The Smirnov's lemma applied to ray theory

R. Portugal

email: *portugal@ime.unicamp.br*

keywords: *Ray theory, ray Jacobian, differential equations, linear algebra*

ABSTRACT

In this paper I present a generalized Smirnov's lemma and its application to ray theory. This lemma can be used to establish the link between the kinematic and dynamic aspects of wave propagation in the high-frequency hypothesis. This is accomplished by a formula which connects traveltime and amplitude along the ray, converting the transport equation into a ordinary differential equation. This formula is obtained by applying the Smirnov lemma to the kinematic ray equations. Moreover, I provide an alternative proof of the lemma.

INTRODUCTION

When wave phenomena is investigated, we can observe that events or disturbances in the medium are independently propagated. Beyond that, if the event has a short period, almost like a delta pulse, this behavior remains throughout the propagation. These observations give a good clue on how the solution should look like, at least under the high-frequency hypothesis. If we look closely to only one event, the solution must take into account two basic aspects: the kinematic and the dynamic. In rough words, for a fixed point inside the medium, the kinematic aspect shows when the pulse reaches the point and the dynamic one shows how the pulse goes through the point.

These two aspects can be assembled into a function representing an approximated solution. These procedure works like the separation of variables method, where the guessed solution is plugged into the equation, generating new equations, hopefully simpler ones, to be solved. In the case of ray theory, the function responsible for the kinematic aspect is called traveltime function and the function which rules the dynamics is called amplitude function. Once the assembled function is plugged into the Helmholtz equation, after some manipulation and additional hypothesis, two new partial differential equations (PDE) are generate. The first one is termed eikonal equation and have only the traveltime function as unknown. The second one is called transport equation and has both traveltime and amplitude functions as the unknowns.

Using the eikonal equation as a starting point, we can derive a system of ordinary differential equations (ODE), called kinematic ray equations. As the name says, this system provides the traveltime solution. However this is only the first half of the game, where the amplitude solution is still missing. Using the traveltime solution already available, we can try to convert the transport equation, which is a PDE, into an ODE which is a version for transport equation valid for the ray.

At some point, this conversion procedure needs to transform a spatial operator acting on the traveltime function into some derivative with respect to the independent variable of the ordinary differential equation (ODE). The well-known method for doing this is to consider a very small ray tube, perform the integration on it, and using of divergence theorem and some geometric arguments to reach the desired result. Another way is to apply the Smirnov's lemma to the equation, reaching the very same result.

REVIEW OF RAY THEORY

When wave phenomena is studied under the high-frequency hypothesis, it is quite natural to transform the acoustic wave equation into the frequency domain by the use of the Fourier transform. This equation is called Helmholtz equation and it is given by

$$\Delta\hat{u}(\mathbf{x},\omega) + \frac{\omega^2}{c(\mathbf{x})^2}\hat{u}(\mathbf{x},\omega) \;=\; -\hat{f}(\omega)\delta(\mathbf{x}-\mathbf{x}_s), \tag{1}$$

where $\mathbf{x} = (x_1, x_2, x_3)$, $c(\mathbf{x})$ is the velocity field, ω is the frequency and $\Delta = \frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2} + \frac{\partial^2}{\partial x_3^2}$ is the laplacian operator. The functions $\hat{f}(\omega)$ and $\hat{u}(\mathbf{x},\omega)$ are the Fourier transforms of the source wavelet and the wave equation solution, respectively.

An approximation for the solution of Helmholtz equation can be assumed to be the ansatz

$$\hat{u}(\mathbf{x},\omega) \;\approx\; A(\mathbf{x})\mathrm{e}^{i\omega\tau(\mathbf{x})}, \tag{2}$$

where $A(\mathbf{x})$ is the amplitude and $\tau(\mathbf{x})$ is the traveltime. This solution is also known as the zero-order high-frequency asymptotic solution and we can insert it in the Helmholtz equation to obtain two partial differential equations: the *eikonal equation*

$$\|\nabla\tau\|^2 \;=\; \left(\frac{\partial\tau}{\partial x_1}\right)^2 + \left(\frac{\partial\tau}{\partial x_2}\right)^2 + \left(\frac{\partial\tau}{\partial x_3}\right)^2 \;=\; \frac{1}{c(\mathbf{x})^2} \tag{3}$$

and the *transport equation*

$$2\,\nabla A(\mathbf{x})\cdot\nabla\tau(\mathbf{x}) + A(\mathbf{x})\Delta\tau(\mathbf{x}) \;=\; 0, \tag{4}$$

where ∇ is the gradient operator and $\cdot$ denotes the inner product.

