

DTIC FILE COPY

AD-A182 752

REPORT DOCUMENTATION PAGE

2a. SECURITY CLASSIFICATION AUTHORITY		1b. RESTRICTIVE MARKINGS	
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE JUL 2 1 1987		3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release distribution unlimited	
4. PERFORMING ORGANIZATION REPORT NUMBER(S) D5		5. MONITORING ORGANIZATION REPORT NUMBER(S) AFOSR-TR- 87-0863	
6a. NAME OF PERFORMING ORGANIZATION Mass. Institute of Technology	6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION AFOSR	
6c. ADDRESS (City, State and ZIP Code) Earth Resources Laboratory Dept of Earth, Atmospheric & Planetary Sciences Cambridge, Massachusetts 02139		7b. ADDRESS (City, State and ZIP Code) Building 410 Bolling AFB, Washington D.C. 20332-6448	
8a. NAME OF FUNDING/SPONSORING ORGANIZATION AFOSR	8b. OFFICE SYMBOL (If applicable) NP	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER F49620-83-C-0038	
8c. ADDRESS (City, State and ZIP Code) Building 410 Bolling AFB, Wash, D.C. 20332-6448		10. SOURCE OF FUNDING NOS.	
		PROGRAM ELEMENT NO. 61102F	TASK NO. A2
		PROJECT NO. 2301	WORK UNIT NO.
11. TITLE (Include Security Classification) EFFECTS OF HETEROGENEITIES ON THE PROPAGATION, SCATTERING AND ATTENUATION OF SEISMIC WAVES AND THE CHARACTERIZATION OF SEISMIC SOURCE (U)			
12. PERSONAL AUTHOR(S) Drs. Keiiti Aki, Vernon F. Cormier, M. Nafi Toksoz			
13a. TYPE OF REPORT FINAL	13b. TIME COVERED FROM 82/12/01 TO 85/11/30	14. DATE OF REPORT (Yr., Mo., Day)	15. PAGE COUNT 266
16. SUPPLEMENTARY NOTATION			
17. COSATI CODES		18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)	
FIELD	GROUP	SUB. GR.	
		Body Wave, Seismology, Seismograms	
		During this reporting period	
19. ABSTRACT (Continue on reverse if necessary and identify by block number) During the first year of funding this title, work was completed on the testing of alternative measures of body wave magnitude. In summary, it was found that alternative measures of body waves magnitude generally often exhibit as much scatter as classical measures of magnitude, although coda magnitudes usually have slightly less scatter than spectral magnitudes and classical magnitudes. In the cases investigated, these differences were usually not statistically significant. Another completed task was an investigation of the intrinsic attenuation of the Earth's mantle selected paths from the Sea of Okhotsk to Regional Seismic Test Network (RSTN) and Global Digital Seismic Network (GDSN) Stations in North America. The conclusion of this study was that the intrinsic attenuation in the mantle beneath eastern north America is both depth and frequency dependent and that spectral and time domain studies of attenuation can be reconciled in the frequency band up to 2Hz. The focus of the project was then divided between source problems related to scattering and wave propagation in three-dimensional, seismic.			
20. DISTRIBUTION/AVAILABILITY OF ABSTRACT UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT. ☒ DTIC USERS ☐		21. ABSTRACT SECURITY CLASSIFICATION Unclassified	
22a. NAME OF RESPONSIBLE INDIVIDUAL JOHN F. PRINCE		22b. TELEPHONE NUMBER (Include Area Code) 202/767-4908	22c. OFFICE SYMBOL NP

(abstract continued)

heterogeneous media. A significant result ~~of this work~~ was that short period and broad-band waveforms can improve the depth resolution determined earthquakes and underground nuclear explosions, forming a powerful discriminant. Scattering was studied theoretically and observationally. The significant result of that work is that the Earth's lithosphere must possess multiple scales of heterogeneity in order to explain both the amplitude and phase fluctuations at large arrays as well as the shapes of local S coda. Scale lengths on the order of 10 km and less than 1 km are needed to explain both data sets. In the final area of research, an asymptotically approximate method of synthesizing seismograms in three-dimensionally varying media was tested and implemented. The results of this study will be important for predicting the amplitude fluctuations induced by larger scale structures (10 km and greater scale lengths) in both source and receiver regions.

AFOSR-TR. 87-0863

FINAL REPORT

ARPA Order - 4397

Program Code - 3D60

Name of Contractor - Earth Resources Laboratory
Department of Earth, Atmospheric, and Planetary Sciences
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139

Start Date of Contract - 1 December 1982

Contract Expiration Date - 30 November 1985

Amount of Contract Dollars - \$455,011

Contract Number - F49620-83-C-0038

Principal Investigators - Keiiti Aki, 617/253-6397
Vernon F. Cormier, 617/253-7862
M. Nafi Toksöz, 617/253-7852

Program Manager - William J. Best

Short Title of Work - Effects of Heterogeneities on the Propagation, Scattering
and Attenuation of Seismic Waves and the Characterization
of Seismic Source

Sponsored by
Advanced Research Projects Agency (DOD)
ARPA Order No. 4397
Monitored by the Air Force Office of Scientific Research
Under Contract No. F49620-83-C-0038

The views and conclusions contained in this document are those of the authors
and should not be interpreted as necessarily representing the official policies,
either expressed or implied, of the Defense Advanced Research Projects Agency
or the U.S. Government.

Approved for public release;
distribution unlimited.

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH (AFSC)
NOTICE OF TRANSMITTAL TO DTIC
This technical report has been reviewed and is
approved for public release IAW AFR 190-12.
Distribution is unlimited.
MATTHEW J. KERPER
Chief, Technical Information Division

87 7 17 045

1. SUMMARY

This three-year project was first initiated under the title "Alternative Measures of Yield, Testing and Implementation of Yield Measures, and the Mantle Attenuation Operator." During the first year of funding under this title, work was completed on the testing of alternative measures of body wave magnitude. The purpose of this study was to determine whether magnitude measures based on integrated coda energy exhibited any less scatter over local and teleseismic networks than classical magnitude measures based on the peak amplitude within several seconds of P onset time. The results of that study were published (Bullitt and Cormier, 1984) and reported in the first year of the project. In summary, it was found that alternative measures of body wave magnitude generally often exhibit as much scatter as classical measures of magnitude, although coda magnitudes usually have slightly less scatter than spectral magnitudes and classical magnitudes. In the cases investigated, these differences were usually not statistically significant. It was concluded that the reason for the marginal improvement of coda magnitudes was due to the fact that a portion of the coda must be focused and defocused by three-dimensional structure in step with the first several cycles of the P wave.

Another completed task related to the first year of the project was an investigation of the intrinsic attenuation of the Earth's mantle along selected paths from the Sea of Okhotsk to Regional Seismic Test Network (RSTN) and Global Digital Seismic Network (GDSN) stations in North America. The conclusion of this study was that the intrinsic attenuation in the mantle beneath eastern North America is both depth and frequency dependent and that spectral and time domain studies of attenuation can be reconciled in the frequency band up to 2 Hz. The effects of scattering in the Earth's upper mantle and crust do not seem to dominate those of intrinsic attenuation in this frequency band. The results of that study were published as a refereed paper (Choy and Cormier, 1986), and are included as part 2 of the current report.

In the second and third year's funding, the scope of the project was expanded to investigate a broader range of scattering, attenuation, and wave propagation problems. The title of the project changed to "Effects of Lateral Heterogeneities on the Propagation, Scattering, and Attenuation



Availability Codes	
Dist	Avail and/or Special
A-1	

of Seismic Waves, and the Characterization of the Seismic Source." Prof. Aki was added as co-investigator along with Prof. Toksöz and Dr. Cormier. The focus of the project was then divided between source problems in source inversion from recorded waveforms and problems related to scattering and wave propagation in three-dimensional, heterogeneous media. The results of source inversion of body waves were reported in previous semi-annual reports and formed the basis of the work that formed the thesis by John Nabelek. A significant result of this work was that short period and broadband waveforms can improve the depth resolution determined earthquakes and underground nuclear explosions, forming a powerful discriminant.

Scattering was studied theoretically and observationally using the results of phase and amplitude studies at seismic arrays, as well as the shapes of coda waves following the S waves of local and regional earthquakes. These studies are included in the thesis by Ru-shan Wu. Chapters of that thesis have been excerpted in earlier semi-annual reports. The significant result of that work is that the Earth's lithosphere must possess multiple scales of heterogeneity in order to explain both the amplitude and phase fluctuations at large arrays as well as the shapes of local S coda. Scale lengths on the order of 10 km and less than 1 km are needed to explain both data sets. Another important result of this work was that a method was proposed for separating intrinsic attenuation and scattering attenuation based on the shape of the coda following the S waves of local and regional events. A data set of deep and shallow focus events in the Hindu Kush region of Asia was studied using this method. Scattering attenuation was found to dominate intrinsic attenuation only in the highest frequency band above 10 Hz.

In the final area of research, an asymptotically approximate method of synthesizing seismograms in three-dimensionally varying media was tested and implemented. This formed the basis of a thesis by Robert Nowack. This work is the only remaining task that has not been included in previous reports, and as such, is included as part 1 of this report. The results of this study will be important for predicting the amplitude fluctuations induced by larger scale structures (10 km and greater scale lengths) in both source and receiver regions. Such structures can be responsible for amplitude fluctuations larger than that induced by regional variations in intrinsic attenuation and are likely to be responsible for much of fluctuations observed in measures of magnitude at teleseismic range.



PUBLICATIONS

Bullitt, J.T., and Cormier, V.F., The relative performance of *m_{subb}* and alternative measures of elastic energy in estimating source size and explosion yield. *Bull. Seism. Soc. Am.*, 74, 1863-1882, 1984.

Choy, G.L., and Cormier, V.F., Direct measurement of the mantle attenuation operator from broadband P and S waveforms, *J. Geophys. Res.*, 91, 7326-7342, 1986.

Nowack, R., and Aki, K., The 2-D Gaussian beam synthetic method: testing and application, *J. Geophys. Res.*, 89, 7797-7819, 1984.

Nowack, R., and Aki, K., Iterative inversion for velocity using waveform data, *Geophys. J. R. Astr. Soc.*, 87, 701-730, 1986.

Wu, R-S., Multiple scattering and energy transfer of seismic waves—separation of scattering effect from intrinsic attenuation—I. Theoretical modeling, *Geophys. J. R. Astr. Soc.*, 82, 57-80, 1985.

Wu, R-S., Gaussian beams, complex rays, and the analytic extension of the Green's function in smoothly inhomogeneous media, *Geophys. J. R. Astr. Soc.*, 83, 93-110, 1985.

Wu, R-S., and Ben-Menahem, A., The elastodynamic near field, *Geophys. J. R. Astr. Soc.*, 81, 609-621, 1985.

Wu, R-S., Fractal dimensions of fault surfaces and the inhomogeneity spectrum of the lithosphere revealed from seismic wave scattering, Proceedings of the conference on Multiple Scattering of Waves in Random Media and Random Rough Surfaces, The Pennsylvania State University, 1985.

PhD. Theses

Nabelek, J.L., Determination of Earthquake Source Parameters from Inversion of Body Waves, M.I.T., 1984.

Nowack, R.L., Wave propagation in laterally varying media and iterative inversion for velocity, M.I.T., 1985.

Wu, R-S., Seismic Wave Scattering and the Small Scale Inhomogeneities in the Lithosphere, M.I.T., 1984.

2. WAVE PROPAGATION IN LATERALLY VARYING MEDIA AND ITERATIVE INVERSION FOR VELOCITY

2.1 INTRODUCTION

In recent years, a variety of techniques have been developed to compute seismograms in laterally varying media. These include the finite difference and finite element methods which are general techniques but expensive. Thus they are limited to small grids and lower frequencies (see for example Boore [1972], Smith [1975], Kelly et.al. [1976]). Perturbation techniques are also useful primarily at lower frequencies (see Kennett [1972a, 1972b], Larner [1970], Aki and Richards [1980]). The boundary integral method is another technique recently investigated in the seismological literature (see Sanchez-Sesma and Esquivel [1979], Sanchez-Sesma et.al. [1982]). The remaining methods to be described are primarily high frequency asymptotic methods.

The ray method is one of the most commonly used techniques for propagation in laterally varying media. It is based on a series solution to the wave equation (elastodynamic equation) in inverse powers of frequency (Kline and Kay [1965], Cerveny et.al. [1977]). In practice, usually only the first term is retained. Examples of seismograms computed using the ray method are found in May and Hron [1978], Cerveny [1979], McMechan and Mooney [1980].

Simple estimates of the validity of the ray method are based on several criterion (Kravtsov and Orlov [1980]). First, the parameters of the medium as well as the amplitude and phase gradient of the wavefront should be slowly varying over the cross section of the Fresnel volume. The boundary

of the Fresnel volume is defined by $\omega\tau_g - \omega\tau_v = \pi$, where τ_g is the phase time for the geometrical ray and τ_v is the phase time for a secondary virtual ray. The Fresnel volume is the physical region around a geometric ray for finite frequencies. Next, the phase difference corresponding to two rays arriving at the same point should be greater than π or a caustic results. Finally, the ray method has the range limitation

$$L \ll \frac{a^2}{\lambda}$$

where L is the range, λ is the wavelength, and a is the inhomogeneity scale. This limitation is a result of the increasing size of the Fresnel radius with distance.

A disadvantage of the ray method even at high frequencies is that the method is invalid at caustics where it incorrectly predicts infinite amplitudes. These singularities can be smoothed over by properly incorporating the effects of adjacent rays into the solution. One way to correct the ray method is to construct a local asymptotic solution in the vicinity of the caustic (see Ludwig [1966]). This method may be used to construct a uniform solution, but may be inconvenient for complicated structures.

Certain diffraction effects can be included in the ray method using an extension known as the geometrical theory of diffraction (Keller [1962]). Keller defined new diffracted ray types which include edge rays and vertex

rays. A simple review of this theory is given by Keller [1978]. Extensions to seismology are given for example by Klem-Musatov and Aizenberg [1984]. In this approach, diffraction coefficients must be determined for different cases.

A simplification of the ray method applied to vector wave equations, such as the elastodynamic equation, was the introduction of the ray centered coordinates by Popov and Psencik [1976,1978]. This is a special orthogonal curvilinear coordinate system which moves along a particular ray. In this coordinate system, the first order transport equations for the P and S displacement components decouple. This results in three scalar transport equations rather than vector equations. Writing the wave equation in this local coordinate system along a given ray is also a fundamental step in the Gaussian beam method.

