Chapter 1

Band-limited Green's functions

1.1 A parametric model of the Green's function

When Green's functions are calculated at a few distinct frequencies they need to be interpolated to all frequencies to be useful in seismic imaging. The interpolation requires a model of the amplitude and phase behavior of the Green's functions.

The most common model is one that parameterizes the Green's function in terms of "arrival time" and "amplitude". This parameterization has the advantage that these quantities can be used directly in time domain Kirchhoff migration or modeling schemes. The Green's function can be characterized by one or more "events" that arrive at each location. Parameterization by a traveltime implies that the phase is a linear function of frequency for each event. This simple model can also be extended to incorporate frequency dispersion and attenuation in the Green's function. In a weakly dispersive model the phase becomes a smooth function of frequency rather than a linear function.

The justification for this event based model can be seen by looking at any seismic section. The data is not a random mish-mash of unrelated amplitudes, it has the appearance of a set of distinct events arriving at different times. The fact that distinct events are visible means that many frequencies are arriving at the same (or nearly the same) time and are constructively combining to produce a band-limited event. This implies that the phases of the different frequencies *for each event* are not random; they must follow an approximately linear trend as a function of frequency.

1.2 Single event models

The simplest model is one that is parameterized by one event. The simplicity of this model is very appealing but it is only valid in very smooth velocity fields. In a complex velocity field there are multiple paths from the source to one subsurface location, and thus multiple events. I will start by considering single event models and then progress to multiple event models.

1.2.1 One non-dispersive event

When there is only one un-dispersed arrival at each location, the Green's function from source location, $\mathbf{s}$, to subsurface location, $\mathbf{x}$, can be represented in terms of the amplitude, $A(\mathbf{s},\mathbf{x})$, traveltime, $\tau(\mathbf{s},\mathbf{x})$, and phase, $\phi_0(\mathbf{s},\mathbf{x})$, of that arrival.

$$G(\mathbf{s},\mathbf{x},\omega) = A(\mathbf{s},\mathbf{x})\, e^{i\phi_0(\mathbf{s},\mathbf{x})}\, e^{i\omega\tau(\mathbf{s},\mathbf{x})} \ .$$

The total phase at any frequency is $\omega\tau(\mathbf{s},\mathbf{x}) + \phi_0(\mathbf{s},\mathbf{x})$. If the Green's function fits this model then linear interpolation of unwrapped phase will exactly recover the phase for all frequencies. The amplitude is constant for all frequencies.

If, for the given velocity field, we know that there is only one non-dispersed arrival, we need only extrapolate two frequencies. The amplitude and unwrapped phase at each location provide enough information to completely specify the Green's function at that location. From the amplitude, $A(\mathbf{s},\mathbf{x},\omega)$, and unwrapped phase, $\zeta(\mathbf{s},\mathbf{x},\omega)$, at two frequencies ω_1 and ω_2 we have:

$$A(\mathbf{s},\mathbf{x}) = (A(\mathbf{s},\mathbf{x},\omega_1) + A(\mathbf{s},\mathbf{x},\omega_2))/2$$

$$\tau(\mathbf{s},\mathbf{x}) = (\zeta(\mathbf{s},\mathbf{x},\omega_2) - \zeta(\mathbf{s},\mathbf{x},\omega_1))/(\omega_2 - \omega_1)$$

$$\phi_0(\mathbf{s},\mathbf{x}) = \zeta(\mathbf{s},\mathbf{x},\omega_1) - \omega_1\tau(\mathbf{s},\mathbf{x})$$

If it can further be assumed that all the arrivals are zero phase then only one frequency is needed.

$$A(\mathbf{s},\mathbf{x}) = A(\mathbf{s},\mathbf{x},\omega_1)$$

$$\tau(\mathbf{s},\mathbf{x}) = \zeta(\mathbf{s},\mathbf{x},\omega_1)/\omega_1.$$

Implicit in many asymptotic schemes is the assumption that phase is a linear function of frequency for all frequencies from zero to very large frequencies. If this is not true the asymptotic methods may give solutions that are inappropriate for the seismic bandwidth. In contrast the simple interpolation of two frequencies in the seismic frequency band only assumes that phase is a linear function over that band.

1.2.2 One dispersed event

If there is still a single event, but it is dispersed, then the phase will not be a linear function of frequency and the amplitude will not be constant. There are two possible classifications for the dispersion. The first is "intrinsic dispersion"; one particular model of intrinsic dispersion, the constant-Q medium, can be modeled by using a complex frequency-dependent velocity in the wavefield extrapolation (**?**). The second type of dispersion is "geometric dispersion" caused by heterogeneity in the velocity model. In a model with a strong velocity gradient the low frequencies "feel" a larger region of the model and thus propagate with a different velocity. The single dispersed

event model will only be valid if the heterogeneity is large enough to cause dispersion but not so large that multiple arrivals are generated.

