

Appendix C

Pseudo-spectral ray tracing with dynamic parameterisation

C.1 Introduction

In this appendix an algorithm is presented for the calculation of seismic ray fields in 2-D smooth media. The method is specifically designed for use in complex but smooth velocity structures where multi-pathing takes place.

The original objective was to make this algorithm part of a perturbation approach that would be able to handle variations in the number of ray arrivals under perturbations of the medium. This line of research was abandoned, however, in favour of the more practical ray field map methods which are central to the main text of this thesis. Nevertheless, the algorithm is interesting enough to be presented here. Its development provided useful insights for the ray field map approach (e.g., Appendix B) and some of its features may be useful in other applications.

The algorithm has two distinctive features related to the numerical representation of the ray field. First, the wave fronts are approximated using a pseudo-spectral expansion. Most ray tracing techniques propagate each ray individually, and sometimes approximate a piece of wave front by local interpolation (e.g., Vinje et al., 1993) or extrapolation (e.g., Beydoun and Keho, 1987). The new algorithm propagates entire wave fronts, represented by their pseudo-spectral expansion coefficients.

Second, it is shown that the parameterisation of the wave front is not necessarily determined by the natural parameterisation of the corresponding ray field, i.e. the parameterisation in terms of initial conditions for the rays. Using the smoothness and continuity of the wave front the parameterisation is adjusted dynamically in

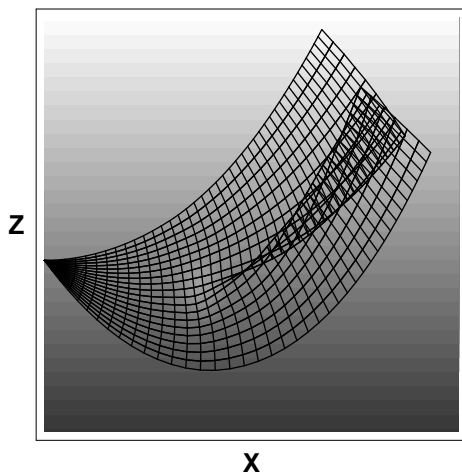


Figure C.1: A ray field emitted from a point source in a simple velocity structure. The velocity structure consists of a constant gradient of the slowness pointing in negative z -direction, and a slow anomaly embedded in it. This anomaly leads to a folding of the ray field and a consequent triplication of the number of arrivals in the region enclosed by the cuspid caustic.

order to control the accuracy of the expansion.

C.2 Pseudo-spectral ray field calculation

C.2.1 Ray field

Subject of study is a central (point source) ray field in a 2-D medium, parameterised by two ray parameters: a flow parameter varying along the rays and a single ensemble parameter for the initial conditions. If the medium is smooth, each ray variable (position, slowness, traveltime, etc.) is a smooth function of the flow parameter. Similarly, if the initial conditions are smooth functions of the ensemble parameter, the entire ray field can be described by smooth functions of the ray parameters. In spatially varying velocity structures, rays within the ray field eventually cross in physical space and the number of rays crossing a particular point of space may change from place to place (see Figure C.1).

The information of interest in seismic imaging algorithms for example is usually a map of ray information on a grid in physical space. If rays cross in physical space, these maps will be multi-valued and discontinuous. It is important to realise that these irregularities are caused by the projection of the ray information onto physical space, not by discontinuities in the ray field. This can be verified by picturing the ray field in a higher dimensional space. The ray field defines a

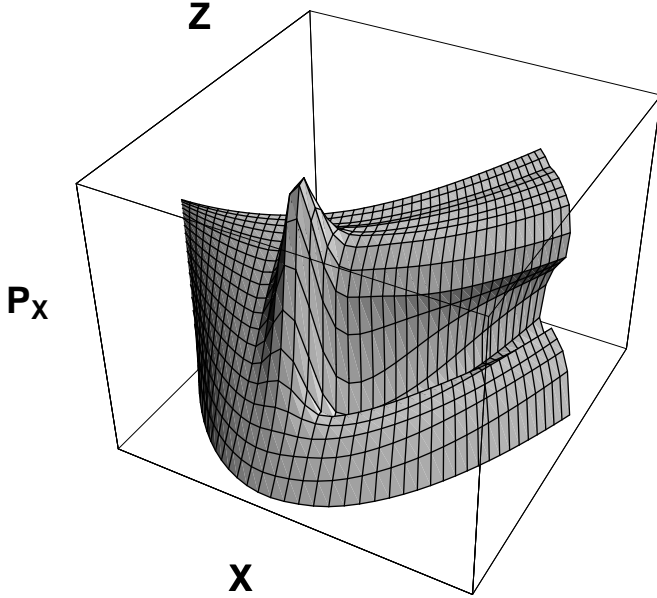


Figure C.2: The same ray field as in Figure C.1, but now in a 3-D projection of the 4D phase space consisting of the physical space at the horizontal plane and the x -component of the slowness along the vertical axis. The rays span a smooth manifold in this subspace, the projection of which onto physical space gives the rays of Figure C.1.

smooth, 2-D manifold in 4-D phase space, a 3-D projection of which can be seen in Figure C.2.

C.2.2 Wave front expansion

In this section and the following the phase space coordinates are written as a single vector \mathbf{y} , which contains both spatial coordinates and slowness components:

$$\mathbf{y} = \begin{pmatrix} \mathbf{x} \\ \mathbf{p} \end{pmatrix}. \quad (\text{C.1})$$

The system of ray equations is summarised by

$$\frac{d\mathbf{y}}{d\sigma} = \mathbf{F}(\mathbf{y}), \quad (\text{C.2})$$

with \mathbf{F} an unspecified right hand side vector (for more details see Chapter 2).

