$, $_0S_6$, $_0S_7$, $_0S_8$, $_1S_8$, $_0S_9$, $_1S_{10}$
$_n\mathscr{G}_2^T$	$_nT_2$, $n = 0, 1, 2, ..., 14$
$_n\mathscr{G}_2^S$	$_nS_2$, $n = 0, 1, 2, ..., 14$
$_n\mathscr{G}_{12}^T$	$_nT_1$, $n = 1, 2, ..., 8$ and $_nT_2$, $n = 0, 1, 2, ..., 8$
$_n\mathscr{G}_{12}^S$	$_nS_1$, $n = 1, 2, ..., 8$ and $_nS_2$, $n = 0, 1, 2, ..., 8$

order n, while $_nS_l$ refers to the corresponding spheroidal mode. The gross Earth functionals in $_n\mathscr{G}_l^S$ were chosen as representative of data likely to become available for the real Earth. The other sets of gross Earth functionals were chosen to test the relative effectiveness of spheroidal and toroidal Q's in giving resolving power at depth, and to test whether kernels G_l with only one angular order l had linear combinations resembling δ functions, or whether kernels for two values of l were required.

For $_n\mathscr{G}_{12}^S$ and $_n\mathscr{G}_l^S$, Fig. 1 shows how the optimal averaging kernel $A(r_0, r)$ at $r_0 = 0.75$ varies with the trough width 2σ in the δ-ness criterion (3.8), (3.9). When $2\sigma^2 > 10^{-2}$, clearly the kernel is not heavily dependent on σ.

Figs 2, 3 and 4 show the kernels $A(r_0, r)$ obtained for various r_0 and for the five sets of gross Earth data listed in Table 1. The δ-ness criterion was equations (3.9) and (4.1). From these figures we see that $_n\mathscr{G}_l^S$, $_n\mathscr{G}_2^S$, $_n\mathscr{G}_{12}^T$ and $_n\mathscr{G}_{12}^S$ produce averaging kernels $A(r_0, r)$ which do resemble $\delta(r-r_0)$, but that the optimal $A(r_0, r)$ obtained from $_n\mathscr{G}_2^T$ can hardly be regarded as an approximation to a delta function. It follows that if we know Q^{-1} for the normal modes described in any of $_n\mathscr{G}_l^S$, $_n\mathscr{G}_2^S$, $_n\mathscr{G}_{12}^T$ or $_n\mathscr{G}_{12}^S$ we can obtain rather good estimates of the local average of $Q^{-1}(r)$ near various depths r_0, simply as linear combinations of the values of Q^{-1} in $\mathscr{G}$. But if we know only the values of Q^{-1} for the normal modes described in $_n\mathscr{G}_2^T$ we cannot obtain good averages of $Q^{-1}(r)$ near any depth other than $r_0 = 0.80$. (We believe we understand why this particular depth is exceptional, but the discussion is both complicated and speculative, so we omit it. However, see Appendix B and Fig. 10.)

For reasons given in Section 7 we conjecture that if Q^{-1} is known for all the toroidal modes with two different fixed angular orders l, or for all the spheroidal modes with a single angular order l, then $Q^{-1}(r)$ is uniquely determined; but that if Q^{-1} is known only for all toroidal modes with a single fixed angular order l then $Q^{-1}(r)$ is not determined.

From our numerical calculations, summarized in Fig. 4, we conclude that if values of Q^{-1} are available for normal modes with low radial orders but for many angular orders then we can obtain averages of $Q^{-1}(r)$ at various depths, but the thickness of the layer over which we average tends to increase with depth. This is the generally accepted view, but we believe that it has not previously received quantitative treatment. We note also that it is subject to one very important exception.
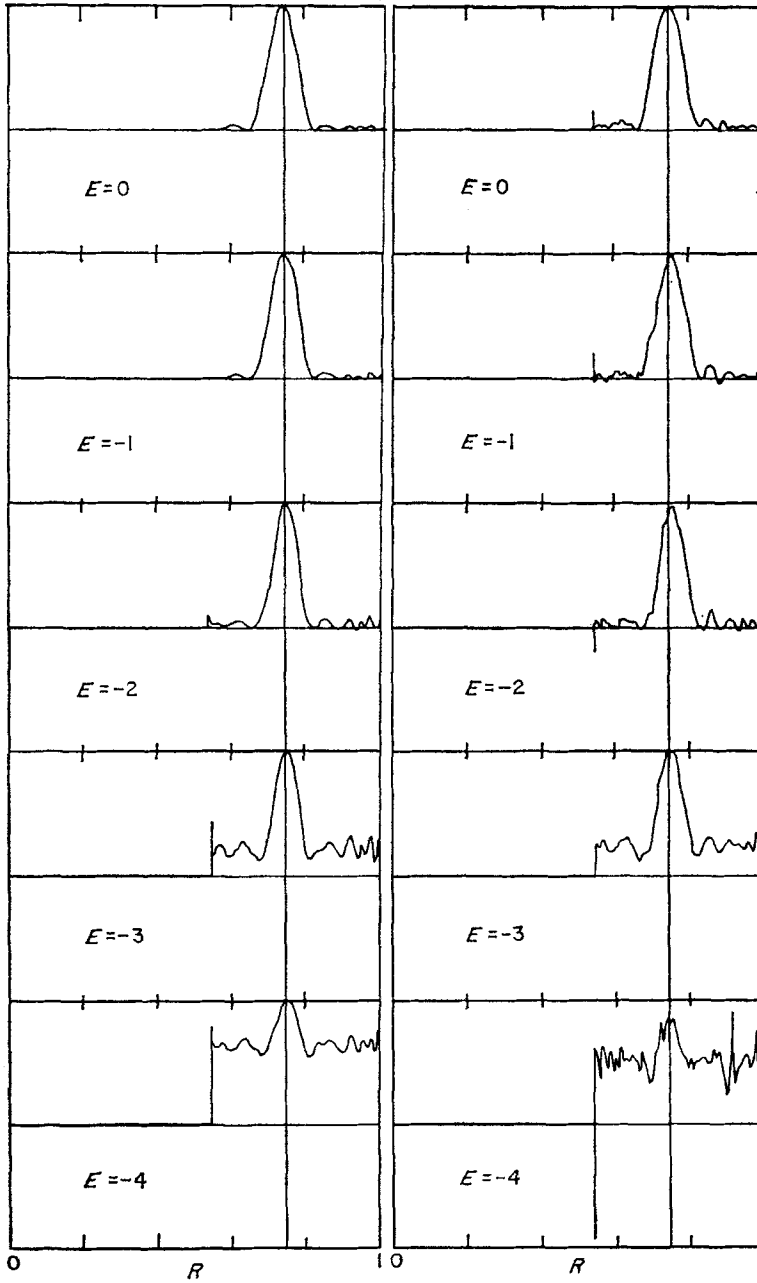
FIG. 1. The optimal averaging kernel $A(r_0, R)$ for $Q^{-1}(R)$ at $r_0 = 0\cdot75$ (fine vertical line) as a function of $2\sigma^2 = 10^E$. The δ-ness criterion is equations (3.8) and (3.9) and the gross Earth data are $_n\mathscr{G}_{12}{}^S$ on the left and $_n\mathscr{G}_1{}^S$ on the right.
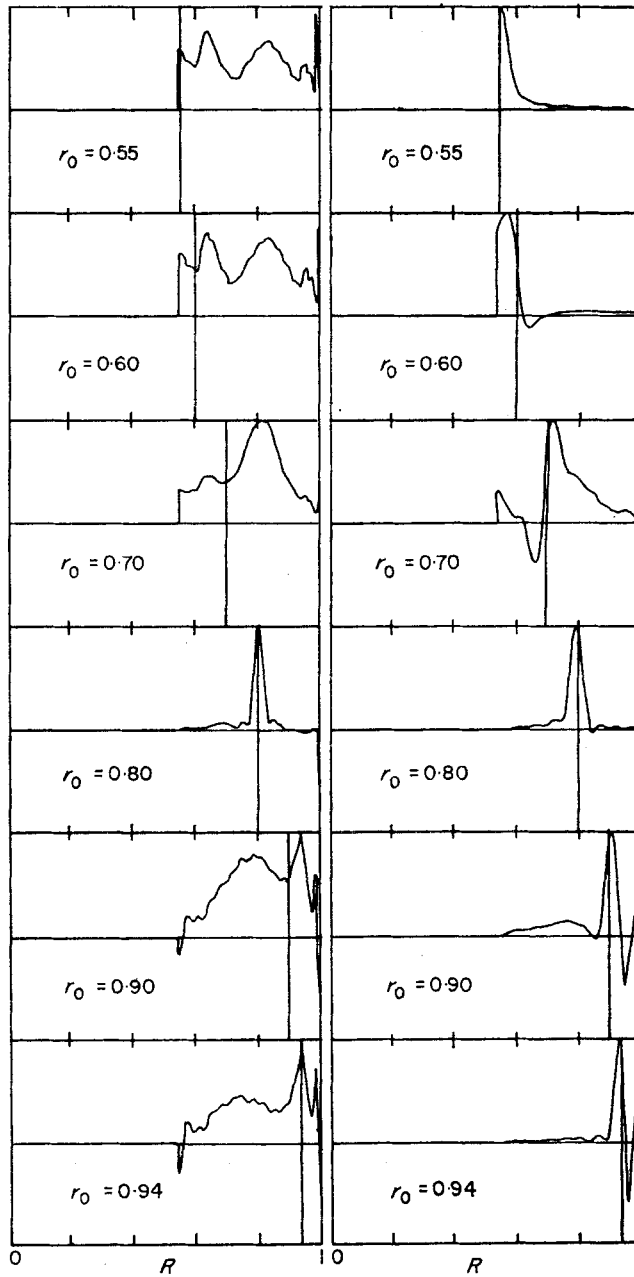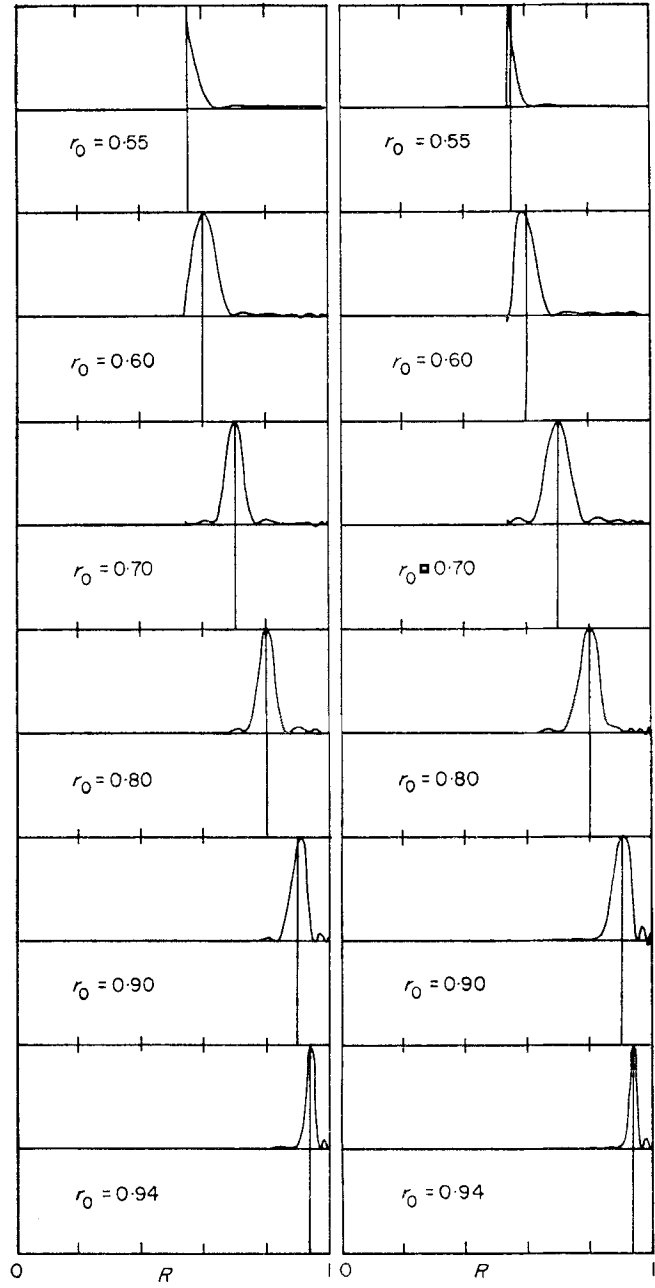
FIG. 2. The optimal averaging kernel $A(r_0, R)$ for $Q^{-1}(R)$ at various r_0 (the fine vertical lines). The δ-ness criterion is equations (3.9) and (4.1) and the gross Earth data are $_n\mathscr{G}_2{}^T$ on the left and $_n\mathscr{G}_{12}{}^T$ on the right.