If the dispersion is mild, it is possible that the amplitude and phase functions can be approximated by piecewise linear functions. In this situation we can calculate the Green's function at several frequencies and linearly interpolate the amplitude and phase between each pair of frequencies. This model assumes that both the amplitude and phase are linear functions of frequency over some bandwidth $\omega_0 \leq \omega \leq \omega_1$. Within this band the Green's function can be modeled by an initial amplitude and phase, along with their gradients. Two Green's functions at frequencies ω_0 and ω_1 are required to fit this model:

$$A_0(x,z) = A(x,z,\omega_0) \tag{1.1}$$

$$\delta A(x,z) = (A(x,z,\omega_1) - A(x,z,\omega_0))/(\omega_1 - \omega_0) \tag{1.2}$$

$$\phi_0(x,z) = \phi(x,z,\omega_0) \tag{1.3}$$

$$\delta\phi(x,z) = (\phi(x,z,\omega_1) - \phi(x,z,\omega_0))/(\omega_1 - \omega_0)\ . \tag{1.4}$$

The Green's function at any frequency within this band can then be calculated by

$$G(x,z,\omega) = A(x,z,\omega)e^{i\phi(x,z,\omega)} \tag{1.5}$$

$$A(x,z,\omega) = A_0(x,z) + \delta A(x,z)(\omega - \omega_0) \tag{1.6}$$

$$\phi(x,z,\omega) = \phi_0(x,z) + \delta\phi(x,z)(\omega - \omega_0)\ . \tag{1.7}$$

This model allows the event to have a constant phase shift (as expected for waves that pass through a focus) and for an amplitude that varies as a function of frequency.

Figure 1.1 shows a time slice through the Greens' function for the same linear velocity gradient medium used in chapter **??**. Three mono-frequency Green's functions were calculated, and then interpolated to all other frequencies. The time domain representation was generated by applying an inverse FFT. The expanding wavefront has overturned waves near the top of the section. Figure 1.2 shows a depth slice through the Green's function. The inflexion on the limbs of the hyperbola corresponds to the offset at which the waves overturn. It is encouraging that all the major features expected in this Green's function have been recovered from an extrapolation of only three frequencies.

If there is strong intrinsic dispersion, then the phase will not be well approximated by a piecewise-linear function. In this case higher order interpolants could be used, e.g. linear interpolation in amplitude and quadratic interpolation in phase. Many other interpolation schemes could also be be employed; this is a possible area for further study.

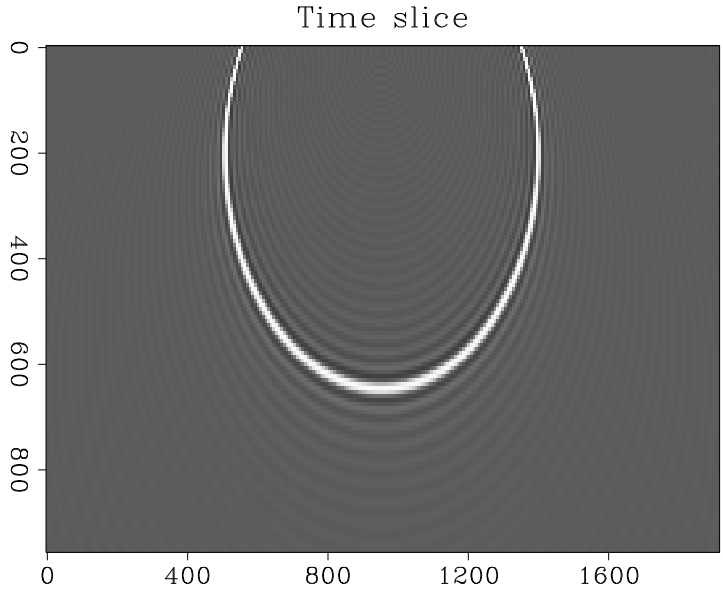

Figure 1.1: Time slice through the full Green's function in a V(z) medium. Note the overturned waves near the top of the panel. TTparam-tsl-greens [ER]

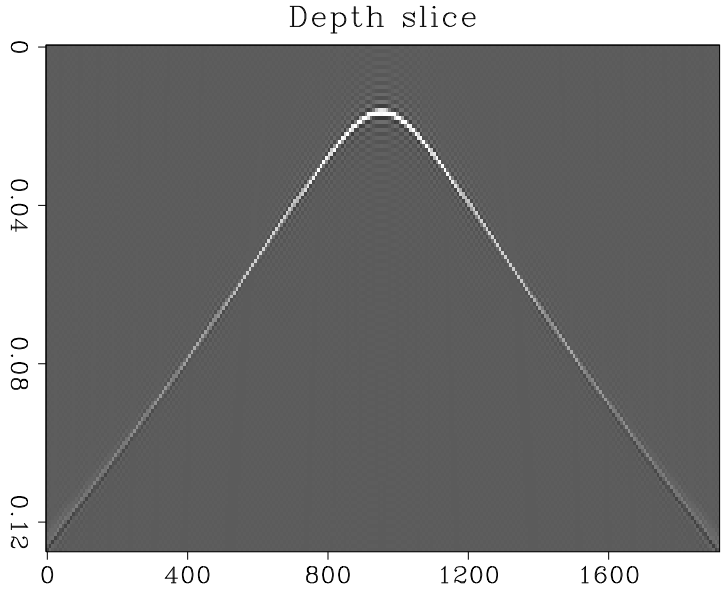

Figure 1.2: Depth slice through the full Green's function in a V(z) medium. TTparam-zsl-greens [ER]

