Theoretical Background for the Inversion of Seismic Waveforms, Including Elasticity and Attenuation

Albert Tarantola

Institut de Physique du Globe, 4 place Jussieu, 75005 Paris, France.

Pure and Applied Geophysics, Vol. 128, Nos. 1/2, 1988, p. 365–399

Abstract—To account for elastic and attenuating effects in the elastic wave equation, the stress-strain relationship can be defined through a general, anisotropic, causal relaxation function $\psi^{ijkl}(\mathbf{x}, \tau)$. Then, the wave equation operator is not necessarily symmetric ('self-adjoint'), but the reciprocity property is still satisfied. The representation theorem contains a term proportional to the history of strain. The dual problem consists of solving the wave equation with *final* time conditions and an anti-causal relaxation function. The problem of interpretation of seismic waveforms can be set as the nonlinear inverse problem of estimating the matter density $\rho(\mathbf{x})$ and all the functions $\psi^{ijkl}(\mathbf{x}, \tau)$. This inverse problem can be solved using iterative gradient methods, each iteration consisting of the propagation of the actual source in the current medium, with causal attenuation, the propagation of the residuals—acting as if they were sources— backwards in time, with anti-causal attenuation, and the correlation of the two wavefields thus obtained.

Key words: Inversion, waveforms, attenuation, Green's function, representation theorem, dual conditions, reciprocity theorem.

1. Introduction

The problem of interpretation of seismic waveforms can be set as the problem of obtaining the earth model which best predicts the actually observed seismograms. This opens two questions: (i) given an earth model, how to solve the forward problem of predicting seismograms? and (ii) how to solve the inverse problem of obtaining the optimum earth model?

The tools for predicting seismograms are the elastic wave equation and the numerical methods developed to obtain solutions, as for instance finite-difference approximations to derivatives. Finite-difference approximations to the wave equation have the advantages of having enough flexibility to be almost model-independent and of accounting, in principle, for a diversity of waves. They are expensive, but nicely adaptable to the newly emerging class of massively parallel computers.

The inverse problem is essentially an optimization problem in a functional space. Difficulties arise because the problem is large sized, and the functional to be minimized is nonquadratic. The modest capabilities of present-day computers prevent the use of true nonquadratic methods of optimization, like Monte Carlo methods. Gradient methods can be used which lead to elegant results.

In this paper I first review the mathematics of the forward problem that are useful for the inverse problem; wave equation, Green's function, and representation theorem. Secondly, I review the mathematics of functional least squares. Finally, I give the solution to the seismic inverse problem, with more generality than in my previous papers, because here I take into account attenuation.

In the problem of interpretation of seismic reflection data the model space has many degrees of freedom (millions to billions), and methods of inversion based on a naïve use of least-squares formulas do not work. In particular, matrix algebra must *not* be used, and partial (or Fréchet) derivatives of data with respect to model parameters should *not* be computed. Much work has to be done analytically, in order to interpret the final formulas of least squares as operations involving only wave propagations, and no linear algebra computations.

For developing a theory for inversion including attenuation we must first choose a model. If in the perfectly elastic approximation it is clear that density $\rho(\mathbf{x})$ and elastic stiffnesses $c^{ijkl}(\mathbf{x})$ (or related quantities) are the right earth parameters to choose, for a more realistic approximation including attenuation, the choice is not so clear, as many models for attenuation exist. I take here the most optimistic point of view: that data sets exist which contain enough information rendering a particular model of attenuation unnecessary, and that the more general parameterization can be chosen: an arbitrary relaxation function $\psi^{ijkl}(\mathbf{x},\tau)$. The goal of inversion is then to obtain the density $\rho(\mathbf{x})$ and the functions $\psi^{ijkl}(\mathbf{x},\tau)$. Of course, some constraints have to be imposed on the relaxation functions, as for instance causality and symmetry conditions. If necessary, some soft constraints can also be imposed, as for instance spatial or temporal smoothness.

Let $u^i(\mathbf{x},t)$ be the i-th component of displacement at point $\mathbf{x}$ and time t. If $\mathbf{x}_r$ $(r = 1, 2, \ldots)$ denote the receiver locations, a possible criterion of goodness of fit between observed and computed seismograms is the minimization of

$$S = \sum_r \int_{t_0}^{t_1} dt \sum_i \left| u^i\big(\mathbf{x}_r,t\big)_{\text{obs}} - \big(\mathbf{x}_r,t\big)_{\text{cal}} \right|, \tag{1a}$$

where 'obs' and 'cal' respectively represent the observed and the calculated displacements from a given earth model. Although results given by this criterion are fairly good, they are difficult to obtain, and this criterion is replaced by the least-squares criterion of minimization of

$$S = \sum_r \int_{t_0}^{t_1} dt \sum_i \Big(u^i\big(\mathbf{x}_r,t\big)_{\text{obs}} - u^i\big(\mathbf{x}_r,t\big)_{\text{cal}} \Big)^2, \tag{1b}$$

which gives less robust results but computations which are manageable with present: day computers.

In Section 2 the rate-of-relaxation function is defined, which will be at the center of our mathematical developments. Section 3 rapidly reviews the fundamental equation: the elastic wave equation with attenuation. In Section 4 I recall the definition of the transposed (and adjoint) of an operator, and in Section 5 the transposed of the wave equation operator is given. The Green function is introduced in Section 6, leading to the general representation theorems of Section 7, which are independent of the reciprocity relations shown in Section 8. Section 9 reviews rapidly the Born approximation, useful for computing the gradient of the least-squares misfit function. A very formal version of the least-squares theory in functional spaces is developed in Section 10, and is applied to the problem of interpretation of seismic waveforms in Section 11. The formula allowing waveform inversion for obtaining the rate-of-relaxation function is, to my knowledge, original. The last Section addresses the results obtained.

2. The Rate-of-Relaxation Function

The most general *linear* relationship between stress, $\sigma^{ij}(\mathbf{x},t)$, and strain $\varepsilon^{ij}(\mathbf{x},t)$, can be described using a kernel $\Psi_1^{ijkl}(\mathbf{x},t;\mathbf{x}',t')$:

$$\sigma^{ij}(\mathbf{x},t) = \int_V dV(\mathbf{x}') \int_{-\infty}^{+\infty} dt' \Psi_1^{ijkl}(\mathbf{x},t;\mathbf{x}',t')\varepsilon^{kl}(\mathbf{x}',t'), \tag{2a}$$

which has to be causal, must have some symmetries, and may be a distribution (containing the delta 'function' and/or its derivatives). This is too general for most seismic purposes. Assuming that the stress-strain relationship is local,

$$\Psi_1^{ijkl}(\mathbf{x},t;\mathbf{x}',t') = \Psi_0^{ijkl}(\mathbf{x};t,t')\delta(\mathbf{x}-\mathbf{x}'),$$

gives

$$\sigma^{ij}(\mathbf{x},t) = \int_{-\infty}^{+\infty} dt' \Psi_0^{ijkl}(\mathbf{x};t,t')\varepsilon^{kl}(\mathbf{x},t'). \tag{2b}$$

If the medium properties do not depend on time,

$$\Psi_0^{ijkl}(\mathbf{x};t,t') = \Psi^{ijkl}(\mathbf{x},t-t'),$$

and

$$\sigma^{ij}(\mathbf{x},t) = \int_{-\infty}^{+\infty} dt' \Psi^{ijkl}(\mathbf{x},t-t')\varepsilon^{kl}(\mathbf{x},t'). \tag{2c}$$

The function $\Psi^{ijkl}(\mathbf{x},\tau)$ has to satisfy causality,

$$\Psi^{ijkl}(\mathbf{x},\tau) = 0 \quad \text{for} \quad \tau < 0, \tag{3}$$

and is assumed to have symmetries:

$$\Psi^{ijkl}(\mathbf{x},\tau) = \Psi^{jikl}(\mathbf{x},\tau) = \Psi^{klij}(\mathbf{x},\tau), \tag{4}$$

(from where it follows $\Psi^{ijkl}(\mathbf{x},\tau) = \Psi^{ijlk}(\mathbf{x},\tau) = \Psi^{jikl}(\mathbf{x},\tau)$). Notice that the causality property allows writing (2c) as

$$\sigma^{ij}(\mathbf{x},t) = \int_{-\infty}^{t^+} dt' \Psi^{ijkl}(\mathbf{x},t-t')\varepsilon^{kl}(\mathbf{x},t').$$

Instead of the function $\Psi^{ijkl}(\mathbf{x},\tau)$ it is customary to use the *creep function* $\phi^{ijkl}(\mathbf{x},\tau)$ defined by

$$\varepsilon^{ij}(\mathbf{x},t) = \int_{-\infty}^{t} dt' \phi^{ijkl}(\mathbf{x},t-t')\dot{\sigma}^{kl}(\mathbf{x},t'), \tag{5}$$

or its inverse, the *relaxation function* $\psi^{ijkl}(\mathbf{x},\tau)$ defined by

$$\sigma^{ij}(\mathbf{x},t) = \int_{-\infty}^{t} dt' \psi^{ijkl}(\mathbf{x},t-t')\dot{\varepsilon}^{kl}(\mathbf{x},t'), \tag{6}$$

where a dot denotes time differentiation. As we have

$$\Psi^{ijkl}(\mathbf{x},\tau) = \dot{\psi}^{ijkl}(\mathbf{x},\tau), \tag{7}$$

$\Psi^{ijkl}(\mathbf{x},\tau)$ can be named the *rate-of-relaxation function.*

***Example* 1.** *Perfect elasticity.* Choosing

$$\Psi^{ijkl}(\mathbf{x},\tau) = c^{ijkl}(\mathbf{x})\delta(\tau), \tag{8}$$

where $\delta(\cdot)$ is the delta 'function', gives Hooke's law

$$\sigma^{ij}(\mathbf{x},t) = c^{ijkl}(\mathbf{x})\varepsilon^{kl}(\mathbf{x},t). \tag{9}$$

For isotropic media,

$$c^{ijkl}(\mathbf{x}) = \lambda_e(\mathbf{x})\delta^{ij}\delta^{kl} + \mu_e(\mathbf{x})\big(\delta^{ik}\delta^{jl} + \delta^{il}\delta^{jk}\big), \tag{10}$$

where $\lambda_e(\mathbf{x})$ and $\mu_e(\mathbf{x})$ are the (elastic) parameters of Lamé, related with the elastic bulk modulus by

$$3k_e(\mathbf{x}) = 3\lambda_e(\mathbf{x}) + 2\mu_e(\mathbf{x}). \tag{11}$$

As a first approximation, for usual rocks, $\lambda_e \simeq \mu_e$.

***Example* 2.** *Elasticity with viscosity.* Choosing

$$\Psi^{ijkl}(\mathbf{x},\tau) = c^{ijkl}(\mathbf{x})\delta(\tau) - d^{ijkl}(\mathbf{x})\dot{\delta}(\tau), \tag{12}$$

gives the Kelvin-Voigt law

$$\sigma^{ij}(\mathbf{x},t) = c^{ijkl}(\mathbf{x})\varepsilon^{kl}(\mathbf{x},t) + d^{ijkl}(\mathbf{x})\dot{\varepsilon}^{kl}(\mathbf{x},t), \tag{13}$$

which corresponds, in a 1D problem, to a perfectly elastic spring and a perfectly viscous dashpot in parallel. For isotropic viscosity,

$$d^{ijkl}(\mathbf{x}) = \lambda_v(\mathbf{x})\delta^{ij}\delta^{kl} + \mu_v(\mathbf{x})\big(\delta^{ik}\delta^{jl} + \delta^{il}\delta^{jk}\big). \tag{14}$$

The viscous bulk modulus is defined by

$$3k_v(\mathbf{x}) = 3\lambda_v(\mathbf{x}) + 2\mu_v(\mathbf{x}). \tag{15}$$

As a first approximation, for usual rocks, $3k_v = 3\lambda_v + 2\mu_v \simeq 0$.

***Example* 3.** *Constant Q*. In a one-dimensional example, KJARTANSSON (1979) shows that the rate-of-relaxation function

$$\Psi(\tau) \propto \frac{1}{\tau^{1+2\gamma}} \quad \text{for} \quad \tau > 0 \tag{16}$$

$$\Psi(\tau) = 0 \quad \text{for} \quad \tau \leq 0 \tag{17}$$

implies a quality factor Q strictly independent on the frequency, and, for a sufficiently small value of the positive parameter γ, fits most of seismic data.

3. The Fundamental Equation

Let us be interested in the description of elastic waves propagating inside a volume $\mathbf{V}$, surrounded by a surface $\mathbf{S}$. Points inside $\mathbf{V}$ will be denoted $\mathbf{x}, \mathbf{x}', \ldots$ while points on $\mathbf{S}$ will be denoted $\xi, \xi', \ldots$. From the fundamental laws of physics we can show (e.g., DAUTRAY and LIONS, 1984) that if we impose to a medium with matter density $\rho(\mathbf{x})$ and rate-of-relaxation function $\Psi^{ijkl}(\mathbf{x}, \tau)$, a volume density of force $\phi^i(\mathbf{x}, t)$, a moment density $M^{ij}(\mathbf{x}, t)$, and a surface traction $\tau^i(\xi, \tau)$, then the displacement field $u^i(\mathbf{x}, t)$ satisfies at any point inside $\mathbf{V}$ the relationship

$$\rho(\mathbf{x}) \frac{\partial^2 u^i}{\partial t^2}(\mathbf{x}, t) - \frac{\partial \sigma^{ij}}{\partial x^j}(\mathbf{x}, t) = \phi^i(\mathbf{x}, t) \tag{18}$$

where

$$\sigma^{ij}(\mathbf{x}, t) = M^{ij}(\mathbf{x}, t) + \int_{-\infty}^{+\infty} dt' \Psi^{ijkl}(\mathbf{x}, t - t') \frac{\partial u^k}{\partial x^l}(\mathbf{x}, t'), \tag{19}$$

and, at the surface,

$$n^j(\xi) \sigma^{ij}(\xi, t) = \tau^i(\xi, t). \tag{20}$$

Notice that $\sigma^{ij}(\mathbf{x}, t)$ is the *total stress* (internal plus external origin) and that in the equations of the previous section the external stress $M^{ij}(\mathbf{x}, t)$ was implicitly assumed to be zero.

4. Mathematical Preliminary: Symmetric and Self-Adjoint Operators

Let $\mathbf{U}$ be a linear space, whose elements, denoted $\mathbf{u}$, are named vectors. For instance each element of $\mathbf{U}$ may be a displacement field $u^i(\mathbf{x}, t)$. A *Linear form* over $\mathbf{U}$ is a linear application from $\mathbf{U}$ into a space of scalars (identical to the real numbers of $\mathbf{R}$ excepted in that the scalars may have a physical dimension). We will say that a linear space $\hat{\mathbf{U}}$ is *a dual* of $\mathbf{U}$ if each element of $\hat{\mathbf{U}}$ defines a linear form over $\mathbf{U}$ (I prefer that definition to the traditional definition of mathematicians where *the* dual is the space of *all* linear forms). Let $\hat{\mathbf{u}}_0 \in \hat{\mathbf{U}}$. The scalar associated to any $\mathbf{u} \in \mathbf{U}$ by $\hat{\mathbf{u}}_0$ is denoted by $\langle \hat{\mathbf{u}}_0, \mathbf{u} \rangle_U$. Alternatively, each element $\mathbf{u}_0 \in \mathbf{U}$ defines a linear form over $\hat{\mathbf{U}}$ through

$$\left\langle \mathbf{u}_0, \hat{\mathbf{u}} \right\rangle_{\hat{U}} = \left\langle \hat{\mathbf{u}}, \mathbf{u}_0 \right\rangle_U. \tag{21}$$

Let $\mathbf{U}$ and Φ be two linear spaces, and $\mathbf{L}$ a linear operator from $\mathbf{U}$ into Φ. For instance, $\mathbf{L}$ may be the wave equation operator defined in Section 5. An operator $\mathbf{L}^t$ mapping $\hat{\Phi}$ into $\hat{\mathbf{U}}$ will be named *the transposed* of $\mathbf{L}$ if for any $\hat{\phi} \in \hat{\Phi}$ and any $\mathbf{u} \in \mathbf{U}$

$$\langle \hat{\phi}, \mathbf{L}\mathbf{u} \rangle_\Phi = \langle \mathbf{L}^t \hat{\phi}, \mathbf{u} \rangle_U. \tag{22}$$

This definition of transposed operators may be compared with the definition of adjoint operators. The adjoint of an operator may only be defined if the spaces in consideration have a scalar product. Let for instance the linear operator $\mathbf{L}$ map the linear space $\mathbf{U}$ into the linear space Φ, and let $(\mathbf{u}_1, \mathbf{u}_2)_U$ and $(\phi_1, \phi_2)_\Phi$ denote respectively the scalar products over $\mathbf{U}$ and Φ. An operator $\mathbf{L}^*$ mapping Φ into $\mathbf{U}$ will be named *the adjoint* of $\mathbf{L}$ if for any Φ and $\mathbf{u}$,

$$(\phi, \mathbf{L}\mathbf{u})_\Phi = (\mathbf{L}^* \phi, \mathbf{u})_U. \tag{23}$$

Details on the mathematical definition of transposed and adjoint operators may be found for instance in TAYLOR and LAY (1980).