The usual approach in ray tracing algorithms is to solve (C.2) for a discrete set of initial conditions parameterised by ensemble parameter ξ . Here, however,

a continuous estimate of the ray field is calculated. This is accomplished by expanding the ξ -dependence into a set of basis-functions $B_j(\xi)$ that are orthonormal with respect to an associated inner product:

$$\int_{\xi_0}^{\xi_1} B_i(\xi)B_j(\xi)W(\xi)d\xi = \delta_{ij}. \quad (\text{C.3})$$

Here $W(\xi)$ is a weighting function, and δ is the Kronecker delta. The ray field may thus be expressed as a series expansion:

$$\mathbf{y}(\sigma, \xi) = \sum_{j=0}^N \mathbf{c}_j(\sigma)B_j(\xi). \quad (\text{C.4})$$

Each vector of coefficients \mathbf{c}_j , which depends on σ , may be calculated using the relation

$$\mathbf{c}_j(\sigma) = \int_{\xi_0}^{\xi_1} \mathbf{y}(\sigma, \xi)B_j(\xi)W(\xi)d\xi. \quad (\text{C.5})$$

In order to trace the coefficients \mathbf{c}_j , the right hand side of equation (C.2) has to be expanded likewise, with coefficients $\mathbf{f}_j(\sigma)$. This yields a differential equation for the coefficients:

$$\frac{d\mathbf{c}_j(\sigma)}{d\sigma} = \mathbf{f}_j(\sigma). \quad (\text{C.6})$$

If the flow parameter σ corresponds to time, then for a constant σ , $\mathbf{x}(\sigma, \xi)$ describes the position of a wave front. Although this does not hold for other types of flow parameter, in the following these ‘‘ray fronts’’ will be referred to as wave fronts as well.

Tracing a progressing wave front with a certain accuracy requires in general a growing number of coefficients. The actual number of coefficients depends on the complexity of the wave front, which in turn is determined by the complexity of the medium traversed so far.

C.2.3 Pseudo-spectral methods

Global expansion methods are popular for the solution of partial differential equations, where they are collectively known as spectral methods. The usual approach is to expand the spatial dependence of the desired solution in a set of suitable basis functions, ending up with a set of ordinary differential equations for the temporal dependence (the method of lines). The current application is slightly different, because the dependence on the initial conditions of an ordinary differential equation

is expanded. Nevertheless, many aspects of both problems are the same, and the existing literature can be used (e.g., Boyd, 2000; Fornberg, 1996). This section provides a summary of the basic principles of spectral methods.

Given an equation

$$Lu = f(x), \tag{C.7}$$

where L is an operator, which is not necessarily linear, the solution is approximated by the series expansion

$$u(x) \approx u_N(x) = \sum_{n=0}^N a_n \phi_n(x). \tag{C.8}$$

Substitution of this series in the equation generates a residual function R :

$$R(x; a_n) = Lu_N - f. \tag{C.9}$$

Spectral methods are designed to find a series of coefficients a_n in such a way that the residual function is made as small as possible. The different spectral and pseudo-spectral techniques differ mainly in their way of minimising R .

The basic ingredients of a spectral method are:

- A suitable set of basis functions. The most desirable properties are completeness, rapid evaluation, rapid derivation, rapid convergence.
- An error criterion (norm), needed to quantify the quality of the approximation.
- An algorithm to determine the coefficients.

The most commonly used basis functions are Chebyshev polynomials, the default choice for functions on finite intervals, and trigonometric functions (Fourier) for periodic functions. Different sets are useful in case of special geometries (e.g. spherical harmonics) or unbounded intervals (e.g. Laguerre functions).

Since the quality of the approximation is usually determined by the maximum error on the entire interval, the norm of choice would be the L_∞ , or sup-norm. The problem with this norm is that it is difficult to deal with analytically. The most practical norm is the L_2 , or least-squares-norm, because it allows fast algorithms for the determination of coefficients, especially if the basis set is orthogonal.

Basically there are two types of algorithms to determine the expansion coefficients. The classical methods (Galerkin, Lanczos “tau”) minimise the residual function by making it orthogonal to as many basis functions as possible. The modern collocation methods minimise the residual by making it zero at a finite set of points. The latter methods are usually referred to as pseudo-spectral methods. Choosing the right set of collocation points makes the pseudo-spectral methods equivalent to the classical methods.

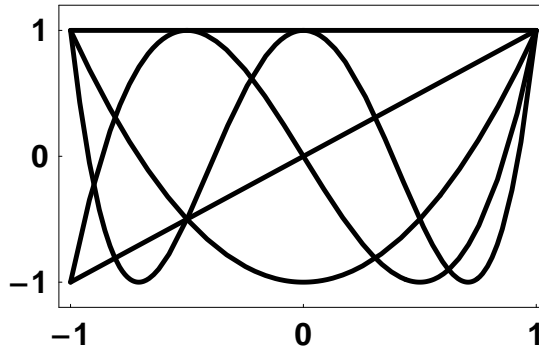


Figure C.3: The first five Chebyshev polynomials. These polynomials oscillate between $+1$ and -1 and form a basis for spectral expansion. For analytic functions the convergence is exponential and uniform over the expansion interval.

In the current application the pseudo-spectral technique is used in combination with Chebyshev polynomials. An advantage of this choice is that it provides uniform exponential convergence in the sup-norm for smooth functions. Algorithms for the Chebyshev expansion may be found in Press et al. (1992). The first five Chebyshev polynomials are displayed in Figure C.3.