1.3 Multiple events

When there are multiple arrivals at a single location the phase is no longer a linear, or even a smoothly varying, function of frequency. Consider the case of two arrivals with amplitudes A_1, A_2 and traveltimes τ_1 and τ_2. The wavefield is a linear superposition of the two arrivals.

$$P(\omega) = A_1 e^{i\omega\tau_1} + A_2 e^{i\omega\tau_2}.$$

However the phase of the combined wavefield is *not* a linear superposition of the two phases.

$$\begin{aligned} P(\omega) &= A_1\cos(\omega\tau_1) + A_2\cos(\omega\tau_2) + i(A_1\sin(\omega\tau_1) + A_2\sin(\omega\tau_2)) \\ \phi(\omega) &= \arctan\left(\frac{A_1\sin(\omega\tau_1) + A_2\sin(\omega\tau_2)}{A_1\cos(\omega\tau_1) + A_2\cos(\omega\tau_2)}\right) \end{aligned}$$

Figure 1.3 shows the unwrapped phase for the combination of two events, one of amplitude 1. at 40ms and one of amplitude 0.9 at 80ms. While the phase curve clearly has an overall linear trend it would not be well approximated by linear interpolation. Indeed it would be difficult to use any low order polynomial fit to a sparse selection of points on the curve.

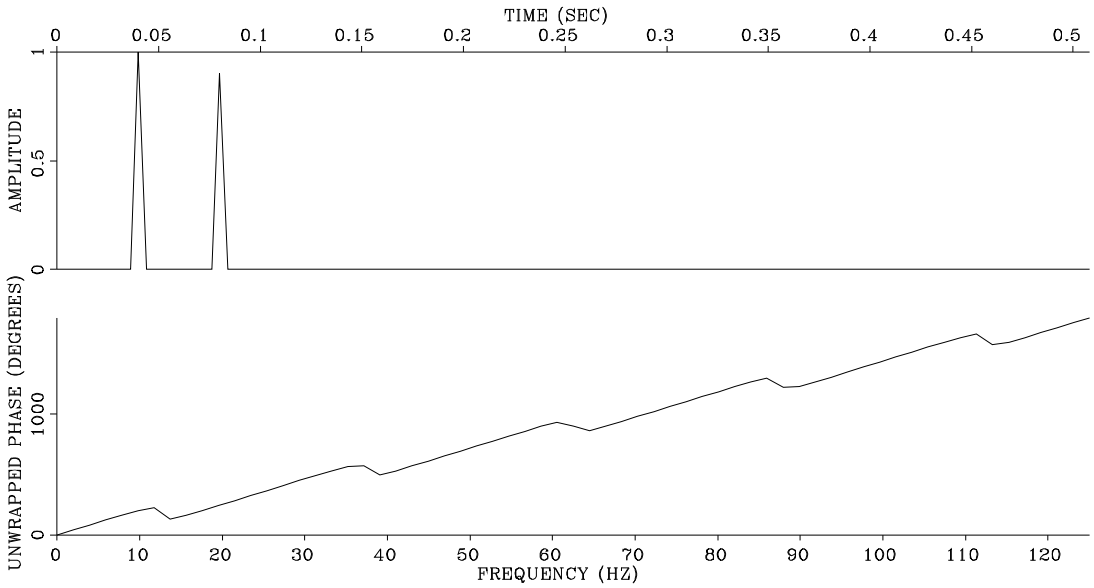

Figure 1.3: Time domain plot and unwrapped phase curve for the superposition of two discrete events. TTparam-twophase [CR]

1.3.1 A non-linear problem

To fit an n-event model to the calculated mono-frequency Green's functions, we must find amplitudes, A_i, phases, ϕ_i, and traveltimes, τ_i, such that the calculated Green's

function, $P(\omega)$, is predicted correctly for all modeled frequencies.

$$\sum_{i=1}^{n} A_i e^{i\phi_i} e^{i\omega\tau_i} = P(\omega)$$

An obvious way to solve this problem is to minimize a norm of the difference between the predicted and observed data:

$$\min_{A,\phi,\tau} || \sum_{i=1}^{n} A_i e^{i\phi_i} e^{i\omega\tau_i} - P(\omega) ||$$

Any of a large number of norms could be chosen but the L_2 norm is the most commonly used. This is a non-linear problem and its solution may be prone to problems associated with multiple minima, slow convergence, etc.

A different approach to the problem can be taken by noticing that it is the dual of a much more familiar problem. In geophysics, we often have a sampled time series and we wish to estimate a sparse, spiky, frequency spectrum. In this case we have a sampled frequency series and we wish to estimate a sparse, spiky, time-domain representation. A large number of existing algorithms could be adapted for this task. Many parametric spectral analysis methods are available. Some have been used in seismic signal deconvolution (**?**; **?**), and others have been used for high resolution spectral analysis in other fields (**?**; **?**; **?**). I chose to first try a very simple method based on the Fourier transform. This scheme was so successful that I did not experiment with the more complex methods. It is possible that the more complex methods can accurately estimate the traveltimes from Green's functions at fewer frequencies. This would be important if the cost of calculating one frequency is large. In 2-D the cost of calculating a few more frequencies is not excessive, in 3-D it might be a significant advantage to calculate fewer frequencies.