FIG. 3. The same as Fig. 2 except that the gross Earth data are $_n\mathscr{G}_2{}^s$ on the left and $_n\mathscr{G}_{12}{}^s$ on the right.
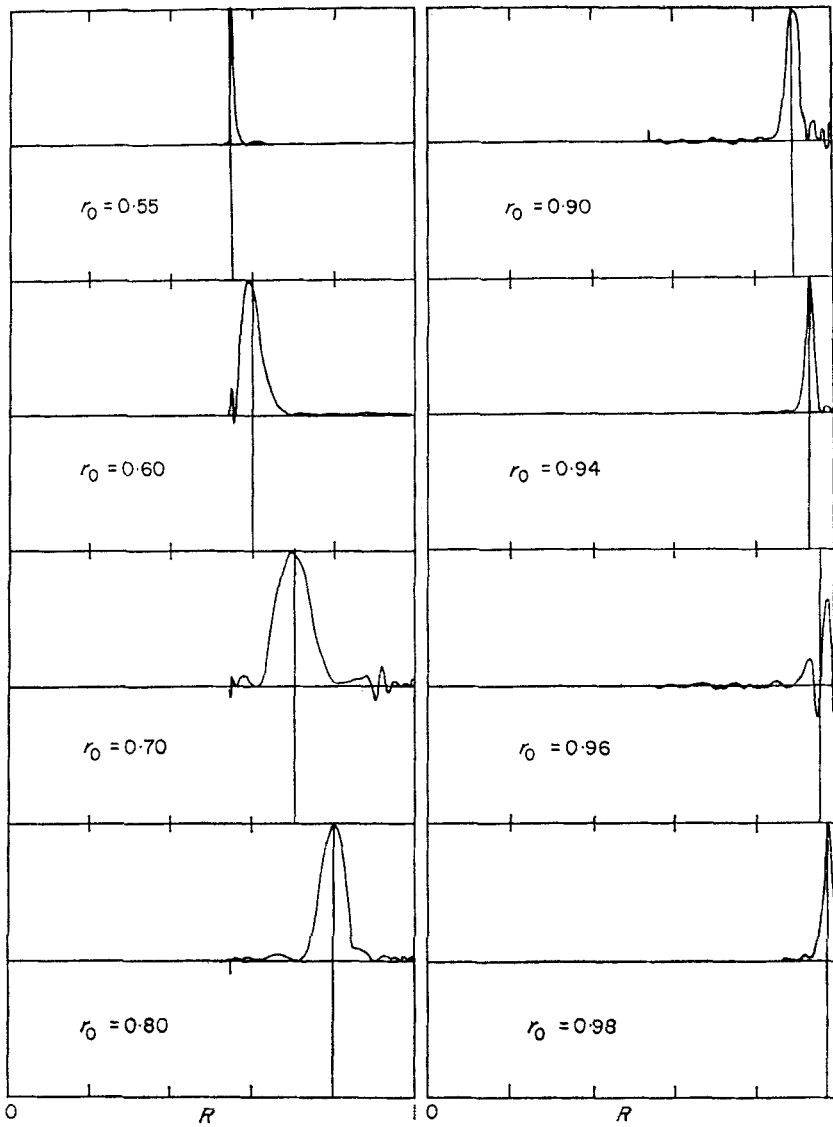
FIG. 4. The same as Fig. 2 except that the gross Earth data are $_n\mathscr{G}_i{}^s$.

For reasons which we do not now understand it is possible to get higher resolving power at the bottom of the mantle than at shallower depths.

A very interesting mathematical question is exactly which infinite sets of normal mode Q's are needed to determine $Q^{-1}(r)$. We emphasize, however, that infinite sets of experimental data are never available. The geophysical problem will always involve only finitely many gross Earth data, and we will never have infinite resolving power. From a practical point of view, what we need is a method of finding the resolving power of a given finite set of data. It seems most unlikely that this practical question will ever be answered analytically for arbitrary sets of gross Earth data; numerical calculations of some sort probably always will be required.

Throughout the discussion of Q in this section we have assumed that the local $Q(r)$ was independent of frequency. This assumption is not essential. The separately observable normal modes cover only about one decade of frequencies. In that

restricted frequency range the local Q might be well represented as a polynomial of low order in ω or $\ln \omega$, say

$$Q(r, \omega)^{-1} = m_1(r) + \omega m_2(r) + \omega^2 m_3(r)$$

or

$$Q(r, \omega)^{-1} = m_1(r) + (\ln \omega)\, m_2(r) + (\ln \omega)^2 m_3(r).$$

Then the space $\mathfrak{M}$ of Earth models would consist of n-tuples $\mathfrak{m}(r) = \big(m_1(r), ..., m_n(r)\big)$ where $n-1$ is the degree of the polynomial used to represent the frequency dependence of the local Q. The techniques outlined at the end of Section 4 would become appropriate.

The foregoing discussion shows what can be learned about local mantle dissipation from good data on the Q's of normal modes. According to Slichter (1967), 'The determination of the Q values of the free vibrations is still in an unsatisfactory state'. Slichter estimates that determinations of Q are probably in error by as much as 20–30 per cent. On the basis of the spin gaps (Gilbert & Backus 1965) and ellipticity gaps (Dahlen 1968) for the degenerate modes, we believe that even Slichter's error estimate is optimistic. Therefore in the present paper we do not propose any list of 'observed' Q's from which $Q^{-1}(r)$ can be estimated.

6. Non-linear gross Earth functionals

When the gross Earth functionals $g_1, ..., g_N$ in $\mathscr{G}$ are non-linear functionals on the linear space $\mathfrak{M}$ of Earth models, estimating the resolving power of the data is more complicated than for linear functionals. As in Section 3, we begin by considering spaces $\mathfrak{M}$ of one-dimensional spherical Earth models, single real-valued functions $m(r)$ of radial distance r. We suppose that somehow, for example as in Inverse I, we have found in $\mathfrak{M}$ a $\mathscr{G}$-acceptable Earth model m_0. Fréchet differentiability of the gross Earth functional g_i at m_0 implies that there is a function $G_i(r)$, depending on m_0, such that if m_1 is any other Earth model then

$$g_i(m_1) - g_i(m_0) = \int_0^1 G_i(r)[m_1(r) - m_0(r)] + O[(m_1 - m_0)^2]. \qquad (6.1)$$

Thus if $m_1 - m_0$ is sufficiently small then $g_i(m_1) - g_i(m_0)$ depends almost linearly on $m_1 - m_0$. Now we define f_i as this linear approximation to δg_i at m_0. That is, for any m in $\mathfrak{M}$,

$$f_i(m) = \int_0^1 G_i(r)\, m(r)\, dr. \qquad (6.2)$$

Then f_i depends on which $\mathscr{G}$-acceptable model m_0 we are using, and if m_0 is fixed then $f_1, ..., f_N$ are well-defined *linear* gross Earth functionals. Therefore we can proceed exactly as in Section 3.

We pick a particular point r_0 between 0 and 1, and a non-pathological δ-ness criterion (3.9), and we find the most δ-like averaging kernel (3.7), that is, the unimodular linear combination $A(r_0, r)$ of $G_1(r), ..., G_N(r)$ which most nearly resembles $\delta(r-r_0)$ when the resemblance is measured by our chosen δ-ness criterion. In a certain limited sense the averaging kernel $A(r_0, r)$ shows us what we can say about the behaviour of a $\mathscr{G}$-acceptable Earth model $m(r)$ when r is near r_0 if our only information about m is its $\mathscr{G}$-acceptability. Suppose that m_1 and m_0 are both $\mathscr{G}$-acceptable Earth models $\big(g_i(m_1) = g_i(m_0) = \gamma_i,\ i = 1, ..., N\big)$ and that m_1 differs only slightly from m_0 (that is, for $i = 1, ..., N$ the quadratic remainder term in

equation (6.1) is much smaller than the term which is linear in $m_1 - m_0$). Then to first order in $m_1 - m_0$ we have

$$\int_0^1 G_i(r)[m_1(r) - m_0(r)]\,dr = 0, \quad i = 1, \ldots, N. \qquad (6.3)$$

But if we multiply the ith of equations (6.3) by $a_i(r_0)$ and sum over i from 1 to N we obtain

$$\int_0^1 A(r_0, r)\,m_1(r)\,dr = \int_0^1 A(r_0, r)\,m_0(r)\,dr \qquad (6.4)$$

with an error of second order in $m_1 - m_0$.

In other words, if we define the average $\langle m \rangle_{r_0}$ by equation (3.3) with the averaging kernel (3.7) then all $\mathscr{G}$-acceptable Earth models m which differ only slightly from m_0 have the same average value near r_0. The resolving length (peak width) of the optimal averaging kernel (6.2) can be thought of as the resolving power of the data for depths r near r_0 when we are examining the $\mathscr{G}$-acceptable Earth models which in the linear space $\mathfrak{M}$ lie close to m_0.

If $m_0'(r)$ is another $\mathscr{G}$-acceptable Earth model which differs so much from $m_0(r)$ that we cannot neglect the quadratic term in the expression (6.1) for $g_i(m_0') - g_i(m_0)$, then m_0' will generate another set of kernels $G_i'(r)$, describing the Fréchet derivatives of the non-linear functionals g_i at m_0'. From these new kernels $G_i'(r)$ we will obtain as a unimodular linear combination a new optimal averaging kernel $A'(r_0, r)$. The average $\langle m' \rangle_{r_0}$ obtained from $A'(r_0, r)$ and m_0' will generally differ from the $\langle m \rangle_{r_0}$ obtained from $A(r_0, r)$, and the resolving lengths of $A(r_0, r)$ and $A'(r_0, r)$ will usually be different. These consequences of the non-linearity of the gross Earth functionals $g_1, \ldots, g_N$ are inherent in the problem.

When the Earth models $\mathfrak{m}$ in $\mathfrak{M}$ are n-tuples of functions of r and the gross Earth functionals in $\mathscr{G}$ are non-linear, the foregoing discussion can be generalized in an obvious way, using Section 4's linear theory for n-dimensional models instead of Section 3's linear theory for one-dimensional models.

7. The density structure in the Earth

As an illustration of the discussion in Section 6, we use the normal mode frequencies to infer the density ρ as a function of radial distance r, assuming that $v_P(r)$ and $v_S(r)$, the compressional and shear velocities, are known. (As remarked at the end of Section 6, we could let ρ, v_P and v_S all be unknown, but we confine ourselves here to the simpler case.) We use five sets of gross Earth functionals, $_n\mathscr{G}_1{}^S$, $_n\mathscr{G}_1{}^T$, $_n\mathscr{G}_2{}^S$, $_n\mathscr{G}_{12}{}^T$ and $_n\mathscr{G}_{12}{}^S$. The notation refers to Table 1 and any one of the foregoing sets of gross Earth functionals consists of the squared angular frequencies of the normal modes listed in Table 1, together with the Earth's total mass (Gutenberg 1959) and total moment of inertia (Jeffreys 1963; King-Hele et al. 1964).