The standard procedure is to solve the eikonal equation and then plug the solution $\tau(\mathbf{x})$ inside (4) to seek the solution $A(\mathbf{x})$. Considering equation (3) as a Hamiltonian to be solved, we can use the method of characteristics, which transforms the first order non-linear PDE into a system of ordinary differential equations:

$$\frac{d\mathbf{x}}{d\sigma} = \lambda\,\mathbf{p}, \tag{5}$$

$$\frac{d\mathbf{p}}{d\sigma} = \frac{-\lambda}{c^3}\,\nabla c, \tag{6}$$

$$\frac{d\tau}{d\sigma} = \frac{\lambda}{c^2}, \tag{7}$$

where $\mathbf{p} = \nabla\tau$ is the slowness vector, σ is a generic increasing parameter and λ is a function which has the role to define explicitly any parameterization. In general, λ can achieve three values, each one giving a special meaning for σ. For $\lambda = 1$, σ is the out-of-plane geometrical spreading along the ray; for $\lambda = c$, σ is the ray arclength and for $\lambda = c^2$, σ is the traveltime along the ray.

If the transport equation (4) is expected to be valid for some domain in space, then it is also valid for any chosen path inside it. In particular, choosing the path as a ray trajectory governed by equations (5)–(7), we can use them to simplify (4), transforming it in another ODE which represents the rule of propagation for amplitudes along a chosen ray.

Multiplying equation (4) by λA, using equation (5) and the definition of the slowness vector we obtain

$$2\,A\,\nabla A \cdot \frac{d\mathbf{x}}{d\sigma} + \lambda A^2 \Delta\tau \;=\; 0. \tag{8}$$

Observing that the first term is a chain rule, after some manipulation we get

$$\frac{d}{d\sigma}\Big[\log(A^2)\Big] \;=\; -\lambda\Delta\tau. \tag{9}$$

At this point, we want to get rid of the Laplacian Δ, because it is a spatial operator and we are interested only in quantities and operators depending on the parameter σ. A traditional way of doing this is to consider a tubular neighborhood of a ray and integrate $\Delta\tau$ on it. Using divergence theorem and geometric arguments, it is possible to show the relation between $\Delta\tau$ and the Jacobian $\mathcal{J}$ along the ray. Bleistein *et al.* (2001) demonstrated this relation for the three dimensional case using determinants properties and algebraic arguments

In this paper I show another algebraic proof of the relation between $\Delta\tau$ and $\mathcal{J}$. I present a general lemma, named Smirnov's lemma after Thomson *et al.*(1985), which can be applied to any n-dimensional system of ODEs. Its demonstration is provided, using elementary matricial notation and an auxiliary linear algebra result. To obtain the desired relation, I simply apply the Smirnov's lemma to the ray equation (5).

SMIRNOV'S LEMMA

Given a solution of a n-dimensional system of ODEs, the Smirnov's lemma (Thomson & Chapman (1985); Smirnov (1964), p.442) provides a relationship between the jacobian of a transformation of variables and the right-hand side of the system. Applying this lemma to the ray equation (5), we can solve the problem of getting rid of the Laplacian Δ in the process to convert the transport equation in some ODE depending on σ.

Smirnov's Lemma: *Let* $\mathbf{x} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$ *be a solution of the system of ordinary differential equations:*

$$\frac{d\mathbf{x}}{d\sigma} \;=\; F(\mathbf{x}) \tag{10}$$

where $F : \mathbb{R}^n \to \mathbb{R}^n$ *is a smooth function. If the transformation of variables*

$$\mathbf{x} = \mathbf{x}(\gamma_1, \gamma_2, \ldots, \gamma_n), \quad \textit{with } \gamma_1 = \sigma, \tag{11}$$

is carried out then

$$\frac{d}{d\sigma}\Big[\log \mathcal{J}\Big] \;=\; \nabla \cdot F, \tag{12}$$

where $\mathcal{J}$ *is the Jacobian of transformation of variables, defined by*

$$\mathcal{J} \;=\; \frac{\partial(x_1, x_2, \ldots, x_n)}{\partial(\gamma_1, \gamma_2, \ldots, \gamma_n)} \;=\; \det\left[\frac{\partial\mathbf{x}}{\partial\gamma_1}, \frac{\partial\mathbf{x}}{\partial\gamma_2}, \ldots, \frac{\partial\mathbf{x}}{\partial\gamma_n}\right]. \tag{13}$$