1.3.2 Event identification in the time domain

One simple way to identify the events is to inverse Fourier transform the data back to the time domain.

$$P(t) = \int_{-\infty}^{\infty} P(\omega) e^{-i\omega t} d\omega$$

Since the Green's function is not calculated for all frequencies, this integral is replaced by a discrete form. The Green's functions is calculated for the set of frequencies, $\omega_k = k\,\delta\omega \;;\; kl \le k \le kh$. The integral then becomes the sum

$$P(t) = \sum_{k=kl}^{kh} P(\omega_k)\, e^{-i\,k\,\delta\omega\,t} \delta\omega \;.$$

The discrete sampling in ω results in a replication in time:

$$P(t + n\,2\pi/\delta\omega) \;=\; \sum_{k=kl}^{kh} P(\omega_k)\, e^{-i\,k\,\delta\omega\,(t+n\,2\pi/\delta\omega)} \;=\; \sum_{k=kl}^{kh} P(\omega_k)\, e^{-i\,k\,\delta\omega\,t} e^{-ikn2\pi} \;=\; P(t)$$

Figure 1.4 shows the result of calculating the inverse transform with a sparse frequency sampling. The Green's function is plotted in the angle/time domain for a fixed radius. There are 128 time samples in each frame so the true, band-limited, Green's function is obtained by using 64 frequencies. In this model the wavefront is triplicating due to the focusing effect of the velocity perturbation; there are three events arriving at some locations. When 32 or 8 frequencies are used the structure of the true time domain response is still clear, even though it is replicated. When four frequencies are used all the different aliases overlap and the true response is hard to determine.

If the different aliases are non-overlapping, the problem of finding the correct time domain Green's function becomes one of choosing the correct alias and windowing out that portion of the data. If they are not distinct the problem is a harder. The correct Green's function is contaminated by the aliases and a more sophisticated scheme would need to be devised to separate out the correct signal.

1.3.3 Choosing maximum amplitude traveltimes

Given the replicated time domain Green's function, I now wish to parameterize the Green's function in terms of one, or more, events. If I only choose one event, it should be the highest energy event. The maximum energy event will be the best possible estimate (in a least-squares sense) that is in the form of a single set of three parameters (amplitude/traveltime/phase).

The maximum energy arrival can be found from the time-domain data by calculating the energy as a function of time and picking the maximum energy location. The energy function will be replicated in time due to the coarse sampling in frequency. However there is a simple way to choose the correct peak. The mono-frequency Green's functions that I have calculated are outgoing waves calculated on a polar grid. This means that if I know the traveltime for an event at one radius, then I know that the traveltimes at the next radius must be greater than the previous traveltimes. I also know that the traveltime difference must be less than the radius increment divided by the maximum velocity.

If there are many events, I can define a lower and upper bound for the traveltime window based on the minimum and maximum traveltimes at the previous radius. The replication interval due to sampling in frequency should be greater than the window size to prevent contamination by aliases. The size of the window and thus the frequency sampling that is used is a function of the difference between first and last traveltimes in the model.

In the current implementation I only pick one event (the maximum energy event) at each radius. I inverse transform the data in a time window centered around the traveltime on the previous radius. The length of the time window is chosen to be half the replication interval caused by frequency sampling. Since different events follow different paths the maximum amplitude traveltime at the current radius may be later or earlier than the traveltime at the previous radius.