The Maslov method is a semi-classical method which can be applied to propagation problems to extend the ray method (see Kravtsov [1968], Chapman and Drummond [1983], and Ziolkowski and Deschamps [1984]). The method is based on the fact that a ray is a construct dependent not just on location, $\vec{x}$, but also on the direction represented by the slowness vector, $\vec{p}$. In the space $(\vec{x}, \vec{p})$ there are no caustics. This can be seen by noting that $(\vec{x}(t), \vec{p}(t))$ is a state vector for the ray system. Thus if two rays have the same state vector for a given time, then they will be the same for all time since this vector contains all the information needed to

construct the ray. Caustics occur upon projection from this extended space to various three dimensional configuration spaces. The caustics will have different locations depending on the configuration space. In the (x,y,z) space, the method reduces to the standard ray method. In the (p_x,y,z) space in a vertically varying medium, the method would be similar to the WKBJ method discussed by Chapman [1978] where the field is the result of an integral over horizontal slowness, p_x , which is then evaluated by an equal phase method. The idea of the Maslov method is to choose the domain of computation which avoids caustics near different receivers. A problem with this method is that the appropriate domain of computation is not known prior to starting the computations. The final solution is the result of piecing together the asymptotic solutions which are found to be valid in different ranges.

Another technique for the propagation in laterally varying media is based on the parabolic approximation to the wave equation. It was first derived by Leontovich and Fock in 1946 for radio wave propagation (see Fock [1965], ch 11). The parabolic approximation assumes small angle forward scattering about some preferred direction of propagation. It results in a one way operator, since back travelling energy is neglected. An advantage of the derived parabolic equation is that a boundary value problem is replaced by a much simpler initial value problem. The method is valid assuming that the fluctuations of the material parameters are small, and slowly varying with respect to the wavelength.

In relation to the ray method, the parabolic method does not require energy to remain bounded within rigid ray tubes. It accounts for the slow lateral spread of ray energy along the wavefront as the wave propagates. The parabolic method thus includes the effects of diffraction due to lateral diffusion along the wavefront (see review by Malyuzhinets [1959]). This explains the similarity in form between the parabolic equation and the diffusion equation. The parabolic equation is also closely related to the Fresnel approximation in optics (see Goodman [1968], Tappert [1976]).

Geophysical applications of the parabolic method include long range underwater acoustics where the global direction of propagation is assumed to be horizontal (Tappert [1976], McDaniel [1976]), and reflection seismology where the global direction is assumed to be vertical for upward travelling reflected waves (Claerbout [1976]). Applications to elastic waves are discussed by Landers and Claerbout [1972], McCoy [1977], Hudson [1980a, 1980b], and Wales and McCoy [1983].

Derivations of one way parabolic operators by the use of splitting matrices can be found in Corones [1975] and McDaniel [1976]. The importance in this approach is the potential estimation of coupling between forward and back travelling energy. The splitting method has been extended to the elastic case by Corones et.al. [1982].

Another method for including non-Fermat information is by using the Kirchhoff integral in which the field is gotten by an integration along some

intermediate wave surface (see Baker and Copson [1950]). For elastic formulations see Pao and Varatharajulu [1976] and Aki and Richards [1980]. For geophysical applications see Trorey [1970,1977], Hilterman [1970,1975], Scott and Helmberger [1983], Hadden and Buchen [1981], and Frazer and Sinton [1984]. Evaluation along many intermediate surfaces leads in the limit to a path integral formulation (Dashen [1979], Flatte et.al. [1979]).

In this thesis, the Gaussian beam method is investigated. It is an extension of both the ray method and the parabolic method in which a global direction must be assumed. The Gaussian beam method has been described in the literature by Popov [1981,1982], and Cerveny et.al. [1982]. These studies were based on the scalar wave equation. The elastic case was first derived by Kirpicnnichova [1972] and more recently by Cerveny [1983] for the 2-D case and Cerveny and Psencik [1983] for the 3-D case.

The Gaussian beam method involves three basic steps. First, a system of rays must be traced from the source to a vicinity of the receiver. Since high frequency energy flows along rays, the system of rays can be thought as the framework on which the wavefield is to be built. This step requires the numerical solution of the kinematic ray tracing system. Next, the wave equation is solved in the 'ray centered' coordinates for each ray using the parabolic approximation. This approximation is a local one related to a particular ray trajectory. These local solutions can be expressed in terms

of Gaussian beams. The final solution is the result of superposing local solutions along individual rays to give an approximate global solution for a given source condition. This step relies on the linearity of the underlying wave equation.

There are several advantages in using the Gaussian beam method. In relation to the ray method, the GB method is finite at caustics. In addition, no prior location of the caustic locations is required as in the Maslov method. The GB method relies on local parabolic approximations and therefore no global direction of propagation is assumed as in the standard parabolic method. Finally, the GB method is fast, comparable to the standard ray method, and possibly faster since no two point ray tracing is required.

This thesis is organized as follows. In section 2.2 the Gaussian beam method is applied to 2-D velocity structures. The validity of the method is tested in two ways. First is the application of the reciprocal theorem for Green's functions in an arbitrary heterogeneous medium. The discrepancy between synthetic seismograms for reciprocal cases is considered as a measure of error. Second, the Gaussian beam method is applied to cases for which solutions are known by other approximate methods. This includes the soft basin problem. The method is then used for two applications. First, it is used to study volcanic earthquakes at Mt. St. Helens. The observed large differences in amplitude and arrival time

between a station inside the crater and those at the flank can be explained by the combined effects of an anomalous velocity structure and shallow focal depths. The method is then applied to scattering of teleseismic P-waves by a lithosphere with randomly fluctuating velocities.

Section 2.3 discusses several topics in Gaussian beam synthesis. This includes the choice of the beam parameter, ϵ , and its interpretation in terms of a complex source point. Asymptotic and exact evaluations of the Gaussian beam integral for a line source are then investigated. Next, the problem connected with an initial point in a gradient is discussed, which in a vertically varying media reduces to the difference between an initial plane wave and an initial Snell wave. A comparison is done with the reflectivity method for the PP phase in the Imperial Valley structure. Then, a comparison is then done between the slant stack of a wavefield and the Gaussian beam decomposition using large beam widths, showing good agreement. Beams reflected at an interface are then investigated using a first order correction to the plane wave reflection coefficient typically used in the Gaussian beam method. Finally, the finite frequency scattering effects from an indented interface are computed using the Gaussian beam method and compared with those computed using the Aki-Larner method.

In the final section, inversion for material parameters using waveform data is investigated. An iterative linearized approach is used in which a linear sensitivity operator must be derived (see Tarantola [1984]). This

can be done economically by using reciprocity of the Green's function. In order to avoid a large matrix inversion, several descent algorithms are described. Data errors and a priori model information are incorporated using covariance operators. A fast and reasonably accurate forward modelling scheme is required, and here we make use of the Gaussian beam method in a slowly varying heterogeneous medium. Different linearizations are possible including a linearization in terms of the field, the Born approximation, and a linearization in terms of the log field, the Rytov approximation. The relative merits are discussed. Finally several numerical examples are performed using a field linearization in order to test the method. Transmitted body waves through a model with sources beneath a heterogeneity and surface stations are used. The results using the waveform data identify the trial structures. A comparison is done with the travel time inversion results.

References

- Aki, K., and P.G. Richards, Quantitative Seismology: Theory and Methods, W.H. Freeman and Company, San Francisco, 1980.
- Babich, V.M., and M.M. Popov, Propagation of concentrated sound beams in a three-dimensional inhomogeneous medium, Soviet Physics-Acoustics, 27, 459-462, 1981.
- Baker, B.B. and E.T. Copson. The mathematical theory of Huygen's principle, Clarendon, Oxford, 1950.
- Boore, D.M., Finite-difference methods for seismic wave propagation in heterogeneous materials, in Math. Methods in Computational Physics, v. 11, B.A. Bolt, editor, Academic Press, New York, 1-37, 1972.
- Cerveny, V., Ray theoretical seismograms for laterally inhomogeneous structures, J. Geophys., 46, 335-342.
- Cerveny, V., I.A. Molotkov, and I. Psencik, Ray Method in Seismology, Karlova Univerzita, Praha, 1977.
- Cerveny, V., M.M. Popov, and I. Psencik, Computation of wave fields in inhomogeneous media - Gaussian beam approach, Geophys. J.R. Astron. Soc., 70, 109-128, 1982.
- Cerveny, V., and I. Psencik, Gaussian beams in two-dimensional elastic inhomogeneous media, Geophys. J.R. Astron. Soc., 72, 417-433, 1983a.
- Cerveny, V., and I. Psencik, Gaussian beams and paraxial ray approximation in three-dimensional elastic inhomogeneous media, J. Geophys., preprint, 1983b.
- Chapman, C.H., A new method for computing synthetic seismograms, Geophys. J.R. Astron. Soc., 54, 481-513, 1978.
- Chapman, C.H., and R. Drummond, Body-wave seismograms in inhomogeneous media using Maslov asymptotic theory, Bull. Seism. Soc. Am., 72, s277-s317, 1982.
- Claerbout, J.F., Fundamentals of Geophysical Data Processing, McGraw-Hill, New York, 1976.
- Corones, J., Bremmer series that correct parabolic approximations, J. Math.

- Anal. and Appl.**, 50, 361-372, 1975.
- Corones, J., B. Defacio, and R.J. Krueger, Parabolic approximations to the time independent elastic wave equation, *J. Math. Phys.*, 23, 577-586, 1982.
- Dashen, R., Path integrals for waves in random media, *J. Math. Phys.*, 20, 894-920, 1979.
- Flatte, S.M., R. Dashen, W.H. Munk, K.M. Watson and F. Zachanasen, Sound transmission through a fluctuating ocean, Cambridge Press, 1979.
- Fock, V.A., Electromagnetic diffraction and propagation problems, Pergamon Press, New York, 1965.
- Frazer, L.N. and J.B. Sinton, A Kirchhoff method for the computation of finite frequency body wave synthetic seismograms in laterally inhomogeneous media, *Geophys. J.R. astr. Soc.*, 78, 413-429, 1984.
- Goodman, J.W., Introduction to Fourier optics, McGraw-Hill, New York, 1968.
- Haddon, R.A.W. and P.W. Buchen, Use of Kirchhoff's formula for body-wave calculations in the Earth, *Geophys. J.*, 67, 1581-1607, 1981.
- Hilterman, R.J., Three-dimensional seismic modelling, *Geophysics*, 35, 1020-1037, 1970.
- Hilterman, R.J., Amplitudes of seismic waves - a quick look, *Geophysics*, 40, 745-762, 1975.
- Hudson, J.A., A parabolic approximation for elastic waves, *Wave Motion*, 2, 207-214, 1980a.
- Hudson, J.A., The parabolic approach for wave propagation as guided modes, *J. phys. D: Appl. phys.*, 13, 145-152, 1980b.
- Keller, J.B., A geometrical theory of diffraction, *J. Opt. Soc. Am.*, 52, 116-130, 1962.
- Keller, J.B., Rays, waves, and asymptotics, *Bull. of the Am. Math. Soc.*, 84, 727-750, 1978.
- Kelly, K.R., R.W. Ward, S. Treitel, and R.M. Alford, Synthetic seismograms: a finite difference approach, *geophysics*, 41, 2-27, 1976.
- Kennett, B.L.N., Seismic waves in laterally varying media, *Geophys. J.R. astr.*

Soc., 26, 37-70, 1972a.

Kennett, B.L.N., Seismic wave scattering by obstacles on interfaces, Geophys. J.R. astr. Soc., 28, 249-266, 1972b.

Kirpichnikova, N.J., Construction of solutions concentrated close to rays for the equations of elasticity theory in an inhomogeneous isotropic space, in Mathematical Problems of Theory of Diffraction and Propagation of Waves, Vol. 1, Nauka, Leningrad, 1971 (in Russian, English translation by American Mathematical Society, 1974)

Kline, M. and I.W. Kay, Electromagnetic theory and geometrical optics, John Wiley and Sons, New York, 1965.

Klem-Musatov, K.D. and A.M. Aizenberg, The ray method and theory of edge waves, Geophys. J.R. astr. Soc., 79, 35-50, 1984.

Kravtsov, Y.A., Two new asymptotic methods in the theory of waves in laterally inhomogeneous media (review), Soviet Physics-Acoustics, 14, 1-17, 1968.

Kravtsov, Y.A., and Y.I. Orlov, Limits of applicability of the method of geometric optics and related problems, Sov. Phys. Usp., 23, 750-762, 1980.

Landers, T. and J.F. Claerbout, Numerical calculations of elastic waves in laterally inhomogeneous media, J. Geophys. Res., 77, 1476-1482, 1972.

Larner K.L., Near-receiver scattering of teleseismic body waves in layered crust mantle models having irregular interfaces, Ph.D. thesis MIT, Cambridge, 1970.

Ludwig, D., Uniform asymptotic expansion at a caustic, Comm. Pure Appl. Math., 29, 215-250, 1966.

Malyuzhinets, G.D., Development in our concepts of diffraction phenomena, Sov. Phys. Usp., 69, 749-758, 1959.

May, B.T., and F. Hron, Synthetic seismogram sections of typical petroleum traps, Geophysics, 43, 119-1147, 1980.

McCoy, J.J., A parabolic theory of stress wave propagation through inhomogeneous linearly elastic solids, J. Appl. Mech., 44, 462-468, 1977.

McDaniel, S.T., Parabolic approximations for underwater sound propagation, J. Acoust. Soc. Am., 58, 1178-1184, 1976.

- McMechan, G.A., and W.D. Mooney, Asymptotic ray theory and synthetic seismograms for laterally varying structures: theory and application to the Imperial Valley, California, Bull. Seism. Soc. Am., 70, 2021-2035, 1980.
- Pao, Y.H. and V. Varatharajulu, Huygen's principle, radiation conditions, and integral formulas for elastic waves, J. Acoust. Soc. Am., 59, 1361-1371, 1976.
- Popov, M.M., A new method of computation of wave fields in high-frequency approximation, in Mathematical Problems of Theory of Propagation of Waves, Vol. 11, 195-216, Nauka, Leningrad, 1981. (in Russian, translated to English in Journal of Soviet Mathematics, 1982)
- Popov, M.M., A new method of computation of wave fields using Gaussian beams, Wave Motion, 4, 85-97, 1982.
- Popov, M.M., and I. Psencik, Ray amplitudes in inhomogeneous media with curved interfaces, Geofyzikalni Sbornik, 24, 118-129, 1976.
- Popov, M.M., and I. Psencik, Computation of ray amplitudes in inhomogeneous media with curved interfaces, Studia Geophys. Geod., 22, 248-258, 1978.
- Sanchez-Sesma, F.J. and J.A. Esquivel, Ground motion on alluvial valleys with incident plane SH waves, Bull. Seism. Soc. Am., 69, 1107-1120, 1979.
- Sanchez-Sesma, F.J., I. Herrera, and J. Aviles, A boundary method for elastic wave diffraction: application to scattering of SH waves by surface irregularities, Bull. Seism. Soc. Am., 72, 473-490, 1982.
- Smith, W.D., The application of finite element analysis to body wave propagation problems, Geophys. J.R. astr. Soc., 42, 747-768, 1975.
- Tappert, F.D., The parabolic approximation method; in Wave Propagation and Underwater Acoustics, Lecture Notes in Physics, 70, 224-287, Springer-Verlag, Berlin, 1976.
- Tarantola, A., Inversion of seismic reflection data in the acoustic approximation, geophysics, 49, 1259-1266, 1984.
- Trorey, A.W., A simple theory of seismic diffraction, Geophysics, 35, 762-784, 1970.
- Trorey, A.W., Diffraction for arbitrary source-receiver location, Geophysics, 42, 1177-1182, 1975.