Transposed operators have an important and very simple property which follows immediately from the definition of the *kernel* of an operator: if two operators are mutually transposed, their integral kernels are identical, excepted in that their variables are 'transposed'.

It can easily be seen that if $\mathbf{W}_U$ and $\mathbf{W}_\Phi$ are the operators defining the scalar products over $\mathbf{U}$ and Φ, respectively:

$$(\mathbf{u}_1, \mathbf{u}_0)_U = \langle \mathbf{W}_U \mathbf{u}_1, \mathbf{u}_0 \rangle_U \tag{24a}$$

$$(\phi_1, \phi_0)_\Phi = \langle \mathbf{W}_\Phi \phi_1, \phi_0 \rangle_\Phi, \tag{24b}$$

then transposed and adjoint are related by

$$\mathbf{L}^* = \mathbf{W}_U^{-1} \mathbf{L}^t \mathbf{W}_\Phi. \tag{25}$$

By definition, if $\mathbf{L}$ maps $\mathbf{U}$ into Φ, then $\mathbf{L}^t$ maps $\hat{\Phi}$ into $\hat{\mathbf{U}}$. In the particular case where Φ, can be identified with a dual of $\mathbf{U}$, both $\mathbf{L}$ and $\mathbf{L}^t$ map $\mathbf{U}$ into $\hat{\mathbf{U}}$. In that case, the definition of transpose (22) can be rewritten

$$\langle \overleftarrow{\mathbf{u}}, \mathbf{L} \overrightarrow{\mathbf{u}} \rangle_{\hat{U}} = \langle \mathbf{L}^t \overleftarrow{\mathbf{u}}, \overrightarrow{\mathbf{u}} \rangle_U, \tag{26}$$

with $\overrightarrow{\mathbf{u}}$ and $\overleftarrow{\mathbf{u}}$ elements of $\mathbf{U}$. If, in that case, $\mathbf{L}$ and $\mathbf{L}^t$ have the same domain of definition, and if for any $\overrightarrow{\mathbf{u}}$ and $\overleftarrow{\mathbf{u}}$ we can replace in (26) $\mathbf{L}^t$ by $\mathbf{L}$:

$$\langle \overleftarrow{\mathbf{u}}, \mathbf{L} \overrightarrow{\mathbf{u}} \rangle_{\hat{U}} = \langle \mathbf{L} \overleftarrow{\mathbf{u}}, \overrightarrow{\mathbf{u}} \rangle_U, \tag{27}$$

we say that $\mathbf{L}$ is *symmetric,* and we write

$$\mathbf{L} = \mathbf{L}^t. \tag{28}$$

By definition, if $\mathbf{L}$ maps $\mathbf{U}$ into itself, then $\mathbf{L}^*$ also maps $\mathbf{U}$ into itself. In that case, the definition of adjoint (23) can be rewritten

$$(\overleftarrow{\mathbf{u}}, \mathbf{L} \overrightarrow{\mathbf{u}})_U = (\mathbf{L}^* \overleftarrow{\mathbf{u}}, \overrightarrow{\mathbf{u}})_U. \tag{29}$$

If, in that case, $\mathbf{L}$ and $\mathbf{L}^*$ have the same domain of definition, and if for any $\overrightarrow{\mathbf{u}}$ and $\overleftarrow{\mathbf{u}}$ we can replace in (29) $\mathbf{L}^*$ by $\mathbf{L}$:

$$(\overleftarrow{\mathbf{u}}, \mathbf{L}\overrightarrow{\mathbf{u}})_U = (\mathbf{L}\overleftarrow{\mathbf{u}}, \overrightarrow{\mathbf{u}})_U, \tag{30}$$

we say that $\mathbf{L}$ is *self-adjoint,* and we write

$$\mathbf{L} = \mathbf{L}^*. \tag{31}$$

In the problem of wave propagation later studied, there is no natural scalar product, so the general concept of transposed operator will be preferred to the more particular concept of adjoint operator.

Practically, when we have an operator $\mathbf{L}$, we can use the two following rules to obtain the formal *transposed:*

(i) a derivative operator is anti-symmetric, i.e., its transposed equals its opposite.

(ii) if $\mathbf{L}$ is an integral operator, we can introduce its integral kernel; the transposed operator is also an integral operator and the kernel of the transposed operator is the transposed of the original kernel.

Once we have the formal transposed, we compute the difference

$$\langle \hat{\phi}, \mathbf{Lu} \rangle_\Phi - \langle \mathbf{L}^t \hat{\phi}, \mathbf{u} \rangle_U,$$

and if necessary, we impose restrictions on the spaces $\mathbf{U}$ and $\hat{\Phi}$ for this difference to vanish. These restrictions are named *dual conditions,* and we will see examples in the next section. Typically, for a differential operator, the dual conditions are boundary conditions. For an integral operator, there are no dual conditions to impose if the integrals defining the linear forms have the same bounds as the integrals defining the operators. If not, we have to impose conditions on the functions outside the bounds.

5. The Wave Equation Operator and its Transpose

Let $\mathbf{U}$ be the space of all conceivable displacement fields $\mathbf{u} = \{u^i(\mathbf{x}, t)\}$, and Φ the space of all conceivable source fields $\phi = \{\phi^i(\mathbf{x}, t)\}$. The *wave equation operator* (with attenuation) is the operator $\mathbf{L}$ mapping $\mathbf{U}$ into Φ defined by

$$(\mathbf{Lu})^i(\mathbf{x}, t) = \rho(\mathbf{x}) \frac{\partial^2 u^i}{\partial t^2}(\mathbf{x}, t) - \frac{\partial}{\partial x^j} \int_{-\infty}^{+\infty} dt' \overrightarrow{\Psi}_0^{jikl}(\mathbf{x}; t, t') \frac{\partial u^k}{\partial x^l}(\mathbf{x}, t'), \tag{32}$$

where $\overrightarrow{\Psi}_0^{jikl}(\mathbf{x}; t, t')$ is causal (the right arrow is to distinguish this function from an anti-causal function to be introduced later). This allows the wave equation with attenuation defined by equations (18)–(19) to be written as:

$$\mathbf{Lu} = \tilde{\phi}, \tag{33}$$

where

$$\tilde{\phi}^i(\mathbf{x}, t) = \phi^i(\mathbf{x}, t) + \frac{\partial \mathbf{M}^{ij}}{\partial x^j}(\mathbf{x}, t). \tag{34}$$

In the definition (32) we choose the function $\overrightarrow{\Psi}_0^{jikl}(\mathbf{x};t,t')$ defined in (2b), rather than the rate-of-relaxation function $\overrightarrow{\Psi}^{jikl}(\mathbf{x},\tau)$ defined in (2c) because it is not more difficult to handle and will later help to clarify the properties of the Green function. To simplify notations, the index $({}_0)$ in $\overrightarrow{\Psi}_0^{jikl}(\mathbf{x};t,t')$ will be dropped in what follows.

For a given $\phi \in \Phi$ we can define, for any $\mathbf{u} \in \mathbf{U}$, the scalar

$$\langle \phi, \mathbf{u} \rangle_U = \int_V dV(\mathbf{x}) \int_{t_0}^{t_1} dt \phi^i(\mathbf{x},t) u^i(\mathbf{x},t), \tag{35}$$

which has the dimension of an *action* (energy × time). As each element of Φ defines a linear form over $\mathbf{U}$, we can say that Φ is a dual of $\mathbf{U}$. Alternatively, for a given $\mathbf{u} \in \mathbf{U}$ we can define, for any $\phi \in \Phi$, the action

$$\langle \mathbf{u}, \phi \rangle_\Phi = \langle \phi, \mathbf{u} \rangle_U = \int_V dV(\mathbf{x}) \int_{t_0}^{t_1} dt \phi^i(\mathbf{x},t) u^i(\mathbf{x},t). \tag{36}$$

The spaces $\mathbf{U}$ and Φ are mutually duals.

Equations (32)–(33) define the linear operator $\mathbf{L}$, mapping $\mathbf{U}$ into Φ. As $\mathbf{U}$ and Φ are mutually duals, the transposed operator $\mathbf{L}^t$ also maps $\mathbf{U}$ into Φ. Let us demonstrate that the transposed of $\mathbf{L}$ is the operator $\mathbf{L}^t$ defined by

$$\big(\mathbf{L}^t \mathbf{u}\big)^i(\mathbf{x},t) = \rho(\mathbf{x}) \frac{\partial^2 u^i}{\partial t^2}(\mathbf{x},t) - \frac{\partial}{\partial x^j} \int_{-\infty}^{+\infty} dt' \overleftarrow{\Psi}_0^{ijkl}(\mathbf{x};t,t') \frac{\partial u^k}{\partial x^l}(\mathbf{x},t'), \tag{37}$$

where $\overleftarrow{\Psi}_0^{ijkl}(\mathbf{x};t,t')$ is the anti-causal. rate-of-relaxation function defined by

$$\overleftarrow{\Psi}_0^{ijkl}(\mathbf{x};t,t') = \overrightarrow{\Psi}_0^{ijkl}(\mathbf{x};t',t), \tag{38}$$

which means that the operator $\mathbf{L}^t$ corresponds to a wave equation with negative attenuation.

For we have to verify that (26) is satisfied. We have

$$\begin{aligned} \langle \overleftarrow{\mathbf{u}}, \mathbf{L}\overrightarrow{\mathbf{u}} \rangle_\Phi - \langle \mathbf{L}^t \overleftarrow{\mathbf{u}}, \overrightarrow{\mathbf{u}} \rangle_U = & \int_V dV(\mathbf{x}) \int_{t_0}^{t_1} dt \overleftarrow{u}_i^j(\mathbf{x},t) (\mathbf{L}\overrightarrow{\mathbf{u}})_i^j(\mathbf{x},t) \\ & - \int_V dV(\mathbf{x}) \int_{t_0}^{t_1} dt \big(\mathbf{L}^t \overleftarrow{\mathbf{u}}\big)^i(\mathbf{x},t) \overrightarrow{\mathbf{u}}^i(\mathbf{x},t). \end{aligned}$$

Inserting (32) and (37), integrating per parts, and using the divergence theorem gives (see Appendix A):

$$\begin{aligned} \langle \overleftarrow{\mathbf{u}}, \mathbf{L}\overrightarrow{\mathbf{u}} \rangle_\Phi - \big\langle \mathbf{L}^t \overleftarrow{\mathbf{u}}, \overrightarrow{\mathbf{u}} \big\rangle_U = & + \int_V dV(\mathbf{x}) \big(\overleftarrow{u}^i(\mathbf{x},t) \overrightarrow{p}^{\,i}(\mathbf{x},t) - \overleftarrow{p}^i(\mathbf{x},t) \overrightarrow{u}^i(\mathbf{x},t) \big) \Big|_{t=t_0}^{t=t_1} \\ & + \int_s dS(\xi) \int_{t_0}^{t_1} dt \big(\overleftarrow{\tau}^i(\xi,t) \overrightarrow{u}^i(\xi,t) - \overleftarrow{u}^i(\xi,t) \overrightarrow{\tau}^i(\mathbf{x},t) \big) \\ & + \int_V dV(\mathbf{x}) \int_{t_0}^{t_1} dt \big(\overleftarrow{\varepsilon}^{ij}(\mathbf{x},t) \overrightarrow{\Sigma}^{ij}(\mathbf{x},t) - \overleftarrow{\Sigma}^{ij}(\mathbf{x},t) \overrightarrow{\varepsilon}^{\,ij}(\mathbf{x},t) \big), \end{aligned} \tag{39}$$

where

$$\overrightarrow{\Sigma}^{ij}(\mathbf{x},t) = \int_{-\infty}^{t_0} dt' \overrightarrow{\Psi}^{ijkl}(\mathbf{x};t,t') \overrightarrow{\varepsilon}^{kl}(\mathbf{x},t')$$
$$\overleftarrow{\Sigma}^{ij}(\mathbf{x},t) = \int_{t_1}^{+\infty} dt' \overleftarrow{\Psi}^{ijkl}(\mathbf{x};t,t') \overleftarrow{\varepsilon}^{kl}(\mathbf{x},t'), \tag{40a}$$

and, where for each sense of the arrows $\rightarrow$ and $\leftarrow$,

$$p^i(\mathbf{x},t) = \rho(\mathbf{x}) \frac{\partial u^i}{\partial t}(\mathbf{x},t) \tag{40b}$$

$$\varepsilon^{ij}(\mathbf{x},t) = \frac{1}{2}\left(\frac{\partial u^i}{\partial x^j}(\mathbf{x},t) + \frac{\partial u^j}{\partial x^i}(\mathbf{x},t)\right) \tag{40c}$$

$$\sigma^{ij}(\mathbf{x},t) = \int_{-\infty}^{+\infty} dt' \Psi^{ijkl}(\mathbf{x};t,t') \varepsilon^{kl}(\mathbf{x},t') \tag{40d}$$

and

$$\tau^i(\xi,t) = n^j(\xi)\sigma^{ij}(\xi,t) \quad \text{for} \quad \xi \in \mathbf{S}. \tag{40e}$$

$\overrightarrow{\Sigma}^{ij}(\mathbf{x},t)$ corresponds to the stress at time t due to the history of the strain $\overrightarrow{\varepsilon}^{ij}(\mathbf{x},t)$ before t_0, while $\overleftarrow{\Sigma}^{ij}(\mathbf{x},t)$ corresponds to the future of the strain $\overleftarrow{\varepsilon}^{ij}(\mathbf{x},t)$ after t_1.

If we restrict the domains of definition of $\mathbf{L}$ and $\mathbf{L}^t$ respectively to subspaces $\overrightarrow{\mathbf{U}}$ and $\overleftarrow{\mathbf{U}}$ such that for any $\overrightarrow{\mathbf{u}} \in \overrightarrow{\mathbf{U}}$ and $\overleftarrow{\mathbf{u}} \in \overleftarrow{\mathbf{U}}$ the right-hand side terms in (39) vanish, then $\mathbf{L}^t$ is the transposed of $\mathbf{L}$. The elements of $\overrightarrow{\mathbf{U}}$ and $\overleftarrow{\mathbf{U}}$ satisfy then *dual conditions.*

First example of dual conditions: If the field $\overrightarrow{\mathbf{u}}$ has a quiescent past:

$$\overrightarrow{u}^i(\mathbf{x},t) = 0 \quad \text{for} \quad t \le t_0, \tag{41a}$$

$$\frac{\partial \overrightarrow{u}^i}{\partial t}\big(\mathbf{x},t_0\big) = 0, \tag{41b}$$

and satisfies a condition of free surface (computed with *positive* attenuation):

$$n^i(\xi) \int_{-\infty}^{+\infty} dt' \overrightarrow{\Psi}_0^{ijkl}(\xi;t,t') \frac{\partial \overrightarrow{u}^k}{\partial x^l}(\xi,t') = 0 \quad \text{for} \quad \xi \in \mathbf{S}, \tag{41c}$$

and the field $\overleftarrow{\mathbf{u}}$ has a quiescent future:

$$\overleftarrow{u}^i\big(\mathbf{x},t_1\big) = 0 \quad \text{for} \quad t \ge t_1, \tag{42a}$$

$$\frac{\partial \overleftarrow{u}^i}{\partial t}\big(\mathbf{x},t_1\big) = 0, \tag{42b}$$

and satisfies a condition of free surface (computed with *negative* attenuation):

$$n^i(\xi) \int_{-\infty}^{+\infty} dt' \overleftarrow{\Psi}_0^{ijkl}(\xi;t,t') \frac{\partial \overrightarrow{u}^k}{\partial x^l}(\xi,t') = 0 \quad \text{for} \quad \xi \in \mathbf{S}, \tag{42c}$$

then, the integrals in (39) vanish: conditions (41)–(42) are *dual conditions.*

Second example of dual conditions: If in the example above we assume conditions of rigid, instead of free surface:

$$\overrightarrow{u}^i(\xi, t) = 0 \quad \text{for} \quad \xi \in \mathbf{S}, \tag{43}$$

$$\overleftarrow{u}^i(\xi, t) = 0 \quad \text{for} \quad \xi \in \mathbf{S}, \tag{44}$$

the integrals in (40) also vanish: conditions (43)–(44), together with the conditions (41a), (41b)–(42a), (42b) are *dual conditions.*

6. The Green Function

Consider the operator $\mathbf{L}_{\text{free}}$ defined by the restriction of the formal definition of $\mathbf{L}$ (32) to the subspace $\overrightarrow{\mathbf{U}} \subset \mathbf{U}$ of fields satisfying the homogeneous initial conditions and the conditions of free surface defined by equations (41). Then, for any ϕ belonging to the source space Φ, the equation

$$\mathbf{L}_{\text{free}} \overrightarrow{\mathbf{u}} = \phi \tag{45}$$

has one solution $\overrightarrow{\mathbf{u}}$, and only one (if we limit our consideration to functions regular enough).