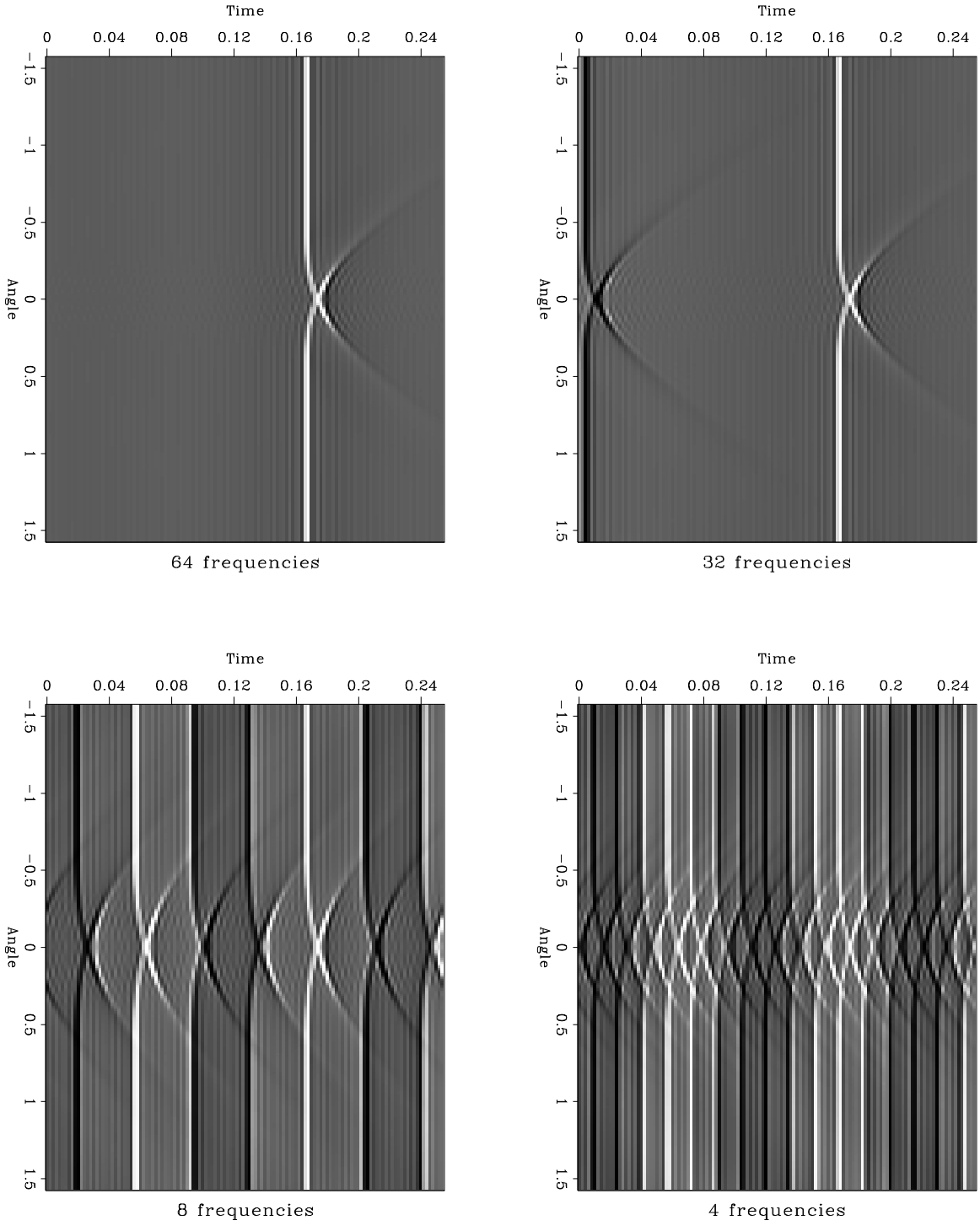

Figure 1.4: The Green's function for one radius, in a velocity model with a low velocity blob. The first frame was calculated from all 64 frequencies, the others used 32, 8, and 4 frequencies spanning the same range. The recovered time domain Green's function is replicated in time. TTparam-4blob [CR]

In practice, the limits imposed by the aliasing criterion may be slightly relaxed when picking the maximum amplitude. There may be aliased events within the search window but, as long as they are lower amplitude than the true maximum amplitude event, they will not be chosen. This makes the choice of number of frequencies a more problem-specific quantity.

The traveltime estimation is started at the initial radius of the calculation. At the radius both the wavefield, and the traveltime are known. The traveltime estimation then proceeds as follows:

1. Calculate the Green's functions for a sparse frequency sampling at the new radius.
2. Choose a time window centered around the traveltime from previous radius.
3. Calculate a sampled time domain representation in the window by slow Fourier transform.
4. Pick the maximum energy sample.
5. Use a quadratic fit to find the traveltime of the local peak of the energy function.
6. Calculate the amplitude, and phase at this traveltime.

The sampling used for the time window is related to the bandwidth over which the Green's functions have been calculated. It needs to be fine enough to resolve separate peaks in the energy function. A slow Fourier transform, rather than an FFT, is used in step 3 because the window typically consists of only a few time samples not centered at zero time.

In step 4 the estimate of maximum amplitude traveltime is refined using a quadratic fit to the energy function. The energy function $E(t)$ is the squared amplitude of the complex time function $P(t)$. Since

$$\begin{aligned} P(t) &= \sum_i P(\omega) e^{-i\omega_i t} \\ &= \left(e^{-i\omega_i t} \quad \cdots \quad e^{-i\omega_n t} \right) \begin{pmatrix} P(\omega_1) \\ \vdots \\ P(\omega_n) \end{pmatrix} , \end{aligned}$$

then

$$\begin{aligned} E(t) &= P(t)^* P(t) \\ &= \left(P^*(\omega_1) \quad \cdots \quad P^*(\omega_n) \right) \begin{pmatrix} e^{i\omega_i t} \\ \vdots \\ e^{i\omega_n t} \end{pmatrix} \left(e^{-i\omega_i t} \quad \cdots \quad e^{-i\omega_n t} \right) \begin{pmatrix} P(\omega_1) \\ \vdots \\ P(\omega_n) \end{pmatrix} \\ &= \left(P^*(\omega_1) \quad \cdots \quad P^*(\omega_n) \right) \mathbf{E} \begin{pmatrix} P(\omega_1) \\ \vdots \\ P(\omega_n) \end{pmatrix} , \end{aligned}$$

where the elements of $\mathbf{E}$ are $\mathbf{E}_{kl} = e^{i(\omega_k - \omega_l)t}$. To perform a quadratic fit to estimate the maximum, we need to calculate the first and second time-derivatives of the energy function at this location. Since the only time dependence is in the matrix $\mathbf{E}$, the derivatives can be simply calculated by taking the time derivative of this matrix.