Wales, S.C. and J.J. McCoy, A comparison of parabolic wave theories for linearly elastic solids, Wave Motion, 5, 99-113, 1983.

Ziolkowski, R.W., and G.A. Deschamps, Asymptotic evaluation of high frequency fields near a caustic: introduction to Maslov's method, Radio Science, 19, 1001-1025, 1984.

2.2 THE 2-D GAUSSIAN BEAM SYNTHETIC METHOD: TESTING AND APPLICATION

2.2.1 INTRODUCTION

In this paper, we will discuss applications of the Gaussian beam method. The Gaussian beam method is a variation of both the asymptotic ray method [Cerveny et. al., 1977] and the parabolic approximation method [Tappert, 1976; Claerbout, 1976]. The Gaussian beam method has been recently described in the literature by Babich and Popov [1981], Popov [1981, 1982] and Cerveny et. al. [1982]. These studies were based on the scalar wave equation. The elastic case was first derived by Kirpicnichova [1971] and more recently by Cerveny and Psencik [1983a] for the 2-D case and Cerveny and Psencik [1983b] for the 3-D case.

There are three basic steps involved in the Gaussian beam method. First, a system of rays must be traced from the source. Since high-frequency energy flows along rays, the system of rays can be thought of as the framework upon which the wavefield is to be built. This step requires the numerical solution of the kinematic ray tracing system. Second, the wave equation (elastodynamic equation) is solved in "ray centered" curvilinear coordinates for each ray using the parabolic approximation. This is now a local parabolic approximation related to a particular ray trajectory. These local solutions can be expressed in terms of Gaussian beams. The final solution is the result of superposing local solutions along individual rays to give an approximate global solution for a given source condition. This step is justified by the linearity of the underlying wave

equation.

There are several advantages to using the Gaussian beam method. In relation to the asymptotic ray method, the Gaussian beam method is always finite at caustics. In addition, no prior knowledge of caustic locations is required as in the Maslov method [Kravtsov, 1988; Chapman and Drummond, 1983]. Since the Gaussian beam method relies on local parabolic approximations, no global direction of propagation must be assumed as in the standard parabolic approximation. Finally, the Gaussian beam method is comparable in cost to ray methods, and possibly even faster since no two-point ray tracing is done.

The organization of this paper is as follows. First, a more detailed description of the Gaussian beam method is given along with computational procedures. Then, various examples are given in order to illustrate different uses of the method as well as to test its validity. Two procedures are used to test the method. First, the reciprocal theorem, valid for an arbitrary heterogeneous medium, is applied. The discrepancy between reciprocal seismograms is considered as a measure of the error in the Gaussian beam solution. Second, the Gaussian beam synthesis is applied to several cases, including the 2-D soft basin problem, for which solutions are known by other approximate methods. The Gaussian beam method is then used for two applications. First, the method is used to study volcanic earthquakes recorded under Mt. St. Helens. Then, the

method is applied to scattering of teleseismic P waves by a lithosphere with randomly fluctuating velocities.

2.2.2 REVIEW OF THE GAUSSIAN BEAM METHOD

In this section, a more detailed description of the Gaussian beam method is given. The discussion is limited to the 2-D case, for which the examples are given, and follows the results of Cerveny et. al. [1982], Cerveny and Psencik [1983a], and Cerveny [1983].

Before solving for the rays and the beam solutions, the velocity model must be parameterized. The velocities are input as discrete points, and then interpolated using cubic splines. This results in continuous velocities as well as first and second derivatives. For velocities given on a mesh in 2-D or 3-D, this involves a tensor product of 1-D splines [De Boor, 1980]. Velocity discontinuities are included by introducing interfaces between layers with smoothly varying velocities. Curved interfaces are input as a discrete set of points and then spline interpolated.

Spline interpolation of the velocity may introduce some oscillation in the interpolated function depending on the input grid values [Azbel et. al., 1980]. Care must be taken to check the interpolated function and to have the grid velocities change slowly. Several methods can be used to avoid the oscillations in the interpolating function. The first is to use smoothed splines as described by Pretlova [1976]. A second approach is to use splines under tension [see Cline, 1981]. For the remainder of this paper, careful use is made of simple bicubic spline interpolation by comparing the desired and interpolated functions.

The first step of the Gaussian beam method is to solve the kinematic ray tracing system. The rays are extremal of Fermat's integral and are determined by solving the ray equations:

$$\frac{dx_i}{ds} = vp_i, \quad \frac{dp_i}{ds} = -v^{-2} \frac{\partial v}{\partial x_i} \quad i = 1, 2 \quad (1)$$

with initial conditions:

$$\vec{x}(s_0) = \vec{x}_0, \quad \vec{p}(s_0) = \vec{p}_0$$

where $\vec{x}$ is the location along the ray and $\vec{p}$ is the slowness vector tangent to the ray. In an elastic medium, the velocity, v , is either α or β . This system can be solved by standard numerical techniques such as the Runge-Kutta method. At an interface, Snell's law is applied locally. These ray equations are appropriate for an isotropic non dispersive media. In an isotropic dispersive media, wave packets will travel at the group velocity, but their trajectories will be determined by the local phase velocity.

The second step of the Gaussian beam method involves solving the wave equation locally in ray-centered coordinates using the parabolic approximation. The ray-centered coordinate system is an orthogonal curvilinear system that follows along a particular ray and was introduced to seismology by Popov and Psencik [1976, 1978]. In a 2-D medium, the ray-centered coordinates can be specified by the unit vector $\vec{t}$ tangent to the ray and the unit vector $\vec{n}$ normal to the ray (see Figure 1). An element of infinitesimal length in the ray-centered coordinates can be written

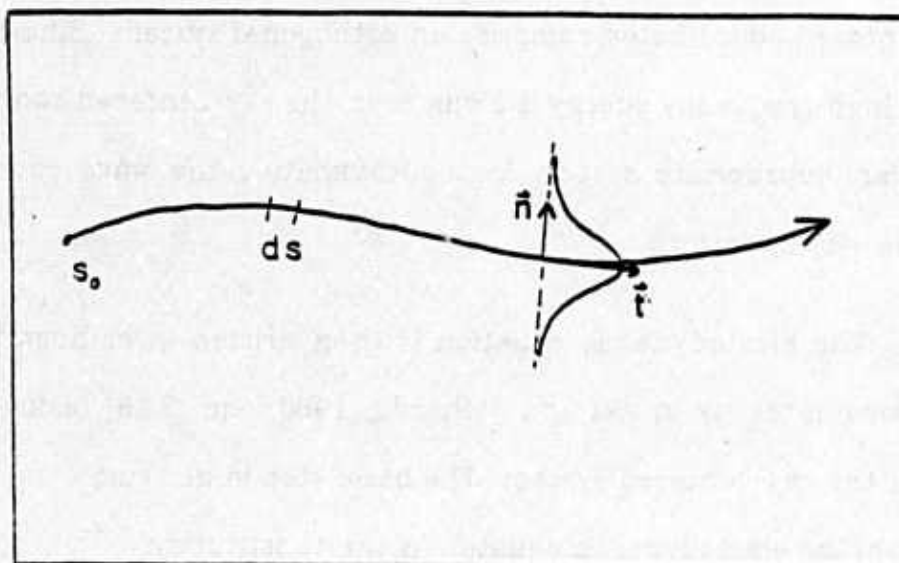


Figure 1

Figure 1 Ray centered coordinate system in 2-D. The beam solution is concentrated near the ray and has a Gaussian shape.

$$|d\vec{r}|^2 = h_1^2 ds^2 + h_2^2 dn^2 \quad (2)$$

where the scale factors, h_1 and h_2 , are given by

$$h_1 = 1 + v^{-1} v_{,n} n \quad h_2 = 1$$

where v and $v_{,n}$ are evaluated on the central ray with $n=0$. The ray-centered coordinates comprise an orthogonal system. Since the local flow of high-frequency energy is along rays, the ray-centered coordinate system is an appropriate system for approximating the wave equation by local one-way operators.

The elastodynamic equation is then written in orthogonal curvilinear coordinates, as in Aki and Richards [1980, eqn. 2.48], using scale factors for the ray-centered system. The basic step in deriving a one-way operator from the elastodynamic equation is the substitution

$$u_j(s, n, \omega, t) = U_j(s, n, \omega) e^{-i\omega(t - \tau(s))} \quad (3)$$

where, $\tau(s) = \int v(s)^{-1} ds$ is the phase delay along the central ray, $v(s) = \alpha$ or $v(s) = \beta$, and U_j is a slowly varying envelope. The envelope is then expanded as $U_j = \sum_{i=0}^{\infty} U_j^{(i)} \omega^{-i/2}$. Note that this is in half powers of ω^{-1} as opposed to the typical ray series expansion in integral powers. In 2-D, U_s and U_n are the in-plane components along the ray and normal to it, and U_b is the out of plane component normal to the ray. After substitution of equation (3) into the elastodynamic equation, the resulting equations are approximated for large ω . The following results are obtained by retaining only terms of

order ω^7 , $\gamma \geq 1$, and letting $\nu = \omega^{1/2} n$ assuming that $\nu = O(1)$. To this order of approximation, the P and S components decouple. For an incident P wave [Cerveny and Psencik, 1983]:

$$\begin{aligned} U_s^0(s, \nu) &= \frac{1}{\sqrt{\alpha(s)\rho(s)}} W^a(s, \nu) \\ U_n^0 &= 0 \\ U_n^1 &= -i\alpha U_{s,\nu}^0 \\ U_b &= 0 \end{aligned} \quad (4)$$

where W^a satisfies the following parabolic equation

$$2i\alpha^{-1} W_s^a + W_{\nu\nu}^a - \alpha^{-3} \alpha_{,nn} \nu^2 W^a = 0 \quad (5)$$

For an incident SV wave

$$\begin{aligned} U_n^0(s, \nu) &= \frac{1}{\sqrt{\beta(s)\rho(s)}} W^b(s, \nu) \\ U_s^0 &= 0 \\ U_s^1 &= i\beta U_{n,\nu}^0 \\ U_b &= 0 \end{aligned} \quad (6)$$

where W^b satisfies the parabolic equation

$$2i\beta^{-1} W_s^b + W_{\nu\nu}^b - \beta^{-3} \beta_{,nn} \nu^2 W^b = 0 \quad (7)$$

and for an incident SH wave

$$U_b^0(s, \nu) = \frac{1}{\sqrt{\beta(s)\rho(s)}} W^b(s, \nu) \quad (8)$$

$$U_s = U_n = 0$$

with W^b satisfying the same parabolic equation as for the SV case. All

velocities and densities are evaluated along the central ray. These results were derived from asymptotic analysis of the elastodynamic equation. The same basic parabolic equation for W can also be obtained from the eikonal equation as shown in appendix 1.

The last term in the parabolic equation (5) or (7) comes from the approximation of order ω

$$h_1 \omega^2 \left[\frac{1}{v(s,n)^2} - \frac{1}{h_1^2 v(s)^2} \right] W \approx -\omega v(s)^{-3} v_{,nn} v^2 W$$

where $v(s)$ and $v_{,nn}$ are evaluated along the central ray, $v(s,n)$ is the velocity in the vicinity of the central ray, and h_1 is the scale factor for the ray-centered coordinates. Without this approximation for the last term, the parabolic equation (5) or (7) is similar in form to those given by Aki and Richards [1980, eqn. 13.154] or Claerbout [1978, 10-3-9]. The parabolic equation (5) or (7) is valid for small angle forward scattering about the central ray direction assuming large scale inhomogeneities with respect to the wavelength. In addition, the validity of the above approximation requires that the velocities be smoothly varying within a certain beam-width of the central ray. Since, as will be seen, the beam solutions to (5) or (7) spread with distance, this will result in a range limitation for the Gaussian beam method.

A particular solution to the parabolic equation (5) or (7) can be written as

$$W(s) = \frac{1}{\sqrt{q(s)}} e^{\frac{i}{2} \frac{p(s)}{q(s)} v^2} \quad (9)$$

where $q(s)$ and $p(s)$ satisfy

$$\frac{d}{ds} \begin{bmatrix} q \\ p \end{bmatrix} = \begin{bmatrix} 0 & v \\ -v & 2v_{,nn} \end{bmatrix} \begin{bmatrix} q \\ p \end{bmatrix} \quad (10)$$

with $v(s)$ and $v_{,nn}$ evaluated along the central ray. These are the same equations as the dynamic ray tracing system for ray theory, except here $p(s)$ and $q(s)$ have to be complex. $p(s)$ and $q(s)$ must also satisfy certain conditions, among them the condition $\text{Im}(p/q) > 0$. In this case the energy will be confined to a vicinity of the central ray, and the solution will be a Gaussian beam (see Figure 1). For ray theory p and q are real, and $q(s)$ has the interpretation of spreading and $v(s) \frac{p(s)}{q(s)}$ has the interpretation of wavefront curvature. The equations in (10) are in transmission-line form, where the ratio p/q is related to the characteristic admittance for the system.

The complete solution to the parabolic equation for a particular central ray is made up of an infinite system of linearly independent beam solutions or modes [Cerveny et.al., 1982] of the form

$$W_k(s, v) = \frac{1}{\sqrt{q(s)}} \left(\frac{q}{q'} \right)^{k/2} H_k(v \sqrt{\text{Im}(p/q)}) e^{\frac{i}{2} \frac{p}{q} v^2} \quad k=0,1,2 \quad (11)$$

where H_k are hermite polynomials. Here, only the zeroth order beam solution, (9), is used for each central ray and no mode coupling between higher order beam solutions is assumed.

The localized solution to the 2-D elastodynamic equation along a particular central ray, specified by a ray index parameter, δ , can be written in the following form [Cerveny and Psencik, 1983a]

$$\vec{u}_\delta(s, \vec{n}, \omega) = \frac{1}{\sqrt{v(s)\rho(s)q(s)}} \vec{r} e^{-i\omega(t - \int_0^s \frac{ds}{v(s)} - \frac{1}{2} \frac{p(s)}{q(s)} n^2)} \quad (12)$$

where for P waves, $v(s) = \alpha(s)$, $\vec{r} = [\vec{t} + \vec{n}\alpha(s) \frac{p(s)}{q(s)} \vec{n}]$;

for SV waves, $v(s) = \beta(s)$ $\vec{r} = [\vec{n} - \vec{t}\beta(s) \frac{p(s)}{q(s)} \vec{n}]$;

and for SH waves, $v(s) = \beta(s)$ $\vec{r} = \vec{\delta}$ (*out of plane component*);

where p and q solve (10) and are in general complex, and $\vec{r}$ gives the orientation of the displacement vector nearby the central ray. Note that the beam solutions are valid only where the ray-centered coordinates are well defined and regular in some region nearby the central ray.