When g_i is ω^2 for the ith normal mode, then $G_i(r)$ is the kernel which gives the first-order change $\delta\omega^2$ produced by a change $\delta\rho$ in the density model when v_P and v_S are kept fixed. This kernel is obtained immediately from equation (49) of Inverse I. In the notation of that paper it is

$$G_i(r) = \frac{4\pi r^2 [R_i' + v_S^2\,M_i' + (v_P^2 - \tfrac{4}{3}v_S^2)\,K_i']}{T_i} \qquad (7.1)$$

where

$$T_i = \int_0^\pi \sin\theta\,d\theta \int_0^{2\pi} d\lambda \int_0^1 r^2\,\rho(r)\,\mathbf{s}_i^2\,dr$$

and where R_i', M_i', K_i' are calculated from the displacement field s_i of the ith normal mode.

In our numerical calculations the functions $v_P(r)$, $v_S(r)$ and the $\rho_0(r)$ which defines the $\mathscr{G}$-acceptable Earth model m_0 were taken from model 5821, described in Appendix C. The δ-ness criterion was equations (3.9) and (4.1), as in Section 5. Figs 5 and 6 give the optimal averaging kernels $A(r_0, r)$ for several values of r_0, which are obtained from $_n\mathscr{G}_2{}^T$, $_n\mathscr{G}_{12}{}^T$, $_n\mathscr{G}_2{}^S$ and $_n\mathscr{G}_{12}{}^S$. Fig. 7 gives the optimal averaging kernels $A(r_0, r)$ obtained from $_n\mathscr{G}_i{}^S$ for $r_0 = 0\cdot1$, $0\cdot2$, $0\cdot3$, $0\cdot4$, $0\cdot5$, $0\cdot54$, $0\cdot55$, $0\cdot6$, $0\cdot7$, $0\cdot8$, $0\cdot94$, $0\cdot98$ together with estimates of the resolving length $L(r_0)$, defined as

$$L(r_0) = \frac{2\int_0^1 |r - r_0|\,|A(r_0, r)|\,dr}{\int_0^1 |A(r_0, r)|\,dr} \tag{7.2}$$

Fig. 8 shows $\rho_0(r)$ together with the local averages

$$\langle \rho_0 \rangle_{r_0} = \int_0^1 \rho_0(r)\,A(r_0, r)\,dr \tag{7.3}$$

for each of the depths $r_0 = 0\cdot05$, $0\cdot10$, $0\cdot15$, ..., $0\cdot85$, $0\cdot90$ and $0\cdot94$ and $0\cdot98$.

The model $\rho_0(r)$ fits all the available spheroidal normal mode frequencies within experimental error. As pointed out in Inverse I, there are many such models, so we cannot say the real Earth must look like $\rho_0(r)$. We can say, however, that if the data are correct and if the real Earth's $\rho(r)$ differs from $\rho_0(r)$ by an amount small enough to permit neglecting the quadratic term in equation (6·1), and if $v_P(r)$, $v_S(r)$ of model 5821 are correct, then the averages (7·3) computed from the real Earth's $\rho(r)$ must agree with the averages (7·3) computed from the model $\rho_0(r)$, at least correct to first order in $(\rho - \rho_0)/\rho_0$, because these averages are the same for all acceptable Earth models close to ρ_0 in $\mathfrak{M}$. The decrease in the weighted averages $\langle \rho \rangle_{r_0}$ for $r_0 = 0\cdot05$, $0\cdot10$ and $0\cdot15$ is obviously a result of the increasingly poor resolution as r_0 approaches 0 (see Fig. 7) and not a property of the function $\rho(r)$.

We note one very important characteristic of $A(r_0, r)$. When r_0 is close to the core-mantle boundary, r_c, the resolution is very high indeed. The integral (3.3) is very nearly

$$\rho(r_c-)\,P_-(r_0) + \rho(r_c+)\,P_+(r_0)$$

where

$$P_-(r_0) = \int_0^{r_c} A(r_0, r)\,dr,$$

$$P_+(r_0) = \int_{r_c}^1 A(r_0, r)\,dr.$$

By giving r_0 two different values near r_c we obtain two different linear combinations of $\rho(r_c-)$ and $\rho(r_c+)$. the densities of core and mantle at the core mantle boundary. We can solve these equations for the two densities separately. If we use $r_0 = 0\cdot54$ and $r_0 = 0\cdot55$ we obtain

$$\rho(r_c-) = 10\cdot04 \pm 0\cdot05 \text{ g/cm}^3.$$

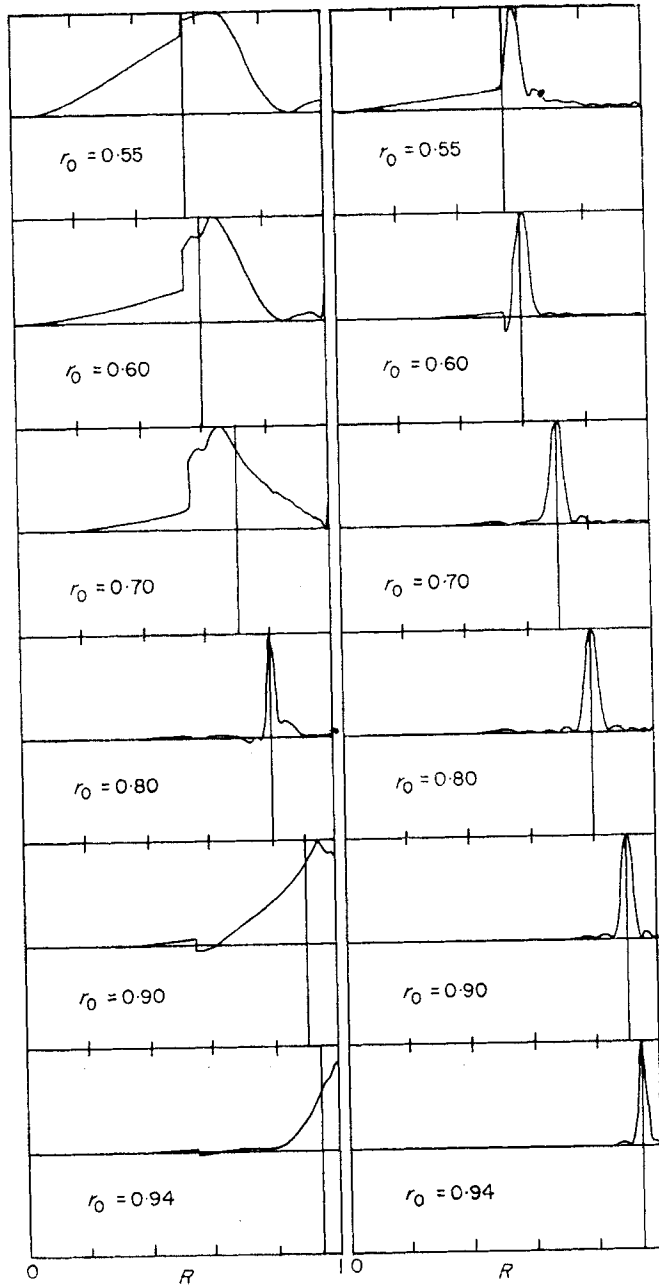$$\rho(r_c+) = 5\cdot72 \pm 0\cdot03 \text{ g/cm}^3.$$

FIG. 5. The optimal averaging kernel $A(r_0, R)$ for $\rho(R)$ at various r_0 (the fine vertical lines). The σ-ness criterion is equations (3.9) and (4.1) and the gross Earth data are $_n\overline{\mathscr{G}}_2{}^T$ on the left and $_n\overline{\mathscr{G}}_{12}{}^T$ on the right.
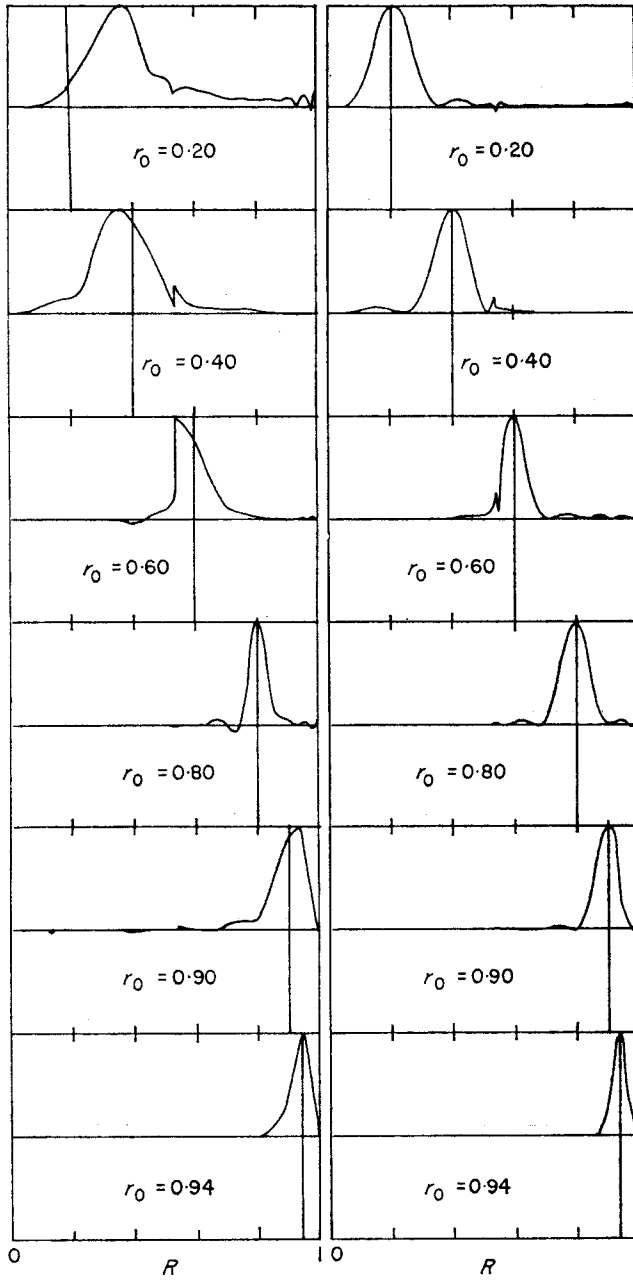
FIG. 6. The same as Fig. 5 except that the gross Earth data are $_n\overline{\mathscr{G}}_2{}^S$ on the left and $_n\overline{\mathscr{G}}_{12}{}^S$ on the right.
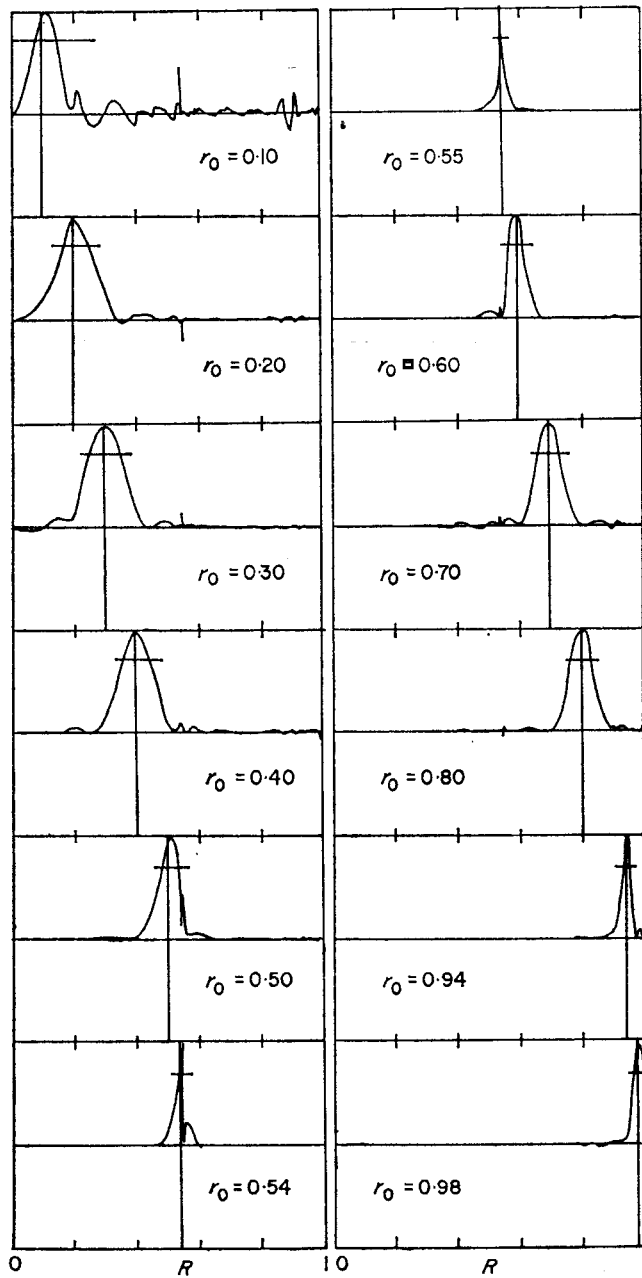
FIG. 7. The same as Fig. 5 except that the gross Earth data are $_n\overline{\mathscr{G}}_1{}^s$. The lengths of the fine horizontal bars are the resolving lengths $L(r_0)$.
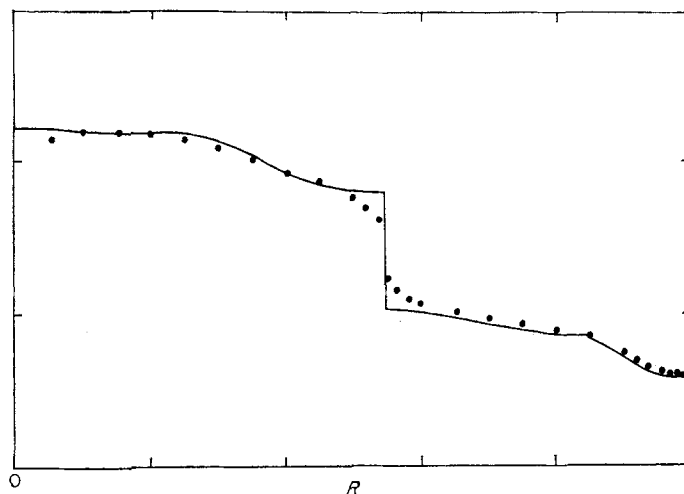
FIG. 8. The solid curve is the density $\rho_0(R)$ in model 5821 and the dots are its local averages at certain R, computed with the averaging kernels shown in Fig. 7. Full scale is 16·551 g/cm³.