This allows $\mathbf{L}_{\text{free}}^{-1}$ to be defined, the inverse of $\mathbf{L}_{\text{free}}$. It is named the Green operator, and is denoted $\overrightarrow{\mathbf{G}}_{\text{free}}$:

$$\overrightarrow{\mathbf{G}}_{\text{free}} = \mathbf{L}_{\text{free}}^{-1}. \tag{46}$$

Equation (45) is then solved formally by

$$\overrightarrow{\mathbf{u}} = \overrightarrow{\mathbf{G}}_{\text{free}} \phi. \tag{47}$$

An integral representation of (47) is written

$$\overrightarrow{u}^i(\mathbf{x}, t) = \int_V dV(\mathbf{x}') \int_{t_0}^{t_1} dt' \overrightarrow{\mathbf{G}}^{ij}_{\text{free}}(\mathbf{x}, t; \mathbf{x}', t') \phi^j(\mathbf{x}', t'), \tag{48}$$

and $\overrightarrow{\mathbf{G}}^{ij}_{\text{free}}(\mathbf{x}, t; \mathbf{x}, t')$, the kernel of the Green operator, is named the Green function.

Consider now the operator $\mathbf{L}^t_{\text{free}}$, defined by the restriction of the formal definition (32) to the subspace $\overleftarrow{\mathbf{U}} \subset \mathbf{U}$ of fields satisfying the homogeneous final conditions and the conditions of free surface (with negative attenuation) defined by equations (42). Again, the equation

$$\mathbf{L}^t_{\text{free}} \overleftarrow{\mathbf{u}} = \phi \tag{49}$$

has one solution, and only one. Let us denote $\overleftarrow{\mathbf{G}}_{\text{free}}$ the inverse of $\mathbf{L}^t_{\text{free}}$:

$$\overleftarrow{\mathbf{G}}_{\text{free}} = \left(\mathbf{L}^t_{\text{free}}\right)^{-1}. \tag{50}$$

Equation (49) is solved formally by

$$\overleftarrow{\mathbf{u}} = \overleftarrow{\mathbf{G}}_{\text{free}} \phi, \tag{51}$$

or, introducing the kernel of $\overleftarrow{\mathbf{G}}_{\text{free}}$,

$$\overleftarrow{u}^i(\mathbf{x}, t) = \int_V dV(\mathbf{x}') \int_{t_0}^{t_1} dt' \overleftarrow{\mathbf{G}}^{ij}_{\text{free}}(\mathbf{x}, t; \mathbf{x}', t') \phi^j(\mathbf{x}', t'). \tag{52}$$

By definition, we have

$$\mathbf{L}_{\text{free}} \overrightarrow{\mathbf{G}}_{\text{free}} = \mathbf{I}, \tag{53a}$$

$$\mathbf{L}^t_{\text{free}} \overleftarrow{\mathbf{G}}_{\text{free}} = \mathbf{I}. \tag{53b}$$

Using the definitions of $\mathbf{L}_{\text{free}}$ and $\mathbf{L}^t_{\text{free}}$ and the kernels of $\overrightarrow{\mathbf{G}}_{\text{free}}$ and $\overleftarrow{\mathbf{G}}_{\text{free}}$, equations (53) take the explicit form

$$\rho(\mathbf{x}) \frac{\partial^2 \overrightarrow{\mathbf{G}}^{ip}_{\text{free}}}{\partial t^2}(\mathbf{x}, t; \mathbf{x}', t') - \frac{\partial}{\partial x^j} \int_{-\infty}^{+\infty} dt'' \overrightarrow{\Psi}^{ijkl}(\mathbf{x}; t, t'') \frac{\partial \overrightarrow{\mathbf{G}}^{kp}_{\text{free}}}{\partial x^l}(\mathbf{x}, t'', \mathbf{x}', t') = \delta^{ip} \delta(\mathbf{x} - \mathbf{x}') \delta(t - t') \tag{54a}$$

$$n^j(\xi) \int_{-\infty}^{+\infty} dt'' \overrightarrow{\Psi}^{ijkl}(\xi; t, t'') \frac{\partial \overrightarrow{\mathbf{G}}^{kp}_{\text{free}}}{\partial x^l}(\xi, t''; \mathbf{x}', t') = 0 \qquad \xi \in \mathbf{S} \tag{54b}$$

$$\overrightarrow{\mathbf{G}}^{ip}_{\text{free}}(\mathbf{x}, t; \mathbf{x}', t') = 0 \quad \text{for} \quad t \le t' \tag{54c}$$

$$\frac{\partial \overrightarrow{\mathbf{G}}^{ip}_{\text{free}}}{\partial t}(\mathbf{x}, t; \mathbf{x}', t') = 0 \quad \text{for} \quad t \le t', \tag{54d}$$

and

$$\rho(\mathbf{x}) \frac{\partial^2 \overleftarrow{\mathbf{G}}^{ip}_{\text{free}}}{\partial t^2}(\mathbf{x}, t; \mathbf{x}', t') - \frac{\partial}{\partial x^j} \int_{-\infty}^{+\infty} dt'' \overleftarrow{\Psi}^{ijkl}(\mathbf{x}; t, t'') \frac{\partial \overleftarrow{\mathbf{G}}^{kp}_{\text{free}}}{\partial x^l}(\mathbf{x}, t'', \mathbf{x}', t') = \delta^{ip} \delta(\mathbf{x} - \mathbf{x}') \delta(t - t') \tag{55a}$$

$$n^j(\xi) \int_{-\infty}^{+\infty} dt'' \overleftarrow{\Psi}^{ijkl} \frac{\partial \overleftarrow{\mathbf{G}}^{kp}_{\text{free}}}{\partial x^l}(\xi, t''; \mathbf{x}', t') = 0 \qquad \xi \in \mathbf{S} \tag{55b}$$

$$\overleftarrow{\mathbf{G}}^{ip}_{\text{free}}(\mathbf{x}, t; \mathbf{x}', t') = 0 \quad \text{for} \quad t \ge t' \tag{55c}$$

$$\frac{\partial \overleftarrow{\mathbf{G}}^{ip}_{\text{free}}}{\partial t}(\mathbf{x}, t; \mathbf{x}', t') = 0 \quad \text{for} \quad t \ge t'. \tag{55d}$$

As the inverse of the transposed equals the transposed of the inverse,

$$\overleftarrow{\mathbf{G}}_{\text{free}} = \left(\overrightarrow{\mathbf{G}}_{\text{free}}\right)^t, \tag{56a}$$

and, as the kernel of the transposed has transposed variables,

$$\overrightarrow{\mathbf{G}}^{ip}_{\text{free}}(\mathbf{x}, t; \mathbf{x}', t') = \overleftarrow{\mathbf{G}}^{pi}_{\text{free}}(\mathbf{x}', t'; \mathbf{x}, t). \tag{56b}$$

Notice that this is *not* a reciprocity relation: it relates $\overrightarrow{\mathbf{G}}_{\text{free}}$ to $\overleftarrow{\mathbf{G}}_{\text{free}}$, but it does not express an internal symmetry of $\overrightarrow{\mathbf{G}}_{\text{free}}$. The reciprocity relationships are analyzed in Section 8.

If instead of $\mathbf{L}_{\text{free}}$ we define $\mathbf{L}_{\text{rigid}}$ as the restriction of $\mathbf{L}$ to the subspace of fields satisfying the homogeneous initial conditions and the conditions of rigid surface defined by equations (43), and $\mathbf{L}^t_{\text{rigid}}$ as the restriction of $\mathbf{L}^t$ to the subspace of fields satisfying the dual conditions defined by equations (42a), (42b) and (44), we can introduce $\overrightarrow{\mathbf{G}}_{\text{rigid}}$ and $\overleftarrow{\mathbf{G}}_{\text{rigid}}$ as above. The equivalent of equations (54)–(55) is

$$\begin{aligned}\rho(\mathbf{x})\frac{\partial^2\overrightarrow{\mathbf{G}}^{ip}_{\text{rigid}}}{\partial t^2}(\mathbf{x},t;\mathbf{x}',t') - \frac{\partial}{\partial x^j}\int_{-\infty}^{+\infty} dt''\,\overrightarrow{\Psi}^{ijkl}(\mathbf{x};t,t'')\frac{\partial\overrightarrow{\mathbf{G}}^{kp}_{\text{rigid}}}{\partial x^l}(\mathbf{x},t'';\mathbf{x}',t')\\ = \delta^{jp}\delta(\mathbf{x}-\mathbf{x}')\delta(t-t')\end{aligned} \tag{57a}$$

$$\overrightarrow{\mathbf{G}}^{ip}_{\text{rigid}}(\xi,t;\mathbf{x}',t') = 0 \qquad \xi\in\mathbf{S} \tag{57b}$$

$$\overrightarrow{\mathbf{G}}^{ip}_{\text{rigid}}(\mathbf{x},t;\mathbf{x}',t') = 0 \quad \text{for} \quad t\le t' \tag{57c}$$

$$\frac{\partial\overrightarrow{\mathbf{G}}^{ip}_{\text{rigid}}}{\partial t}(\mathbf{x},t;\mathbf{x}',t') = 0 \quad \text{for} \quad t\le t', \tag{57d}$$

and

$$\begin{aligned}\rho(\mathbf{x})\frac{\partial^2\overleftarrow{\mathbf{G}}^{ip}_{\text{rigid}}}{\partial t^2}(\mathbf{x},t;\mathbf{x}',t') - \frac{\partial}{\partial x^j}\int_{-\infty}^{+\infty} dt'\,\overleftarrow{\Psi}^{ijkl}(\mathbf{x};t,t'')\frac{\partial\overleftarrow{\mathbf{G}}^{kp}_{\text{rigid}}}{\partial x^l}(\mathbf{x},t'';\mathbf{x}',t')\\ = \delta^{jp}\delta(\mathbf{x}-\mathbf{x}')\delta(t-t')\end{aligned} \tag{58a}$$

$$\overleftarrow{\mathbf{G}}^{jp}_{\text{rigid}}(\xi,t;\xi',t') = 0 \qquad \xi\in\mathbf{S} \tag{58b}$$

$$\overleftarrow{\mathbf{G}}^{jp}_{\text{rigid}}(\mathbf{x},t;\mathbf{x}',t') = 0 \quad \text{for} \quad t> t' \tag{58c}$$

$$\frac{\partial\overleftarrow{\mathbf{G}}^{ip}_{\text{rigid}}}{\partial t}(\mathbf{x},t;\mathbf{x}',t') = 0 \quad \text{for} \quad t> t', \tag{58d}$$

and we also have

$$\overrightarrow{\mathbf{G}}^{ip}_{\text{rigid}}(\mathbf{x},t;\mathbf{x}',t') = \overleftarrow{\mathbf{G}}^{pi}_{\text{rigid}}(\mathbf{x}',t',\mathbf{x},t). \tag{59}$$

7. Representation Theorems

Let $\overleftarrow{\mathbf{G}}^{ij}(\mathbf{x},t;\mathbf{x}',t')$ be any Green's function satisfying the wave equation associated with the dual problem (i.e., with negative attenuation in our case):

$$\begin{aligned}\rho(\mathbf{x})\frac{\partial^2\overleftarrow{\mathbf{G}}^{ip}}{\partial t^2}(\mathbf{x},t;\mathbf{x}',t') - \frac{\partial}{\partial x^j}\int_{-\infty}^{+\infty} dt''\,\overleftarrow{\Psi}^{ijkl}(\mathbf{x};t,t'')\frac{\partial\overleftarrow{\mathbf{G}}^{kp}}{\partial x^l}(\mathbf{x},t'';\mathbf{x}',t')\\ = \delta^{ip}\delta(\mathbf{x}-\mathbf{x}')\delta(t-t'),\end{aligned} \tag{60a}$$

and with final conditions of rest:

$$\overleftarrow{\mathbf{G}}^{ip}(\mathbf{x},t;\mathbf{x}',t') = 0 \quad \text{for} \quad t \geq t' \tag{60b}$$

$$\frac{\partial \overleftarrow{\mathbf{G}}^{ip}}{\partial t}(\mathbf{x},t;\mathbf{x}',t') = 0 \quad \text{for} \quad t \geq t'. \tag{60c}$$

As no surface conditions have yet been specified, there exists an infinity of such Green's functions.

Let $u^i(\mathbf{x},t)$ be an arbitrary field, not necessarily satisfying a wave equation. For any $t', t_0 < t' < t_1$, we have

$$u^p(\mathbf{x}',t') = \int_V dV(\mathbf{x}) \int_{t_0}^{t_1} dt \delta^{ip} \delta(\mathbf{x}-\mathbf{x}') \delta(t-t') u^i(\mathbf{x},t). \tag{61}$$

Inserting (60a) into (61), using the final conditions (60b)–(60c), integrating per parts, and using the divergence theorem gives (see Appendix B)

$$\begin{aligned} u^p(\mathbf{x}',t') = &\int_V dV(\mathbf{x}) \int_{t_0}^{t_1} dt \overleftarrow{\mathbf{G}}^{ip}(\mathbf{x},t;\mathbf{x}',t') \tilde{\phi}^i(\mathbf{x},t) \\ &+ \int_s dS(\xi) \int_{t_0}^{t_1} dt \overleftarrow{\mathbf{G}}^{ip}(\xi,t;\mathbf{x}',t') \tilde{\tau}^i(\xi,t) \\ &- \int_s dS(\xi) \int_{t_0}^{t_1} dt \left[n^j(\xi) \int_{-\infty}^{+\infty} dt'' \overleftarrow{\Psi}^{ijkl}(\xi,t,t'') \frac{\partial \overleftarrow{\mathbf{G}}^{kp}}{\partial x^l}(\xi,t'',\mathbf{x}',t') \right] u^i(\xi,t) \\ &+ \int_V dV(\mathbf{x}) \rho(\mathbf{x}) \left[\overleftarrow{\mathbf{G}}^{ip}(\mathbf{x},t_0;\mathbf{x}',t') \frac{\partial u^i}{\partial t}(\mathbf{x},t_0) - \frac{\partial \overleftarrow{\mathbf{G}}^{ip}}{\partial t}(\mathbf{x},t_0;\mathbf{x}',t') u^i(\mathbf{x},t_0) \right] \\ &- \int_V dV(\mathbf{x}) \int_{t_0}^{t_1} dt \frac{\partial \overleftarrow{\mathbf{G}}^{ip}}{\partial x^j}(\mathbf{x},t;\mathbf{x}',t') \Sigma^{ij}(\mathbf{x},t) \end{aligned} \tag{62}$$

where

$$\tilde{\phi}^i(\mathbf{x},t) = \rho(\mathbf{x}) \frac{\partial^2 u^i}{\partial t^2}(\mathbf{x},t) - \frac{\partial}{\partial x^j} \int_{-\infty}^{+\infty} dt'' \overrightarrow{\Psi}^{ijkl}(\mathbf{x};t,t'') \frac{\partial u^k}{\partial x^l}(\mathbf{x},t''), \tag{63}$$

$$\Sigma^{ij}(\mathbf{x},t) = \int_{-\infty}^{t_0} dt'' \overrightarrow{\Psi}^{ijkl}(\xi;t,t'') \frac{\partial u^k}{\partial x^l}(\mathbf{x},t''), \tag{64}$$

and

$$\tilde{\tau}^i(\xi t) = n^j(\xi) \int_{-\infty}^{+\infty} dt'' \overrightarrow{\Psi}^{ijkl}(\xi;t,t'') \frac{\partial u^k}{\partial x^l}(\xi,t''). \tag{65}$$

Equation (62) is a quite general representation theorem, but remember that the Green function is not defined uniquely. Notice that it is because we have imposed a *negative* attenuation in the definition (60) of $\overleftarrow{\mathbf{G}}$ that we have in (63) and (65) expression which correspond to the usual source field and surface tractions of a field propagating with *positive* attenuation.

Example. Let us be interested in a field $\overrightarrow{u}^i(\mathbf{x},t)$ defined, for $t \in [t_0, t_1]$, by

$$\rho(\mathbf{x})\frac{\partial^2 \overrightarrow{u}^i}{\partial t^2}(\mathbf{x},t) - \frac{\partial \overrightarrow{\sigma}^{ij}}{\partial x^j}(\mathbf{x},t) = \phi^i(\mathbf{x},t) \tag{66a}$$

$$\overrightarrow{\sigma}^{ij}(\mathbf{x},t) = \mathbf{M}^{ij}(\mathbf{x},t) + \int_{-\infty}^{+\infty} dt' \,\overrightarrow{\Psi}^{ijkl}(\mathbf{x};t,t')\frac{\partial \overrightarrow{u}^k}{\partial x^l}(\mathbf{x},t') \tag{66b}$$

$$n^j(\xi)\overrightarrow{\sigma}^{ij}(\xi t) = \tau^i(\xi,t) \quad \text{for} \quad \xi \in \mathbf{S} \tag{66c}$$

and assume that the history of the field is known for $t < t_0$.Then, choosing for $\overleftarrow{\mathbf{G}}$ the operator $\overleftarrow{\mathbf{G}}_{\text{free}}$, satisfying free boundary conditions (55), introducing the transposed operator $\overleftarrow{\mathbf{G}}_{\text{free}}$, using (62) we obtain, using equation (56b) and the divergence theorem, and relabelling variables,

$$\begin{aligned}
\overrightarrow{u}^i(\mathbf{x},t) = &\int_V dV(\mathbf{x}') \int_{t_0}^{t_1} dt' \overrightarrow{\mathbf{G}}^{ij}_{\text{free}}(\mathbf{x},t;\mathbf{x}',t')\phi^j(\mathbf{x}',t') \\
&+ \int_s dS(\xi') \int_{t_0}^{t_1} dt' \overrightarrow{\mathbf{G}}^{ij}_{\text{free}}(\mathbf{x},t;\xi',t')\tau^j(\xi',t') \\
&- \int_V dV(\mathbf{x}') \int_{t_0}^{t_1} dt' \frac{\partial \overrightarrow{\mathbf{G}}^{ij}_{\text{free}}}{\partial x'^k}(\mathbf{x},t;\mathbf{x}',t')\big(\mathbf{M}^{jk}(\mathbf{x}',t') + \Sigma^{jk}(\mathbf{x}',t')\big) \\
&+ \int_V dV(\mathbf{x}')\rho(\mathbf{x}') \Bigg[\overrightarrow{\mathbf{G}}^{ij}_{\text{free}}\big(\mathbf{x},t;\mathbf{x}',t'=t_0\big)\frac{\partial \overrightarrow{u}^j}{\partial t'}\big(\mathbf{x}',t'=t_0\big) \\
&\qquad - \frac{\partial \overrightarrow{\mathbf{G}}^{ij}_{\text{free}}}{\partial t'}\big(\mathbf{x},t;\mathbf{x}',t'=t_0\big)\,\overrightarrow{u}^j\big(\mathbf{x}',t'=t_0\big)\Bigg],
\end{aligned} \tag{67}$$

where $\Sigma^{ij}(\mathbf{x},t)$ is the stress due to the strain for $t < t_0$ not already relaxed:

$$\Sigma^{ij}(\mathbf{x},t) = \int_{-\infty}^{t_0} dt' \,\overrightarrow{\Psi}^{ijkl}(\mathbf{x};t,t')\frac{\partial \overrightarrow{u}^k}{\partial x^l}(\mathbf{x},t'). \tag{68}$$

Notice that the partial derivatives of the Green function are with respect to the source space-time coordinates.