The beam solutions must be modified in the presence of any interfaces. Following Popov [1982], it is assumed that to first approximation the beam solutions, (11), for a given ray remain uncoupled at a smoothly curved interface. Then, the particular beam solution, (9), is only modified by a possibly complex reflection or transmission coefficient and the change in $p(s)$ and $q(s)$ due to the interface. Formulas for how real-valued $p(s)$ and $q(s)$ transform at an interface are given by Cerveny and Psencik [1979], and for curvature by Cerveny and Hron [1980]. Formulas for how

complex-valued $p(s)$ and $q(s)$ transform at an interface are given by Cervený [1983c], where the equations are written so that $q(s)$ is a relative spreading across each interface. This avoids writing an additional factor in the beam amplitude.

For Gaussian beams, $p(s)$ and $q(s)$ can be written as [see Cervený et. al., 1982]

$$q(s) = \epsilon q_1(s) + q_2(s) \quad (13)$$

$$p(s) = \epsilon p_1(s) + p_2(s)$$

where $p(s)$ and $q(s)$ satisfy equation (10), ϵ is a complex valued parameter to be determined, and $\begin{bmatrix} q_1 & q_2 \\ p_1 & p_2 \end{bmatrix}$ is the fundamental matrix of linearly independent real solutions of (10) with initial conditions $\begin{bmatrix} 1 & 0 \\ 0 & 1/v_0 \end{bmatrix}$. From these initial conditions

$$K_1(s_0) = v_0 \frac{p_1(s_0)}{q_1(s_0)} = 0 \quad \rightarrow \quad \text{initial plane wave at source}$$

$$K_2(s_0) = v_0 \frac{p_2(s_0)}{q_2(s_0)} = \infty \quad \rightarrow \quad \text{initial line source}$$

where $v_0 = v(s_0)$. Thus, the complex solution is a linear combination of real solutions for a plane wave and line source weighted by ϵ as in (13). ϵ is a factor chosen so that a) $q(s) \neq 0$; resulting in no singularities in amplitude along the ray, and b) $\text{Im}(\frac{p(s)}{q(s)}) > 0$, implying that the solution is concentrated near the ray. Following Cervený et. al. [1982], ϵ is written

in the form

$$\varepsilon = S_0 - i \frac{\omega}{2v_0} L_M^2 \quad (14)$$

where, geometrically, L_M is the beam-width at a specific location along the beam known as the beam waist, and S_0 is a parameter which shifts the location of the beam waist away from the source location, s_0 , along the ray. Cerveny et. al.[1982] showed that conditions a) and b) above are both satisfied when $L_M \neq 0$. This result depends on the fact that the determinant of the fundamental matrix above is a nonzero constant along the ray, and therefore, $q_1(s)$ and $q_2(s)$ cannot simultaneously be zero for a given s . Thus, the method is constructed to ensure no singularities in amplitude.

The fundamental beam solution can now be written from (12) as

$$\vec{u}_\delta = \frac{1}{\sqrt{v(s)\rho(s)q(s)}} \vec{r} e^{-i\omega(t - \int_0^s \frac{ds}{v(s)}) + \frac{i\omega}{2v} K(s)n^2 - \frac{n^2}{L(s)}} \quad (15)$$

where

$$q(s) = [S_0 - i \frac{\omega}{2v_0} L_M^2] q_1 + q_2 = \text{complex spreading} \quad (16)$$

$$K(s) = v(s) \text{Re} \left[\frac{p(s)}{q(s)} \right] = \text{phase front curvature of the beam} \quad (17)$$

$$K(s) = \frac{v(s) [(S_0 p_1 + p_2)(S_0 q_1 + q_2) + (\frac{\omega L_M^2}{2v_0})^2 p_1 q_1]}{[(S_0 q_1 + q_2)^2 + (\frac{\omega L_M^2}{2v_0})^2 q_1^2]}$$

and

$$L(s) = \left[\frac{\omega}{2} \operatorname{Im} \left(\frac{p(s)}{q(s)} \right) \right]^{-1/2} = \text{beam halfwidth} \quad (18)$$

$$L(s) = [L_M^2 q_1^2 + (\frac{2v_0}{\omega L_M})^2 (S_0 q_1 + q_2)^2]^{1/2}$$

For convenience, the parameter L_0 will sometimes be used to specify the initial beam-width at the beam waist. L_0 is specified by $L_M = \frac{2v_0}{\omega}^{1/2} L_0$, and is independent of frequency since from (18), $L_M = L(s_0 - S_0) = O(\omega^{-1/2})$.

The properties of a Gaussian beam in a homogeneous medium are briefly reviewed in Appendix 2. An example of a Gaussian beam in a homogeneous medium with $L_M = 1$ and $\lambda = \frac{\pi}{4}$ is shown in Figure 2. The beam waist is located at $s = 0$. The solid lines show the spread of the beam with distance and the dashed straight lines are the far field asymptotes. The distance s_c separates the columnated near zone from the diverging far zone of the beam and is proportional to the initial beam width squared over the wavelength. The near zone has planar phase fronts while the far zone diverges as if from a point. By adjusting the initial beam-width, the beam-width observed at a given station range can be changed. Adjusting the initial beam-width so that the station distance is just s_c results in the smallest possible beam-width at the station. This will be called the critical initial beam-width for a given station range. It also places the station just between the columnated planar part of the beam and the far field spreading part. For a homogeneous medium the critical

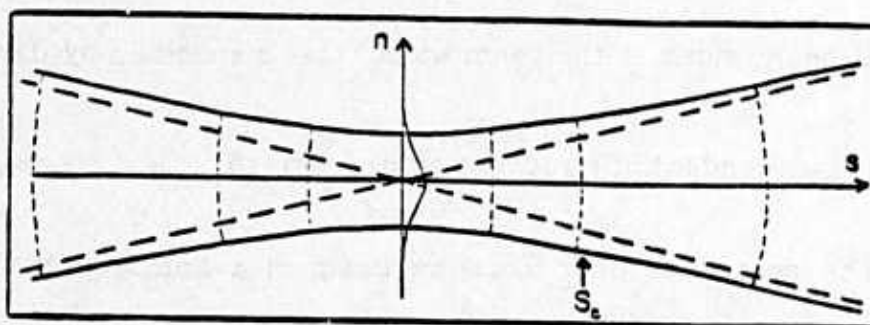


Figure 2

Figure 2 A Gaussian beam in a homogeneous medium with $L_M = 1$. and $\lambda = \pi/4$.

initial beam-width is $L_{Mcr} = \sqrt{\lambda s / \pi}$. For the general inhomogeneous case

$$L_{Mcr} = \left(\frac{2\nu_0}{\omega} \right)^{1/2} \left(\left| \frac{S_0 q_1 + q_2}{q_1} \right| \right)^{1/2} \quad (19)$$

[see Cerveny et.al., 1982]

An alternative way to generate a Gaussian beam in a homogeneous medium is to use an initial planar Gaussian amplitude distribution and diffract it using the Fresnel diffraction integral [see Marcuse, 1982]. Since the Gaussian beam is a known solution to the Fresnel integral for a given initial amplitude distribution, the field from an arbitrary initial amplitude distribution can be approximated as a sum of basis Gaussian beams.

The final step of the Gaussian beam method is a superposition of beam solutions to represent a given source using the proper weight functions, $\Phi(\delta)$, where δ is the ray index parameter. For an initial line source, δ represents an initial takeoff angle for a given ray. For an initial plane wave, δ represents the distance along the wavefront for a given ray. Each beam solution solves the wave equation in the vicinity of a ray, and together, the weighted superposition approximately satisfies a source condition for a given source. Thus, the following integral over ray index parameter and frequency can be written

$$\bar{u}(\bar{x}, t) = \frac{1}{\pi} \text{Re} \int_0^{\omega_M} \int_{\delta_0}^{\delta_M} \Phi(\delta) F(\omega) \bar{u}_\delta(s, \pi, \omega, t) i \delta d\omega \quad (20)$$

where $F(\omega)$ is the source spectra, and δ_0, δ_M define the range of the ray

index parameter. δ_0 and δ_M must be chosen to adequately cover the area of interest with rays. Using the slowness approach [Chapman, 1978]

$$\bar{u}(\bar{x}, t) = \int_{\delta_0}^{\delta_M} \Phi(\delta) \bar{g}(\bar{x}, t, \delta) d\delta \quad (21)$$

where $\bar{g}(\bar{x}, t)$ is a wave packet [see Cervený, 1983]. In the examples, $F(\omega)$ is the source spectra for the following damped cosine wavelet [see Cervený et.al., 1977]

$$f(t) = \exp(-(\omega t / \gamma)^2) \cos(\omega t + \varphi) \quad (22)$$

Using this source wavelet, the wave packet $\bar{g}(\bar{x}, t, \delta)$ is approximately Gaussian in both time and space [Cervený, 1983]. Equation (21) is then approximated by the finite sum

$$\bar{u}(\bar{x}, t) = \sum_{i=0}^N \Phi(\delta_i) \bar{g}(\bar{x}, t, \delta_i) \Delta\delta \quad (23)$$

where $\Delta\delta$ must be chosen small enough to smoothly interpolate the beam solutions and to adequately sample the medium.

The weight factors $\Phi(\delta)$ for a given source must then be specified. The weight function for an acoustic line source is given by Popov [1982] and Cervený et. al. [1982], and for a plane wave by Cervený [1982]. In Appendix 3, the weight function for a 2-D point force in an elastic medium is obtained by comparing the Gaussian beam superposition with the exact plane wave decomposition.

There are thus several free parameters when computing the Gaussian beam solution for a given problem: δ_0 , δ_M , and $\Delta\delta$ for the finite sum, the initial beam-width at the beam waist L_M , and the shift of the beam waist away from the source given by the parameter S_0 . The sensitivity of the method on these parameters, particularly L_M , will be investigated in several of the following examples.

2.2.3 APPLICATIONS

In this section, several applications of the Gaussian beam method are given in order to illustrate the uses of the method and test its range of applicability. The computer program for the Gaussian beam synthesis was modified from the ray tracing program, RAY81, written by I. Psencik [1983]. For all examples, the slowness method is used in which the frequency integral is evaluated first assuming Gaussian packets and a damped cosine source wavelet [Cerveny, 1983].

Example A

In this example, an initial plane wave in a homogeneous medium is decomposed into a number of centered Gaussian beams. Each beam is propagated along the straight line rays shown in Figure 3a. The initial plane wave is located at $z=-50$ km. The station is shown by the triangle at $x=25$ km, $z=0$ km. The velocity is 6 km/sec and the source wavelet parameters are, $f=2$ Hz, $\gamma=4$, $\varphi_0=0$ (see eqn. (22)). The contribution of each beam to the resulting sum is shown for 3 different initial beam-widths in Figures 3b-d. The initial beam-width is represented by L_0 , a frequency independent parameter where $L_M = \left(\frac{2v_0}{\omega}\right)^{1/2} L_0$. L_M has units of km and L_0 has units of $\text{km}^{1/2}$. For this example, $L_M = 98L_0$. The value $L_0 = 7 \text{ km}^{1/2}$ represents the critical initial beam-width that gives the smallest beam-

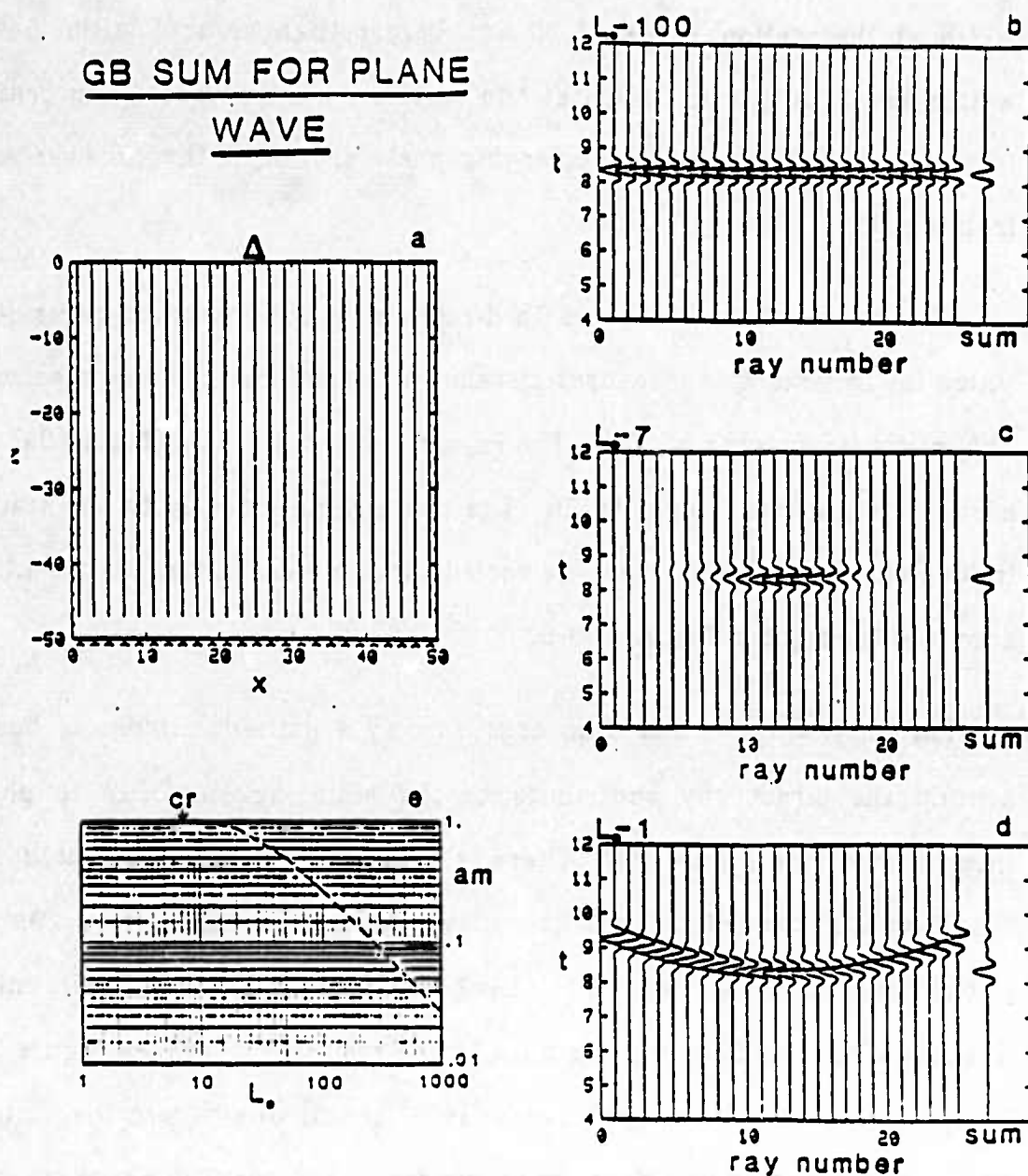


Figure 3

Figure 3 Superposition of Gaussian beams to represent an initial plane wave in a homogeneous medium. a) ray diagram for initial plane wave at $z = -50$ km. Individual beam contributions indexed by ray number along the wavefront and the resulting sum for three initial beam-widths b) $L_0 = 100 \text{ km}^{1/2}$, c) $L_0 = 7 \text{ km}^{1/2}$, and d) $L_0 = 1 \text{ km}^{1/2}$. e) The amplitude of the Gaussian beam sum as a function of L_0 .

width at the station range of 50 km. Larger than critical initial beam widths result in planar beams at the receiver, while smaller than critical initial beam-widths result in diverging phase fronts at the receiver as if from a point.