The error estimates were obtained simply by using other values of r_0 near 0·55 and are meaningful only if $v_P(r)$ and $v_S(r)$ describe the real Earth.

No scientific measurement is complete without a statement about its uncertainties We believe that in geophysical inverse problems where an Earth model is constructed from gross Earth data $\mathscr{G}$, one proper description of the uncertainties is by means of figures like Figs 7 and 8. The $\mathscr{G}$-acceptable model m_0, the optimal averaging kernels $A(r_0, r)$, and the result $\langle m \rangle_{r_0}$ of averaging m_0 with these kernels are all items of information which are essential to any assessment of the relation between the proposed model m_0 and the real Earth. It is only the local averages $\langle m \rangle_{r_0}$ which are in any sense determined by the data, and the resolving power of the averages is best made clear by inspection of the averaging kernels $A(r_0, r)$, although a single number like the $L(r_0)$ defined in equation (7.2) does give a rough idea of the local resolving power.

We think it important to emphasize that many different sets of gross Earth functionals will produce essentially the same averaging kernels obtained in this section. In Appendix D we give the averaging kernels obtained with a set of 26 gross Earth functionals which have only 17 members in common with the set $_n\mathscr{G}_l^S$.

8. Speculative discussion of Borg's theroem

Suppose the space $\mathfrak{M}$ of Earth models is the Hilbert space of all square-integrable functions on $0 \leqslant r \leqslant 1$, and that $g_1, g_2, \ldots$ is an *infinite* set of *linear* gross Earth functionals. If the kernels $G_1(r), G_2(r), \ldots$ which generate $g_1, g_2, \ldots$ constitute a complete set of functions in $\mathfrak{M}$ then the values of $g_1(m), g_2(m), \ldots$ uniquely determine the function m. (Strictly speaking, m is determined except on a set of Lebesgue measure 0, but if m is piece-wise continuous it is uniquely determined except at the single isolated points of discontinuity where its value is without physical meaning.) Suppose we denote by $\mathfrak{M}_N$ the linear space consisting of all functions which are linear combinations of $G_1, G_2, \ldots, G_N$. Then, as N increases, $\mathfrak{M}_N$ will contain better and better approximations to delta functions. If the converse is formulated carefully it is also true, but this is not the place for such a discussion. In any event, for *infinite* sets of kernels, completeness is essentially equivalent to the possibility of constructing δ-like unimodular linear combinations.

Of course, it is geophysically unrealistic to discuss infinite sets of gross Earth data; we will never have such an infinite set of observations. Nevertheless, it is of some mathematical interest to learn which infinite sets of gross Earth data do uniquely determine the Earth's structure. Our results for finite sets of gross Earth functionals can only suggest conjectures about the mathematical problem, but they do suggest a definite conjecture. Suppose we know $v_P(r)$ and $v_S(r)$ and are trying to determine $\rho(r)$ from the total mass and moment of inertia and the frequencies of some infinite set of normal modes. If the set of normal modes we use consists of all the spheroidal modes $_nS_l$ belonging to one fixed angular order l, Fig. 6 suggests that $\rho(r)$ will be uniquely determined. If our normal modes consist of all the toroidal modes $_nT_l$ for one fixed l, Fig. 5 suggests that $\rho(r)$ will not be uniquely determined; but ρ will be determined if we have all the modes $_nT_l$ for *two* different values of l.

In Inverse I we made a conjecture based on an unproved extension of Borg's (1945) theorem: given the total mass, moment, $v_P(r)$, and $v_S(r)$, then all the eigenfrequencies of $_nT_l$ for two different l would determine $\rho(r)$ uniquely, but all for one l would not. Spheroidal modes were conjectured to behave like toroidal modes in this respect. The evidence from Figs 5 and 6 supports our conjecture for toroidal but not for spheroidal modes. Perhaps the difference is connected with the fact that Borg's theorem was proved only for Sturm-Liouville equations of second order, and only the toroidal modes are governed by such an equation. Very crudely, Borg's theorem may be phrased thus: if the spectra of a single Schrödinger equation are known for two different sets of homogeneous boundary conditions then the potential function in the Schrödinger equation is uniquely determined. Our conjecture in Inverse I was based on the idea that perhaps the two spectra could be obtained in other ways than from two different sets of boundary conditions. When the frequencies of all $_nT_l$ for two different l are known, we have only one set of boundary conditions, but we have the spectrum for two different potential functions whose difference is known.

The evidence of Figs 5 and 6 suggests that if this conjectural extension of Borg's theorem is correct, then the eigenfrequencies of $_nS_l$ for a *single* l should be regarded as *two* spectra. Such a suggestion is not completely ridiculous. In a uniform slab the analogue of the spheroidal modes is obtained by coupling P and $S-V$ waves; for a given horizontal wave number, as the vertical wave number increases this set of modes decouples into two separate sets of organ-pipe modes, one corresponding to vertically travelling P waves and the other to vertically travelling S waves (Ewing, Jardetsky & Press 1957). In any case, it is evidently dangerous to conjecture extensions of Borg's second-order theorem to systems of fourth and sixth order in the naive manner of Inverse I.

9. Geographical Earth structure

All the detailed discussion so far has been directed at spherical Earth models, n-tuples of functions of radial distance r. The method for calculating local averages generalizes in an obvious way to aspherical Earth models, n-tuples of functions of the position vector $\mathbf{r}$. For simplicity of exposition we restrict the discussion to the case $n = 1$, in which the Earth models are single real-valued functions $m(\mathbf{r})$. We denote by E the region in three-space occupied by the Earth. First we consider *linear* gross Earth functionals.

If g_i is a linear Fréchet-differentiable gross Earth functional, then in direct analogy with equation (3.1) we have for any Earth model m

$$g_i(m) = \int\limits_E G_i(\mathbf{r})\, m(\mathbf{r})\, dV \qquad (9.1)$$

where G_i is a known function of position $\mathbf{r}$ and dV is the volume element in E. The measured gross Earth data $\gamma_1, ..., \gamma_N$ are three-dimensional generalized moments of the unknown Earth model m:

$$\gamma_i = \int_E G_i(\mathbf{r})\, m(\mathbf{r})\, dV. \tag{9.2}$$

We try to calculate from the measured values of $\gamma_1, ..., \gamma_N$ some estimate of a local average of $m(\mathbf{r})$ in some neighbourhood of a given point $\mathbf{r}_0$. We consider only linear averages,

$$\langle m \rangle_{\mathbf{r}_0} = \int_E A(\mathbf{r}_0, \mathbf{r})\, m(\mathbf{r})\, dV \tag{9.3}$$

where

$$\int_E A(\mathbf{r}_0, \mathbf{r})\, dV = 1 \tag{9.4}$$

(any function satisfying equation (9.4) we call unimodular). Ideally we would like to have $A(\mathbf{r}_0, \mathbf{r}) = \delta(\mathbf{r} - \mathbf{r}_0)$, the three-dimensional Dirac delta function.

Since $\langle m \rangle_{\mathbf{r}_0}$ is supposed to depend only on $g_1(m), ..., g_N(m)$ and to be linear in m, Appendix A shows that it must be linear in $g_1(m), ..., g_N(m)$. Hence there exist constants $a_1(\mathbf{r}_0), ..., a_N(\mathbf{r}_0)$ such that for any Earth model m

$$\langle m \rangle_{\mathbf{r}_0} = \sum_{i=1}^{N} a_i(\mathbf{r}_0)\, g_i(m) \tag{9.5}$$

while for any $\mathscr{G}$-acceptable Earth model

$$\langle m \rangle_{\mathbf{r}_0} = \sum_{i=1}^{N} a_i(\mathbf{r}_0)\, \gamma_i. \tag{9.6}$$

From equation (9.1) we see that $\langle m \rangle_{\mathbf{r}_0}$ has the form (9.3) with

$$A(\mathbf{r}_0, \mathbf{r}) = \sum_{i=1}^{N} a_i(\mathbf{r}_0)\, G_i(\mathbf{r}). \tag{9.7}$$

For a given $\mathbf{r}_0$, the problem is to choose that unimodular linear combination of $G_1(\mathbf{r}), ..., G_N(\mathbf{r})$ which most nearly approximates $\delta(\mathbf{r} - \mathbf{r}_0)$. To do so we need a three-dimensional δ-ness criterion. We choose any function $J(\mathbf{r}_0, \mathbf{r})$ which vanishes when $\mathbf{r} = \mathbf{r}_0$ and increases monotonically as $|\mathbf{r} - \mathbf{r}_0|$ increases away from zero. Examples are

$$J(\mathbf{r}_0, \mathbf{r}) = |\mathbf{r} - \mathbf{r}_0|^2$$

and

$$J_\sigma(\mathbf{r}_0, \mathbf{r}) = 2\sigma^2 \{1 - \exp[-|\mathbf{r} - \mathbf{r}_0|^2 / 2\sigma^2]\}. \tag{9.8}$$

Then for any function $A(\mathbf{r})$ we can define

$$\Delta_J(A) = \int_E J(\mathbf{r}_0, \mathbf{r})\, A(\mathbf{r})^2\, dV. \tag{9.9}$$

Then just as in Section 3 we insert equation (9.7) into equation (9.9) and minimize the resulting positive-definite quadratic form in $a_1, ..., a_N$ subject to the side-condition (9.4).