$$\begin{aligned}
\frac{dE}{dt} &= \begin{pmatrix} P^*(\omega_1) & \cdots & P^*(\omega_n) \end{pmatrix} \mathbf{E}' \begin{pmatrix} P(\omega_1) \\ \vdots \\ P(\omega_n) \end{pmatrix} \\
\frac{d^2E}{dt^2} &= \begin{pmatrix} P^*(\omega_1) & \cdots & P^*(\omega_n) \end{pmatrix} \mathbf{E}'' \begin{pmatrix} P(\omega_1) \\ \vdots \\ P(\omega_n) \end{pmatrix}
\end{aligned}$$

Where the elements of $\mathbf{E}'$ are $\mathbf{E}'_{kl} = i(\omega_k - \omega_l)e^{i(\omega_k - \omega_l)t}$ and the elements of $\mathbf{E}''$ are $\mathbf{E}'_{kl} = -(\omega_k - \omega_l)^2 e^{i(\omega_k - \omega_l)t}$. Given these quantities we can find the local maximum of the energy function at time

$$t_{\max} = t - \frac{dE}{dt} \Big/ \frac{d^2E}{dt^2} \quad .$$

This can be interpreted either as a quadratic interpolation of the energy function or the first step of a Newton-Raphson iterative solution for the zero of the gradient. If a quadratic interpolation is not effective, more iterations could be performed. In the problems I have tested so far, one step seems to be sufficient.

Once the peak of the energy function has been picked, giving the traveltime, the amplitude and phase at that time are calculated from the complex trace, $P(t)$.

Figure 1.5 shows the traveltime, amplitude and phase calculated for the "velocity-blob" model, used in chapter **??**. This model produces a triplicating wavefront. The wavefront passes through a caustic at depth 950m, and there is a significant focusing of energy at this point. After the caustic there are three branches to the wavefront. In many locations the back branch (the last arriving) is the highest energy arrival. The phase plot shows a large phase change in the regions where a later arrival is picked. The waves that pass through the focus have a 90° phase shift in 2-D. When the events are very close in traveltime their energy peaks merge and the traveltime/amplitude/phase triplet that is picked is a combination of multiple events. This can be seen in the center of the triplicating region where an arrival of mixed phase is picked.

Figure 1.6 shows the traveltime contours of the maximum energy traveltime field superimposed on a snapshot of the full outgoing Green's function taken at a traveltime of .18 sec. The .18 sec contour overlays the wavefield exactly.

Figure 1.7 compares time slices of the Green's function estimated by outward extrapolation of all frequencies with the Green's function re-interpolated from the maximum energy traveltimes. The amplitude/traveltime/phase triplets have been converted back to a full time domain Green's function. Figure 1.8 compares depth slices of the two functions. The maximum energy traveltimes follow the back branch

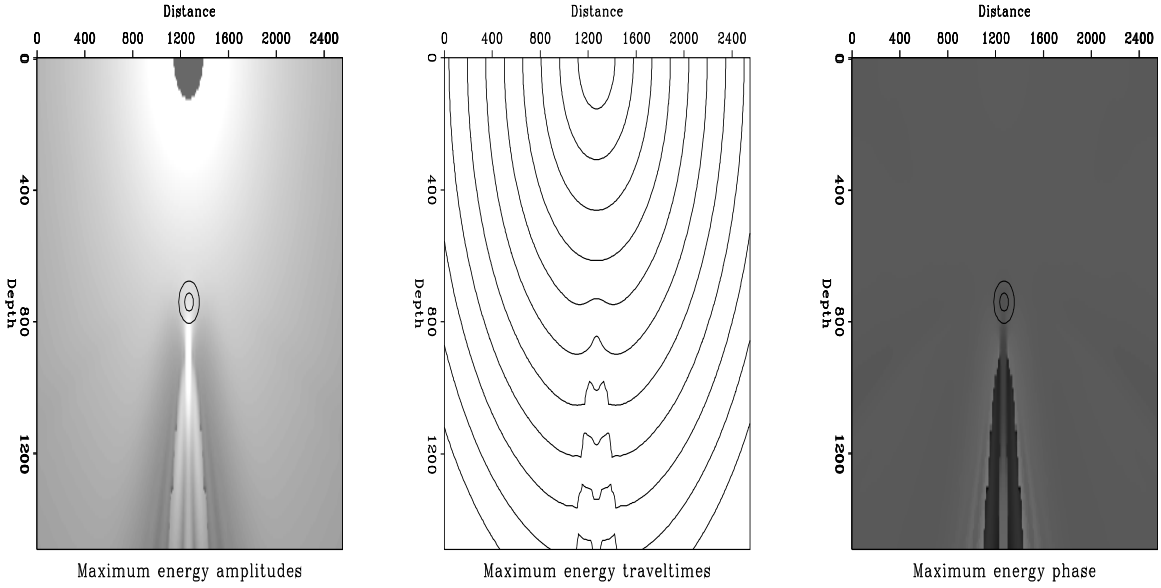

Figure 1.5: Amplitude, traveltime and phase plots for the "blob" model. The amplitude and phase plots have the contours of the velocity model superimposed on the plot. Note the focusing below the velocity anomaly in the amplitude plot and the phase changes in the phase plot TTparam-blob-ttamp [CR]