C.2.4 Examples

Figure C.4 shows the application of the ray field expansion technique to the same model studied in Figures C.1 and C.2. Time is used as flow parameter and the integration of the coefficients in time was performed using a fourth order Runge-Kutta scheme. The wave fronts are the same as in Figure C.1. The rays, however, are now replaced by small ray segments. Although no individual rays were traced, the segments are plotted to provide insight into the structure of the ray field. The number of ray segments connecting two wave front is equal to the number of terms used in the expansion of the second of the two fronts. The number of coefficients increases during propagation.

Figures C.5 through C.7 show an example for a ray field traced in a low velocity layer. Figure C.5 depicts the ray field for a point source, for a scalar Hamiltonian ray tracing system with the choice $n = 2$ (see Section 2.7). Figure C.6 shows how the number of coefficients used depends on the complexity of the wave front, and may decrease as well as increase. Figure C.7 shows the logarithmic magnitude of the expansion coefficients for the z -coordinate of the last wave front in Figure C.5, and its linear trend reveals exponential convergence. This exponential convergence behaviour was initially considered an asset of the method, but unfortunately it is only observed for simple models like this one.

A problem that may occur is that the length scales of variations of the wave

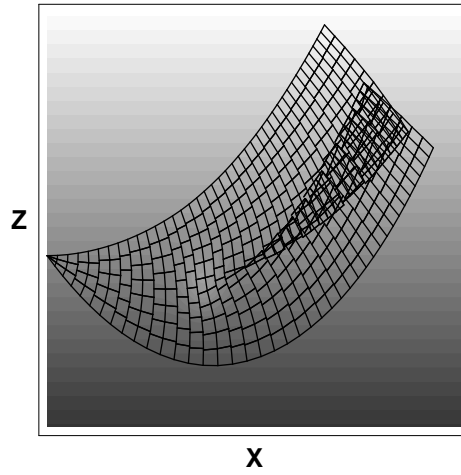


Figure C.4: The same ray field as in Figure C.1, but now traced using the ray field expansion technique.

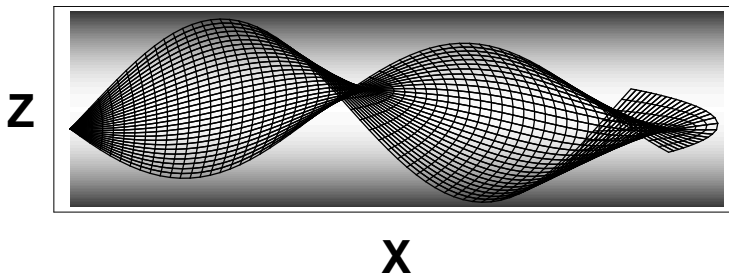


Figure C.5: A fan of rays in a low velocity layer calculated using the expansion approach. The rays are drawn for reference.

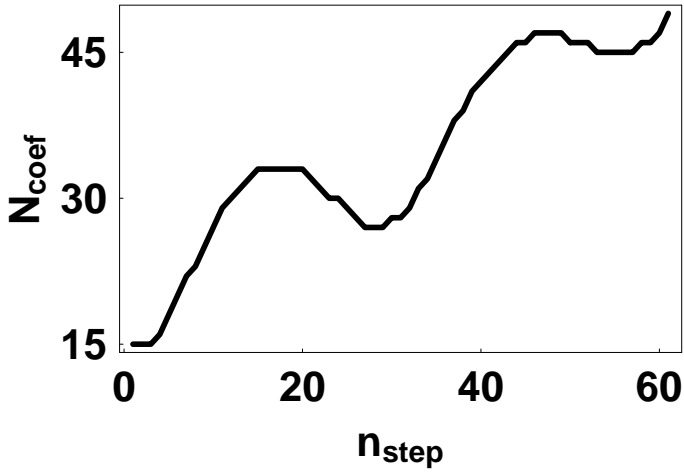


Figure C.6: The number of coefficients is related to the complexity of the wave front. In special circumstances, as in the ray field of Figure C.5, the wave front may get simpler in time which is reflected in the number of coefficients used.

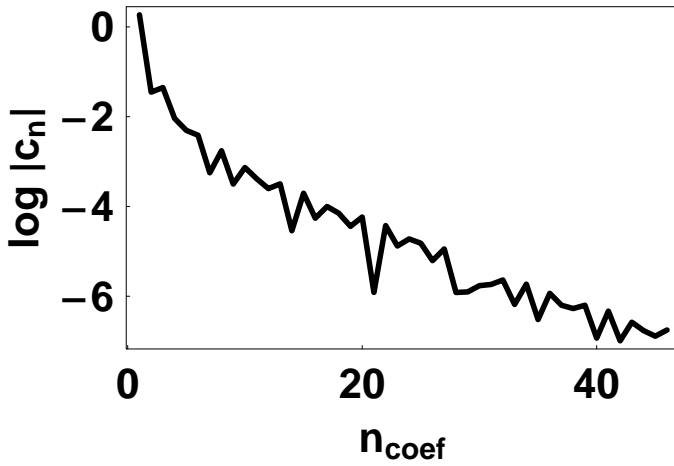


Figure C.7: The Chebyshev spectrum for z in the last wave front of Figure C.5 clearly shows the property of exponential convergence.

front may vary much over the range of ξ . Especially in regions of large geometrical spreading, the variations on a wave front may be large for a small range of take-off angles. The sampling density needed to approximate the wave front adequately is determined by those ranges with the smallest length scales of variations. Since the sampling is homogeneously distributed over the range of ξ — actually the sampling distribution is proportional to the weighting function $W(\xi)$ mentioned in Section C.2.2, as in Gaussian quadrature — this results in oversampling and thus inefficiency in regions of slower variations.