In equation (67) the fields ϕ, τ, $\mathbf{M}$, Σ, $\mathbf{u}_0$, and v_0, can be named 'generalized sources' of the field $\mathbf{u}$. Then this equation defines a linear operator from the space of generalized sources into the space of displacements. This operator can be named the 'generalized Green operator'.

8. Reciprocity Theorems

In the previous definition of Green functions, we have used a function $\Psi^{ijkl}(\mathbf{x};t,t')$: the medium parameters may depend on time. Then there is no reciprocity relation satisfied.

If we assume that the medium parameters do not depend on time,

$$\Psi^{ijkl}(\mathbf{x};t,t') = \Psi^{ijkl}(\mathbf{x};t-t',0), \tag{69}$$

then, changing t by $-t$ switches from the dual problem defined by (55) into the primal problem (54), and the dual problem (58) into the primal problem (57). Then

$$\overrightarrow{\mathbf{G}}^{ij}_{\text{free}}(\mathbf{x}, t; \mathbf{x}', t') = \overleftarrow{\mathbf{G}}^{ji}_{\text{free}}(\mathbf{x}, -t; \mathbf{x}', -t'), \tag{70}$$

and

$$\overrightarrow{\mathbf{G}}^{ij}_{\text{rigid}}(\mathbf{x}, t; \mathbf{x}', t') = \overleftarrow{\mathbf{G}}^{ji}_{\text{rigid}}(\mathbf{x}, -t; \mathbf{x}', -t'). \tag{71}$$

As the density $\rho(\mathbf{x})$ is also independent on time, the whole wave equation is invariant by translation on time. Then

$$\overrightarrow{\mathbf{G}}^{ij}_{\text{free}}(\mathbf{x}, t; \mathbf{x}', t') = \overrightarrow{\mathbf{G}}^{ij}_{\text{free}}(\mathbf{x}, t - t', \mathbf{x}', 0), \tag{72}$$

and

$$\overrightarrow{\mathbf{G}}^{ij}_{\text{rigid}}(\mathbf{x}, t; \mathbf{x}', t') = \overrightarrow{\mathbf{G}}^{ij}_{\text{rigid}}(\mathbf{x}, t - t'; \mathbf{x}', 0). \tag{73}$$

From (70) and (72) it follows the reciprocity relation for $\mathbf{G}_{\text{free}}$:

$$\boxed{\overrightarrow{\mathbf{G}}^{ij}_{\text{free}}(\mathbf{x}, \tau; \mathbf{x}', 0) = \overrightarrow{\mathbf{G}}^{ji}_{\text{free}}(\mathbf{x}', \tau; \mathbf{x}', 0) \quad ,} \tag{74}$$

while from (71) and (73) it follows the reciprocity relation for $\mathbf{G}_{\text{rigid}}$:

$$\boxed{\overrightarrow{\mathbf{G}}^{ij}_{\text{rigid}}(\mathbf{x}, \tau; \mathbf{x}', 0) = \overrightarrow{\mathbf{G}}^{ji}_{\text{rigid}}(\mathbf{x}', \tau; \mathbf{x}, 0) \quad .} \tag{75}$$

The response at point $\mathbf{x}$ along the i-th axis for a source at point $\mathbf{x}'$ along the j-axis, equals the response at point $\mathbf{x}'$ along the j-axis for a source at point $\mathbf{x}$ along the i-th axis. For both experiments, the source starts at 0 and we record at τ.

Notice that to any Green's function $\overrightarrow{\mathbf{G}}^{ij}(\mathbf{x}, t; \mathbf{x}', t')$ we can associate the Green function of the dual problem $\overleftarrow{\mathbf{G}}^{ij}(\mathbf{x}, t; \mathbf{x}', t')$, that they satisfy necessarily the property

$$\overrightarrow{\mathbf{G}}^{ij}(\mathbf{x}, t; \mathbf{x}', t') = \overleftarrow{\mathbf{G}}^{ji}(\mathbf{x}', t'; \mathbf{x}, t), \tag{76}$$

but that to satisfy a reciprocity property we need more structure: an equivalence between primal and dual problems under some change of variables.

9. The Born Approximation

Let us consider the field $u^i(\mathbf{x},t)$ defined by the system of equations

$$\rho(\mathbf{x})\frac{\partial^2 u^i}{\partial t^2}(\mathbf{x},t) - \frac{\partial \sigma^{ij}}{\partial x^j}(\mathbf{x},t) = \phi^i(\mathbf{x},t), \tag{77a}$$

$$\sigma^{ij}(\mathbf{x},t) = \mathbf{M}^{ij}(\mathbf{x},t) + \int_{-\infty}^{+\infty} dt' \Psi^{ijkl}(\mathbf{x},t-t')\frac{\partial u^k}{\partial x^l}(\mathbf{x},t'), \tag{77b}$$

$$n^j(\xi)\sigma^{ij}(\xi,t) = \tau^i(\xi,t) \quad \text{for} \quad \xi \in \mathbf{S}, \tag{77c}$$

$$u^i\big(\mathbf{x},t_0\big) = \alpha^i(\mathbf{x}), \tag{77d}$$

$$\frac{\partial u^i}{\partial t}\big(\mathbf{x},t_0\big) = \beta^i(\mathbf{x}), \tag{77e}$$

$$u^i(\mathbf{x},t) = \gamma^i(\mathbf{x},t) \quad \text{for} \quad t < t_0. \tag{77f}$$

A perturbation of the model parameters

$$\rho(\mathbf{x}) \longrightarrow \rho(\mathbf{x}) + \delta\rho(\mathbf{x}) \tag{78a}$$

$$\Psi^{ijkl}(\mathbf{x},\tau) \longrightarrow \Psi^{ijkl}(\mathbf{x},\tau) + \delta\Psi^{ijkl}(\mathbf{x},\tau) \tag{78b}$$

leads to a perturbation of the displacement field

$$u^i(\mathbf{x},t) \longrightarrow u^i(\mathbf{x},t) + \delta u^i(\mathbf{x},t). \tag{78c}$$

For the use of gradient methods, we need the first order approximation to $\delta u^i(\mathbf{x},t)$. Inserting (78) into (77), subtracting (77), and dropping second-order terms we arrive easily at

$$\rho(\mathbf{x})\frac{\partial^2 \delta u^i}{\partial t^2}(\mathbf{x},t) - \frac{\partial \delta\sigma^{ij}}{\partial x^j}(\mathbf{x},t) = \delta\phi^i(\mathbf{x},t), \tag{79a}$$

$$\partial\sigma^{ij}(\mathbf{x},t) = \delta\mathbf{M}^{ij}(\mathbf{x},t) + \int_{-\infty}^{+\infty} dt' \Psi^{jikl}(\mathbf{x},t-t')\frac{\partial \delta u^k}{\partial x^l}(\mathbf{x},t'), \tag{79b}$$

$$n^j(\xi)\delta\sigma^{ij}(\xi,t) = 0 \quad \text{for} \quad \xi \in \mathbf{S}, \tag{79c}$$

$$\delta u^i\big(\mathbf{x},t_0\big) = 0, \tag{79d}$$

$$\frac{\partial \delta u^i}{\partial t}\big(\mathbf{x},t_0\big) = 0, \tag{79e}$$

$$\delta u^i(\mathbf{x},t) = 0 \quad \text{for} \quad t < t_0, \tag{79f}$$

where $\delta\phi$ and $\delta\mathbf{M}$ are the 'secondary Born sources'

$$\delta\phi^i(\mathbf{x},t) = -\delta\rho(\mathbf{x})\frac{\partial^2 u^i}{\partial t^2}(\mathbf{x},t), \tag{80a}$$

and

$$\delta \mathbf{M}^{ij}(\mathbf{x},t) = \int_{-\infty}^{+\infty} dt' \delta\Psi^{ijkl}(\mathbf{x};t,t') \frac{\partial u^k}{\partial x^l}(\mathbf{x},t'). \tag{80b}$$

The field $\delta\mathbf{u}$ defined by this system of equations corresponds to the Born approximation to displacement perturbation. The intuitive interpretation of equations (79) is as follows. The field $\delta\mathbf{u}$ propagates in the unperturbed medium (because $\rho(\mathbf{x})$ and $\Psi^{ijkl}(\mathbf{x};t,t')$ appear in the left-hand side, but not $\delta\rho(\mathbf{x})$ and $\delta\Psi^{ijkl}(\mathbf{x};t,t')$. Sources for this field exist where the medium has been perturbed. They are proportional to the perturbations $\delta\rho(\mathbf{x})$ and $\delta\Psi^{ijkl}(\mathbf{x};t,t')$, and to the reference field $u^i(\mathbf{x},t)$. By comparison with equations (18) to (20), we see that the source corresponding to the density perturbation is a force density, while the one corresponding to the perturbation of the visco-elastic parameters is a moment density.

Using the representation theorem (67) we obtain

$$\begin{aligned} \delta u^p(\mathbf{x}',t') = & \int_V dV(\mathbf{x}) \int_{t_0}^{t_1} dt \overrightarrow{\mathbf{G}}^{pi}_{\text{free}}(\mathbf{x}',t';\mathbf{x},t)\delta\phi^i(\mathbf{x},t) \\ & - \int_V dV(\mathbf{x}) \int_{t_0}^{t_1} dt \frac{\partial \overrightarrow{\mathbf{G}}^{pi}_{\text{free}}}{\partial x^j}(\mathbf{x}',\tau';\mathbf{x},t)\delta\mathbf{M}^{ij}(\mathbf{x},t). \end{aligned} \tag{81}$$

Equations (80)–(81) give the explicit expression to the Born approximation.

10. Least Squares in Functional Spaces

Assume that using sources $\phi^i(\mathbf{x},t), \tau^i(\xi,t)$ and $\mathbf{M}^{ij}(\mathbf{x},t)$ (usually of only one type) we generate a displacement field $\overrightarrow{u}^i(\mathbf{x},t)$ in a medium described by the parameters $\rho(\mathbf{x})$ and $\Psi^{ijkl}(x,\tau)$, and that we measure the field $\overrightarrow{u}^i(\mathbf{x},t)$ at some receiver locations $\mathbf{x}_r$ $(r = 1,2,\ldots)$. We wish to use the observations $\overrightarrow{u}^i(\mathbf{x}_r,t)_{\text{obs}}$ to infer the values of the parameters $\rho(\mathbf{x})$ and $\Psi^{ijkl}(\mathbf{x},\tau)$ describing the medium.

We assume here that the field $\overrightarrow{u}^i(\mathbf{x},t)$ satisfies homogeneous initial conditions, and propagates with a free surface. In this section, the notations $\mathbf{L}$ and $\mathbf{G}$ will stand respectively for $\mathbf{L}_{\text{free}}$ and $\overrightarrow{\mathbf{G}}_{\text{free}}$ (or the corresponding generalized operators introduced in Section 7). The field $\overrightarrow{\mathbf{u}}$ is then defined by the equation $\mathbf{L}\overrightarrow{\mathbf{u}} = \psi$, where ψ denotes the generalized sources (representing $\phi^i(\mathbf{x},t), \tau^i(\xi,t)$, and/or $\mathbf{M}^{ij}(\mathbf{x},t)$). The operator $\mathbf{L}$ is a function of the medium parameters. To make this dependence explicit, we write $\mathbf{L}[\mathbf{m}]$, where $\mathbf{m}$ represents a *model* of the medium, i.e., a set of functions $\{\rho(\mathbf{x}), \Psi^{ijkl}(\mathbf{x},t)\}$. The models of the medium belong to the 'model space' $\mathbf{M}$.

Then, $\overrightarrow{\mathbf{u}}$ is defined by

$$\mathbf{L}[\mathbf{m}]\overrightarrow{\mathbf{u}} = \psi. \tag{82}$$

The observed values $\overrightarrow{u}^i(\mathbf{x}_r,t)$ will be denoted by $\mathbf{d}_{\text{obs}}$. The values $\overrightarrow{u}^i(\mathbf{x}_r,t)$ calculated from a model $\mathbf{m}$ will be denoted by $\mathbf{d}_{\text{cal}}$ or $\mathbf{d}[\mathbf{m}]$. The data vectors $\mathbf{d}$ belong to a 'data space' $\mathbf{D}$.

The aim of least-squares inversion (TARANTOLA and VALETTE, 1982a, 1982b; TARANTOLA, 1987) is to obtain the model $\mathbf{m}$ minimizing the misfit function.

$$\begin{aligned} S[\mathbf{m}] &= \frac{1}{2}(\|\mathbf{d}[\mathbf{m}] - \mathbf{d}_{\text{obs}}\|^2 + \|\mathbf{m} - \mathbf{m}_{\text{prior}}\|^2) \\ &= \frac{1}{2}\left[\langle \mathbf{C}_D^{-1}(\mathbf{d}[\mathbf{m}] - \mathbf{d}_{\text{obs}}), (\mathbf{d}[\mathbf{m}] - \mathbf{d}_{\text{obs}})\rangle + \langle \mathbf{C}_M^{-1}(\mathbf{m} - \mathbf{m}_{\text{prior}}), (\mathbf{m} - \mathbf{m}_{\text{prior}})\rangle\right], \end{aligned} \tag{83}$$

where $\mathbf{C}_D$ is the covariance operator describing data uncertainties, $\mathbf{m}_{\text{prior}}$ is some *a priori* model, and $\mathbf{C}_M$ is the covariance operator describing uncertainties in $\mathbf{m}_{\text{prior}}$.

The gradient $\hat{\gamma}$ of the misfit function is defined by the first-order development

$$S(\mathbf{m} + \delta\mathbf{m}) = S(\mathbf{m}) + \langle \hat{\gamma}, \delta\mathbf{m} \rangle + O(\|\delta\mathbf{m}\|^2). \tag{84}$$

It is an element of the dual of the model space (identified with the model space weighted by $\mathbf{C}_M^{-1}$).

The direction of steepest ascent is then (TARANTOLA, 1987a)

$$\gamma = \mathbf{C}_M \hat{\gamma}, \tag{85}$$

and the algorithm of steepest descent for the minimization of $S(\mathbf{m})$ is

$$\mathbf{m}_{n+1} = \mathbf{m}_n - \alpha_n \gamma_n, \tag{86}$$

where α_n is a constant sufficiently small to ensure

$$S(\mathbf{m}_{n+1}) < S(\mathbf{m}_n). \tag{87}$$

Let us now formally compute the gradient of the misfit function. As the term $\langle \mathbf{C}_M^{-1}(\mathbf{m} - \mathbf{m}_{\text{prior}}), (\mathbf{m} - \mathbf{m}_{\text{prior}})$ is quadratic in $\mathbf{m}$, it makes no problem, and is dropped (the reader will easily correct for it).

Formally, $\mathbf{d}$ is obtained by projecting the field $\vec{u}^i(\mathbf{x}, t)$ into the observation points $\mathbf{x}_r$:

$$\mathbf{d} = \mathbf{P}\vec{\mathbf{u}}, \tag{88}$$

where $\mathbf{P}$ is the projector ($\mathbf{P}^2 = \mathbf{P}$) defined by

$$(\mathbf{P}\vec{\mathbf{u}})^i(\mathbf{x}_r, t) = \vec{u}(\mathbf{x}_r, t). \tag{89}$$

The reader may easily verify that the transposed of $\mathbf{P}$ is the operator defined by

$$(\mathbf{P}^t\hat{\mathbf{d}})^i(\mathbf{x}, t) = \sum_r \delta(\mathbf{x} - \mathbf{x}_r)\hat{d}^i(\mathbf{x}_r, t), \tag{90}$$

where $\hat{\mathbf{d}}$ is an element of the dual of the data space (identified with the data space weighted by $\mathbf{C}_D^{-1}$).