If the resulting optimal kernel $A(\mathbf{r}_0, \mathbf{r})$ does not have a high peak for $\mathbf{r}$ near $\mathbf{r}_0$, with small magnitude elsewhere, we conclude that the given gross Earth data do not

permit calculation of a local average of m near r_0. (In drawing this conclusion we must be careful to avoid pathologically small values of σ in equation (9.8).) If the optimal $A(r_0, r)$ does resemble a blurred $\delta(r - r_0)$, then $\langle m \rangle_{r_0}$ is given by equation (9.6).

The generalization to non-linear gross Earth functionals proceeds exactly as in Section 6.

Now we note an extremely important fact. The deviations of density and seismic velocities from sphericity and the deviations of the local stress–strain relation from isotropy are all probably small enough to be treated by first-order perturbation theory. But in that case, their contributions to the measured gross Earth data are *linear* gross Earth functionals. As we have seen, there is at present no technique which we are sure *a priori* will establish uniqueness of the local averages when the measured gross Earth data are non-linear. But when they are linear this uniqueness is automatic. In other words, if we can somehow establish that the real Earth differs only slightly from a certain isotropic spherical Earth model, then the local averages of that difference, if they exist, are uniquely determined by the data, even if they represent aspherical deviations, anhydrostatic stresses, or anisotropies in the stress–strain relation.

The generalization of the δ-ness criterion for the aspherical case to n-dimensional Earth models is accomplished exactly as in Section 4 for the spherical case.

10. Conclusions

The principal result of our work is that it is possible to draw rigorous conclusions about the internal structure of the Earth from a finite set of gross Earth data, and that if the data are properly chosen the conclusions are not only rigorous but geophysically interesting.

The test of whether a set $\mathscr{G}$ of gross Earth functionals $g_1, ..., g_N$ is 'properly chosen' is whether the corresponding kernels G_i have unimodular linear combinations which are good approximations to Dirac delta functions. If the data are properly chosen, then among the weighted averages of real Earth structure which are computable directly from the data there appear averages which concentrate their weight in narrow intervals of depth. Thus local average values of the functions $m_1(r), ..., m_n(r)$ which define the Earth model (density, compressional and shear velocities, etc.) are calculable from the observed data.

When the gross Earth functionals are linear, these local averages are determined completely by the observed values of the corresponding gross Earth functionals. When the gross Earth data are non-linear, the local averages of Earth structure are determined for all Earth models which resemble any particular acceptable model. However, for non-linear gross Earth functionals there remains the possibility that two widely different models will be acceptable. This paper does not deal with that problem. The only approach to it known to us is the one discussed in Section 3 of Inverse I.

There is one serious practical limitation to the ideas put forward in this paper. With the computing facilities presently available to us we cannot easily handle sets $\mathscr{G}$ of gross Earth functionals with more than about 40 members. If we could handle more data, we would presumably get better resolution than that shown in Figs 2–7. We can never get infinite resolution, but with higher resolution the techniques described here and in Inverse I should be capable of giving rigorous answers to a number of qualitative questions about the interior of the Earth, such as whether there are low velocity zones or density inversions in the mantle.

Acknowledgments

Institute of Geophysics and Planetary Physics,
  University of California,
    San Diego,
      California.

1968 *February.*

References

Anderson, D. L. & Archambeau, C. B., 1964. The anelasticity of the Earth, *J. geophys. Res.*, **69**, 2071–2084.

Backus, G. & Gilbert, F., 1961. The rotational splitting of the free oscillations of the Earth, *Proc. natn. Acad. Sci.*, **47**, 362–371.

Backus, G. & Gilbert, F., 1967. Numerical applications of a formalism for geophysical inverse problems, *Geophys. J. R. astr. Soc.*, **13**, 247–276.

Borg, G., 1945. Eine Umkehrung der Sturm-Liouvilleschen Eigenwertaufgabe, *Acta math.*, **78**, 1–96.

Courant, R. & Hilbert, D., 1953. *Methods of Mathematical Physics—I*, Interscience, New York.

Dahlen, A., 1968. The normal modes of a rotating, elliptical Earth, *Geophys. J. R. astr. Soc.*, to be published.

Dunford, N. & Schwartz, J., 1958. *Linear Operators—I*, Interscience, New York.

Ewing, M., Jardetzky, W. S., & Press, F., 1957. *Elastic Waves in Layered Media*, McGraw-Hill, New York.

Gibbs, W. & Wilson, E., 1901. *Vector Analysis*, Yale University Press.

Gilbert, F. & Backus, G., 1965. The rotational splitting of the free oscillations of the Earth—2, *Rev. Geophys.*, **3**, 1–9.

Gilbert, F. & Backus, G., 1968. Approximate solutions to the inverse normal mode problem, *Bull. seism. Soc. Am.*, **58**, 103–131.

Gutenberg, B., 1959. *The Physics of the Earth's Interior*, Academic Press, New York.

Jeffreys, H., 1963. The hydrostatic theory of the figure of the Earth, *Geophys. J. R. astr. Soc.*, **8**, 196–202.

Kato, T., 1966. *Perturbation Theory for Linear Operators*, Springer-Verlag, New York.

King-Hele, D. C., Cook, G. E. & Watson, H. M., 1964. Even zonal harmonics in the Earth's gravitational potential, *Nature, Lond.*, **202**, 996.

Slichter, L. B., 1967. Free oscillations of the Earth, *International Dictionary of Geophysics*, ed. by S. K. Runcorn, Pergamon, Oxford.

APPENDIX A

Linearity of averages. We have a linear space $\mathfrak{M}$ of Earth models and a finite set $\mathscr{G}$ of linearly independent linear gross Earth functionals $g_1, ..., g_N$. We also have an ' averaging operation ' which assigns to each $\mathfrak{m}$ in $\mathfrak{M}$ a real number $\langle \mathfrak{m} \rangle$ which depends linearly on $\mathfrak{m}$ and can be calculated from $g_1(\mathfrak{m}), ..., g_N(\mathfrak{m})$ alone. We want to prove that $\langle \mathfrak{m} \rangle$ depends linearly on $g_1(\mathfrak{m}), ..., g_N(\mathfrak{m})$.

Let $\mathfrak{R}$ denote the one-dimensional linear space of real numbers, while $\mathfrak{R}^N$ denotes the N-dimensional linear space of ordered N-tuples of real numbers. Then

we may think of the whole collection $\mathscr{G}$ of gross Earth functionals as a single function ϕ which assigns to each $\mathfrak{m}$ in $\mathfrak{M}$ an ordered N-tuple in $\mathfrak{R}^N$:

$$\phi(\mathfrak{m}) = (g_1(\mathfrak{m}), \ldots, g_N(\mathfrak{m})). \qquad (\mathrm{A}.1)$$

Because each functional g_i is linear, $\phi : \mathfrak{M} \to \mathfrak{R}^N$ is also linear.

We are also given that the 'average', $\langle \mathfrak{m} \rangle$, can be calculated from $g_1(\mathfrak{m}), \ldots, g_N(\mathfrak{m})$. That is, there is a function $\psi : \mathfrak{R}^N \to \mathfrak{R}$ which assigns to each ordered N-tuple $(x_1, \ldots, x_N)$ of real numbers a real number $\psi(x_1, \ldots, x_N)$, and for any $\mathfrak{m}$ in $\mathfrak{M}$ we have

$$\psi(g_1(\mathfrak{m}), \ldots, g_N(\mathfrak{m})) = \langle \mathfrak{m} \rangle. \qquad (\mathrm{A}.2)$$

Combining equations (A.1) and (A.2) we see that

$$\langle \mathfrak{m} \rangle = \psi\phi(\mathfrak{m}), \qquad (\mathrm{A}.3)$$

so that the composite mapping $\psi\phi : \mathfrak{M} \to \mathfrak{R}$ is linear.

Now we have the mappings ϕ and ψ shown in the following diagram:

$$\mathfrak{M} \overset{\phi}{\to} \mathfrak{R}^N \overset{\psi}{\to} \mathfrak{R}$$

and we know that ϕ and the composite mapping $\psi\phi$ are linear. We want to conclude that ψ is linear. First we observe that ϕ maps $\mathfrak{M}$ *onto* $\mathfrak{R}^N$; that is, every member of $\mathfrak{R}^N$ is $\phi(\mathfrak{m})$ for at least one $\mathfrak{m}$ in $\mathfrak{M}$. This follows from the linear independence of $g_1, \ldots, g_N$. The set $\phi(\mathfrak{M})$ of all members of $\mathfrak{R}^N$ which are $\phi(\mathfrak{m})$ for some $\mathfrak{m}$ in $\mathfrak{M}$ is a linear subspace of $\mathfrak{R}^N$ because ϕ is linear. If $\phi(\mathfrak{M})$ is not all of $\mathfrak{R}^N$, there is a non-zero member of $\mathfrak{R}^N$, say $(a_1, \ldots, a_N)$, which is orthogonal to the space $\phi(\mathfrak{M})$. But then for every $\mathfrak{m}$ in $\mathfrak{M}$

$$a_1 g_1(\mathfrak{m}) + \ldots + a_N g_N(\mathfrak{m}) = 0,$$

contradicting our hypothesis that $g_1, \ldots, g_N$ are linearly independent.

Our conclusion that ψ is linear now follows from a

THEOREM: Suppose X, Y, Z are linear spaces and $\phi : X \to Y$ and $\psi : Y \to Z$. Suppose that ϕ and $\psi\phi$ are linear and that ϕ is onto. Then ψ is linear.

Proof. Let $\mathbf{y}_1$ and $\mathbf{y}_2$ be any two vectors in Y, and let a_1 and a_2 be any two real numbers. By hypothesis, there are vectors $\mathbf{x}_1$ and $\mathbf{x}_2$ in X such that $\mathbf{y}_i = \phi(\mathbf{x}_i)$, $i = 1, 2$. Then

$$\psi(a_1 \mathbf{y}_1 + a_2 \mathbf{y}_2) = \psi(a_1 \phi(\mathbf{x}_1) + a_2 \phi(\mathbf{x}_2))$$
$$= \psi\phi(a_1 \mathbf{x}_1 + a_2 \mathbf{x}_2)$$

because ϕ is linear. But $\psi\phi$ is also linear, so

$$\psi(a_1 \mathbf{y}_1 + a_2 \mathbf{y}_2) = a_1(\psi\phi)(\mathbf{x}_1) + a_2(\psi\phi)(\mathbf{x}_2).$$

From the definition of a composite mapping, $(\psi\phi)(\mathbf{x}_i) = \psi[\phi(\mathbf{x}_i)] = \psi(\mathbf{y}_i)$, so

$$\psi(a_1 \mathbf{y}_1 + a_2 \mathbf{y}_2) = a_1 \psi(\mathbf{y}_1) + a_2 \psi(\mathbf{y}_2)$$

as asserted in the theorem.