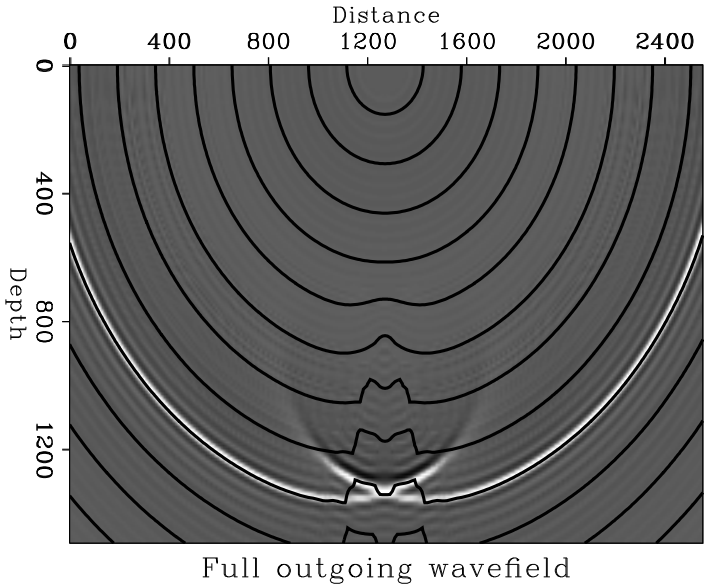

Figure 1.6: Maximum energy traveltime contours superimposed on a snapshot of the full outgoing Green's function. TTparam-blob-over [CR]

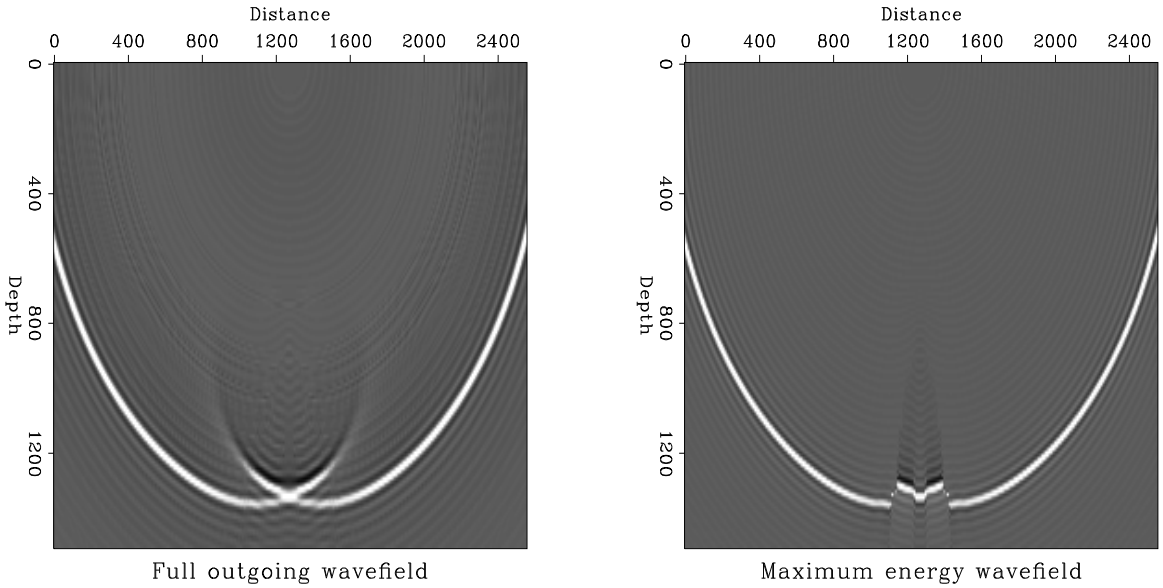

Figure 1.7: Time slice at .18sec of the full outgoing Green's function and the Green's function parameterized by the maximum energy traveltime field. TTparam-blob-compt [CR]