We have

$$S(\mathbf{m}) = \frac{1}{2}\langle \mathbf{C}_D^{-1}(\mathbf{P}\overrightarrow{\mathbf{u}} - \mathbf{d}_{\text{obs}}), (\mathbf{P}\overrightarrow{\mathbf{u}} - \mathbf{d}_{\text{obs}})\rangle, \tag{91}$$

where

$$\mathbf{L}[\mathbf{m}]\overrightarrow{\mathbf{u}} = \psi. \tag{92}$$

A perturbation of the medium parameters leads to

$$\mathbf{m} \rightarrow \mathbf{m} + \delta\mathbf{m} \tag{93}$$

leads to

$$S\big(\mathbf{m} + \delta\mathbf{m}\big) = \frac{1}{2}\big\langle \mathbf{C}_D^{-1}\big(\mathbf{P}\big(\overrightarrow{\mathbf{u}} + \Delta\overrightarrow{\mathbf{u}} - \mathbf{d}_{\text{obs}}\big), \big(\mathbf{P}\big(\overrightarrow{\mathbf{u}} + \Delta\overrightarrow{\mathbf{u}}\big) - \mathbf{d}_{\text{obs}}\big)\big\rangle, \tag{94}$$

where $\Delta\overrightarrow{\mathbf{u}}$ is defined by

$$\mathbf{L}[\mathbf{m} + \delta\mathbf{m}](\overrightarrow{\mathbf{u}} + \Delta\overrightarrow{\mathbf{u}}) = \phi, \tag{95}$$

and depends (nonlinearly) on $\delta\mathbf{m}$. Let $\delta\overrightarrow{\mathbf{u}}$ denote the first order approximation to $\Delta\overrightarrow{\mathbf{u}}$:

$$\Delta\overrightarrow{\mathbf{u}} = \delta\overrightarrow{\mathbf{u}} + O\big(\|\delta\mathbf{m}\|^2\big). \tag{96}$$

Then

$$\begin{aligned} S(\mathbf{m} + \delta\mathbf{m}) &= \frac{1}{2}\Big\langle \mathbf{C}_D^{-1}\big(\mathbf{P}(\overrightarrow{\mathbf{u}} + \delta\overrightarrow{\mathbf{u}}) - \mathbf{d}_{\text{obs}}\big), \big(\mathbf{P}(\overrightarrow{\mathbf{u}} + \delta\overrightarrow{\mathbf{u}}) - \mathbf{d}_{\text{obs}}\big)\Big\rangle \\ &= S(\mathbf{m}) + \frac{1}{2}\big\langle \mathbf{C}_D^{-1}\big(\mathbf{P}\overrightarrow{\mathbf{u}} - \mathbf{d}_{\text{obs}}\big), \mathbf{P}\delta\overrightarrow{\mathbf{u}}\big\rangle + \frac{1}{2}\big\langle \mathbf{C}_D^{-1}\mathbf{P}\delta\overrightarrow{\mathbf{u}}, \big(\mathbf{P}\overrightarrow{\mathbf{u}} - \mathbf{d}_{\text{obs}}\big)\big\rangle \\ &\quad + O\big(\|\delta\mathbf{m}\|^2\big); \end{aligned}$$

as covariance operators are symmetric,

$$S(\mathbf{m} + \delta\mathbf{m}) = S(\mathbf{m}) + \Big\langle \mathbf{C}_D^{-1}\big(\mathbf{P}\overrightarrow{\mathbf{u}} - \mathbf{d}_{\text{obs}}\big), \mathbf{P}\delta\overrightarrow{\mathbf{u}}\Big\rangle + O\big(\|\delta\mathbf{m}\|^2\big);$$

and, introducing the transpose of the projector $\mathbf{P}$,

$$S(\mathbf{m} + \delta\mathbf{m}) = S(\mathbf{m}) + \Big\langle \mathbf{P}^t\mathbf{C}_D^{-1}\big(\mathbf{P}\overrightarrow{\mathbf{u}} - \mathbf{d}_{\text{obs}}\big), \delta\overrightarrow{\mathbf{u}}\Big\rangle + O\big(\|\delta\mathbf{m}\|^2\big). \tag{97}$$

The field $\delta\overrightarrow{\mathbf{u}}$, first order approximation to $\Delta\overrightarrow{\mathbf{u}}$ corresponds to the Born approximation to $\Delta\overrightarrow{\mathbf{u}}$ (see Section 9). It corresponds then to the field created by some Born secondary generalized sources $\delta\psi$ and propagated into the unperturbed medium $\mathbf{m}$:

$$\delta\overrightarrow{\mathbf{u}} = \mathbf{G}[\mathbf{m}]\delta\psi. \tag{98}$$

Equation (97) then becomes

$$\begin{aligned} S(\mathbf{m} + \delta\mathbf{m}) &= S(\mathbf{m}) + \Big\langle \mathbf{P}^t\mathbf{C}_D^{-1}\big(\mathbf{P}\overrightarrow{\mathbf{u}} - \mathbf{d}_{\text{obs}}\big), \mathbf{G}\delta\psi\Big\rangle + O\big(\|\delta\mathbf{m}\|^2\big) \\ &= S(\mathbf{m}) + \langle \overleftarrow{\mathbf{u}}, \delta\psi\rangle + O\big(\|\mathbf{m}\|^2\big), \end{aligned} \tag{99}$$

where $\overleftarrow{\mathbf{u}}$ is defined by

$$\overleftarrow{\mathbf{u}} = \mathbf{G}^t \mathbf{P}^t \mathbf{C}_D^{-1} \big(\mathbf{P} \overrightarrow{\mathbf{u}} - \mathbf{d}_{\text{obs}} \big), \tag{100a}$$

i.e.,

$$\mathbf{L}^t \overleftarrow{\mathbf{u}} = \mathbf{P}^t \mathbf{C}_D^{-1} \big(\mathbf{P} \overrightarrow{\mathbf{u}} - \mathbf{d}_{\text{obs}} \big). \tag{100b}$$

The field $\overleftarrow{\mathbf{u}}$ is created by the sources $\mathbf{P}^t \mathbf{C}_D^{-1} (\mathbf{P} \overrightarrow{\mathbf{u}} - \mathbf{d}_{\text{obs}})$, and satisfies conditions dual to those satisfied by $\overrightarrow{\mathbf{u}}$.

The Born secondary sources $\delta\psi$ depend linearly on a $\overrightarrow{\mathbf{u}}$ and on $\delta\mathbf{m}$. Introducing the notation

$$\delta\psi = (\mathbf{A} \overrightarrow{\mathbf{u}}) \delta\mathbf{m} \tag{101}$$

leads to

$$S(\mathbf{m} + \delta\mathbf{m}) = S(\mathbf{m}) - \big\langle \overleftarrow{\mathbf{u}}, (\mathbf{A} \overrightarrow{\mathbf{u}}) \delta\mathbf{m} \big\rangle + O\big(\|\delta\mathbf{m}\|^2 \big),$$

and properly introducing the transpose of the operator $(\mathbf{A} \overrightarrow{\mathbf{u}})$,

$$S(\mathbf{m} + \delta\mathbf{m}) = S(\mathbf{m}) + \big\langle (\mathbf{A} \overrightarrow{\mathbf{u}})^t \overleftarrow{\mathbf{u}}, \delta\mathbf{m} \big\rangle + O\big(\|\delta\mathbf{m}\|^2 \big). \tag{102}$$

By comparison with (87), the last equation gives the gradient of the least squares misfit functional S:

$$\boxed{\hat{\gamma} = (\mathbf{A} \overrightarrow{\mathbf{u}})^t \overleftarrow{\mathbf{u}} \quad .} \tag{103}$$

Equation (88) gives then

$$\gamma = \mathbf{C}_M (\mathbf{A} \overrightarrow{\mathbf{u}})^t \overleftarrow{\mathbf{u}}, \tag{104}$$

and (89) finally gives

$$\mathbf{m}_{n+1} = \mathbf{m}_n - \alpha_n \mathbf{C}_M (\mathbf{A} \overrightarrow{\mathbf{u}}_n)^t \overleftarrow{\mathbf{u}}_n. \tag{105}$$

All the partial steps needed for an iteration of the steepest descent algorithm are:

$\mathbf{L}\big[\mathbf{m}_n\big] \overrightarrow{\mathbf{u}}_n = \psi$	(solve for $\mathbf{u}_n$)
$\delta\mathbf{d}_n = \mathbf{P} \overrightarrow{\mathbf{u}}_n - \mathbf{d}_{\text{obs}}$	(compute data residuals)
$\mathbf{C}_D \delta\hat{\mathbf{d}}_n = \delta\mathbf{d}_n$	(solve to obtain the weighted residuals)
$\delta\phi_n = \mathbf{P}^t \delta\hat{\mathbf{d}}_n$	(consider these as sources)
$\mathbf{L}^t \overleftarrow{\mathbf{u}}_n = \delta\phi_n$	(solve for $\overleftarrow{\mathbf{u}}_n$, i.e., propagate the sources with dual conditions, solving the dual problem)
$\hat{\gamma}_n = \big(\mathbf{A} \overrightarrow{\mathbf{u}}_n\big)^t \overleftarrow{\mathbf{u}}_n$	(compute the gradient $\hat{\gamma}_n$, where $\mathbf{A}$ has been defined in (101))
$\gamma_n = \mathbf{C}_M \hat{\gamma}_n$	(unweight the gradient)
$\mathbf{m}_{n+1} = \mathbf{m}_n - \alpha_n \gamma_n$	(update the model)

(106)

The above formulas correspond to a crude steepest descent method. Current implementations of gradient methods for the inversion of seismic waveforms (GAUTHIER *et al.*, 1968; KOLB, 1986; MORA, 1987; PICA, 1987) are rather based in conjugate gradients (e.g., FLETCHER, 1981; SCALES, 1985), which converge more rapidly.

11. The Inverse Problem of Interpretation of Seismic Waveforms

This section applies the results of the previous section to the problem of interpretation of seismic reflection data. Typically, a source is fired consecutively at different locations $\mathbf{x}_s$ $(s = 1, 2, \ldots)$, and, for each source position, the displacement u^i is observed at some locations $\mathbf{x}_r$ $(r = 1, 2, \ldots)$. In all rigor, the time variable t runs from $-\infty$ to $+\infty$ and there is only one source, concentrated at different points at different times. More intuitively, we can consider that the time variable is reset to $t = t_0$ at each new shot, and we record the earth's surface displacements until $t = t_1$. In that case, it can easily be seen that the gradient of the misfit function for an experiment with different sources is simply the sum of the gradients corresponding to each source. We see thus that, with any of the two points of view, we can limit our consideration to an experiment with a single source.

The observed seismograms are denoted by $u^i(\mathbf{x}, t)_{\text{obs}}$, while the computed seismograms corresponding to the n-th earth model are denoted $u^i(\mathbf{x}, t)_n$. The recording time belongs to the interval $[t_0, t_1)$.

Let us take in order all the steps (106).

Equation (106)(a): $\mathbf{L}[\mathbf{m}_n]\overrightarrow{\mathbf{u}}_n = \psi$. Let $\rho(\mathbf{x})_n$ and $\Psi^{ijkl}(\mathbf{x}, \tau)_n$ denote the current earth model. If the sources of seismic waves are described by the force density $\phi^i(\mathbf{x}, t)$ the surface traction $\tau^i(\mathbf{x}, t)$ and/or the moment density $M^{ij}(\mathbf{x}, t)$ then the current displacement field $\overrightarrow{u}^i(\mathbf{x}, t)_n$ is defined by

$$\rho(\mathbf{x})_n \frac{\partial^2 \overrightarrow{u}^i}{\partial t^2}(\mathbf{x}, t)_n - \frac{\partial \overrightarrow{\sigma}^{ij}}{\partial x^j}(\mathbf{x}, t)_n = \phi^i(\mathbf{x}, t), \tag{107a}$$

$$\overrightarrow{\sigma}^{ij}(\mathbf{x}, t)_n = M^{ij}(\mathbf{x}, t) + \int_{-\infty}^{+\infty} dt' \, \overrightarrow{\Psi}^{ijkl}(\mathbf{x}, t - t')_n \frac{\partial \overrightarrow{u}^k}{\partial x^l}(\mathbf{x}, t')_n, \tag{107b}$$

$$n^j(\xi)\overrightarrow{\sigma}(\xi, t)_n = \tau^i(\xi, t) \quad \text{for} \quad \xi \in \mathbf{S}, \tag{107c}$$

$$\overrightarrow{u}^i\big(\mathbf{x}, t_0\big)_n = 0, \tag{107d}$$

$$\frac{\partial \overrightarrow{u}^i}{\partial t}\big(\mathbf{x}, t_0\big)_n = 0, \tag{107e}$$

$$\overrightarrow{u}^i(\mathbf{x}, t)_n = 0 \quad \text{for} \quad t < t_0, \tag{107f}$$

where a free surface and homogeneous initial conditions are assumed. The computation of the field $\overrightarrow{u}^i(\mathbf{x}, t)_n$ can be performed using any numerical method, as, for instance, finite-differences (ALTERMAN and KARAL, 1968; VIRIEUX, 1986).

Equation (106)*(b)*: $\delta\mathbf{d}_n = \mathbf{P}\overrightarrow{\mathbf{u}}_n - \mathbf{d}_{\text{obs}}$. This simply corresponds to the definition of the residuals at the receiver locations:

$$\delta d^i\big(\mathbf{x}_r, t\big)_n = \overrightarrow{u}^i\big(\mathbf{x}_r, t\big)_n - \overrightarrow{u}^i\big(\mathbf{x}_r, t\big)_{\text{obs}}. \tag{108}$$

Equation (106)*(c)*: $\mathbf{C}_D \delta\hat{\mathbf{d}}_n = \delta\mathbf{d}_n$. As an example, assume independent and uniform uncertainties. Then,

$$\delta\hat{d}^i\big(\mathbf{x}_r, t\big)_n = \frac{1}{\sigma^2}\delta d^i\big(\mathbf{x}_r t\big)_n. \tag{109}$$

Of course, other more realistic choices of the covariance operator describing experimental uncertainties can be made.

Equation (106)*(d)*: $\delta\phi_n = \mathbf{P}^t \delta\hat{\mathbf{d}}_n$. From equation (90) we obtain

$$\delta\phi^i(\mathbf{x}, t)_n = \sum_r \delta\big(\mathbf{x} - \mathbf{x}_r\big)\delta\hat{d}^i\big(\mathbf{x}_r, t\big)_n. \tag{110}$$

This corresponds to a composite source, one point source at each receiver location, radiating the weighted residuals in phase.

Equation (106)*(e)*: $\mathbf{L}^t \overleftarrow{\mathbf{u}}_n = \delta\phi_n$. As demonstrated in Section 5, to compute the field $\overleftarrow{u}^i(\mathbf{x}, t)_n$, solution of $\mathbf{L}^t \overleftarrow{\mathbf{u}}_n = \delta\phi_n$, means to solve the differential system

$$\rho(\mathbf{x})_n \frac{\partial^2 \overleftarrow{u}^i}{\partial t^2}(\mathbf{x}, t)_n - \frac{\partial \overleftarrow{\sigma}^{ij}}{\partial x^j}(\mathbf{x}, t)_n = \delta\phi^i(\mathbf{x}, t)_n, \tag{111a}$$

$$\overleftarrow{\sigma}^{ij}(\mathbf{x}, t)_n = \int_{-\infty}^{+\infty} dt' \overleftarrow{\Psi}^{ijkl}(\mathbf{x}, t - t')_n \frac{\partial \overleftarrow{u}^k}{\partial x^l}(\mathbf{x}, t')_n, \tag{111b}$$

$$n^j(\xi)\overleftarrow{\sigma}(\xi, t) = 0 \quad \text{for} \quad \xi \in \mathbf{S}, \tag{111c}$$

$$\overleftarrow{u}^i\big(\mathbf{x}, t_1\big)_n = 0, \tag{111d}$$

$$\frac{\partial \overleftarrow{u}^i}{\partial t}\big(\mathbf{x}, t_1\big)_n = 0, \tag{111e}$$

$$\overleftarrow{u}^i(\mathbf{x}, t)_n = 0 \quad \text{for} \quad t > t_1, \tag{111f}$$

where instead of a quiescent past, the field has a quiescent future, and where the attenuation is negative (anti-causal). To solve this problem, we can use the same computer code needed to solve the system (107), reversing the time (see GAUTGIER *et al.,* 1986) and changing $\overrightarrow{\Psi}^{ijkl}(\mathbf{x}, \tau)$ by $\overleftarrow{\Psi}^{ijkl}(\mathbf{x}, \tau) = \overrightarrow{\Psi}^{ijkl}(\mathbf{x}, -\tau)$. Notice that a reverse-time propagation with negative attenuation is numerically as stable as a forward time propagation with positive attenuation.