The phenomenon is illustrated by the example of Figures C.8 and C.9. The model consists of two Gaussian shape velocity anomalies. The first anomaly that the wave front meets is fast and causes a shadow zone with very large geometrical spreading. In order to maintain sufficient accuracy to actually see the second anomaly behind the first one, the algorithm needs to use a large amount of expansion terms. This results in severe oversampling in the areas of less geometrical spreading as is seen from the density of ray segments at the sides of the ray ensemble.

Figure C.9 shows the graphs of the ray variables as functions of the ensemble parameter ξ for the last wave front in Figure C.8. It is clear that the variations in these functions are mostly concentrated in a small region of ξ and thus it is this part that determines the number of coefficients to be used in the spectral expansion. This problem is dealt with in the following two sections.

C.3 Dynamic parameterisation

In the examples of Section C.2.4 take-off angles at the source have been used as ensemble parameters. It turns out that this is not a very efficient choice, because the variations of the various ray field variables are not distributed homogeneously over the range of take-off angles. The question is now what parameterisation to use in order to obtain maximal efficiency of the expansion technique.

The highest efficiency is achieved if the number of expansion terms needed for a certain level of accuracy is minimal. Since the function to be approximated is in fact a vector function, the best configuration is such that each vector component needs approximately the same number of terms to achieve its respective accuracy goal. Another way to put this is that all components of the vector should be comparably smooth.

To be able to deal with this analytically, a smoothness or roughness measure for vector functions has to be defined. The roughness of a single-valued function is often measured by an integral over the square of the derivative. An obvious generalisation of this is to define the roughness of a vector function by an integral over a weighted sum of the squared derivatives of all components.

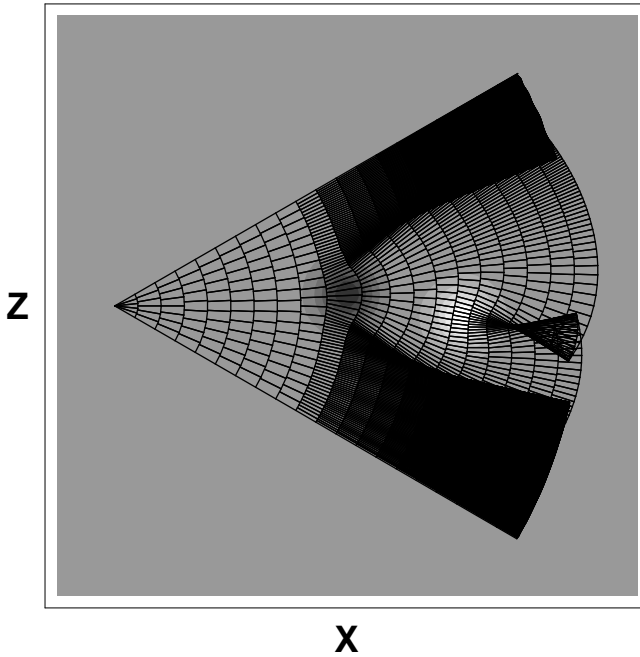


Figure C.8: An example of how large variations in geometrical spreading lead to a non-uniform sampling in the ray field expansion approach.

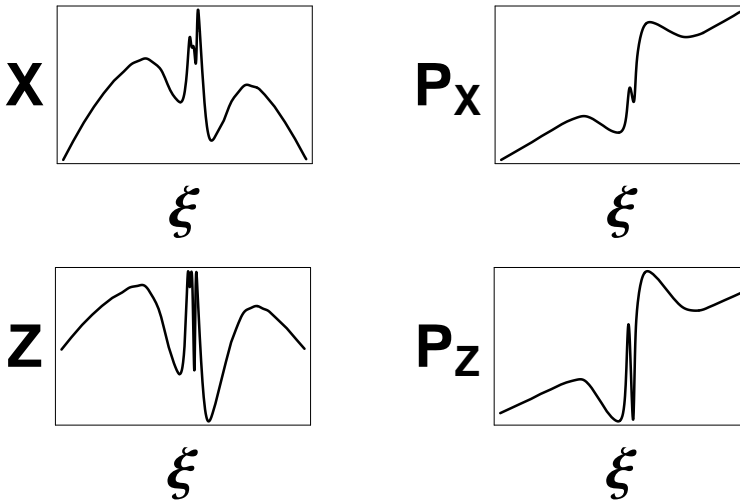


Figure C.9: Graphs of the ray field variables as a function of ensemble parameter ξ . The graphs correspond to the last wave front plotted in Figure C.8

C.3.1 Phase space metric

In a formal sense this is realised by defining a metric, or norm, in the vector space on which the function is defined. In our case, the vector space is the phase space, and the metric may be defined by

$$dl^2 = \langle d\mathbf{y}, d\mathbf{y} \rangle_{\mathbf{A}} = d\mathbf{y} \cdot \mathbf{A} \cdot d\mathbf{y}, \quad (\text{C.10})$$

where $d\mathbf{y}$ is a differential in phase space and \mathbf{A} is a positive definite matrix containing weights. For a one parameter curve $\mathbf{y}(\gamma)$ in phase space this metric determines the length of an infinitesimal segment of the curve by

$$\left(\frac{dl}{d\gamma} \right)^2 = \left\langle \frac{d\mathbf{y}}{d\gamma}, \frac{d\mathbf{y}}{d\gamma} \right\rangle_{\mathbf{A}}. \quad (\text{C.11})$$

The roughness measure R for such a one parameter curve in phase space may then be defined as

$$R = \int_{\gamma_0}^{\gamma_1} \left(\frac{dl}{d\gamma} \right)^2 d\gamma. \quad (\text{C.12})$$

A proper definition of the phase space metric would be such that the parameterisation γ that minimises the roughness R , also minimises the number of expansion terms needed for any component of \mathbf{y} . Since the optimal metric will probably be different for each situation and cannot be determined without extensive calculations, the most practical thing to do is to define a parametric form for the metric that is adequate for most circumstances, and that may be adapted to specific situations by a rule-of-thumb.