APPENDIX B

Dirichlet kernels. Section 4 was a very brief discussion of different δ-ness criteria. In this Appendix we describe an entirely different approach to the attempt to construct δ-like linear combinations of a given finite set of functions or n-tuples of functions.

To motivate the present approach, first we apply our integral δ-ness criterion to the well-understood set of functions $G_i(r) = \sin i\pi r$, $i = 1, ..., N$, on the interval $0 \leqslant r \leqslant 1$. Fig. 9 shows the Dirichlet kernel for those functions,

$$A_D(r_0, r) = \frac{2}{\pi} \sum_{i=1}^{N} \sin i\pi r_0 \sin i\pi r, \qquad (A.4)$$

and also their most δ-like linear combination, selected according to the δ-ness criterion (3.8), (3.9) with $2\sigma = 1$. All kernels are plotted for $N = 17$ and $r_0 = 0.5$ only.
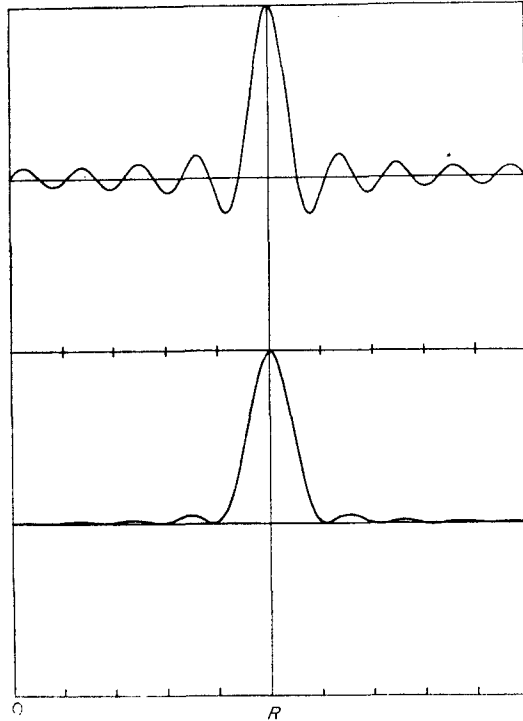
FIG. 9. Averaging kernels $A(r_0, R)$ on $0 \leqslant R \leqslant 1$ computed at $r_0 = 0.5$ from $\sin \pi iR$, $i = 1, .., 17$. Above is the Dirichlet kernel and below is the optimal kernel given by δ-ness criterion equations (3.9) and (4.1).

We see from Fig. 9 that the averaging kernel produced from $\sin \pi r$, $\sin 2\pi r$, ..., $\sin 17\pi r$ by the δ-ness criterion (3.8), (3.9) has a resolving length about twice as long as that of the Dirichlet kernel and sidebands much smaller than those of the Dirichlet kernel. In addition, for reasons unknown to us at this writing, our δ-ness kernel in Fig. 9 is non-negative throughout $0 \leqslant r \leqslant 1$. Considerable work by one of us (G.B.) over the past year on integral δ-ness criteria for generating averaging kernels on the sphere from spherical harmonics indicates empirically that kernels so generated usually have poorer resolution and smaller sidebands than the Dirichlet kernel.

If this empirical indication generalizes to the much less regular kernels $G_1, ..., G_N$ considered here, it suggests that we might gain resolution at the expense of larger sidebands if we could find the analogue of the Dirichlet kernel for an arbitrary finite set of linearly independent functions $G_1(r), ..., G_N(r)$ on $0 \leqslant r \leqslant 1$. To this end, we reformulate the problem of interpreting a finite number of linear gross Earth

data. Nothing is lost by discussing from the outset the general case, with n-dimensional Earth models.

Let $\mathfrak{M}$ be the Hilbert space of n-tuples of real functions square-integrable on $0 \leqslant r \leqslant 1$. A typical member of $\mathfrak{M}$ is

$$\mathbf{m}(r) = \big(m_1(r), \ldots, m_n(r)\big).$$

The inner product of two members of $\mathfrak{M}$ is defined as

$$(\mathbf{m}, \mathbf{m}') = \sum_{v=1}^{n} \int_0^1 m_v(r)\, m_v'(r)\, dr.$$

Let g be a continuous linear functional on $\mathfrak{M}$. Then there is a member of $\mathfrak{M}$,

$$\mathbf{G}(r) = \big(G_1(r), \ldots, G_n(r)\big),$$

such that for any $\mathbf{m}$ in $\mathfrak{M}$,

$$g(\mathbf{m}) = (\mathbf{m}, \mathbf{G}) = \sum_{v=1}^{n} \int_0^1 m_v(r)\, G_v(r)\, dr.$$

Now suppose we are given N linearly independent linear functionals $g_1, \ldots, g_N$ on $\mathfrak{M}$. We would like to say as much as we can about an unknown Earth model $\mathbf{m}$ in $\mathfrak{M}$ simply from the known values of $g_1(\mathbf{m}), \ldots, g_N(\mathbf{m})$, that is, from the known inner products

$$g_i(\mathbf{m}) = (\mathbf{m}, \mathbf{G}_i) = \sum_{v=1}^{n} \int_0^1 m_v(r)\, G_{iv}(r)\, dr, \qquad (A.5)$$

with $i = 1, \ldots, N$. Let $\mathfrak{M}_N$ denote the N-dimensional linear subspace of $\mathfrak{M}$ consisting of all linear combinations of $\mathbf{G}_1, \ldots, \mathbf{G}_N$. Let $\mathbf{H}_1, \ldots, \mathbf{H}_N$ be any orthonormal basis for $\mathfrak{M}_N$. Then $\mathbf{H}_1, \ldots, \mathbf{H}_N$ are linear combinations of $\mathbf{G}_1, \ldots, \mathbf{G}_N$ and vice-versa, so the problem is to say what we can about $\mathbf{m}$ given only $(\mathbf{m}, \mathbf{H}_1), \ldots, (\mathbf{m}, \mathbf{H}_N)$.

We can always find an infinite orthonormal basis for $\mathfrak{M}$ whose first N members are $\mathbf{H}_1, \ldots, \mathbf{H}_N$ (Dunford & Schwartz 1958), and which therefore can be written $\mathbf{H}_1, \mathbf{H}_2, \ldots$. Then any member of $\mathfrak{M}$ can be expanded in the form

$$\mathbf{m} = \sum_{i=1}^{\infty} (\mathbf{m}, \mathbf{H}_i)\, \mathbf{H}_i. \qquad (A.6)$$

The series (A.6) consists of two parts,

$$\langle \mathbf{m} \rangle = \sum_{i=1}^{N} (\mathbf{m}, \mathbf{H}_i)\, \mathbf{H}_i \qquad (A.7)$$

and

$$\mathbf{m}_\perp = \sum_{i=N+1}^{\infty} (\mathbf{m}, \mathbf{H}_i)\, \mathbf{H}_i. \qquad (A.8)$$

As is well known (Dunford & Schwartz 1958), the two parts $\langle \mathbf{m} \rangle$ and $\mathbf{m}_\perp$ are determined uniquely by $\mathbf{m}$ and $\mathfrak{M}_N$ and do not depend on $\mathbf{H}_1, \ldots, \mathbf{H}_N$; $\langle \mathbf{m} \rangle$ is the orthogonal projection of $\mathbf{m}$ onto $\mathfrak{M}_N$ and $\mathbf{m}_\perp$ is the component of $\mathbf{m}$ orthogonal to $\mathfrak{M}_N$. This decomposition of $\mathbf{m}$ is ideal for our purposes. If we know only $(\mathbf{m}, \mathbf{H}_1), \ldots, (\mathbf{m}, \mathbf{H}_N)$, then we know $\langle \mathbf{m} \rangle$ exactly and we know nothing whatever about $\mathbf{m}_\perp$.

Now $\langle \mathbf{m} \rangle$ summarizes succinctly all the data we have about $\mathbf{m}$. Can $\langle \mathbf{m} \rangle$ be regarded in any sense as an estimate of $\mathbf{m}$? To answer this question, we must calcu-

late $\langle \mathbf{m} \rangle$ in some detail. First we note that in the notation of Gibbs & Wilson (1901) equation (A.7) can be written

$$\langle \mathbf{m} \rangle = \mathbf{m} \cdot \mathbf{I}_N$$

where $\mathbf{I}_N$ is the second-order identity tensor on $\mathfrak{M}_N$:

$$\mathbf{I}_N = \sum_{i=1}^{N} \mathbf{H}_i \mathbf{H}_i,$$

juxtaposition here indicating a tensor product. But if $\mathbf{G}^1, ..., \mathbf{G}^N$ is the basis for $\mathfrak{M}_N$ dual to $\mathbf{G}_1, ..., \mathbf{G}_N$ then

$$\mathbf{I}_N = \sum_{i=1}^{N} \mathbf{G}^i \mathbf{G}_i,$$

so

$$\langle \mathbf{m} \rangle = \sum_{i=1}^{N} (\mathbf{m}, \mathbf{G}^i) \, \mathbf{G}_i. \tag{A.9}$$

To find the dual basis we proceed as usual. We denote the $N \times N$ matrix (G_i, G_j) by g_{ij} and its inverse by g^{ij}. Then

$$\mathbf{G}^i = \sum_{j=1}^{N} g^{ij} \mathbf{G}_j. \tag{A.10}$$

Therefore

$$\langle \mathbf{m} \rangle = \sum_{i,j=1}^{N} g^{ij} (\mathbf{m}, \mathbf{G}_j) \, \mathbf{G}_i. \tag{A.11}$$

Now we must interpret (A.11) in terms of averaging kernels. We note that $\langle \mathbf{m} \rangle$ is an n-tuple of functions of r, say $(\langle \mathbf{m} \rangle_{1r}, ..., \langle \mathbf{m} \rangle_{nr})$. Then the μth component of equation (A.11) is

$$\langle \mathbf{m} \rangle_{\mu r_0} = \sum_{i,j=1}^{N} g^{ij} \left[\sum_{\nu=1}^{n} \int_0^1 m_\nu(r) \, G_{j\nu}(r) \, dr \right] G_{\mu i}(r_0). \tag{A.12}$$

Thus

$$\langle \mathbf{m} \rangle_{\mu r_0} = \sum_{\nu=1}^{n} \int_0^1 A_{\mu\nu}(r_0, r) \, m_\nu(r) \, dr \tag{A.13}$$

where

$$A_{\mu\nu}(r_0, r) = \sum_{i,j=1}^{N} g^{ij} G_{\mu i}(r_0) \, G_{\nu j}(r). \tag{A.14}$$

Equation (A.13) has exactly the form of equation (4.5). When $n = 1$ (one-dimensional Earth models) and $G_i(r) = \sin i\pi r$, $i = 1, ..., N$, then equation (A.14) is exactly the Dirichlet kernel (A.4), so we have here a generalization of the Dirichlet kernel to arbitrary sets of n-tuples of functions.