Proof: We simply differentiate the Jacobian in equation (13) with respect to σ and use the rule for differentiation of determinants (Golub & Van Loan (1996); Bleistein *et al.* (2001)), obtaining

$$\begin{aligned} \frac{d\mathcal{J}}{d\sigma} &= \frac{d}{d\sigma}\det\left[\frac{\partial\mathbf{x}}{\partial\gamma_1}, \frac{\partial\mathbf{x}}{\partial\gamma_2}, \ldots, \frac{\partial\mathbf{x}}{\partial\gamma_n}\right] \\ &= \sum_{k=1}^{n}\det\left[\frac{\partial\mathbf{x}}{\partial\gamma_1}, \ldots, \frac{\partial\mathbf{x}}{\partial\gamma_{k-1}}, \frac{d}{d\sigma}\left(\frac{\partial\mathbf{x}}{\partial\gamma_k}\right), \frac{\partial\mathbf{x}}{\partial\gamma_{k+1}}, \ldots, \frac{\partial\mathbf{x}}{\partial\gamma_n}\right]. \end{aligned} \tag{14}$$

By the use of equation (10), we have

$$\frac{d}{d\sigma}\left(\frac{\partial\mathbf{x}}{\partial\gamma_k}\right) \;=\; \frac{\partial}{\partial\gamma_k}\left(\frac{d\mathbf{x}}{d\sigma}\right) \;=\; \frac{\partial}{\partial\gamma_k}F(\mathbf{x}) \;=\; F'(\mathbf{x})\frac{\partial\mathbf{x}}{\partial\gamma_k}, \tag{15}$$

where $F'(\mathbf{x})$ is the Jacobian matrix defined by

$$\Big[F'(\mathbf{x})\Big]_{ij} \;=\; \frac{\partial F_i(\mathbf{x})}{\partial x_j}. \tag{16}$$

Substituting equation (15) in (14), we obtain

$$\frac{d\mathcal{J}}{d\sigma} \;=\; \sum_{k=1}^{n}\det\left[\frac{\partial\mathbf{x}}{\partial\gamma_1}, \ldots, \frac{\partial\mathbf{x}}{\partial\gamma_{k-1}}, F'(\mathbf{x})\frac{\partial\mathbf{x}}{\partial\gamma_k}, \frac{\partial\mathbf{x}}{\partial\gamma_{k+1}}, \ldots, \frac{\partial\mathbf{x}}{\partial\gamma_n}\right] \tag{17}$$

$$= \left(\sum_{k=1}^{n}\frac{\partial F_k}{\partial x_k}\right)\det\left[\frac{\partial\mathbf{x}}{\partial\gamma_1}, \frac{\partial\mathbf{x}}{\partial\gamma_2}, \ldots, \frac{\partial\mathbf{x}}{\partial\gamma_n}\right]. \tag{18}$$

where I have used the linear algebra result (25) given in Appendix A

Therefore,

$$\frac{d\mathcal{J}}{d\sigma} \;=\; (\nabla \cdot F)\,\mathcal{J}, \tag{19}$$

which leads to equation (12).

APPLICATION

Using the Smirnov's lemma, for $F(\mathbf{x}) = \lambda\mathbf{p} = \lambda\nabla\tau$,

$$\frac{d}{d\sigma}\Big[\log\mathcal{J}\Big] \;=\; \nabla\cdot(\lambda\nabla\tau), \tag{20}$$

or, equivalently,

$$\frac{d}{d\sigma}\Big[\log\mathcal{J}/\lambda\Big] \;=\; \lambda\Delta\tau, \tag{21}$$

where τ is the traveltime solution from the ray equations (5)–(7) and the Jacobian $\mathcal{J}$ is given by

$$\mathcal{J} \;=\; \frac{\partial(x_1,x_2,x_3)}{\partial(\gamma_1,\gamma_2,\gamma_3)} \;=\; \det\left[\frac{\partial\mathbf{x}}{\partial\gamma_1},\,\frac{\partial\mathbf{x}}{\partial\gamma_2},\,\frac{\partial\mathbf{x}}{\partial\gamma_3}\right], \tag{22}$$

with $\gamma_1 = \sigma$.

Using equation (21) in equation (9), we are able to complete the conversion of the transport equation (9) into an ordinary differential equation,

$$\frac{d}{d\sigma}\Big[\log(A^2)\Big] \;=\; -\frac{d}{d\sigma}\Big[\log(\mathcal{J}/\lambda)\Big]. \tag{23}$$

The above equation can be directly integrated from σ_0 to σ, resulting in

$$A(\sigma) \;=\; A(\sigma_0)\sqrt{\frac{\lambda(\sigma)\mathcal{J}(\sigma_0)}{\lambda(\sigma_0)\mathcal{J}(\sigma)}}\,, \tag{24}$$

which is the explicit formula for the amplitude along the ray.