Equation (106)*(f)*: $\hat{\gamma}_n = (\mathbf{A}\overrightarrow{\mathbf{u}}_n)^t \overleftarrow{\mathbf{u}}_n$. Once the field $\overleftarrow{u}^i(\mathbf{x}, t)_n$ has been computed, we can turn to the computation of the gradient. The operator $\mathbf{A}$ is defined by equation (101):

$$\delta\psi_n = \big(\mathbf{A}\overrightarrow{\mathbf{u}}_n\big)\delta\mathbf{m}, \tag{101 again}$$

where $\delta\psi_n$ is the Born secondary generalized source corresponding to the perturbation $\delta\mathbf{m} =$

$\{\delta\rho, \delta\Psi^{ijkl}\}$. Using the results of Section 9 gives

$$\begin{aligned}\left\langle \overleftarrow{\mathbf{u}}_n,\, (\mathbf{A}\overrightarrow{\mathbf{u}}_n)\delta\mathbf{m}\right\rangle &= \langle \overleftarrow{\mathbf{u}}_n,\, \delta\psi_n\rangle \\ &= + \int_V dV(\mathbf{x}) \left[\int_{t_0}^{t_1} dt \overleftarrow{v}^i(\mathbf{x},t)\overrightarrow{v}^i(\mathbf{x},t)\right]\delta\rho(\mathbf{x}) \\ &\quad - \int_V dV(\mathbf{x}) \int_{-\infty}^{+\infty} dt' \left[\int_{t_0}^{t_1} dt \overleftarrow{\varepsilon}^{ij}(\mathbf{x},t)\overrightarrow{\varepsilon}^{kl}(\mathbf{x},t-t')\right]\delta\psi(\mathbf{x},t').\end{aligned} \tag{112}$$

As, by definition of transpose,

$$\left\langle \overleftarrow{\mathbf{u}}_n,\, (\mathbf{A}\overrightarrow{\mathbf{u}}_n)\delta\mathbf{m}\right\rangle = \left\langle (\mathbf{A}\overrightarrow{\mathbf{u}}_n)^t \overleftarrow{\mathbf{u}}_n,\, \delta\mathbf{m}\right\rangle, \tag{113}$$

the components of the gradient follow using (103):

$$\boxed{\hat{\gamma}_\rho(\mathbf{x})_n = \overleftarrow{v}^i(\mathbf{x},t)_n \otimes \overrightarrow{v}^i(\mathbf{x},t)_n\Big|_{t=0} \quad ,} \tag{114}$$

and

$$\boxed{\hat{\gamma}_\Psi^{ijkl}(\mathbf{x},t)_n = -\overleftarrow{\varepsilon}^{ij}(\mathbf{x},t)_n \otimes \overrightarrow{\varepsilon}^{kl}(\mathbf{x},t)_n \quad ,} \tag{115}$$

where the time correlation between two functions $f(t)$ and $g(t)$ has been defined by

$$f(t) \otimes g(t) = \int_{t_0}^{t_1} dt' f(t')g(t'-t). \tag{116}$$

In Appendix C, I give a more direct demonstration of formulas (114)–(115).

If we are not interested in the attenuating properties of the medium, but only in the elastic properties, we make the hypothesis

$$\Psi^{ijkl}(\mathbf{x},\tau) = c^{ijkl}(\mathbf{x})\delta(\tau). \tag{8 again}$$

Instead of modifying the theory to compute the gradient with respect to the elastic stiffness, it is clear that only the correlations (115) at zero lag can contribute. Then,

$$\hat{\gamma}_c^{ijkl}(\mathbf{x})_n = \overleftarrow{\varepsilon}^{ij}(\mathbf{x},t)_n \otimes \overrightarrow{\varepsilon}^{kl}(\mathbf{x},t)_n\Big|_{t=0}, \tag{117}$$

which corresponds to the result demonstrated by TARANTOLA (1987b).

Equation (106)*(g)*: $\gamma_n = \mathbf{C}_M\hat{\gamma}$. The operator $\mathbf{C}_M$, describes the confidence we have on our *a priori* model. Assuming for instance that the uncertainties on density are independent on the uncertainties on the rate-of-relaxation function,

$$\mathbf{C}_M = \begin{bmatrix} \mathbf{C}_\rho & 0 \\ 0 & \mathbf{C}_\Psi \end{bmatrix}, \tag{118}$$

gives, introducing the corresponding covariance functions,

$$\gamma_\rho(\mathbf{x})_n = \int_V dV(\mathbf{x}')\mathbf{C}_\rho(\mathbf{x},\mathbf{x}')\hat{\gamma}_\rho(\mathbf{x}')_n, \tag{119a}$$

and

$$\gamma_\Psi^{ijkl}(\mathbf{x},\tau)_n = \int_V dV(\mathbf{x}') \int_{-\infty}^{+\infty} dt' C_\Psi^{ijklpqrs}(\mathbf{x},\tau;\mathbf{x}',\tau')\hat{\gamma}_\Psi^{pqrs}(\mathbf{x}',\tau')_n. \tag{119b}$$

Equation (106)*(h)*: $\mathbf{m}_{n+1} = \mathbf{m}_n - \alpha_n\gamma_n$. This gives, finally,

$$\rho(\mathbf{x})_{n+1} = \rho(\mathbf{x})_n - \alpha_n\gamma_\rho(\mathbf{x})_n, \tag{120}$$

and

$$\Psi^{ijkl}(\mathbf{x},\tau)_{n+1} = \Psi^{ijkl}(\mathbf{x},\tau)_n - \alpha_n\gamma_\Psi^{ijkl}(\mathbf{x},\tau)_n. \tag{121}$$

12. Discussion and Conclusion

For the purposes of inversion of seismic waveforms it seems better to introduce a completely general rate-of-relaxation function $\Psi^{ijkl}(\mathbf{x},\tau)$ rather than to use a particular model (Standard linear solid, Constant Q, etc.). Some seismic data sets at least seem to contain enough information to give useful constraints on $\Psi^{ijkl}(\mathbf{x},\tau)$, as for instance seismic reflection data.

The representation theorems are as simple with this general linear model as for the particular perfectly elastic model. Only a supplementary stress term appears, which is due to the history of deformation before the initial time. Reciprocity is also satisfied.

The formulas given in Section 11 correspond to a steepest descent method. Practically, considerable modification must be introduced. For instance, conjugate gradients may be used in stead of steepest descent. But more fundamentally, I suggest not using at the outset all the components of the gradient that may be computed. First invert for elastic parameters, and, after convergence, allow attenuation to be introduced. This means that, at the beginning, only the values for zero lag of the correlations (115) should be taken into account.

In fact, as suggested by TARANTOLA (1986) the elastic parameters themselves are highly hierachisized. For instance, with seismic reflection surface data, the long wavelengths of the P-wave velocity have to be first computed, then the P-wave impedance (product of density by velocity). When a good model has been obtained, the S-wave velocity and S-wave impedance can be computed.

When all these things have satisfactorily converged, then formula (115) is allowed to one one step further, and to obtain the beat model of attenuation.

Appendix A: Transposed of the Wave Equation Operator. Dual Conditions

The operators $\mathbf{L}$ and $\mathbf{L}^t$ have been defined by:

$$(\mathbf{L}\mathbf{u})^i(\mathbf{x},t) = \rho(\mathbf{x})\frac{\partial^2 u^i}{\partial t^2}(\mathbf{x},t) - \frac{\partial}{\partial x^j}\int_{-\infty}^{+\infty} dt'\,\overrightarrow{\Psi}^{ijkl}(\mathbf{x};t,t')\frac{\partial u^k}{\partial x^l}(\mathbf{x},t'), \tag{A-1}$$

and

$$(\mathbf{L}^t\mathbf{u})^i(\mathbf{x},t) = \rho(\mathbf{x})\frac{\partial^2 u^i}{\partial t^2}(\mathbf{x},t) - \frac{\partial}{\partial x^j}\int_{-\infty}^{+\infty} dt' \overleftarrow{\Psi}^{ijkl}(\mathbf{x};t,t')\frac{\partial u^k}{\partial x^l}(\mathbf{x},t'), \tag{A-2}$$

where $\overrightarrow{\Psi}^{jikl}(\mathbf{x};t,t')$ is a causal function,

$$\overrightarrow{\Psi}^{ijkl}(\mathbf{x};t,t') = 0 \quad \text{for} \quad t < t', \tag{A-3}$$

and

$$\overleftarrow{\Psi}^{ijkl}(\mathbf{x};t,t') = \overrightarrow{\Psi}^{ijkl}(\mathbf{x};t',t). \tag{A-4}$$

We have

$$\langle \overleftarrow{\mathbf{u}}, \mathbf{L}\overrightarrow{\mathbf{u}}\rangle_\Phi - \left\langle \mathbf{L}^t\overleftarrow{\mathbf{u}}, \overrightarrow{\mathbf{u}}\right\rangle_U = \int_V dV(\mathbf{x})\int_{t_0}^{t_1} dt\, \overleftarrow{u}^i(\mathbf{x},t)(\mathbf{L}\overrightarrow{\mathbf{u}})^i(\mathbf{x},t) - \int_V dV(\mathbf{x})\int_{t_0}^{t_1} dt \big(\mathbf{L}^t\overleftarrow{\mathbf{u}}\big)^i(\mathbf{x},t)\overrightarrow{\mathbf{u}}^i(\mathbf{x},t),$$

and, inserting (A-1) and (A-2),

$$\langle \overleftarrow{\mathbf{u}}, \mathbf{L}\overrightarrow{\mathbf{u}}\rangle_\Phi - \left\langle \mathbf{L}^t\overleftarrow{\mathbf{u}}, \overrightarrow{\mathbf{u}}\right\rangle_U = A + B, \tag{A-5}$$

where

$$A = \int_V dV(\mathbf{x})\rho(\mathbf{x})\int_{t_0}^{t_1} dt\left[\overleftarrow{u}^i(\mathbf{x},t)\frac{\partial^2\overrightarrow{u}^i}{\partial t^2}(\mathbf{x},t) - \frac{\partial^2\overleftarrow{u}^i}{\partial t^2}(\mathbf{x},t)\overrightarrow{u}^i(\mathbf{x},t)\right], \tag{A-6}$$

and

$$\begin{aligned} B = &-\int_V dV(\mathbf{x})\int_{t_0}^{t_1} dt\,\overleftarrow{u}^i(\mathbf{x},t)\left[\frac{\partial}{\partial x^j}\int_{-\infty}^{+\infty} dt'\overrightarrow{\Psi}^{ijkl}(\mathbf{x};t,t')\frac{\partial\overrightarrow{u}^k}{\partial x^l}(\mathbf{x},t')\right] \\ &+\int_V dV(\mathbf{x})\int_{t_0}^{t_1} dt\left[\frac{\partial}{\partial x^j}\int_{-\infty}^{+\infty} dt'\overleftarrow{\Psi}^{ijkl}(\mathbf{x};t,t')\frac{\partial\overleftarrow{u}^k}{\partial x^l}(\mathbf{x},t')\right]\overrightarrow{u}^i(\mathbf{x},t). \end{aligned} \tag{A-7}$$

As

$$\frac{\partial}{\partial t}\left[\overleftarrow{u}^i\frac{\partial\overrightarrow{u}^i}{\partial t} - \frac{\partial\overleftarrow{u}^i}{\partial t}\overrightarrow{u}^i\right] = \overleftarrow{u}^i\frac{\partial^2\overrightarrow{u}^i}{\partial t^2} - \frac{\partial^2\overleftarrow{u}^i}{\partial t^2}\overrightarrow{u}^i,$$

we have

$$A = \int_V dV(\mathbf{x})\rho(\mathbf{x})\left[\overleftarrow{u}^i(\mathbf{x},t)\frac{\partial^2\overrightarrow{u}^i}{\partial t^2}(\mathbf{x},t) - \frac{\partial^2\overleftarrow{u}^i}{\partial t^2}(\mathbf{x},t)\overrightarrow{u}^i(\mathbf{x},t)\right]\Bigg|_{t=t_0}^{t=t_1}. \tag{A-8}$$

Equation (A-7) can be rewritten

$$B = C + D, \tag{A-9}$$

where

$$\begin{aligned} C = &-\int_V dV(\mathbf{x})\int_{t_0}^{t_1} dt\frac{\partial}{\partial x^j}\left[\overleftarrow{u}^i(\mathbf{x},t)\left[\int_{-\infty}^{+\infty} dt'\overrightarrow{\Psi}^{ijkl}(\mathbf{x};t,t')\frac{\partial\overrightarrow{u}^k}{\partial x^l}(\mathbf{x},t')\right]\right] \\ &+\int_V dV(\mathbf{x})\int_{t_0}^{t_1} dt\frac{\partial}{\partial x^j}\left[\left[\int_{-\infty}^{+\infty} dt'\overleftarrow{\Psi}^{ijkl}(\mathbf{x};t,t')\frac{\partial\overleftarrow{u}^k}{\partial x^l}(\mathbf{x},t')\right]\overrightarrow{u}^i(\mathbf{x},t)\right], \end{aligned} \tag{A-10}$$

and

$$\begin{aligned} D &= +\int_V dV(\mathbf{x}) \int_{t_0}^{t_1} dt \frac{\partial \overleftarrow{u}^i}{\partial x^j}(\mathbf{x},t) \left[\int_{-\infty}^{+\infty} dt' \overrightarrow{\Psi}^{ijkl}(\mathbf{x};t,t') \frac{\partial \overrightarrow{u}^k}{\partial x^l}(\mathbf{x},t') \right] \\ &= \int_V dV(\mathbf{x}) \int_{t_0}^{t_1} dt \left[\int_{-\infty}^{+\infty} dt' \overleftarrow{\Psi}^{ijkl}(\mathbf{x};t,t') \frac{\partial \overleftarrow{u}^k}{\partial x^l}(\mathbf{x},t') \right] \frac{\partial \overrightarrow{u}^i}{\partial x^j}(\mathbf{x},t). \end{aligned} \tag{A-11}$$

Using the divergence theorem gives

$$\begin{aligned} C = &-\int_s dS(\xi) \int_{t_0}^{t_1} dt \overleftarrow{u}^i(\xi,t) \left[n^j(\xi) \int_{-\infty}^{+\infty} dt' \overrightarrow{\Psi}^{ijkl}(\xi;t,t') \frac{\partial \overrightarrow{u}^k}{\partial x^l}(\xi,t') \right] \\ &+ \int_s dS(\xi) \int_{t_0}^{t_1} dt \left[n^j(\xi) \int_{-\infty}^{+\infty} dt' \overleftarrow{\Psi}^{ijkl}(\xi;t,t') \frac{\partial \overleftarrow{u}^k}{\partial x^l}(\xi,t') \right] \overrightarrow{u}^i(\xi,t). \end{aligned} \tag{A-12}$$

The term D can be rewritten

$$\begin{aligned} D = &+\int_V dV(\mathbf{x}) \int_{t_0}^{t_1} dt \frac{\partial \overleftarrow{u}^i}{\partial x^j}(\mathbf{x},t) \left[\left[\int_{-\infty}^{t_0} dt' + \int_{t_0}^{t_1} dt' + \int_{t_1}^{+\infty} dt' \right] \overrightarrow{\Psi}^{ijkl}(\mathbf{x};t,t') \frac{\partial \overrightarrow{u}^k}{\partial x^l}(\mathbf{x},t') \right] \\ &- \int_V dV(\mathbf{x}) \int_{t_0}^{t_1} dt \left[\left[\int_{-\infty}^{t_0} dt' + \int_{t_0}^{t_1} dt' \right. \right. \\ &\qquad \left. \left. + \int_{t_1}^{+\infty} dt' \right] \overleftarrow{\Psi}^{ijkl}(\mathbf{x};t,t') \frac{\partial \overleftarrow{u}^k}{\partial x^l}(\mathbf{x},t') \right] \frac{\partial \overrightarrow{u}^i}{\partial x^j}(\mathbf{x},t), \end{aligned}$$

or, using the causality and anti-causality properties $\overrightarrow{\Psi}^{ijkl}(\mathbf{x};t,t')$ and $\overleftarrow{\Psi}^{ijkl}(\mathbf{x};t,t')$ respectively,

$$\begin{aligned} D = &\int_V dV(\mathbf{x}) \int_{t_0}^{t_1} dt \frac{\partial \overleftarrow{u}^i}{\partial x^j}(\mathbf{x},t) \left[\left[\int_{-\infty}^{t_0} dt' + \int_{t_0}^{t_1} dt' \right] \overrightarrow{\Psi}^{ijkl}(\mathbf{x};t,t') \frac{\partial \overrightarrow{u}^k}{\partial x^l}(\mathbf{x},t') \right] \\ &- \int_V dV(\mathbf{x}) \int_{t_0}^{t_1} dt \left[\left[\int_{t_0}^{t_1} dt' + \int_{t_1}^{+\infty} dt' \right] \overleftarrow{\Psi}^{ijkl}(\mathbf{x};t,t') \frac{\partial \overleftarrow{u}^k}{\partial x^l}(\mathbf{x},t') \right] \frac{\partial \overrightarrow{u}^i}{\partial x^j}(\mathbf{x},t), \end{aligned}$$

i.e.,

$$D = E + F, \tag{A-13}$$

where

$$\begin{aligned} E = &+\int_V dV(\mathbf{x}) \int_{t_0}^{t_1} dt \int_{t_0}^{t_1} dt' \frac{\partial \overleftarrow{u}^i}{\partial x^j}(\mathbf{x},t) \overrightarrow{\Psi}^{ijkl}(\mathbf{x};t,t') \frac{\partial \overrightarrow{u}^k}{\partial x^l}(\mathbf{x},t') \\ &- \int_V dV(\mathbf{x}) \int_{t_0}^{t_1} dt \int_{t_0}^{t_1} dt' \overleftarrow{\Psi}^{ijkl}(\mathbf{x};t,t') \frac{\partial \overleftarrow{u}^k}{\partial x^l}(\mathbf{x},t') \frac{\partial \overrightarrow{u}^i}{\partial x^j}(\mathbf{x},t), \end{aligned} \tag{A-14}$$

and

$$\begin{aligned} F = &+\int_V dV(\mathbf{x}) \int_{t_0}^{t_1} dt \frac{\partial \overleftarrow{u}^i}{\partial x^j}(\mathbf{x},t) \left[\int_{-\infty}^{t_0} dt' \overrightarrow{\Psi}^{ijkl}(\mathbf{x};t,t') \frac{\partial \overrightarrow{u}^k}{\partial x^l}(\mathbf{x},t') \right] \\ &- \int_V dV(\mathbf{x}) \int_{t_0}^{t_1} dt \left[\int_{t_1}^{+\infty} dt' \overleftarrow{\Psi}^{ijkl}(\mathbf{x};t,t') \frac{\partial \overleftarrow{u}^k}{\partial x^l}(\mathbf{x},t') \right] \frac{\partial \overrightarrow{u}^i}{\partial x^j}(\mathbf{x},t). \end{aligned} \tag{A-15}$$

Using (A-4) and the symmetries between indexes ij and kj gives

$$E = 0. \tag{A-16}$$

This ends the proof for equation (39) of the text.