To see whether equation (A.13) can be interpreted as a local average of $m_\mu(r)$ near $r = r_0$ we must examine whether in equation (A.14) $A_{\mu\mu}(r_0, r)$ has a large peak near $r = r_0$ and small sidebands elsewhere, and whether $A_{\mu\nu}(r_0, r)$ is very small for all r when $\nu \neq \mu$. If these desiderata are present, equation (A.13) is indeed a local average of $m_\mu(r)$ with r near r_0, contaminated by contributions from $m_\mu(r)$ at distant r and by contributions from $m_\nu(r)$ with $\nu \neq \mu$. We know of no better way to test for these desiderata of δ-ness than to examine the graphs $A_{\mu\nu}(r_0, r)$. And there is no reason *a priori* to prefer equation (A.14) to the averaging kernels (4.6) obtained from integral δ-ness criteria. We suspect, however, that if δ-like averaging kernels are available at all, then equation (A.14) will be such a kernel and will have better resolution but larger sidebands than (4.6).
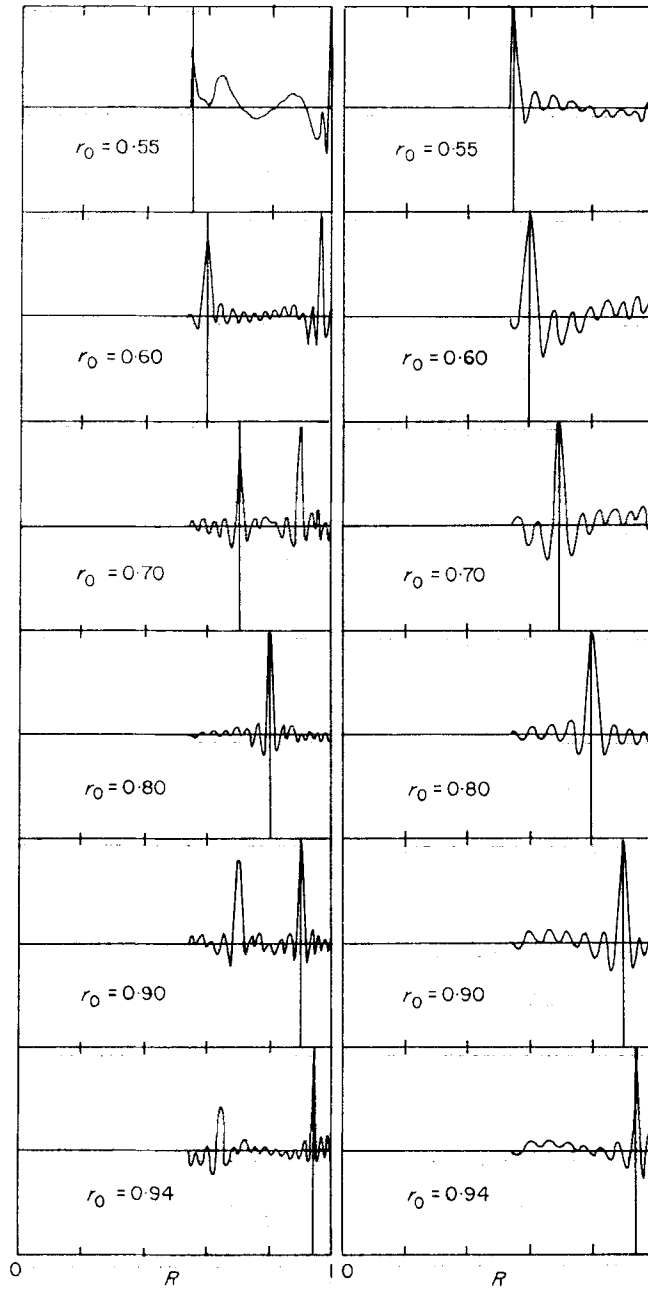
FIG. 10. Dirichlet kernels $A(r_0, R)$ for $Q^{-1}(R)$ at various r_0 (the fine vertical lines). The gross Earth data are ${}_n\mathcal{G}_2{}^T$ on the left and ${}_n\mathcal{G}_{12}{}^T$ on the right.
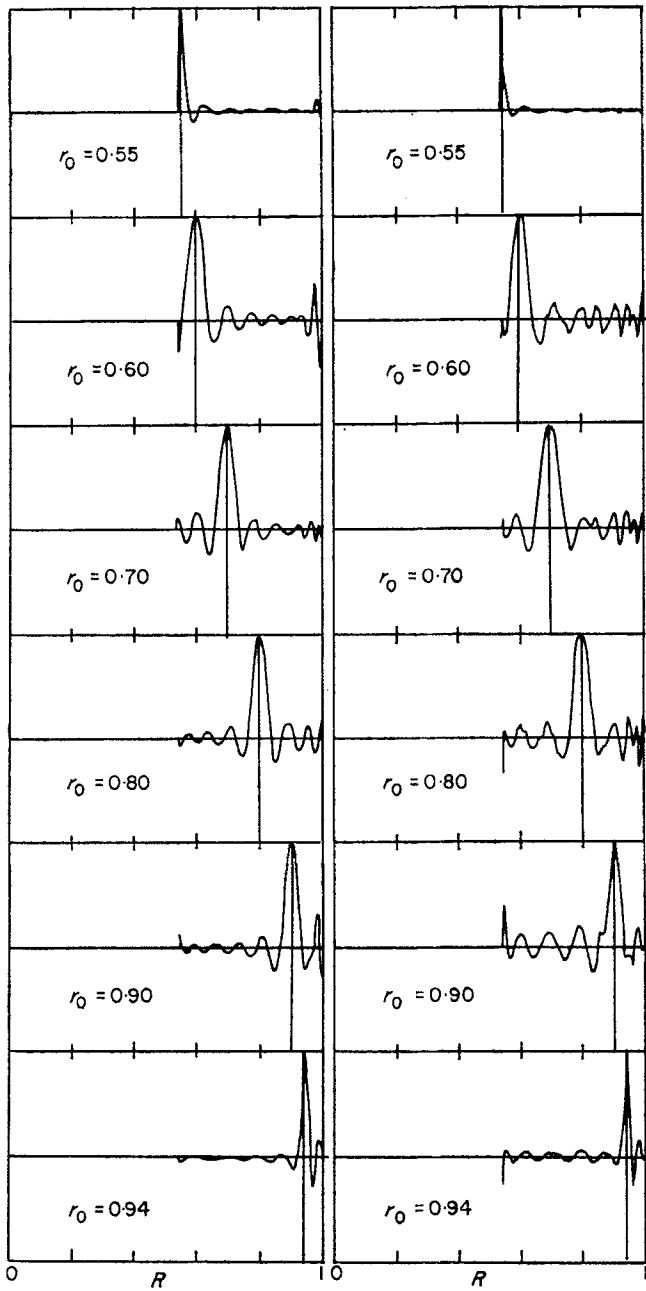
FIG. 11. The same as Fig. 10 except that the gross Earth data are $_n\mathcal{G}_2{}^S$ on the left and $_n\mathcal{G}_{12}{}^S$ on the right.
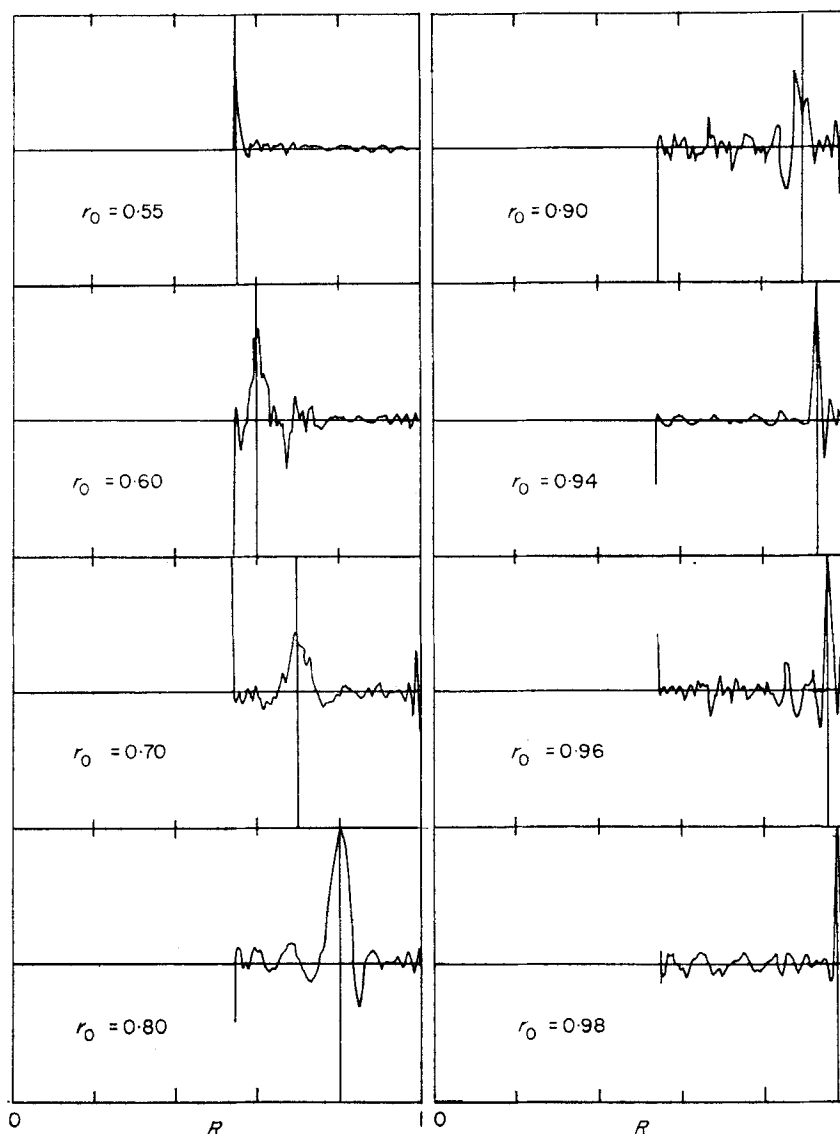
FIG. 12. The same as Fig. 10 except that the gross Earth data are $_n\mathscr{G}_l{}^S$.

In Figs 10–12 we give the Dirichlet averaging kernels appropriate to the sets of linear gross Earth functionals discussed in Section 5. These figures are to be compared directly with Figs 2–4, which apply to the same underlying dissipationless Earth model and give the most δ-like averaging kernals for $Q^{-1}(r)$ obtained by the method of Section 3. In the present case, the Dirichlet averaging kernels do not have much better resolution and they have much higher sidebands than the kernels obtained from integral δ-ness criteria. The latter averaging kernels seem to be preferable.