The 'ray number' in Figures 3b-d represents the ray index parameter, which for this example measures distance along the initial plane wavefront. The initial ray spacing is 2 km. The rays in the ray diagram (Figure 3a) are assigned a number from 1 to 25. The beam going directly to the station follows along ray number 13. The scaled sum for each initial beam-width is shown to the right in Figures 3b-d.

For $L_0 = 1 \text{ km}^{1/2}$, less than critical, only a limited number of beams around the direct ray contribute to the resulting sum due to phase interference (see Figure 3d). There is also a spurious end effect in the resulting sum caused by lack of cancellation of the end beams. As the initial beam-width is changed to $L_0 = 7 \text{ km}^{1/2}$, near critical, again only a limited number of beams contribute to the resulting sum (see Figure 3c). This is now caused by amplitude decay of beams away from the station. The spurious end effects have been avoided in the resulting sum by using the critical initial beam-width. Finally for $L_0 = 100 \text{ km}^{1/2}$, larger than critical, the beams are more planar at the station distance and no amplitude or phase interference limits the number of rays contributing to the resulting sum (Figure 3b). The sum will then depend on the ray

aperture used.

Figure 3e shows the amplitude of the resulting sum as a function of initial beam-width. The critical initial beam-width is shown. The true amplitude at the receiver is one. For larger than critical initial beam-widths, the amplitude of the Gaussian beam sum falls off since a finite aperture of rays is used.

The following comments can be made concerning decomposing an initial plane wave in a homogeneous medium into Gaussian beams. For larger than critical initial beam-widths, the amplitude is not stable using a finite ray aperture. For near critical initial beam-widths, the result is stable by amplitude decay of beams away from the receiver. The end effects are avoided here by amplitude decay. For smaller than critical initial beam-widths, a stable result occurs by phase interference, but spurious end effects are noticeable in the resulting sum. This is similar to a decomposition of the initial wavefront into point-like sources as viewed at the receiver.

Example B

In this example an initial half plane is decomposed into Gaussian beams. The half plane is located at a depth of 2.5 km and terminates at $x=15$ km as seen in Figure 1a. Receivers are located on the surface from

GB REPRESENTATION OF HALFPLANE

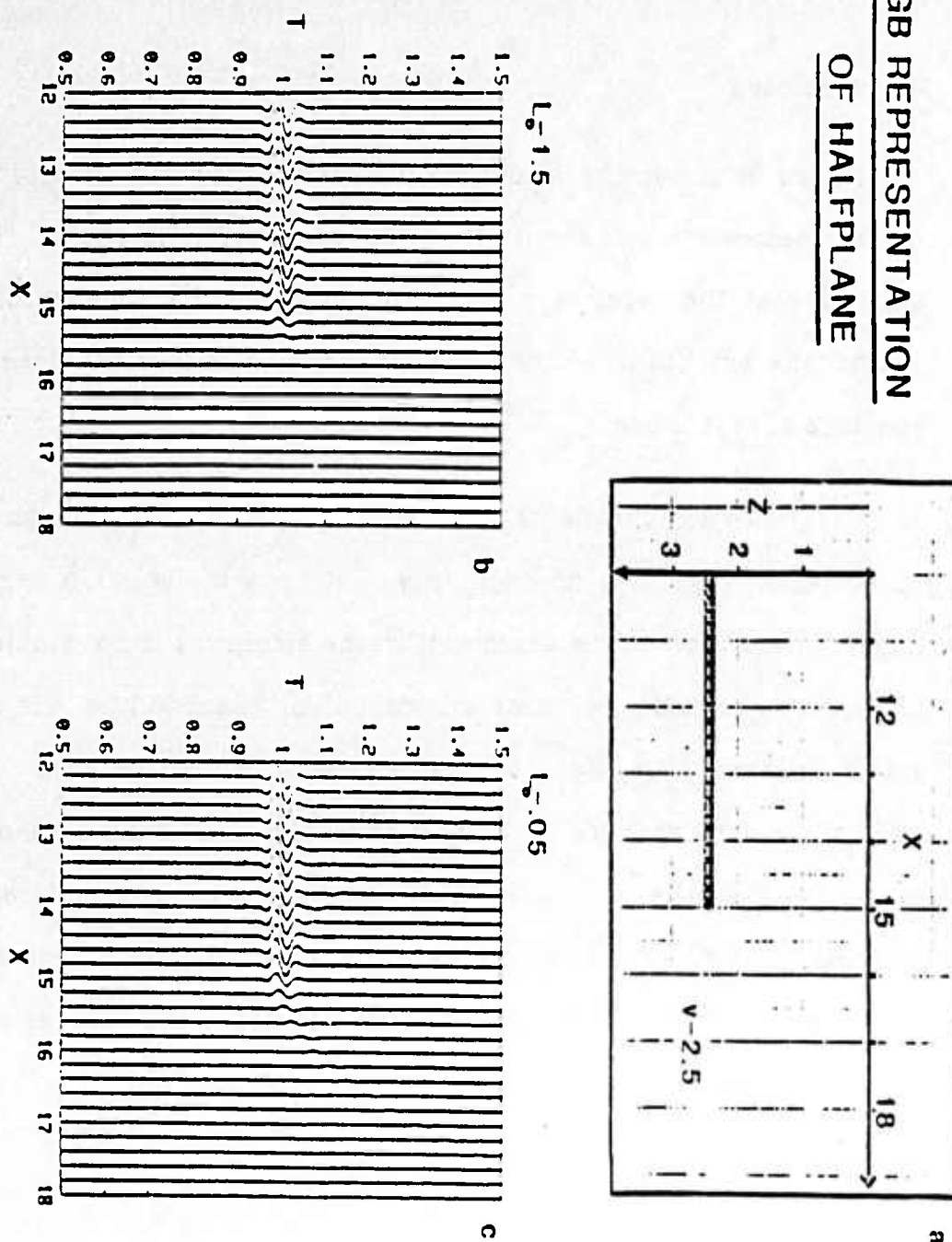


Figure 4

Figure 4 Superposition of Gaussian beams to represent an initial half plane. a) Half plane located at $z = -2.5$ km and terminating at $x = 15$ km. Stations located on the surface from $x = 12$ km to $x = 18$ km. b) Gaussian beam seismograms for $L_0 = 1.5 \text{ km}^{1/2}$. c) Gaussian beam seismograms for $L_0 = 0.05 \text{ km}^{1/2}$.

12 to 18 km with a spacing of 0.2 km. The velocity is 2.5 km/sec and the source wavelet parameters are, $f=20$ hz, $\gamma=4$, $\varphi_0=\pi/2$.

The initial beam-width, L_M , is related to L_0 by $L_M = 0.2L_0$. The critical value of L_0 for a receiver range of 2.5 km is $L_{0cr} = \sqrt{s} = 1.58 \text{ km}^{1/2}$. Figure 4b shows the resulting seismograms for $L_0 = 1.5 \text{ km}^{1/2}$. Since the initial amplitude on the edge of the half plane goes to zero gradually over 0.3 km, the amplitude at the receivers smoothly goes to zero at $x=15$ km with no evident edge diffraction. When L_0 is reduced to $.05 \text{ km}^{1/2}$, the initial amplitude on the edge of the half plane goes to zero much more sharply. A diffracted wave from the edge of the half plane can now be seen (Figure 4c).

Example C

In this example, the decomposition of a line source in a homogeneous medium into Gaussian beams is investigated. Each beam is propagated along the rays shown in the ray diagram in Figure 5a, for a source at $x=25$ km, $z=-50$ km. The station at which the resulting sum is evaluated is shown by the triangle located at $x=25$ km, $z=0$ km. The velocity is 6 km/sec and the source wavelet parameters are, $f=2$ hz, $\gamma=4$, $\varphi_0=0$. The contribution of each beam to the resulting sum at the station is shown for three initial beam-widths in Figures 5b-d. The initial beam-width is represented by L_0 where, $L_M = 98L_0$. The 'ray number' corresponds to the ray index

parameter which for this case is the initial takeoff angle at the source. The rays in the ray diagram (Figure 4a) are assigned a number from 1 to 18. The beam going directly to the station is represented by ray number 10. The scaled sum for each initial beam-width is shown to the right in Figures 5b-d.

The initial beam-width of $L_0 = 10 \text{ km}^{1/2}$ is slightly larger than the critical value of $L_0 = 7 \text{ km}^{1/2}$. For a given station distance, L_0 equal to critical separates planar beams from the beams that diverge as if from a point. For $L_0 = 100 \text{ km}^{1/2}$, larger than critical, only a limited number of beams contribute around the direct ray due to phase interference (see Figure 5b). A noticeable end effect from the finite ray aperture can be seen in the sum. For $L_0 = 10 \text{ km}^{1/2}$, again only a limited number of rays about the direct ray contribute to the sum (see Figure 5c). This is now due to amplitude decay of beams away from the station. The end effects have been avoided in this case. For $L_0 = 1 \text{ km}^{1/2}$, smaller than critical, neither amplitude or phase interference limits the number of beams contributing about the direct ray (see Figure 5d). The result then will depend on the ray aperture used.

The amplitude of the Gaussian beam seismogram as a function of initial beam-width is shown in Figure 5e. The amplitude of 1.0 in Figure 5e corresponds to the amplitude of the far field approximation for a line source, where $U(D) \approx -1/4\pi\sqrt{\lambda/r}e^{i2\pi r/v}e^{i\pi/4}$. For $r = 50 \text{ km}$,

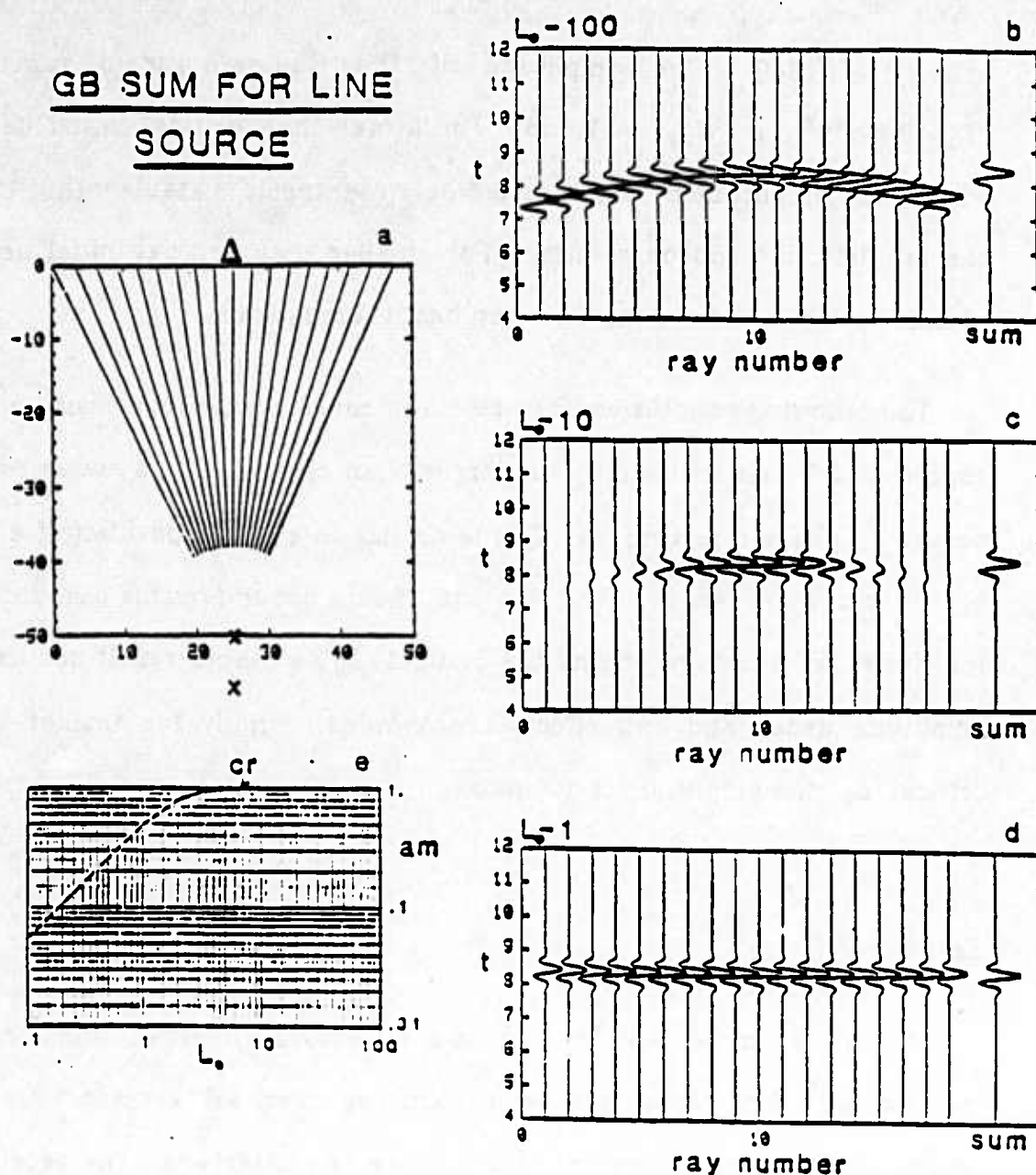


Figure 5

Figure 5 Superposition of Gaussian beams to represent a line source in a homogeneous medium a) ray diagram for line source at $x=25$, $z=-50$ Individual beam contributions indexed by ray number representing initial takeoff angle and the resulting sum for b) $L_0=100 \text{ km}^{1/2}$, c) $L_0=10 \text{ km}^{1/2}$, and d) $L_0=1 \text{ km}^{1/2}$ e) The amplitude of the Gaussian beam sum as a function of L_0

$|U(D)| = .01949$. The amplitude of the Gaussian beam sum for $L_0 = 10 \text{ km}^{1/2}$ is $|U_{GB}| = .01958$. For larger than critical initial beamwidths, the amplitude of the Gaussian beam synthesis is stable within 1% of the far field line source solution. For smaller than critical initial beamwidths, the amplitude of the Gaussian beam sum falls off.