The definition of a metric in phase space implies the summation of coordinate and slowness terms, which have a different physical meaning and hence different ranges of values. In order to compensate this at least one degree of freedom is required in the metric. The weight matrix can be defined as

$$\mathbf{A} = \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \alpha \mathbf{I} \end{pmatrix}, \quad (\text{C.13})$$

with α , the *phase space geometry factor*, the desired degree of freedom. This yields the metric

$$dl^2 = \langle d\mathbf{y}, d\mathbf{y} \rangle_{\mathbf{A}} = d\mathbf{x} \cdot d\mathbf{x} + \alpha d\mathbf{p} \cdot d\mathbf{p}. \quad (\text{C.14})$$

The challenge now is to find the parameterisation that minimises the roughness of a wave front. To find it, it is important to recognise the following two facts. First, the smoothness of the Lagrangian manifold in phase space ensures that for any non-degenerate parameterisation γ , the metric derivative $dl/d\gamma$ is strictly positive. Second, the integral of this metric derivative is equal to the length of the wave front

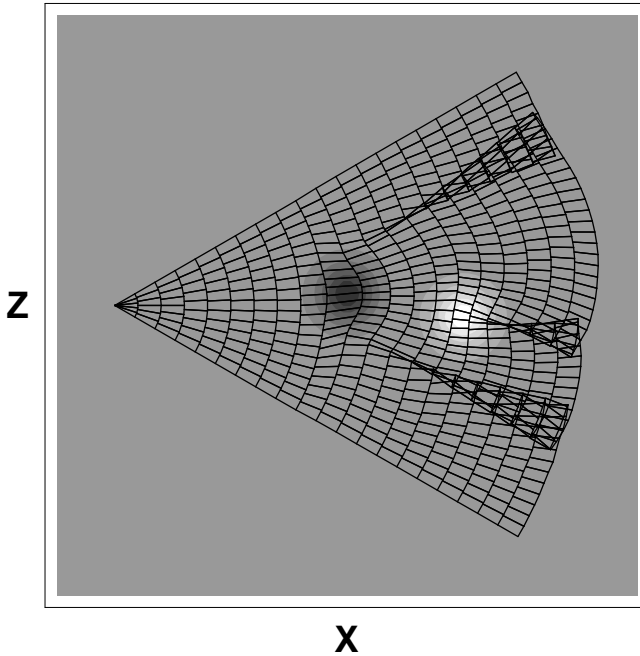


Figure C.10: The same ray field as in Figure C.8, now using the continuous reparameterisation formalism. The ray field is now homogenously sampled.

in phase space corresponding to the chosen metric, and hence is independent of the actual parameterisation. Combination of these facts suggests that the desired parameterisation is the one that renders the length of a differential wave front segment constant:

$$\frac{dl}{d\gamma} = C(\sigma). \quad (\text{C.15})$$

This can be shown by means of the calculus of variations (see Appendix C.A.1). A possible approach to determine a value for α is explained in Appendix C.A.2.

After choosing this optimal parameterisation the next challenge is how to achieve it in practice. Suppose that criterion (C.15) is met for a given wave front, then simply using the system (C.2) to propagate to the next value of σ will violate it instantly. In order to maintain optimal parameterisation throughout the ray field integration, a dynamic reparameterisation formalism is presented in Section C.3.2.

C.3.2 Dynamic wave front parameterisation

Since an ensemble of rays is traced at once, for each value of σ not only the derivative along the ray ($\partial \mathbf{y} / \partial \sigma$) is known, but also the derivative along the wave front ($\partial \mathbf{y} / \partial \xi$). This makes it possible to integrate the ray field not along the ray paths themselves, but along more general paths. This is accomplished by adding an extra term to the ray equations which is proportional to the derivative along the wave front. In essence this implies a reparameterisation of the ray field, which is represented by the new symbol $\tilde{\mathbf{y}}$:

$$\tilde{\mathbf{y}}(\lambda, \gamma) = \mathbf{y}(\sigma(\lambda, \gamma), \xi(\lambda, \gamma)). \quad (\text{C.16})$$

The new parameters λ and γ may be compared to σ and ξ respectively. Lines of constant λ parameterise the wave fronts, while lines of constant γ do not correspond to rays anymore, since the γ -dependence of $\tilde{\mathbf{y}}$ is basically a reparameterisation of the ξ -dependence of \mathbf{y} . For an arbitrary function $f(\lambda, \gamma)$ the following holds:

$$\frac{\partial \tilde{\mathbf{y}}}{\partial \lambda} = \mathbf{F}(\tilde{\mathbf{y}}) + f(\lambda, \gamma) \frac{\partial \tilde{\mathbf{y}}}{\partial \gamma}, \quad (\text{C.17})$$

in which \mathbf{y} now depends on λ and γ as in the right hand side of (C.16). Equation (C.17) may be used as a basis for an expansion approach as in Section C.2. Appendix C.A.3 contains a derivation of (C.17).