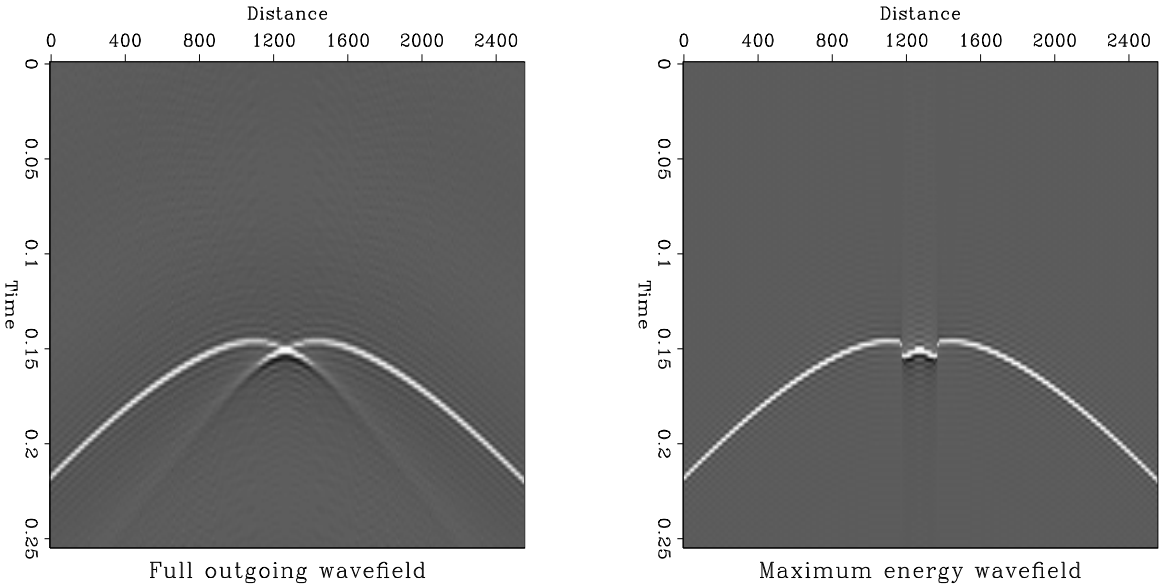

Figure 1.8: Depth slice at 1100m of the full outgoing Green's function and the wavefield reconstructed from the Green's function parameterized by the maximum energy traveltime field. TTparam-blob-compz [CR]

of the triplication at places where the energy is higher than the first arrival. The $\pi/2$ phase shift of the later arrival is correctly predicted. At the center, the two crossing events generate a higher amplitude than the back branch, so the combination of these two events is picked. This can be clearly seen in Figure 1.5.

In this simple model I could have chosen to pick three events at each location. This would have enabled me to completely reconstruct the outgoing Green's function. However, this requires knowledge of the number of arrivals. While I know the right number of events to choose for this simple model, it is unlikely to be so for a real earth model. Some of the more sophisticated algorithms can be used to estimate both the number of events and the times of the events. In this work I restrict myself to picking one event and choose the maximum energy arrival as that one event. In some models this may neglect significant energy in the Green's function and produce a poorer image of the data than might otherwise be possible. In practice I have found that one event is sufficient in the cases I have tested.

1.4 Comparison with other traveltime estimation methods

The two most common classes of methods for estimating traveltimes are those based on ray tracing, and those based on finite difference solutions to the Eikonal equation.

Finite difference solutions to the eikonal are very fast but they only calculate first arriving traveltimes. A first arriving traveltime field has the advantage that it is guaranteed to be continuous (its derivative is not necessarily continuous). The finite difference solution does not explicitly calculate raypaths but these may be infered from the gradient of the traveltime field. However the errors in the gradient are larger than the errors in the traveltimes and this can make the raypaths deduced from the traveltimes unreliable. A major advantage of the finite difference method is that it gives a solution everywhere on the computational grid. There are no "shadow-zones".

Ray tracing methods explicitly calculate the rays and then use the ray coordinate frame to calculate the traveltimes, amplitudes, and phases. Ray methods can calculate multiple arrivals if enough rays are traced. One major problem with ray tracing is difficulty in obtaining full coverage of the subsurface. There are shadow zones in many models where no rays will penetrate. The second major problem is associated with caustics in the wavefield. At the caustic the asymptotic amplitude is infinite and very near the caustic the ray methods can calculate unreasonably large amplitudes. These problems can be ameliorated by using a method such as paraxial ray tracing that interpolates the traveltimes and amplitudes using a paraxial approximation. When enough rays are traced the paraxial approximation can find a solution in the full subsurface. If the shadow zones are large the paraxial approximation may be poor but this may not matter too much as the amplitudes in these zones will be very weak. Care must be taken to avoid rays that pass directly through caustics. Amplitudes are often limited to some reasonable maximum value so that rays that pass close to a caustic do not give excessively high amplitudes.

Since ray tracing can model multiple arrivals the user has the choice of using first arrivals, maximum-amplitude arrivals or all arrivals in their imaging scheme. A maximum-amplitude Green's function has a continuous amplitude field but a discontinuous traveltime field. In contrast a first arriving Green's function has a continuous traveltime field but a discontinuous amplitude field. These discontinuities in the traveltime field may exacerbate aliasing problems in the imaging algorithms but the use of maximum amplitude events ensures that the most significant energy is imaged. In the next chapter I compare maximum amplitude traveltimes calculated by paraxial ray tracing and maximum energy traveltimes calculated using my band-limited approximation.

My method shares desirable features with both of these schemes. It is calculated on a regular computational mesh so there are no shadow zones, and a solution is calculated everywhere. However, unlike a finite difference solution to the eikonal equation, it calculates the maximum energy arrival rather than the first arrival. Like a maximum amplitude field from ray tracing, my method will produce a traveltime field that is discontinuous and an amplitude field that is continuous. Both of the other methods use a high frequency approximation to the wave equation, my method calculates a solution using the wave equation in the same bandwidth as the seismic data.