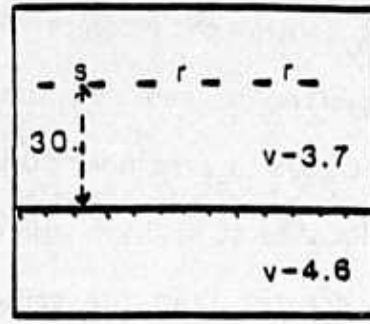
The following conclusions can be made concerning decomposing a line source into Gaussian beams. For larger than critical L_0 , a stable result occurs by phase interference. This is similar to a decomposition of a line source into plane waves. Spurious end effects occur for this case due to the finite ray aperture. Using the critical L_0 , a stable result occurs by amplitude decay and end effects are avoided. Finally for smaller than critical L_0 , the amplitude is not stable.

Example D

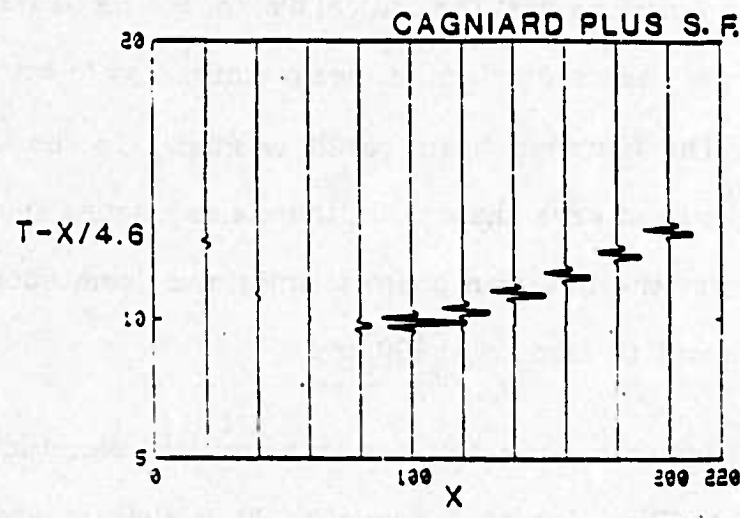
In this example, an SH line source above a welded interface is investigated. The velocities are 3.7 km/sec over 4.6 km/sec with the source and receivers located 30 km above the interface. The receivers range from 20 km to 200 km with a spacing of 20 km. The model is shown in Figure 6a.

Figure 6b shows the seismograms computed using the Cagniard method [see Aki and Richards, 1980]. The Cagniard result is convolved with

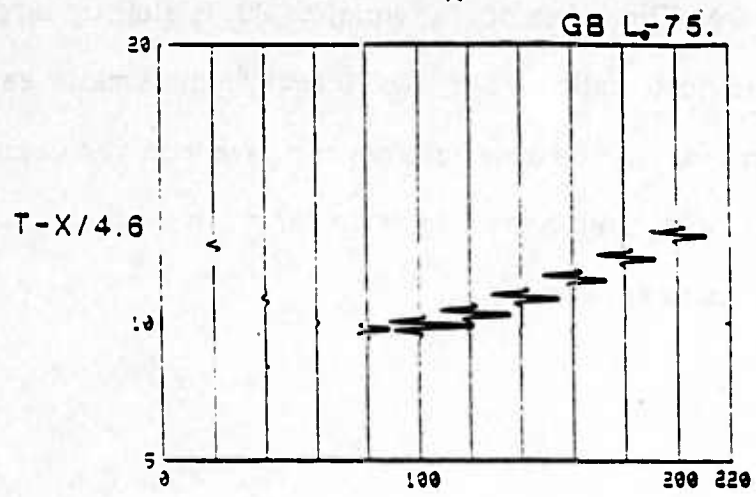
SH LINE
SOURCE



a



b



c

Figure 6

Figure 6 SH line source above a welded interface. a) model geometry with source and receivers 30 km above interface. b) Cagniard seismograms c) Gaussian beam seismograms using $L_0=75. km^{1/2}$

a source wavelet with parameters $f=3 \text{ Hz}$, $\gamma=4$, $\varphi_0=0$, and the seismograms are plotted in reduced time, $T-X/4.6$. The precritical reflection coefficient goes to zero near 60 km and then changes sign. The critical distance is located at 81.2 km, and the amplitude becomes large at a distance slightly greater than the critical distance. For post-critical distances, total reflection occurs with a phase shift. A small head wave can also be seen emerging past the critical distance. The Gaussian beam result using a larger than critical initial beam-width, $L_0=75 \text{ km}^{1/2}$, is shown in Figure 6c. The Gaussian beam result is similar to the Cagniard result, including the head wave phase. But, there also exists a spurious end effect cutting across the Gaussian beam seismograms from about 6 seconds at 100 km to about 12 seconds at 200 km.

The Gaussian beam result for three initial beam-widths is shown in Figures 7a-c. The value of L_0 equal to 20 is slightly larger than critical for the farthest station. For $L_0=20 \text{ km}^{1/2}$, the small head wave phase is not generated, but the other phases compare with the Cagniard result. For larger L_0 , the head wave begins to form. In addition, an unwanted end effect begins to appear.

Example E

In this example, the effect of a single caustic in a layer over a gradient is reviewed. This is a standard example, but illustrates several features of

SH LINE
SOURCE

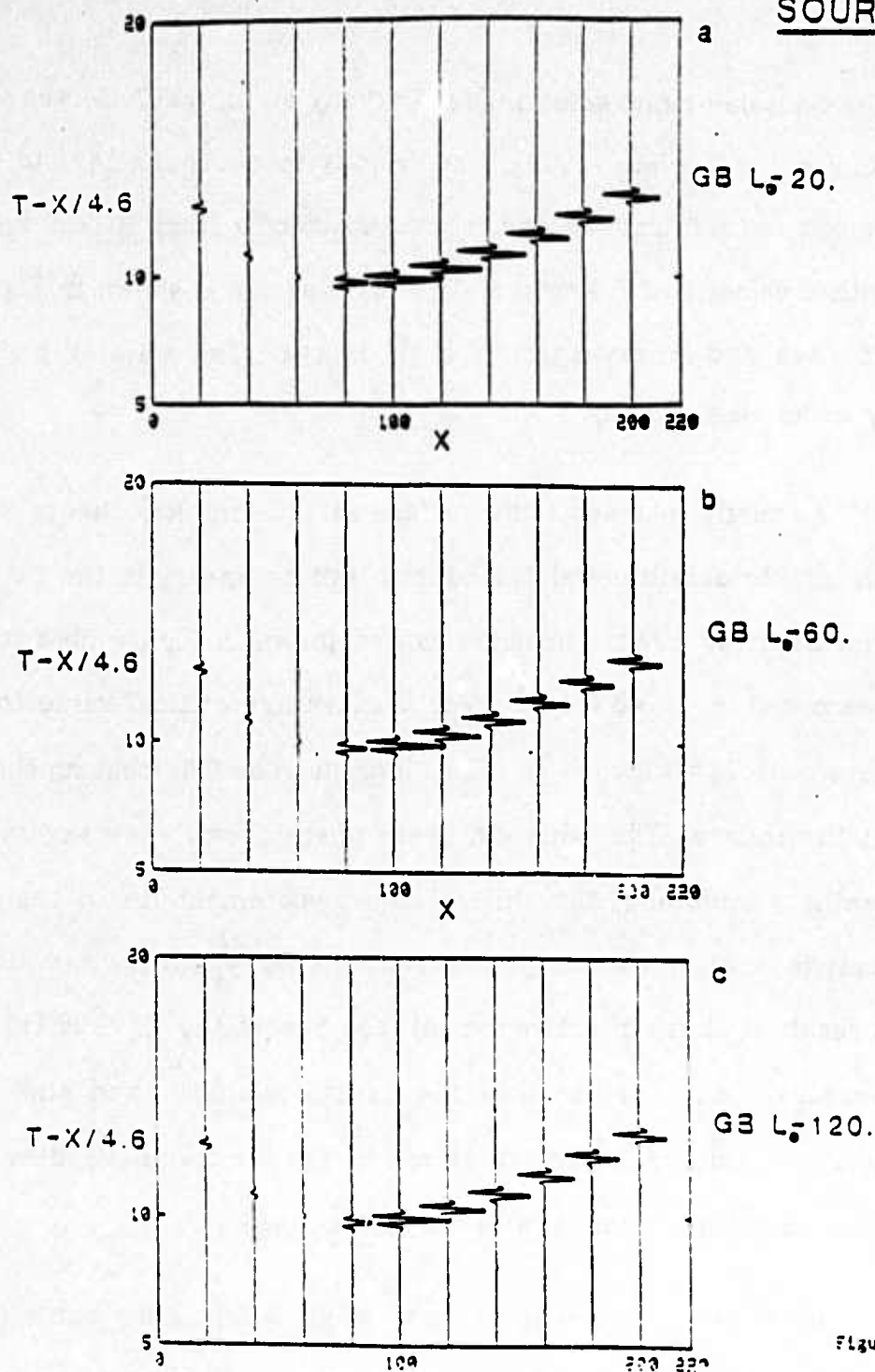


Figure 7

Figure 7 Gaussian beam seismograms for SH line source above welded interface for a) $L_0 = 20 \text{ km}^{1/2}$, b) $L_0 = 60 \text{ km}^{1/2}$, and c) $L_0 = 120 \text{ km}^{1/2}$

the Gaussian beam solution [see Cervený et. al., 1977; Cervený et. al., 1982; Klimes and Cervený, 1983]. The velocity model is constant to 15 km with a velocity of 5.6 km/sec and increases linearly from 15 km down to 40 km with a velocity of 8 km/sec. The ray diagram is shown in Figure 8a using 36 rays and a ray spacing of 1 degree. The wavelet parameters are $f = 4 \text{ Hz}$, $\gamma = 3$, $\varphi_0 = 0.0$.

A caustic intersects the surface at 120 km. Ray theory would predict an infinite amplitude at the caustic and no energy in the caustic shadow. The Gaussian beam seismograms are shown in Figure 8b-c for two initial beam-widths, $L_0 = 8 \text{ km}^{1/2}$ near the average critical value for the set of rays, and $L_0 = 128 \text{ km}^{1/2}$. S_0 is chosen to be 0.0, locating the beam-waist at the source. The Gaussian beam seismograms show several interesting features including, the shift of the peak amplitude to the right of the caustic location and the penetration of energy into the caustic shadow. As a result of phase interference between beams, the $L_0 = 128 \text{ km}^{1/2}$ does not penetrate any further into the caustic shadow than the $L_0 = 8 \text{ km}^{1/2}$ solution. But, spurious phases can be seen in the $L_0 = 128 \text{ km}^{1/2}$ solution which don't quite cancel with the ray set used.

In order to investigate how each beam contributes to the final seismograms, the individual beam solutions are plotted for several station locations and initial beam-widths. The results for a station at 140 km are shown in Figure 9. Figure 9a shows the ray diagram using 18 rays and a

LAYER OVER GRADIENT

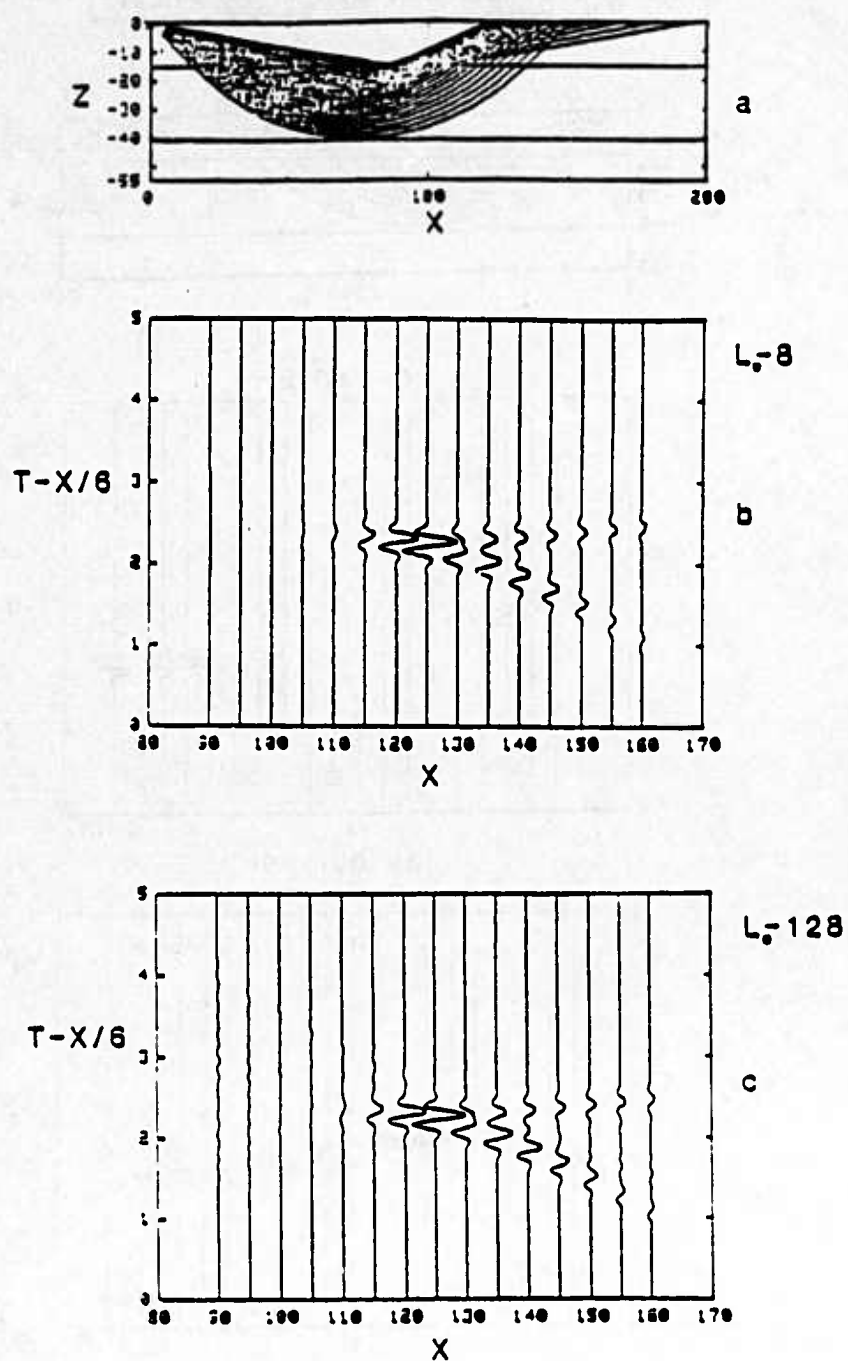


Figure 8

Figure 8 Layer over a gradient. a) ray diagram where caustic intersects surface at $x=120$ km. b) Gaussian beam seismograms for $L_0=8$ km^{1/2}. c) Gaussian beam seismograms for $L_0=128$ km^{1/2}.