The trick is now to choose $f(\lambda, \gamma)$ in (C.17) in such a way that for the ray field solution equation (C.15) holds for every value of λ . To find the proper constraint on $f(\lambda, \gamma)$ it is required that:

$$\frac{d}{d\lambda} \left(\frac{dl}{d\gamma} \right) = D(\lambda), \quad (\text{C.18})$$

where the value of $D(\lambda)$ is the derivative of $C(\lambda)$, which in turn is the equivalent to $C(\sigma)$ from equation (C.15). Hence the meaning of $D(\lambda)$ is the derivative of the length of the wave front with λ . Considering

$$\frac{\partial}{\partial \lambda} \left\langle \frac{\partial \tilde{\mathbf{y}}}{\partial \gamma}, \frac{\partial \tilde{\mathbf{y}}}{\partial \gamma} \right\rangle_{\mathbf{A}} = 2 \left\langle \frac{\partial \tilde{\mathbf{y}}}{\partial \gamma}, \frac{\partial^2 \tilde{\mathbf{y}}}{\partial \gamma \partial \lambda} \right\rangle_{\mathbf{A}}, \quad (\text{C.19})$$

for any choice of metric, using (C.11) and (C.17) leads to

$$\frac{\partial}{\partial \lambda} \left(\frac{dl}{d\gamma} \right) = \left(\frac{dl}{d\gamma} \right)^{-1} \left[\left\langle \frac{\partial \tilde{\mathbf{y}}}{\partial \gamma}, \frac{\partial \mathbf{F}(\tilde{\mathbf{y}})}{\partial \gamma} \right\rangle_{\mathbf{A}} + f(\lambda, \gamma) \left\langle \frac{\partial \tilde{\mathbf{y}}}{\partial \gamma}, \frac{\partial^2 \tilde{\mathbf{y}}}{\partial \gamma^2} \right\rangle_{\mathbf{A}} + \left(\frac{\partial f(\lambda, \gamma)}{\partial \gamma} \right) \left(\frac{dl}{d\gamma} \right)^2 \right]. \quad (\text{C.20})$$

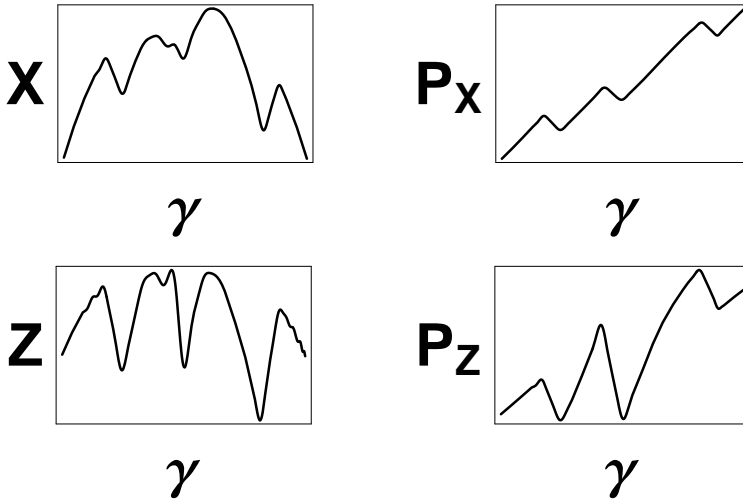


Figure C.11: Graphs of the ray field variables as a function of ensemble parameter γ . The graphs correspond to the last wave front plotted in Figure C.10, and are much smoother than the corresponding ones in Figure C.9

Since the magnitude of $d\tilde{\mathbf{y}}/d\gamma$ is constant, the second term in the right hand side vanishes. This is easily verified if by recognising

$$\left\langle \frac{\partial \tilde{\mathbf{y}}}{\partial \gamma}, \frac{\partial^2 \tilde{\mathbf{y}}}{\partial \gamma^2} \right\rangle_{\mathbf{A}} = \left(\frac{d\mathbf{l}}{d\gamma} \right) \frac{\partial}{\partial \gamma} \left(\frac{d\mathbf{l}}{d\gamma} \right). \quad (\text{C.21})$$

Together, equations (C.18) and (C.20) yield an expression for the derivative of $f(\lambda, \gamma)$, which should be satisfied everywhere to achieve optimal parameterisation:

$$\frac{\partial f(\lambda, \gamma)}{\partial \gamma} = \frac{D(\lambda)}{C(\lambda)} - \left\langle \frac{\partial \tilde{\mathbf{y}}}{\partial \gamma}, \frac{\partial \mathbf{F}(\tilde{\mathbf{y}})}{\partial \gamma} \right\rangle_{\mathbf{A}} C(\lambda)^{-2}. \quad (\text{C.22})$$

Hence $f(\lambda, \gamma)$ is determined except for two constants which can be used to make $f(\lambda, \gamma)$ vanish at the boundaries of the ensemble. Although the right hand side of (C.17) is more expensive to calculate than the original one (C.2), the new method is more efficient because the wave fronts require fewer expansion terms.

The dynamic reparameterisation is illustrated in Figures C.10 and C.11. These Figures show the same model as Figures C.8 and C.9 from Section C.2.4, but now computed using the reparameterisation approach. It is obvious both from the ray field plot and from the graphs that the reparameterisation does a good job in reducing the number of expansion terms needed.