ACKNOWLEDGMENTS

This work was partially supported by FAPESP (grant 97/12318-0) and the sponsors of the *Wave Inversion Technology (WIT) Consortium*. I am especially grateful to Lúcio Santos for the fruitful discussions.

REFERENCES

Bleistein, N., Cohen, J. K., and Stockwell Jr., J. W. (2001). *Mathematics of multidimensional seismic imaging, migration and inversion*, volume 13 of *IAM*. Springer-Verlag, New York.

Golub, G. H. and Van Loan, C. F. (1996). *Matrix computations*. The Johns Hopkins University Press, Baltimore, 3rd edition.

Smirnov, V. I. (1964). *A course in higher mathematics*, volume 4. Pergamon Press, Oxford. Translated by Brown, D. E.

Thomson, C. J. and Chapman, C. H. (1985). An introduction to Maslov's asymptotic method. *Geophys. J. R. astr. Soc.*, 85:143–168.

APPENDIX A

Linear algebra identity: *Given two square matrices* $\mathbf{M}$ *and* $\mathbf{X} = \left[\mathbf{x}^1, \mathbf{x}^2, \ldots, \mathbf{x}^n\right]$*, then*

$$\begin{aligned} tr(\mathbf{M})\det(\mathbf{X}) &= \det\left[\mathbf{M}\mathbf{x}^1,\, \mathbf{x}^2,\, \ldots,\, \mathbf{x}^n\right] + \det\left[\mathbf{x}^1,\, \mathbf{M}\mathbf{x}^2,\, \ldots,\, \mathbf{x}^n\right] + \cdots \\ &\quad \cdots + \det\left[\mathbf{x}^1,\, \mathbf{x}^2,\, \ldots,\, \mathbf{M}\mathbf{x}^n\right], \end{aligned} \tag{25}$$

where $\mathrm{tr}(\mathbf{M})$ *is the trace of* $\mathbf{M}$*, defined by*

$$\mathrm{tr}(\mathbf{M}) = \sum_{k=1}^{n} \mathbf{M}_{kk}. \tag{26}$$

Proof: First, we prove the result for a non-singular matrix $\mathbf{X}$. Let $\mathbf{Y}$ be a similar matrix to $\mathbf{M}$, defined by

$$\mathbf{Y} = \mathbf{X}^{-1}\mathbf{M}\mathbf{X}. \tag{27}$$

Therefore $\mathbf{M}$ and $\mathbf{Y}$ have the same eigenvalues, which implies

$$\mathrm{tr}(\mathbf{M}) = \sum_{k=1}^{n} \mathbf{M}_{kk} = \sum_{k=1}^{n} \mathbf{Y}_{kk} = \mathrm{tr}(\mathbf{Y}). \tag{28}$$

We can rewrite equation (27) as the solution of linear systems

$$\mathbf{X}\mathbf{y}^k = \mathbf{M}\mathbf{x}^k, \qquad k = 1, 2, \ldots, n, \tag{29}$$

where $\mathbf{x}^k$ and $\mathbf{y}^k$ are the columns of $\mathbf{X}$ and $\mathbf{Y}$, respectively. By the use of Cramer's rule, we have

$$y_j^k = \frac{\det\left[\mathbf{x}^1, \ldots, \mathbf{x}^{j-1}, \mathbf{M}\mathbf{x}^k, \mathbf{x}^{j+1}, \ldots, \mathbf{x}^n\right]}{\det(\mathbf{X})}. \tag{30}$$

Substituting equation (30) back into (28), we finally have

$$\mathrm{tr}(\mathbf{M}) = \mathrm{tr}(\mathbf{Y}) = \sum_{k=1}^{n} y_k^k = \frac{1}{\det(\mathbf{X})} \sum_{k=1}^{n} \det\left[\mathbf{x}^1, \ldots, \mathbf{x}^{k-1}, \mathbf{M}\mathbf{x}^k, \mathbf{x}^{k+1}, \ldots, \mathbf{x}^n\right], \tag{31}$$

proving (25). For a singular matrix $\mathbf{X}$, it is possible to construct a sequence of non-singular matrices $\{\mathbf{X}^{(k)}\}_{k=1}^{\infty}$ which converges to $\mathbf{X}$. After applying all the above results to each member of the sequence, we evaluate the limit, since all operations are continuous. Therefore, in the limit, the result (25) is also valid for $\mathbf{X}$ singular.