LAYER OVER GRADIENT

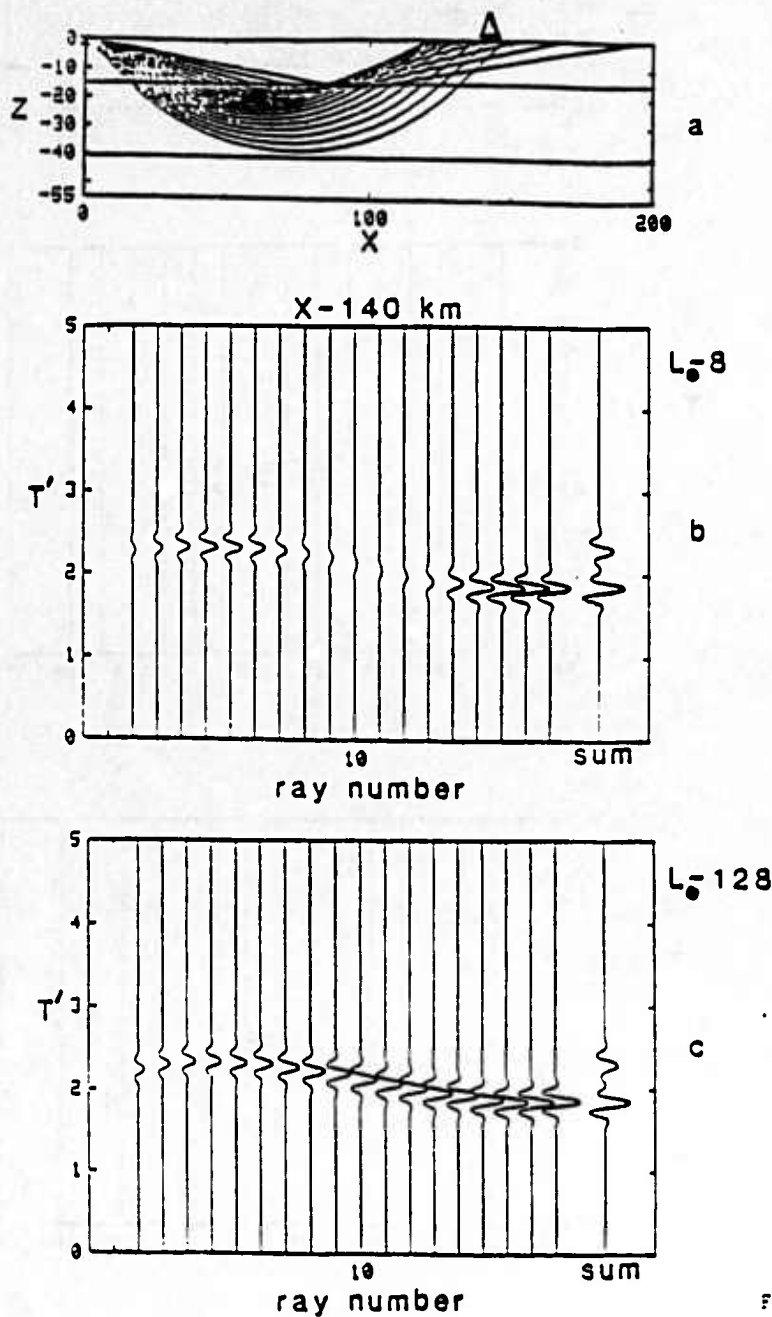


Figure 9

Figure 9 Layer over a gradient. a) ray diagram with station at $x=140$ km. b) individual beam contributions and resulting sum for $L_0=8 \text{ km}^{1/2}$. c) individual beam contributions and resulting sum for $L_0=128 \text{ km}^{1/2}$.

ray spacing of 2 degrees. The station is located by the triangle, and two geometric arrivals are seen to arrive at this station. Figures 9b-c shows the individual beam contributions for $L_0=8 \text{ km}^{1/2}$ and $L_0=128 \text{ km}^{1/2}$. The 'ray number' represents the ray index parameter. The rays are here simply indexed from 1 to 18 from a shallow initial takeoff angle to a steep initial takeoff angle from horizontal. The two geometric arrivals are seen in the scaled sum to the right in Figures 9b-c.

For $L_0=8 \text{ km}^{1/2}$, near the average critical value, the energy is concentrated around the two direct beams to the station with ray numbers 4 and 17. The individual beam wavelets in Figure 9b map out a curve which has stationary points corresponding to the direct beams to the station. For $L_0=128 \text{ km}^{1/2}$, the energy is more evenly distributed among the beams as seen in Figure 9c. The beam solutions away from the stationary points now phase interfere giving a similar resulting sum as for the $L_0=8 \text{ km}^{1/2}$ case.

For very large L_0 , each beam solution is effectively an initial plane wave contribution. For a large L_0 and a one dimensional medium, the individual beam contributions in phase time and ray number are similar to plane wave decompositions in phase time, $\Theta = \tau + px$, and horizontal ray parameter [see Chapman, 1978].

Figure 10 shows the results for a station at $x=120 \text{ km}$ where the caustic intersects the surface. Here the two geometric arrivals merge as

LAYER OVER GRADIENT

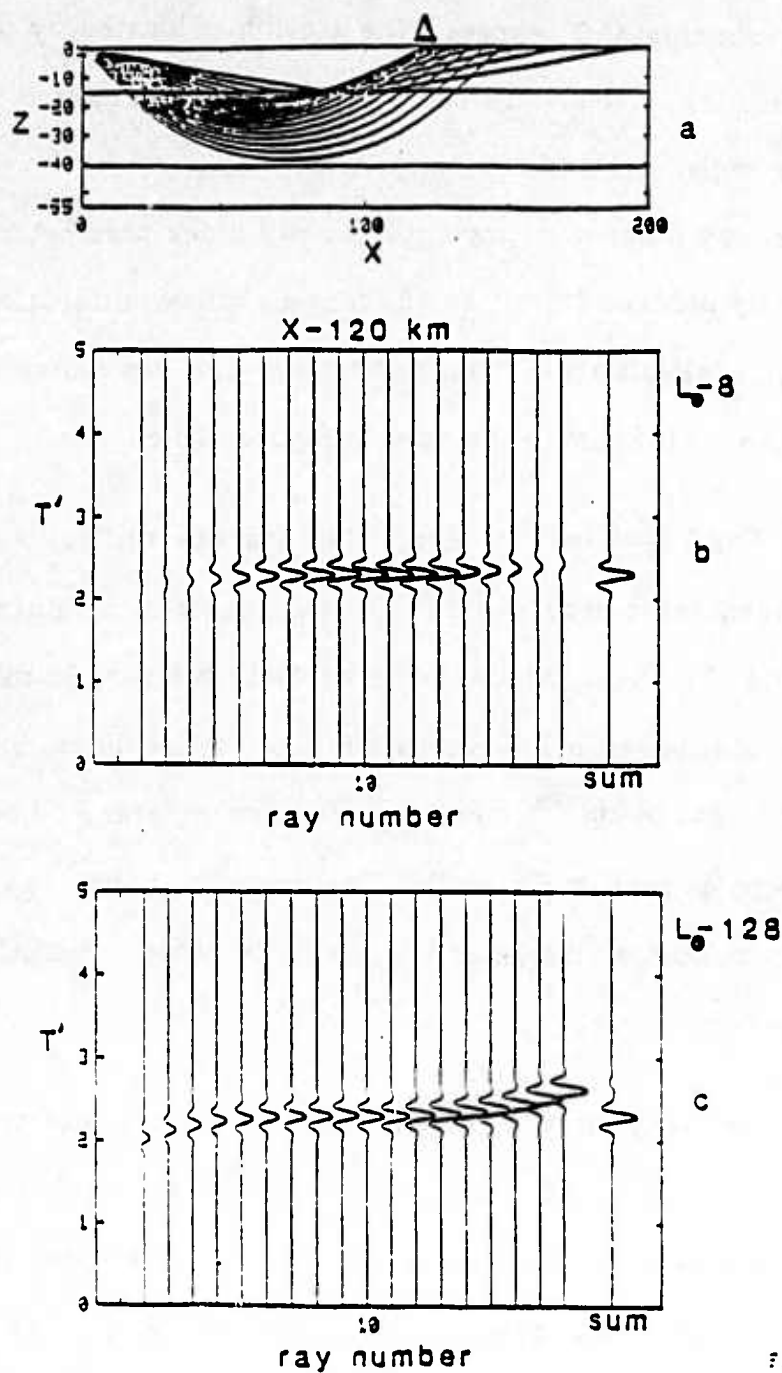


Figure 10

Figure 10 Layer over gradient. a) ray diagram with station at $x=120 \text{ km}$. b) individual beam contributions and resulting sum for $L_0=8 \text{ km}^{1/2}$. c) individual beam contributions and resulting sum for $L_0=128 \text{ km}^{1/2}$.

one. Figure 10b shows the individual beam contributions for $L_0 = 8 \text{ km}^{1/2}$. Again the 'ray number' represents the ray index parameter and indexes the rays from 1 to 18. There is now only one stationary point in the curve represented by the locus of individual beam wavelets, located at ray number 10. The result for $L_0 = 128 \text{ km}^{1/2}$ is shown in Figure 10c. The sum shows a spurious end effect caused by the lack of cancellation of the end beams. The $L_0 = 8 \text{ km}^{1/2}$ case avoids this end effect by amplitude decay of beams away from the station.

In Figure 11, the beam contributions are shown for a station at $x = 110 \text{ km}$ in the caustic shadow. Here no geometric arrival occurs and the individual beam wavelets shown in Figures 11b-c have no stationary points. The $L_0 = 8 \text{ km}^{1/2}$ case shows a clean resulting sum, whereas the $L_0 = 128 \text{ km}^{1/2}$ case shows spurious end effects which are the same order as the actual signal.

Example F

In this example, the effect of caustics on $q(s)$ is investigated. The caustics result from multiple bounces of rays in a surface waveguide. The velocity model consists of a 0.1 km homogeneous surface layer having a velocity of 1.8 km/sec over a linear velocity gradient from 1.8 km/sec to 5.6 km/sec at 50 km. The ray diagram for a source in the thin surface layer is shown in Figure 12a. The complex function $q^*(s)$, is shown in

LAYER OVER GRADIENT

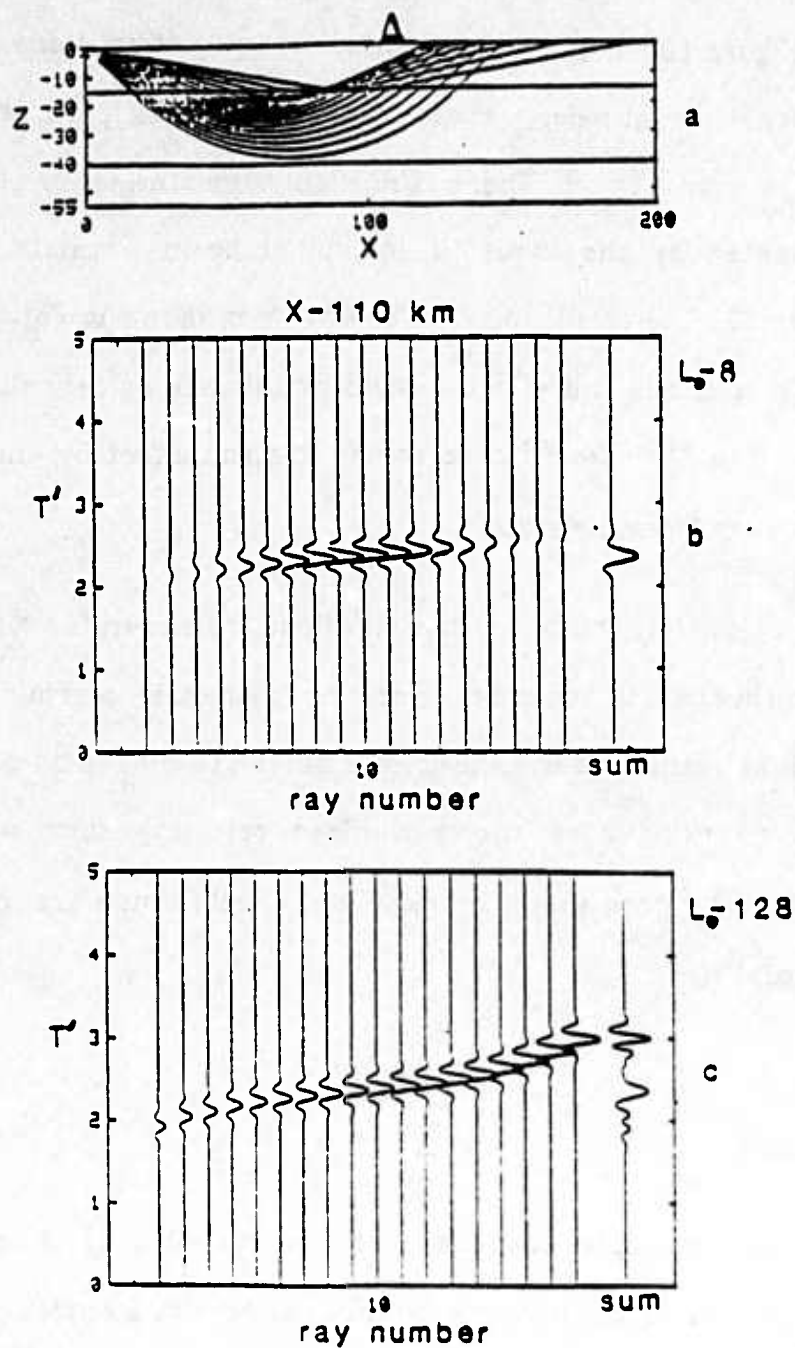


Figure 11

Figure 11 Layer over a gradient a) ray diagram with station at $x=110$ km b) individual beam contributions and resulting sum for $L_0=8 \text{ km}^{1/2}$ c) individual beam contributions and resulting sum for $L_0=128 \text{ km}^{1/2}$.