C.4 Ray field mapping

The integration of the ray tracing system (C.17) yields ray-related variables such as position, slowness, amplitude, traveltime as a function of the ray parameters (λ, γ) . Finding out which rays arrive at a certain grid-position \mathbf{x}^* means solving the non-linear system

$$\mathbf{x}(\lambda, \gamma) = \mathbf{x}^*, \quad (\text{C.23})$$

the ray field map discussed in Chapter 3. There may be multiple solutions, each corresponding to a distinct arrival. The total number of solutions is unknown a priori. In general this means that finding all solutions in such a system is an exhaustive search. The problem may be attacked in three ways.

The first way and conceptually simplest way is to divide the ray field into cells in a way similar to the wave front construction techniques (e.g., Lambaré et al., 1996), followed by inverse interpolation inside the cells, see also Chapter 4.

Another way is to locate the caustics in the ray field. The caustics, expressed in terms of the ray parameters, separate different branches of the ray field, each of which has a single solution to the projection algorithm.

Finally, it is possible to use the fact that the total number of real and complex solutions to a polynomial system of equations is known a priori. The real solutions correspond to the ray-geometric arrivals. The choice of Chebyshev polynomials to expand the ray field allows to find all solutions. For example if one of the coordinates is used as independent parameter λ , each arrival is found by solving numerically a single polynomial equation. If another independent parameter is used, its dependence should also be fitted by a polynomial, which then yields a system of two bivariate polynomial equations. Similarly in 3-D this gives a system of either two or three polynomial equations in as many variables.

Although solving a system of polynomial equations is numerically demanding, the work may be reduced by using the solutions of one grid point as starting values for next. The polynomial system for neighbouring grid points is only slightly different, which allows the use of perturbation methods.

C.5 Discussion and conclusions

A new method for calculating ray-fields in 2-D smooth media was developed. The essence of the method is to consider the complete ray field as a single object. The dependence of the rays on the initial conditions is approximated using a pseudo-spectral expansion in terms of Chebyshev polynomials. The result of a calculation is a continuous analytic representation of the ray-field which may subsequently be used to project ray information onto a grid defined in the medium, to generate tables of possibly multi-valued travel times and amplitudes for use in imaging and source location algorithms.

In order to overcome difficulties of the method related to variations in geometrical spreading over the range of initial conditions a technique for dynamic reparameterisation was developed. It allows a more efficient calculation of the ray-field and its projections by abandoning the ray itself as the basis of the expansion.

The proposed technique is more complex and less flexible than for example the wave front construction methods and it is not easily extended to 3-D. It is not likely that it will be able to compete with wave front construction for speed in the ray field calculation. The primary purpose was to use the resulting analytic description of the ray field as a basis for perturbation techniques. These techniques were not developed, however, because the methods proposed in the following chapters are more powerful.

C.A Appendix

C.A.1 Optimal parameterisation

For a general function $f(x)$, which increases monotonically between x_0 and x_1 , a reparameterisation is sought in terms of y ($y_0 \leq y \leq y_1$) such that the roughness condition of equation (C.12) is minimised. Write

$$\tilde{f}(y) = f(x(y)), \quad (\text{C.24})$$

with $\tilde{f}(y_0) = f(x_0)$ and $\tilde{f}(y_1) = f(x_1)$. For the roughness integral this yields

$$R_y = \int_{y_0}^{y_1} \left(\frac{df(x)}{dx} \frac{dx(y)}{dy} \right)^2 dy. \quad (\text{C.25})$$

Using calculus of variations (e.g., Lanczos, 1986) the requirement $\delta R_y = 0$ gives the Euler-Lagrange equation

$$\frac{d}{dy} \left(\frac{df(x)}{dx} \frac{dx(y)}{dy} \right) = \frac{d}{dy} \left(\frac{d\tilde{f}(y)}{dy} \right) = 0. \quad (\text{C.26})$$

This shows that the derivative of the function in the new parameterisation is constant. The constant is determined by the prescribed values of the functions on the boundaries of the domain.

C.A.2 Determination of phase space metric

The purpose of the phase space metric and the phase space geometry factor α (C.3.1) is to be able to relate variations in slowness to variations in spatial coordinates. Comparable smoothness of these quantities along the range of γ may only be achieved if the metric is more or less equally sensitive to variations in both.

The goal is to choose α in such a way that both terms on the right hand side of (C.3.1) are of similar scale. A rule-of-thumb to determine α can be found through scale analysis. The slowness of the medium $u(\mathbf{x})$ is approximated by a constant U , and the gradient of the slowness by

$$\nabla u(\mathbf{x}) \approx \frac{\delta U}{L_U}, \quad (\text{C.27})$$

where δU is representative for the variations in the slowness, and L_U the corresponding length scale. Now first consider the case where all space coordinates and slowness components are dependent variables. The square of the metric derivative is then given by

$$\left(\frac{dl}{d\gamma}\right)^2 = \frac{d\mathbf{x}}{d\gamma} \cdot \frac{d\mathbf{x}}{d\gamma} + \alpha \frac{d\mathbf{p}}{d\gamma} \cdot \frac{d\mathbf{p}}{d\gamma}. \quad (\text{C.28})$$

The slowness may be represented by

$$\mathbf{p} = u(\mathbf{x})\hat{\mathbf{n}}, \quad (\text{C.29})$$

where the dependence on σ is implicit, and

$$\hat{\mathbf{n}}(\gamma) = \begin{pmatrix} \cos \theta(\gamma) \\ \sin \theta(\gamma) \end{pmatrix}, \quad (\text{C.30})$$

with $\theta(\gamma)$ the angle of \mathbf{p} with the x -axis. This gives

$$\frac{d\mathbf{p}}{d\gamma} \cdot \frac{d\mathbf{p}}{d\gamma} = \left(\nabla u(\mathbf{x}) \cdot \frac{d\mathbf{x}}{d\gamma}\right)^2 + u(\mathbf{x})^2 \left(\frac{d\theta}{d\gamma}\right)^2. \quad (\text{C.31})$$