Appendix B: Representation Theorem

Let $\overleftarrow{G}^{ij}(\mathbf{x}, t; \mathbf{x}', t')$ be defined by

$$\rho(\mathbf{x}) \frac{\partial^2 \overleftarrow{G}^{ip}}{\partial t^2}(\mathbf{x}, t; \mathbf{x}', t') - \frac{\partial}{\partial x^j} \int_{-\infty}^{+\infty} dt'' \overrightarrow{\Psi}^{ijkl}(\mathbf{x}; t, t'') \frac{\partial \overleftarrow{G}^{kp}}{\partial x^l}(\mathbf{x}, t''; \mathbf{x}', t') = \delta^{ip} \delta(\mathbf{x} - \mathbf{x}') \delta(t - t'), \tag{B-1a}$$

$$\overleftarrow{G}^{ip}(\mathbf{x}, t; \mathbf{x}', t') = 0 \quad \text{for} \quad t \geq t' \tag{B-1b}$$

$$\frac{\partial \overleftarrow{G}^{ip}}{\partial t}(\mathbf{x}, t; \mathbf{x}', t') = 0 \quad \text{for} \quad t \geq t', \tag{B-1c}$$

where $\overleftarrow{\Psi}^{ijkl}(\mathbf{x}; t, t')$ is anti-causal:

$$\overleftarrow{\Psi}^{ijkl}(\mathbf{x}; t, t') = 0 \quad \text{for} \quad t > t'. \tag{B-2}$$

For an arbitrary function $u^i(\mathbf{x}, t)$ we have, for any t', $t_0 < t' < t_1$,

$$u^p(\mathbf{x}', t') = \int_V dV(\mathbf{x}) \int_{t_0}^{t_1} dt \delta^{ip} \delta(\mathbf{x} - \mathbf{x}') \delta(t - t') u^i(\mathbf{x}, t). \tag{B-3}$$

Inserting (B-1a), into (B-3) gives

$$u^p(\mathbf{x}', t') = A + B, \tag{B-4}$$

where

$$A = \int_V dV(\mathbf{x}) \rho(\mathbf{x}) \int_{t_0}^{t_1} dt \frac{\partial^2 \overleftarrow{G}^{ip}}{\partial t^2}(\mathbf{x}, t; \mathbf{x}', t') u^i(\mathbf{x}, t), \tag{B-5}$$

and

$$B = -\int_V dV(\mathbf{x}) \int_{t_0}^{t_1} dt \left[\frac{\partial}{\partial x^j} \int_{-\infty}^{+\infty} dt'' \overleftarrow{\Psi}^{ijkl}(\mathbf{x}; t, t'') \frac{\partial \overleftarrow{G}^{kp}}{\partial x^l}(\mathbf{x}, t''; \mathbf{x}', t') \right] u^i(\mathbf{x}, t). \tag{B-6}$$

As

$$\frac{\partial^2 \overleftarrow{G}^{ip}}{\partial t^2} u^i = \frac{\partial}{\partial t} \left[\frac{\partial \overleftarrow{G}^{ip}}{\partial t^2} u^i - \overleftarrow{G}^{ip} \frac{\partial u^i}{\partial t} \right] + \overleftarrow{G}^{ip} \frac{\partial^2 u^i}{\partial t^2},$$

and, as $\overleftarrow{G}^{ip}(\mathbf{x},t;\mathbf{x}',t')$ is anti-causal, we have

$$\begin{aligned} A = &\int_V dV(\mathbf{x}) \int_{t_0}^{t_1} dt \overleftarrow{G}^{ip}(\mathbf{x},t;\mathbf{x}',t')\rho(\mathbf{x}) \frac{\partial^2 u^i}{\partial t^2}(\mathbf{x},t) \\ &- \int_V dV(\mathbf{x})\rho(\mathbf{x}) \left[\frac{\partial \overleftarrow{G}^{ip}}{\partial t}\big(\mathbf{x},t_0;\mathbf{x}',t'\big) u^i\big(\mathbf{x},t_0\big) - \overleftarrow{G}^{ip}\big(\mathbf{x},t_0;\mathbf{x}',t'\big) \frac{\partial u^i}{\partial t}\big(\mathbf{x},t_0\big) \right] . \end{aligned} \tag{B-7}$$

Equation (B-6) can be rewritten

$$B = C + D, \tag{B-8}$$

where

$$C = -\int_V dV(\mathbf{x}) \int_{t_0}^{t_1} dt \frac{\partial}{\partial x^j} \left[\int_{-\infty}^{+\infty} dt'' \overleftarrow{\Psi}^{ijkl}(\mathbf{x};t,t'') \frac{\partial \overleftarrow{G}^{kp}}{\partial x^l}(\mathbf{x},t'';\mathbf{x}',t') \right] u^i(\mathbf{x},t), \tag{B-9}$$

and

$$D = +\int_V dV(\mathbf{x}) \int_{t_0}^{t_1} dt \left[\int_{-\infty}^{+\infty} dt'' \overleftarrow{\Psi}^{ijkl}(\mathbf{x};t,t'') \frac{\partial \overleftarrow{G}^{kp}}{\partial x^l}(\mathbf{x},t'';\mathbf{x}',t') \right] \frac{\partial u^i}{\partial x^j}(\mathbf{x},t). \tag{B-10}$$

Using the divergence theorem gives

$$C = -\int_s dS(\xi) \int_{t_0}^{t_1} dt \left[n^j(\xi) \int_{-\infty}^{+\infty} dt'' \overleftarrow{\Psi}^{ijkl}(\mathbf{x};t,t'') \frac{\partial \overleftarrow{G}^{kp}}{\partial x^l}(\mathbf{x},t'';\mathbf{x}',t') \right] u^i(\mathbf{x},t). \tag{B-11}$$

Using

$$\overleftarrow{\Psi}^{ijkl}(\mathbf{x};t,t'') = \overrightarrow{\Psi}^{ijkl}(\mathbf{x};t'',t) \tag{B-12}$$

changing $t \leftrightarrow t''$, $ij \leftrightarrow kl$ and reordering gives

$$D = \int_V dV(\mathbf{x}) \int_{-\infty}^{+\infty} dt \frac{\partial \overleftarrow{G}^{ip}}{\partial x^j}(\mathbf{x},t;\mathbf{x}',t') \int_{t_0}^{t_1} dt'' \overrightarrow{\Psi}^{ijkl}(\mathbf{x};t,t'') \frac{\partial u^k}{\partial x^l}(\mathbf{x},t''). \tag{B-13}$$

The anti-causality property of $\overleftarrow{G}^{ip}(\mathbf{x},t;\mathbf{x}',t')$ and the causality property of $\overrightarrow{\Psi}^{ijkl}(\mathbf{x};t,t')$ allow to change the bounds of integration:

$$D = \int_V DV(\mathbf{x}) \int_{t_0}^{t_1} dt \frac{\partial \overleftarrow{G}^{ip}}{\partial x^j}(\mathbf{x},t;\mathbf{x}',t') \int_{t_0}^{+\infty} dt'' \overrightarrow{\Psi}^{ijkl}(\mathbf{x};t,t'') \frac{\partial u^k}{\partial x^l}(\mathbf{x},t''), \tag{B-14}$$

i.e.,

$$D = \int_V dV(\mathbf{x}) \int_{t_0}^{t_1} dt \frac{\partial \overleftarrow{G}^{ip}}{\partial x^j}(\mathbf{x},t;\mathbf{x}',t') \left[\int_{-\infty}^{+\infty} dt'' - \int_{-\infty}^{t_0} dt'' \right] \overrightarrow{\Psi}^{ijkl}(\mathbf{x};t,t'') \frac{\partial u^k}{\partial x^l}(\mathbf{x},t''). \tag{B-15}$$

This gives

$$D = E + F, \tag{B-16}$$

where

$$E = +\int_V dV(\mathbf{x}) \int_{t_0}^{t_1} dt \frac{\partial \overleftarrow{G}^{ip}}{\partial x^j}(\mathbf{x}, t; \mathbf{x}', t') \int_{-\infty}^{+\infty} dt'' \overrightarrow{\Psi}^{ijkl}(\mathbf{x}; t, t'') \frac{\partial u^k}{\partial x^l}(\mathbf{x}, t''), \tag{B-17}$$

and

$$F = -\int_V dV(\mathbf{x}) \int_{t_0}^{t_1} dt \frac{\partial \overleftarrow{G}^{ip}}{\partial x^j}(\mathbf{x}, t; \mathbf{x}', t') \int_{-\infty}^{t_0} dt'' \overrightarrow{\Psi}^{ijkl}(\mathbf{x}; t, t'') \frac{\partial u^k}{\partial x^l}(\mathbf{x}, t''). \tag{B-18}$$

Equation (B-17) can be written

$$E = G + H, \tag{B-19}$$

where

$$G = +\int_V dV(\mathbf{x}) \int_{t_0}^{t_1} dt \frac{\partial}{\partial x^j} \left[\overleftarrow{G}^{ip}(\mathbf{x}, t; \mathbf{x}', t') \int_{-\infty}^{+\infty} dt'' \overrightarrow{\Psi}^{ijkl}(\mathbf{x}; t, t'') \frac{\partial u^k}{\partial x^l}(\mathbf{x}, t'') \right], \tag{B-20}$$

and

$$H = -\int_V dV(\mathbf{x}) \int_{t_0}^{t_1} dt \overleftarrow{G}^{ip}(\mathbf{x}, t; \mathbf{x}', t') \left[\frac{\partial}{\partial x^j} \int_{-\infty}^{+\infty} dt'' \overrightarrow{\Psi}^{ijkl}(\mathbf{x}; t, t') \frac{\partial u^k}{\partial x^l}(\mathbf{x}, t'') \right]. \tag{B-21}$$

Using the divergence theorem gives

$$G = +\int_s dS(\xi) \int_{t_0}^{t_1} dt \overleftarrow{G}^{ip}(\xi, t; \mathbf{x}', t') \left[n^j(\xi) \int_{-\infty}^{+\infty} dt'' \overrightarrow{\Psi}^{ijkl}(\xi; t, t'') \frac{\partial u^k}{\partial x^l}(\xi, t'') \right]. \tag{B-22}$$

This ends the proof for equation (62) of the text.

Appendix C: Alternative Computation of the Gradient

The calculated displacement at the receiver locations $\mathbf{d} = \{\overrightarrow{u}^i(\mathbf{x}_r, t)\}$, is a nonlinear function of the model parameters $\mathbf{m} = \{\rho(\mathbf{x}),\ \Psi^{ijkl}(\mathbf{x}, \tau)\}$. A perturbation of the model parameters

$$\rho(\mathbf{x}) \longrightarrow \rho(\mathbf{x}) + \delta\rho(\mathbf{x}) \tag{C-1a}$$

$$\Psi^{ijkl}(\mathbf{x}, \tau) \longrightarrow \Psi^{ijkl}(\mathbf{x}, \tau) + \delta\Psi^{ijkl}(\mathbf{x}, \tau) \tag{C-1b}$$

leads to a perturbation of the displacement field

$$\overrightarrow{u}^i\big(\mathbf{x}_r, t\big) \longrightarrow \overrightarrow{u}^i\big(\mathbf{x}_r, t\big) + \delta \overrightarrow{u}^i\big(\mathbf{x}_r, t\big). \tag{C-1c}$$

Writing the first-order approximation of $\delta\mathbf{d} = \{\delta \overrightarrow{u}^i(\mathbf{x}_r, t)\}$ as

$$\delta\mathbf{d} = \mathbf{A}\delta\rho + \mathbf{B}\delta\Psi, \tag{C-2a}$$

or, explicitly

$$\overrightarrow{u}^i\big(\mathbf{x}_r, t\big) = \int_V dV(\mathbf{x}) A^i\big(\mathbf{x}_r, t; \mathbf{x}\big)\delta\rho(\mathbf{x}) + \int_V dV(\mathbf{x}) \int_{-\infty}^{+\infty} d\tau B^{ijklm}\big(\mathbf{x}_r, t; \mathbf{x}, \tau\big)\delta\Psi^{jklm}(\mathbf{x}, \tau), \tag{C-2b}$$

defines the derivative operators of calculated displacements with respect to model parameters and their kernels.

As demonstrated in Section 9,

$$\begin{aligned}\delta\overrightarrow{u}^p(\mathbf{x}',t') = &-\int_V dV(\mathbf{x})\int_{t_0}^{t_1} dt\,\overrightarrow{G}^{pi}_{\text{free}}(\mathbf{x}',t';\mathbf{x},t)\delta\rho(\mathbf{x})\frac{\partial^2\overrightarrow{u}^i}{\partial t^2}(\mathbf{x},t)\\ &-\int_V dV(\mathbf{x})\int_{t_0}^{t_1} dt\frac{\partial\overrightarrow{G}^{pi}_{\text{free}}}{\partial x^j}(\mathbf{x}',t';\mathbf{x},t)\int_{-\infty}^{+\infty} dt'\delta\Psi^{ijkl}(\mathbf{x};t,t')\frac{\partial\overrightarrow{u}^k}{\partial x^l}(\mathbf{x},t'),\end{aligned} \tag{C-3}$$

which can also be written, at the receiver locations,

$$\begin{aligned}\delta\overrightarrow{u}^i\big(\mathbf{x}_r,t\big) = &\int_V dV(\mathbf{x})\left[-\int_{t_0}^{t_1} dt'\,\overrightarrow{G}^{ij}_{\text{free}}\big(\mathbf{x}_r,t;\mathbf{x},t'\big)\frac{\partial^2\overrightarrow{u}^j}{\partial t'^2}(\mathbf{x},t')\right]\delta\rho(\mathbf{x})\\ &+\int_V dV(\mathbf{x})\int_{-\infty}^{+\infty} d\tau\left[-\int_{t_0}^{t_1} dt'\frac{\overrightarrow{G}^{ij}_{\text{free}}}{\partial x^k}\big(\mathbf{x}_r,t;\mathbf{x},t'\big)\frac{\partial\overrightarrow{u}^l}{\partial x^m}(\mathbf{x},t'-\tau)\right]\delta\Psi^{jklm}(\mathbf{x},\tau).\end{aligned} \tag{C-4}$$

By comparison with (C-2) this gives

$$A^i\big(\mathbf{x}_r,t;\mathbf{x}\big) = -\int_{t_0}^{t_1} dt'\,\overrightarrow{G}^{ij}_{\text{free}}\big(\mathbf{x}_r,t;\mathbf{x},t'\big)\frac{\partial^2\overrightarrow{u}^j}{\partial t'^2}(\mathbf{x},t'), \tag{C-5a}$$

and

$$B^{ijklm}\big(\mathbf{x}_r,t;\mathbf{x},\tau\big) = -\int_{t_0}^{t_1} dt'\frac{\partial\overrightarrow{G}^{ij}_{\text{free}}}{\partial x^k}\big(\mathbf{x}_r,t;\mathbf{x},t'\big)\frac{\partial\overrightarrow{u}^l}{\partial x^m}(\mathbf{x},t'-\tau). \tag{C-5b}$$