In Fig. 13 we give the Dirichlet kernels for $\rho(r)$ appropriate to the gross Earth data $_n\overline{\mathscr{G}}_l{}^S$ used in Section 7, and referring to the same $\rho_0(r)$, $v_P(r)$, $v_S(r)$ as in Fig. 7. The kernels in Figs 7 and 13 are directly comparable, and again the Dirichlet kernels give not much better resolution but much higher sidebands.
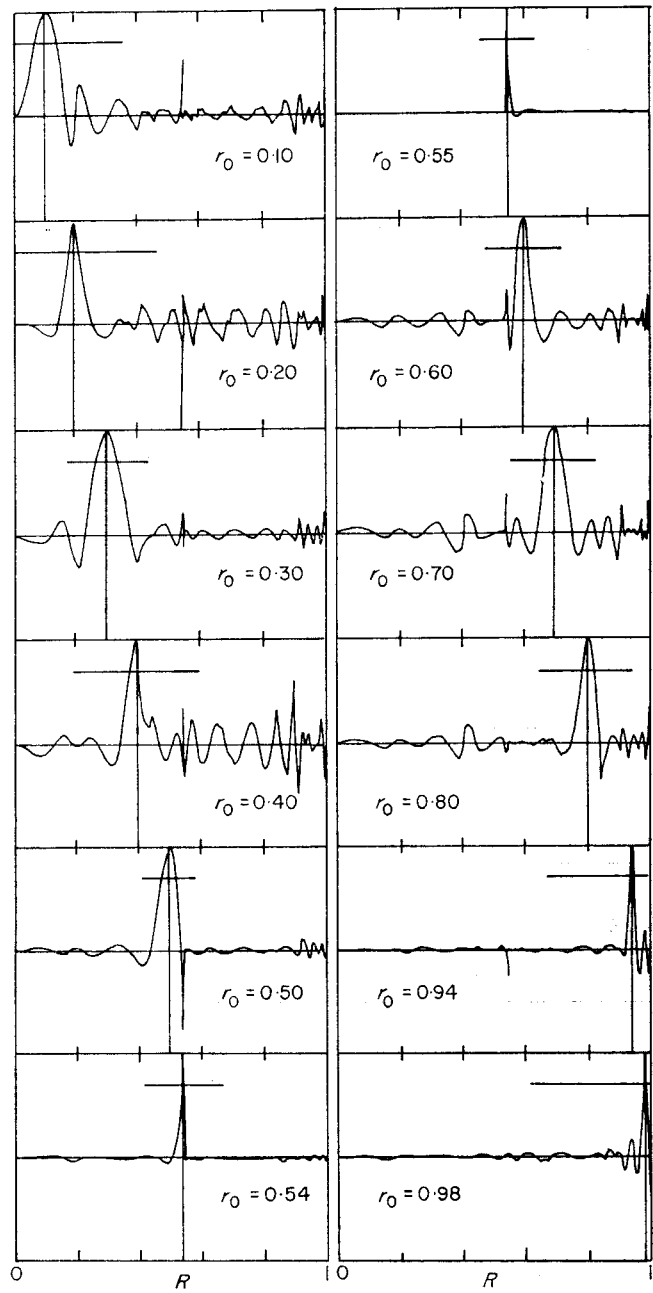
Fig. 13. Dirichlet kernels $A(r_0, R)$ for $\rho(R)$ at various values of r_0 (the fine vertical lines). The lengths of the fine horizontal bars are the resolving lengths $L(r_0)$. Gross Earth data are $_n\mathscr{G}_l{}^S$.
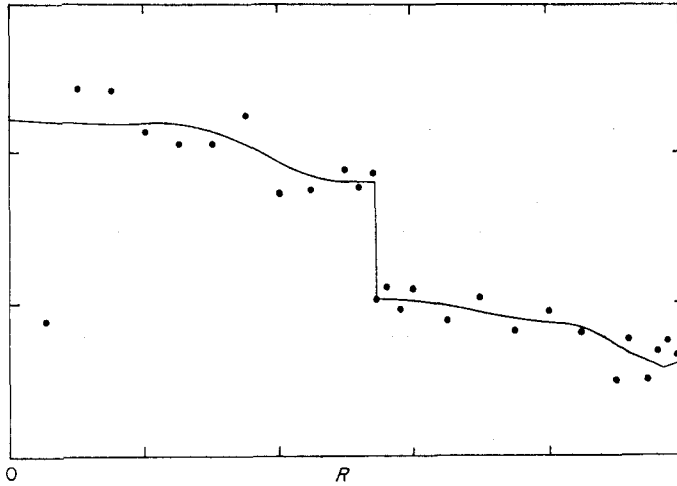
FIG. 14. The same as Fig. 8 except that the averaging kernels in Fig. 13 are used
instead of those in Fig. 7.

Fig. 14 gives the local averages $\langle \rho_0 \rangle_{r_0}$ obtained by applying to $\rho_0(r)$ the Dirichlet
averaging kernels of Fig. 13 rather than the optimally δ-like kernels of Fig. 7. The
differences between Figs 8 and 14 arise not from any *errors* of observation or inter-
pretation but simply because the two figures represent local averages of the same
function $\rho_0(r)$ taken with different averaging kernels. The sidebands of the Dirichlet
kernels in Fig. 13 are so large that the 'local' averages (Fig. 14) obtained from them
would be much less interesting to most geophysicists than the local averages (Fig. 8)
obtained from the kernels of Fig. 7.

In any case, if $v_P(r)$ and $v_S(r)$ describe the real Earth, and if $\rho_0(r)$ is close enough
to the density of the real Earth to permit neglecting the second order terms in
equation (6.1), then the averages of $\rho_0(r)$ calculated with either the kernels of Fig. 7
or the kernels of Fig. 13 are equal to the corresponding averages for the real Earth,
with an error of second order in the deviation of the real Earth from ρ_0.
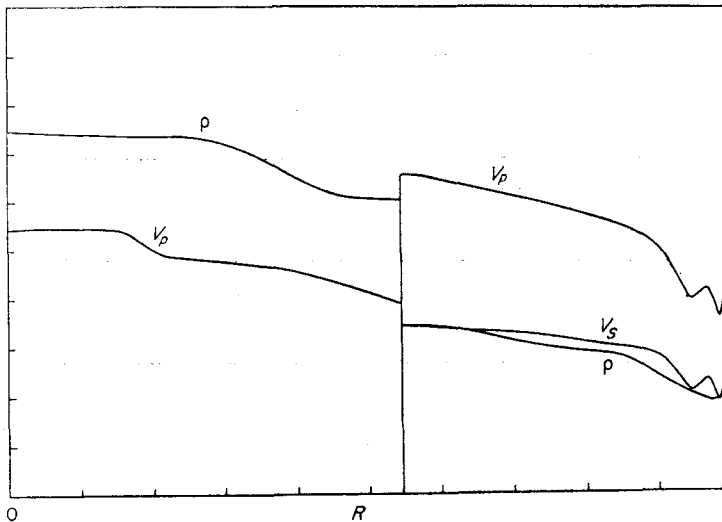

FIG. 15. Earth model 5821. Full scale is 16·551 g/cm³ for the density ρ, and
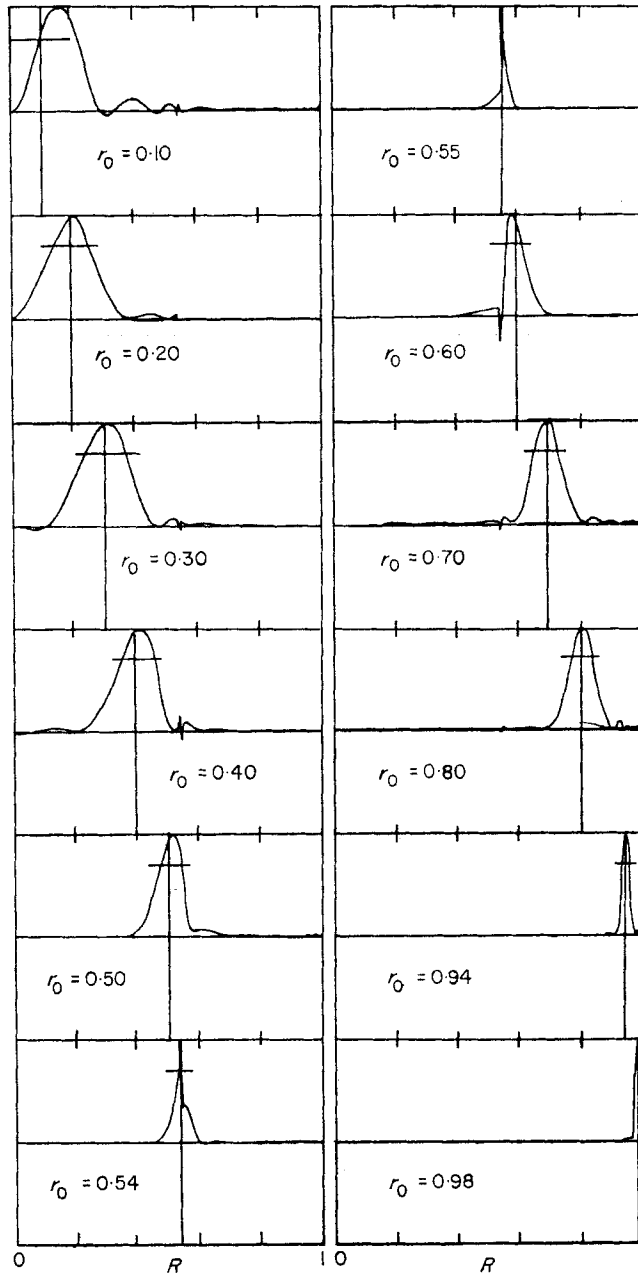20·551 km s⁻¹ for compressional velocity v_P and shear velocity v_S.

FIG. 16. The same as Fig. 7 except that the gross Earth data are $_n\bar{\mathscr{G}}_i{}^{ST}$ instead of $_n\bar{\mathscr{G}}_i{}^{S}$.

APPENDIX C

The density model used for averaging. Model 5821 is shown in Fig. 15. It is similar to model GM of Gilbert & Backus (1968). For 216 eigenfrequencies the relative r.m.s. error between the observed and computed values is 0·23 per cent and the maximum error is 0·77 per cent ($_0S_5$). The observed eigenfrequencies were corrected foı perturbations due to rotation and ellipticity using Table I of Dahlen (1968).

APPENDIX D

Stability of averaging kernels. As remarked at the end of Section 7, if we know $v_P(r)$ and $v_S(r)$ and are trying to obtain local averages of the density $\rho(r)$, we seem to have considerable latitude in our choice of the set $\mathscr{G}$ of gross Earth data used in the computation. In this appendix we exhibit the averaging kernels for $\rho(r)$ obtained from the set $_n\overline{\mathscr{G}}_l^{ST}$ consisting of the Earth's mass and moment and the squared eigenfrequencies of the 24 normal modes $_0S_0$, $_1S_0$, $_2S_0$, $_3S_0$, $_1S_1$, $_2S_1$, $_0S_2$, $_2S_2$, $_1S_3$, $_0S_4$, $_1S_4$, $_2S_4$, $_4S_4$, $_0S_7$, $_1S_8$, $_0S_{25}$, $_0S_{49}$, $_0S_{73}$, $_0S_{97}$, $_0T_7$, $_0T_{14}$, $_0T_{27}$, $_0T_{53}$, $_0T_{105}$. The $\mathscr{G}$-acceptable model $\rho_0(r)$ from which we start is again model 5821. The averaging kernels $A(r_0, r)$ obtained from the δ-ness criterion (3.9), (4.1), together with their resolving lengths $L(r_0)$ (see equation (7.2)), are shown in Fig. 16, while the resulting local averages of $\rho_0(r)$ are shown in Fig. 17. The set $_n\overline{\mathscr{G}}_l^{ST}$ has poorer resolution in the core and better resolution in the mantle than the set $_n\overline{\mathscr{G}}_l^{S}$. As a result, $_n\overline{\mathscr{G}}_l^{ST}$ partly resolves the density inversion in the upper mantle. Therefore, if this inversion were real, and if $v_P(r)$ and $v_S(r)$ were correct for the real Earth, we would already have enough resolving power in the 26 gross Earth data of $_n\overline{\mathscr{G}}_l^{ST}$ to detect the density inversion.
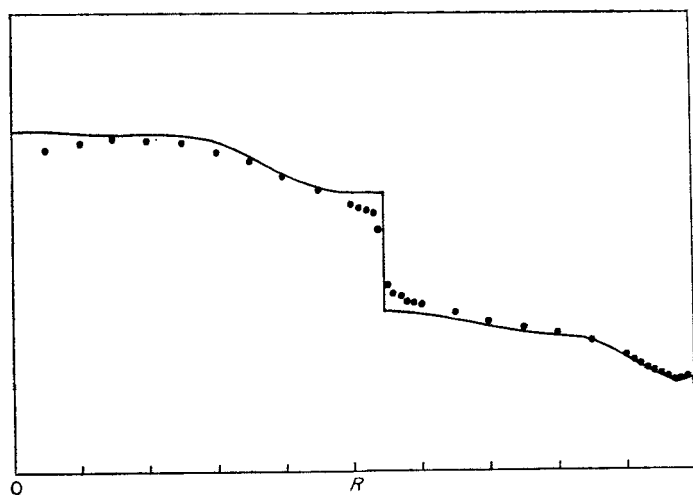

FIG. 17. The same as Fig. 8 except that the averaging kernels in Fig. 16 are used instead of those in Fig. 7.