For the scale analysis each derivative to γ is substituted by an estimate, as in the case of \mathbf{x} :

$$\left|\frac{d\mathbf{x}}{d\gamma}\right| \approx \frac{\Delta r}{\Delta\gamma}, \quad (\text{C.32})$$

which gives

$$\left(\frac{dl}{d\gamma}\right)^2 \approx \left(\frac{\Delta r}{\Delta\gamma}\right)^2 + \alpha \left[\left(\frac{\delta U}{L_U} \frac{\Delta r}{\Delta\gamma}\right)^2 + U^2 \left(\frac{\Delta\theta}{\Delta\gamma}\right)^2 \right]. \quad (\text{C.33})$$

A value for α is now found by equating the first and the second term of the right hand side:

$$\begin{aligned} \alpha &\approx \left(\frac{\Delta r}{\Delta\gamma}\right)^2 \left[\left(\frac{\delta U}{L_U} \frac{\Delta r}{\Delta\gamma}\right)^2 + U^2 \left(\frac{\Delta\theta}{\Delta\gamma}\right)^2 \right]^{-1} \\ &= \left(\frac{\delta U^2}{L_U^2} + \frac{U^2}{R^2}\right)^{-1}, \end{aligned} \quad (\text{C.34})$$

where R , defined as

$$R = \frac{\Delta r}{\Delta \theta}, \quad (\text{C.35})$$

is a representative value for the radius of curvature of the wave front. It depends on the medium and the initial conditions which of the terms in (C.34) dominates.

If one of the space coordinates is used as independent parameter the case is slightly different. If x is independent for example, and y is the dependent variable with slowness component $u(\mathbf{x}) \sin \theta$, the metric turns out to be

$$\left(\frac{dl}{d\gamma}\right)^2 = \left(\frac{dy}{d\gamma}\right)^2 + \alpha \left(\frac{du(\mathbf{x})}{dy} \frac{dy}{d\gamma} \sin \theta + u(\mathbf{x}) \cos \theta \frac{d\theta}{d\gamma}\right)^2. \quad (\text{C.36})$$

This yields for α :

$$\alpha \approx \left(\frac{\delta U}{L_U} \sin \theta + \frac{U}{R} \cos \theta\right)^{-2}, \quad (\text{C.37})$$

where it should be noted that θ is often small, to allow x to be used as an independent parameter. The most difficult part of these estimations of α is the a priori determination of a characteristic radius of curvature.

C.A.3 Derivation of Equation (C.17)

Equation (C.17) is derived in a more formal way by considering the change of parameterisation from (σ, ξ) to (λ, γ) . The total differential for the new parameterisation $\tilde{\mathbf{y}}$ in terms of the old \mathbf{y} gives:

$$\begin{aligned} d\tilde{\mathbf{y}} = & \left[\left(\frac{\partial \mathbf{y}}{\partial \sigma}\right) \left(\frac{\partial \sigma}{\partial \lambda}\right) + \left(\frac{\partial \mathbf{y}}{\partial \xi}\right) \left(\frac{\partial \xi}{\partial \lambda}\right) \right] d\lambda \\ & + \left[\left(\frac{\partial \mathbf{y}}{\partial \sigma}\right) \left(\frac{\partial \sigma}{\partial \gamma}\right) + \left(\frac{\partial \mathbf{y}}{\partial \xi}\right) \left(\frac{\partial \xi}{\partial \gamma}\right) \right] d\gamma. \end{aligned} \quad (\text{C.38})$$

Since there is no reason to adjust the parameterisation in the direction along the rays, choose

$$\frac{\partial \sigma}{\partial \lambda} = 1, \quad \text{and} \quad \frac{\partial \sigma}{\partial \gamma} = 0, \quad (\text{C.39})$$

such that λ is equivalent to σ , and lines of constant λ will coincide with lines of constant σ . Combining equations (C.38) and (C.2) gives

$$\frac{\partial \tilde{\mathbf{y}}}{\partial \lambda} = \mathbf{F}(\mathbf{y}) + \left(\frac{\partial \mathbf{y}}{\partial \xi}\right) \left(\frac{\partial \xi}{\partial \lambda}\right), \quad \text{and} \quad (\text{C.40})$$

$$\frac{\partial \tilde{\mathbf{y}}}{\partial \gamma} = \left(\frac{\partial \mathbf{y}}{\partial \xi}\right) \left(\frac{\partial \xi}{\partial \gamma}\right). \quad (\text{C.41})$$

To be able to compute the ray field in its new parameterisation the partial derivative to ξ in equation (C.40) is eliminated using (C.41):

$$\frac{\partial \tilde{\mathbf{y}}}{\partial \lambda} = \mathbf{F}(\mathbf{y}) + \left(\frac{\partial \xi}{\partial \gamma} \right)^{-1} \left(\frac{\partial \xi}{\partial \lambda} \right) \frac{\partial \tilde{\mathbf{y}}}{\partial \gamma}, \quad (\text{C.42})$$

and since the dependence of ξ on λ and γ is still free, the choice

$$f(\lambda, \gamma) = \left(\frac{\partial \xi}{\partial \gamma} \right)^{-1} \left(\frac{\partial \xi}{\partial \lambda} \right), \quad (\text{C.43})$$

leads to equation (C.17).