The operators **A** and **B** defined by (C-5) map the density and rate-of-relaxation spaces into the data space. Their transposes map the dual of the data space into the duals of the density and rate-of-relaxation spaces, i.e., for fixed $\delta\hat{\mathbf{d}}$ they give

$$\delta\hat{\rho} = \mathbf{A}^t\delta\hat{\mathbf{d}}, \tag{C-6a}$$

and

$$\delta\hat{\Psi} = \mathbf{B}^t\delta\hat{\mathbf{d}}, \tag{C-6b}$$

or, introducing their kernels,

$$\delta\hat{\rho}(\mathbf{x}) = \sum_r\int_{t_0}^{t_1} dt\big(A^t\big)^i\big(\mathbf{x},\mathbf{x}_r,t\big)\delta\hat{u}^i\big(\mathbf{x}_r,t\big), \tag{C-7a}$$

and

$$\delta\hat{\Psi}^{jklm}(\mathbf{x},\tau) = \sum_r\int_{t_0}^{t_1} dt\big(B^t\big)^{jklmi}\big(\mathbf{x},\tau,\mathbf{x}_r,t\big)\delta\hat{u}^i\big(\mathbf{x}_r,t\big), \tag{C-7b}$$

but, as the kernels of the transposes equal the transposes of the kernels, we can write

$$\delta\hat{\rho}(\mathbf{x}) = \sum_r\int_{t_0}^{t_1} dt A^i\big(\mathbf{x}_r,t,\mathbf{x}\big)\delta\hat{u}^i\big(\mathbf{x}_r,t\big), \tag{C-8a}$$

and

$$\delta\hat{\Psi}^{jklm}(\mathbf{x},\tau)=\sum_r\int_{t_0}^{t_1}dtB^{ijklm}\big(\mathbf{x}_r,t,\mathbf{x},\tau\big)\delta\hat{u}^i\big(\mathbf{x}_r,t\big), \tag{C-8b}$$

i.e.,

$$\delta\hat{\rho}(\mathbf{x})=-\sum_r\int_{t_0}^{t_1}dt\int_{t_0}^{t_1}dt'\overrightarrow{G}^{ij}_{\text{free}}\big(\mathbf{x}_r,t;\mathbf{x},t'\big)\frac{\partial^2\overrightarrow{u}^j}{\partial t'^2}(\mathbf{x},t')\delta\hat{u}^i\big(\mathbf{x}_r,t\big), \tag{C-9a}$$

and

$$\delta\hat{\Psi}^{jklm}(\mathbf{x},\tau)=-\sum_r\int_{t_0}^{t_1}dt\int_{t_0}^{t_1}dt'\frac{\partial\overrightarrow{G}^{ij}_{\text{free}}}{\partial x^k}\big(\mathbf{x}_r,t;\mathbf{x},t'\big)\frac{\partial\overrightarrow{u}^l}{\partial x^m}(\mathbf{x},t'-\tau)\delta\hat{u}^i\big(\mathbf{x}_r,t\big). \tag{C-9b}$$

Defining

$$\overleftarrow{u}^j(\mathbf{x},t')=\sum_r\int_{t_0}^{t_1}dt\overrightarrow{G}^{ij}\big(\mathbf{x}_r,t;\mathbf{x},t'\big)\delta\hat{u}^i\big(\mathbf{x}_r,t\big) \tag{C-10}$$

gives

$$\delta\hat{\rho}(\mathbf{x})=-\int_{t_0}^{t_1}dt'\overleftarrow{u}^j(\mathbf{x},t')\frac{\partial^2\overrightarrow{u}^j}{\partial t'^2}(\mathbf{x},t'), \tag{C-11a}$$

and

$$\delta\hat{\Psi}^{jklm}(\mathbf{x},\tau)=-\int_{t_0}^{t_1}dt'\frac{\partial\overleftarrow{u}^j}{\partial x^k}(\mathbf{x},t')\frac{\partial\overrightarrow{u}^l}{\partial x^m}(\mathbf{x},t'-\tau). \tag{C-11b}$$

Using (56b) of the text, the definition (C-10) can be rewritten

$$\overleftarrow{u}^j(\mathbf{x},t')=\sum_r\int_{t_0}^{t_1}dt\overleftarrow{G}^{ji}\big(\mathbf{x},t';\mathbf{x}_r,t'\big)\delta\hat{u}^i\big(\mathbf{x}_r,t\big), \tag{C-12}$$

i.e.,

$$\overleftarrow{u}^i(\mathbf{x},t)=\int_V dV(\mathbf{x}')\int_{t_0}^{t_1}dt'\overleftarrow{G}^{ij}(\mathbf{x},t;\mathbf{x}',t')\delta\phi^j(\mathbf{x}',t'), \tag{C-13}$$

where

$$\delta\phi^i(\mathbf{x},t)=\sum_r\delta\big(\mathbf{x}-\mathbf{x}_r\big)\delta\hat{u}^i\big(\mathbf{x}_r,t\big). \tag{C-14}$$

The representation theorem allows then the interpretation of $\overleftarrow{u}^i(\mathbf{x},t)$ as a field satisfying final conditions of rest, propagating with anti-causal attenuation, and due to the sources (C-14).

Using an integration per parts, the causality of $\overrightarrow{u}^i(\mathbf{x},t)$ and the anti-causality of $\overleftarrow{u}^i(\mathbf{x},t)$, equation (C-11a) can also be written

$$\delta\hat{\rho}(\mathbf{x})=\int_{t_0}^{t_1}dt'\frac{\partial\overleftarrow{u}^j}{\partial t'}(\mathbf{x},t')\frac{\partial u^j}{\partial t'}(\mathbf{x},t'). \tag{C-11c}$$

Formulas (C-11b) and (C-11c) correspond to formulas (114)–(115) of the text.

Acknowledgements

I thank my colleague G. Jobert for many helpful suggestions. This research has been sponsored by the European Commission, the French CNRS, Amoco, Aramco, C.G.G., Convex Computers, Elf-Aquitaine, I.F.P., Mobil, Shell, Sun Microsystems and Total-C.F.P.

References and References for General Reading

AKI, K. and RICHARDS, P. G. (1980), *Quantitative Seismology* (2 volumes) (Freeman and Co.).

ALTERMAN, Z. S. and KARAL, F. C., Jr. (1968), *Propagation of elastic waves in layered media by finite difference methods*, Bull. Seismol. Soc. of America 58, 367–398.

BACKUS, G. (1970), *Inference from inadequate and inaccurate data:* I, Proceedings of the National Academy of Sciences 65, 1, 1–105.

BACKUS, G. (1970), *Inference from inadequate and inaccurate data:* II, Proceedings of the National Academy of Sciences 65, 2, 281–287.

BACKUS, G. (1970), *Inference from inadequate and inaccurate data:* III, Proceedings of the National Academy of Sciences 67, 1, 282–289.

BACKUS, G. and GILBERT, F. (1967), *Numerical applications of a formalism for geophysical inverse problems,* Geophys. J. R. Astron. Soc. 13, 247–276.

BACKUS, G. and GILBERT, F. (1968), *The resolving power of gross Earth data,* Geophys. J. R. Astron. Soc. 16, 169–205.

BACKUS, G. and GILBERT, F. (1970), *Uniqueness in the inversion of inaccurate gross Earth data,* Philos. Trans. R. Soc. London 266, 123–192.

BACKUS, G. (1971), *Inference from inadequate and inaccurate data, Mathematical problems in the Geophysical Sciences:* Lecture in Applied Mathematics, 14, American Mathematical Society, Providence, Rhode Island.

BALAKRISHNAN, A. V., *Applied Functional Analysis* (Springer-Verlag, 1976).

BAMBERGER, A., CHAVENT, G., HEMON, Ch., and LAILLY, P. (1982), *Inversion of normal incidence seismograms,* Geophysics 47, 757–770.

BEN-MENAHEM, A. and SINGH, S. J., *Seismic Waves and Sources* (Springer-Verlag, 1981).

BENDER, C. M. and ORSZAG, S. A., *Advanced Mathematical Methods for Scientists and Engineers* (McGraw-Hill, 1978).

BERKHOUT, A. J., *Seismic Migration, Imaging of Acoustic Energy by Wave Field Extrapolation* (Elsevier, 1980).

BEYDOWN, W. (1985), *Asymptotic wave methods in heterogeneous media,* Ph.D Thesis, Massachusetts Institute of Technology.

BLAND, D. R., *The Theory of Linear Viscoelasticity* (Pergamon Press, 1960).

BLEISTEIN, *N., Mathematical Methods for Wave Phenomena* (Academic Press, 1984).

BOURBIE, T., COUSSY, O., and ZINSZNER, B. (1986), *Acoustique des milieux poreux,* Technip, 1986.

BOSCHI, E. (1973), *Body force equivalents for dislocations in visco-elastic media,* J. Geophys. Res. 78, 32, 7727–8590.

CLAERBOUT, J. F. (1971), *Toward a unified theory of reflector mapping,* Geophysics 36, 467–481.

CLAERBOUT, J. F., *Fundamentals of Geophysical Data Processing* (McGraw-Hill, 1976).

CLAERBOUT, J. F., *Imaging the Earth's Interior* (Blackwell, 1985).

CLAERBOUT, J. F. and MUIR, F. (1973), *Robust modelling with erratic data,* Geophysics 38, 5, 826–844.

COURANT, R. and HILBERT, D., *Methods of Mathematical Physics* (Interscience Publishers, 1953).

DAUTRAY, R. and LIONS, J. L., *Analyse mathématique et calcul numérique pour les sciences et les techniques* (3 volumes) (Masson, Paris, 1984 and 1985).

DAVIDON, W. C. (1959), *Variable metric method for minimization,* AEC Res. and Dev., Report ANL-5990 (revised).

DEVANEY, A. J. (1984), *Geophysical diffraction tomography,* IEEE Trans. Geos. Remote Sensing GE-22, No. 1.

FLETCHER, R., *Practical Methods of Optimization, Volume 1: Unconstrained Optimization* (Wiley, 1980).

FLETCHER, R., *Practical Methocis of Optimization, Volume 2: Constrained Optimization* (Wiley, 1981).

FRANKLIN, J. N. (1970), *Well posed stochastic extensions of ill posed linear problems,* J. Math. Anal. Applic. 31, 682–716.

GAUTHIER, O., VIRIEUX, J., and TARANTOLA, A. (1986), *Two-dimensional inversion of seismic waveforms: Numerical results,* Geophysics 51, 1387–1403.

HAMILTON, E. L. (1972), *Compressional-wave attenuation in marine sediments,* Geophysics 37, 620–646.

HERMAN, G. T., *Image reconstruction from projections, the fundamentals of computerized tomography* (Academic Press, 1980).

HSI-PING Liu, ANDERSON, D. L., and KANAMORI, H. (1976), *Velocity dispersion due to anelasticity; implications for seismology and mantle composition,* Geophys. J. R. Astr. Soc. 47, 41–58.

IKELLE, L. T., DIET, J. P., and TARANTOLA, A. (1986), *Linearized inversion of multi offset seismic reflection data in the w-k domain,* Geophysics 51, 1266–1276.

JACOBSON, R. S. (1987), *An investigation into the fundamental relationships between attenuation, phase dispersion, and frequency using seismic refraction profiles over sedimentary structures,* Geophysics 52, 72–87.

JOHNSON, L. R. (1974), *Green's function for Lamb's problem,* Geophys. J. R. Astr. Soc. 37, 99–131.

KENNETT, B,. L. N. (1974), *On variational principles and matrix methods in elastodynamics,* Geophys. J. R. Astr. Soc. 37, 391–405.

KENNETT, B. L. N. (1978) *Some aspects of non-linearity in inversion,* Geophys. J. R. Astr. Soc. 55, 373–391.

KJARTANSSON, E. (1979), *Constant Q-wave propagation and attenuation,* J. Geoph. Res. 84, B9, 4737–4748.

KOLB, P., COLLINO, F., and LAILLY, P. (1986), *Pre-stack inversion of a ID medium,* Proceedings of the IEEE 74, 3, 498–508.

LICKNEROWICZ, A., *Eléments de Culcul Tensoriel* (Armand Collin, Paris, 1960).

LIONS, J. L., *Contrôle optimal de systèmes gouvernés par des équations aux dérivées partielles* (Dunod, Paris, 1968). English translation: *Optimal control of systems governed by partial dtflerential equations* (Springer, 1971).

Liu, H. P., ANDERSON, D. L., and KANAMORI, H. (1976), *Velocity dispersion due to anelasticity; implications for seismology and mantle composition,* Geophys. J. R. Astr. Soc. 47, 41–58.

MILNE, R. D., *Applied Functional Analysis* (Pitman Advanced Publishing Program, Boston, 1980).

MISNER, Ch. W., THORNE, K. S., and WHEELER, J. A., Gravitation (Freeman, 1973).

MORA, P. (1988), *Elastic wavefield inversion of reflection and transmission data,* Geophysics 53, 750–759.

MORITZ, H., *Advanced Physical Geodesy* (Herbert Wichmann Verlag, Karlsruhe, Abacus Press, Tunbridge Wells, Kent, 1980).

MORSE, P. M. and FESHBACH, H., *Methods of Theoretical Physics* (McGraw-Hill, 1953).

NERCESSIAN, A., HIRN, A., and TARANTOLA, A. (1984) *Three-dimensional seismic transmission prospecting of the Mont-Dore volcano, France,* Geophys. J. R. Astr. Soc. 76, 307–315.

NOLET, G. (1985), *Solving or resolving inadequate and noisy tomographic systems,* J. Comp. Phys. 61, 463–482.

O'CONNELL, R. J. and BUDIANSKY, B. (1978), *Measures of dissipation in viscoelastic media,* Geophysical Research Letters 5, 1, 5–8.

PARK, J. and GILBERT, F. (1986), *Coupled free oscillations of an aspherical, dissipative, rotating Earth: Galerkin theory,* J. Geoph. Res. 91, B7, 7241–7260.

PARKER, R. L. (1975), *The theory of ideal bodies for gravity interpretation,* Geophys, J. R. Astron. Soc. 42, 315–334.

PARKER, R. L. (1977), *Understanding inverse theory,* Ann. Rev. Earth Plan. Sci. 5, 35–64.

PICA, A., DIET, J. P., and TARANTOLA, A. (1988), *Practice of nonlinear inversion of real seismic reflection data in a laterally invariant medium,* submitted to Geophysics.

POLACK, E. and RIBIÈRE, G. (1969), *Note sur la convergence de méthodes de directions conjuguées,* Revue Fr. Inf. Rech. Oper. 16-R1, 35–43.

POWELL, M. J. D., *Approximation Theory and Methods* (Cambridge University Press, 1981).

RAZAVY, M. and LEONACH, B. (1986), *Reciprocity principle and the approximate solution of the wave equation,* Bull. Seism. Soc. of America 76, 1776–1789.

RICKER, N. H., *Transient Waves in Visco-elastic Media* (Elsevier, 1977).

SCALES, L. E., *Introduction to Non-linear Optimization* (Macmillan, 1985).

SCHOENBERGER, M. and LEVIN, F. K. (1974), *Apparent attenuation due to intrubed multiples,* Geophysics 39, 278–291.

SCHWARTZ, L., *Méthodes mathématiques pour les sciences physiques* (Hermann, Paris, 1965).

SCHWARTZ, L., *Théorie des distributions* (Hermann, Paris, 1966).

SCHWARTZ, L., *Analyse (topologie générale et analyse fonctionelle)* (Hermann, Paris, 1970).

TANIMOTO, A. (1985), *The Backus-Gilbert approach to the three-dimensional structure in the upper mantle.—I. Lateral variation of surface wave phase velocity with its error and resolution,* Geophys. J. R. Astr. Soc. 82, 105–123.

TARANTOLA, A. (1984a), *Linearized inversion of seismic reflection data,* Geophysical Prospecting 32, 998–1015.

TARANTOLA, A. (1984b), *Inversion of seismic reflection data in the acoustic approximation,* Geophysics 49, 1259–1266.

TARANTOLA, A., *The seismic reflection inverse problem,* in *Inverse Problems of Acoustic and Elastic Waves* (edited by: F. Santosa, Y.-H. Pao, W. Symes, and Ch. Holland) (SIAM, Philadelphia, 1984).

TARANTOLA, A. (1986), *A strategy for nonlinear elastic inversion of seismic reflection data,* Geophysics 51, 1893–1903.

TARANTOLA, A., *Inverse Problem Theory; Methods for Data Fitting and Model Parameter Estimation* (Elsevier, 1987a).

TARANTOLA, A., *Inversion of travel time and seismic waveforms,* in *Seismic Tomography* (edited by G. Nolet) (Reidel, 1987b).

TARANTOLA, A., JOBERT, G., TRÉZÉGUET, D., and DENELLE, E. (1988), *Inversion of seismic waveforms can either be performed by time or by depth extrapolation,* Geophysical Prospecting 36, 383–416.

TARANTOLA, A. and NERCESSIAN, A. (1984), *Three-dimensional inversion without blocks,* Geophys. J. R. Astr. Soc. 76, 299–306.

TARANTOLA, A., and VALETTE, B. (1982a), *Inverse Problems = Quest for Information,* J. Geophys. 50, 159–170.

TARANTOLA, A. and VALETTE, B. (1982b), *Generalized nonlinear inverse problems solved using the least squares criterion,* Rev. Geophys. Space Phys. 20, 2, 219–232.

TAYLOR, A. E. and LAY, D. C., *Introduction to Functional Analysis* (Wiley, 1980).

VIRIEUX, J. (1986), *P-SV wave propagation in heterogeneous media; velocity-stress finite-difference method,* Geophysics 51, 889–901.

WALSH, G. R., *Methods of Optimization* (Wiley, 1975).

WATSON, G. A., *Approximation Theory and Numerical Methods* (Wiley, 1980).

WOODHOUSE, J. H. (1974), *Surface waves in a laterally varying layered structure,* Geophys. J. R. Astr. Soc. 37, 461–490.

WOODHOUSE, J. H. and DZIEWONSKI, A. M. (1984), *Mapping the upper mantle: Three-dimensional modeling of Earth structure by inversion of seismic waveforms,* Journal of Geophysical Research 89, B7, 5953–5986.

WU, R.-S., and BEN-MENAHEM, A. (1985), *The elastodynamic near field,* Geophys. J. R. Astr. Soc. 81, 609–621.

(Received June 10, 1987, revised January 10, 1988, accepted January 11, 1988)