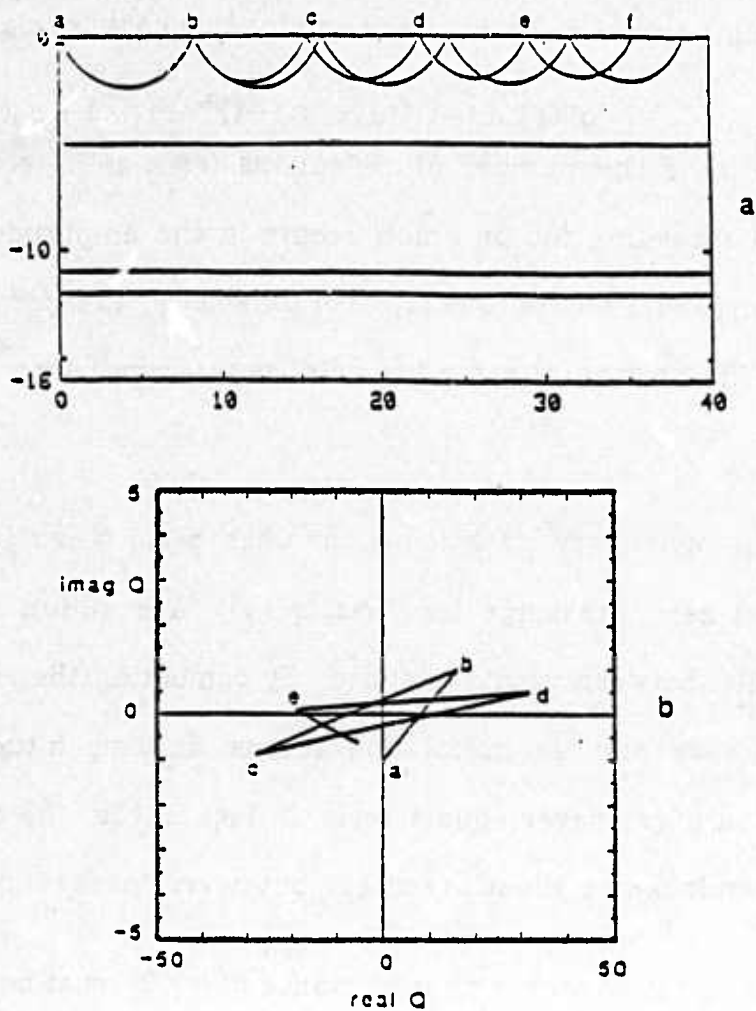


Figure 12

Figure 12 rays in a surface waveguide a) ray diagram b) graph of the complex function $q'(s)$.

Figure 12b as a function of distance along the ray. $q^r(s)$ is a relative spreading across each interface and for this case is related to $q(s)$ by

$$q^r(s) = (-1)^{\alpha_r} q(s) = (-1)^{\alpha_r} [\epsilon q_1(s) + q_2(s)]$$

where α_r is the number of reflections from an interface. $q^r(s)$ is the actual spreading factor which occurs in the amplitude of the beam [see Cerveny and Psencik, 1983c]. For simplicity, the parameter ϵ is set to $-i$. The horizontal scale for $q^r(s)$ is ten times larger than the vertical scale.

For ray theory, caustics occur when $q_2(s) = \text{Real}(q(s))$ goes to zero. Several zero crossings for $\text{Real}(q^r(s))$ are shown in Figure 12b, for example, between points b and c. By combining the results for an initial plane wave and an initial line source, the magnitude of the complex function $q^r(s)$ never equals zero. In Figure 12b, the plot of $q^r(s)$ cycles counterclockwise about the origin, but never passes through it.

Using ray theory, a phase advance of $\pi/2$ must be introduced at each caustic. This is done using the KMAH index. Using a complex $q^r(s)$, the phase of $q^r(s)$ automatically changes by π between zero crossings of $\text{real}(q^r(s))$. Thus, as long as the phase of $q^r(s)$ is kept continuous, taking a square root results in the proper number of $\pi/2$ phase shifts.

Example G

In this example the effect of the scale of a velocity inhomogeneity on the Gaussian beam solution is investigated. A single heterogeneity is considered with an incident plane wave impinging from below. The background velocity is 6 km/sec and the wavelet parameters are, $f = 1$ Hz, $\gamma = 4$, $\varphi_0 = 0$. The velocity heterogeneity is circular and the velocity varies smoothly in a Gaussian manner. The radius of the heterogeneity ranges from 15 to 25 km with a velocity contrast of 10% lower than the background. The Gaussian beam solution is then compared with the finite difference solution of the 45 degree parabolic equation typically used in seismic exploration and forward scattering simulations [see Claerbout, 1976]. The ray diagram for a 25 km radius heterogeneity is shown in Figure 13.

Figure 14 shows the seismograms for an inhomogeneity with a radius of 25 km, a velocity contrast of 10%, a distance from the initial wavefront to the heterogeneity of 50 km, and a distance from the initial wavefront to the stations of 120 km. The finite difference solution of the parabolic equation, Parab45, is shown in Figure 14a. Figures 14b-d show the Gaussian beam solutions for three initial beam-widths. The critical initial beam-width for the station distance in a homogeneous medium is $L_{0cr} = 11 \text{ km}^{1/2}$. From the results of Example A, the decomposition of an initial plane wave in a homogeneous medium requires initial beam-widths comparable or smaller than the critical initial beam-width for the receiver distance. The maximum amplitude for each solution is shown in Figure 14

SINGLE INHOMOGENEITY

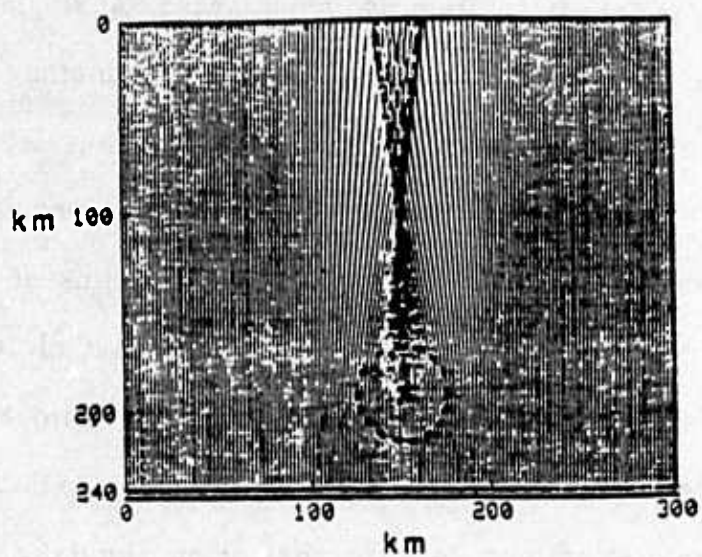


Figure 13

Figure 13 Ray diagram for a single heterogeneity with a radius of 25 km and a velocity contrast of 10%.

SINGLE INHOMOGENEITY

a-25 km A1
v 6-6 km/sec

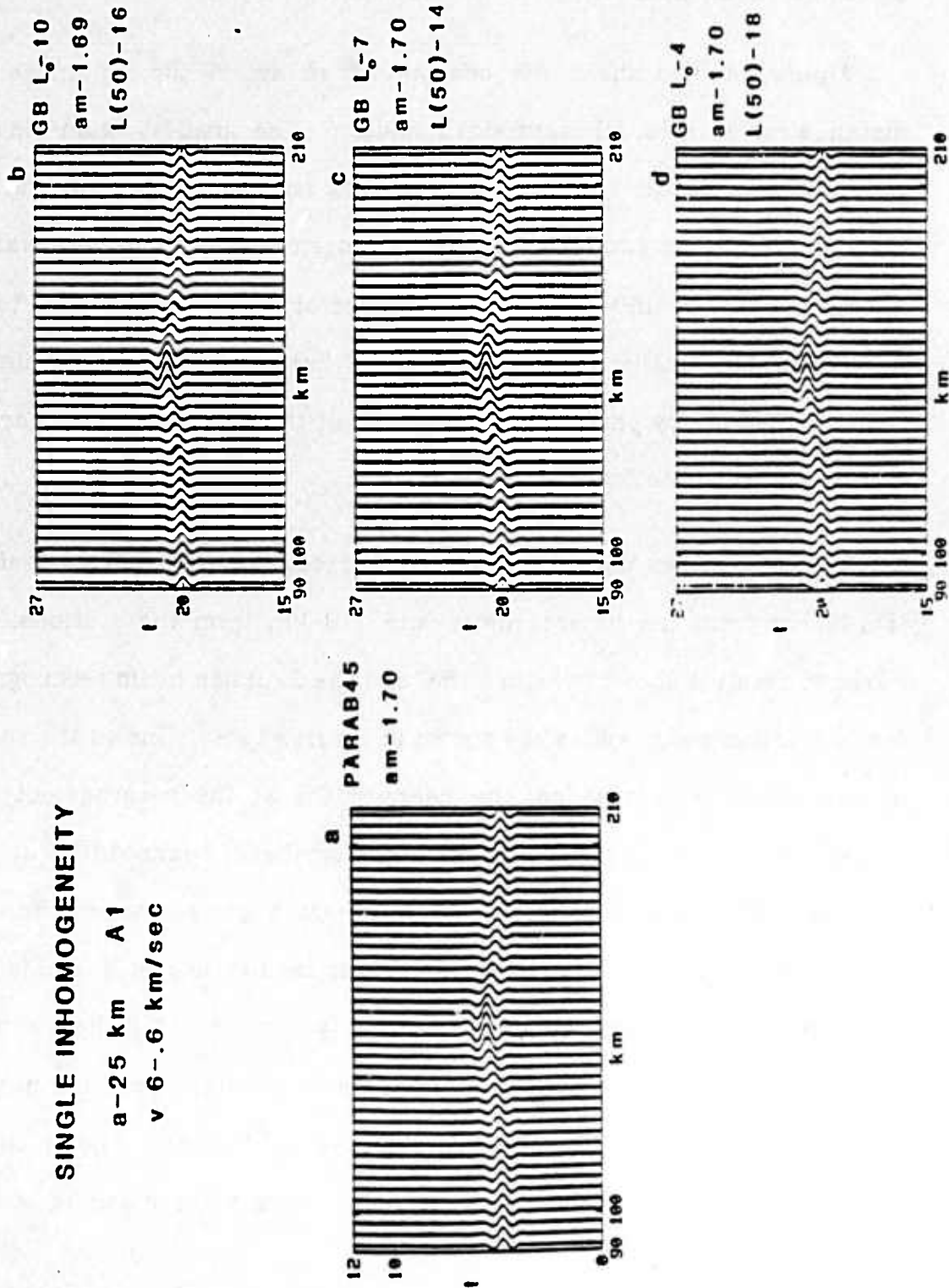


Figure 14 Single velocity heterogeneity with a radius of 25 km, a velocity contrast of 10%, the distance from the initial wavefront to the heterogeneity of 50 km, and the distance from the initial wavefront to the stations of 140 km. a) Parab45 seismograms Gaussian beam seismograms for b) $L_0=10 \text{ km}^{1/2}$, c) $L_0=7 \text{ km}^{1/2}$, and d) $L_0=4 \text{ km}^{1/2}$.

The Gaussian beam results for the different initial beam-widths are seen to be stable and compare with the Parab45 result.

Figure 14 also shows the beam-width in km at the inhomogeneity distance for each initial beam-width solution. The smallest beam-width at the heterogeneity distance is 14 km resulting from an initial beam width of $L_0 = 7 \text{ km}^{1/2}$, and is smaller than the heterogeneity radius. From example A, this smallest beam-width defines the zone of contributing beams to the final sum. For smaller than critical initial beam-widths a similar limited region is defined by phase interference about the stationary direct arrival and is related to the first Fresnel zone.

Figure 15 shows the same experiment, except now the initial wavefront is 140 km from the heterogeneity and 210 km from the stations. The Parab45 result is shown in Figure 15a, and the Gaussian beam seismograms for four initial beam-widths are shown in Figures 15b-e. Due to the spread of the beams with distance, the beam-widths at the heterogeneity are larger than the previous case. The smallest beam-width at the heterogeneity distance is 23 km and is just slightly smaller than the heterogeneity radius. The Gaussian beam results are still stable and compare with the Parab45 result. The $L_0 = 4 \text{ km}^{1/2}$ result has a much larger beam-width of 50 km at the heterogeneity distance but the number of beams contributing is now limited by phase interference. The amplitude for this case is fairly stable but a spurious secondary phase can be seen.

SINGLE INHOMOGENEITY

a-25 km B1
v 6-.6 km/sec

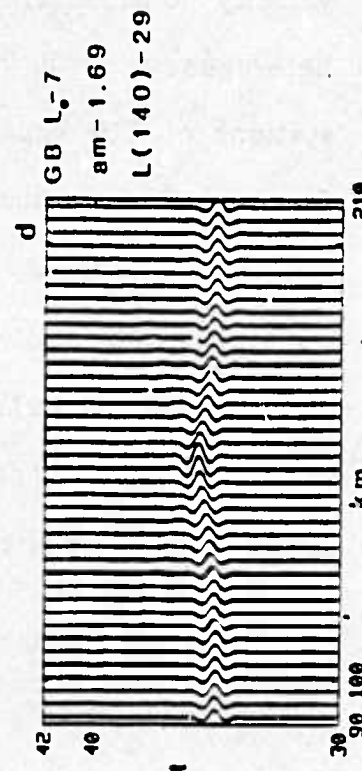
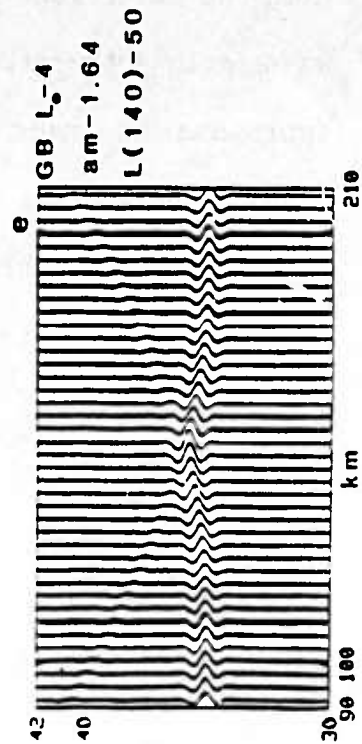
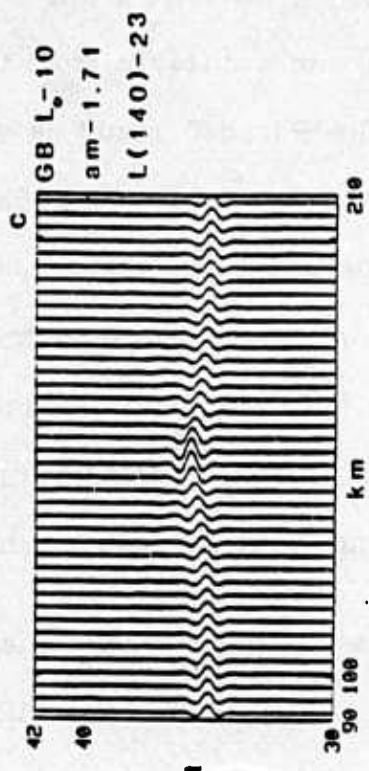
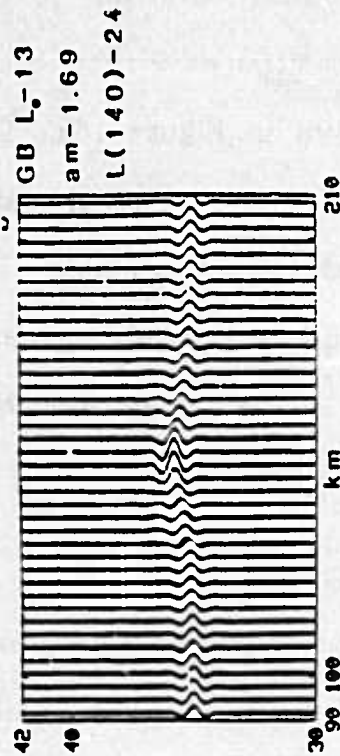
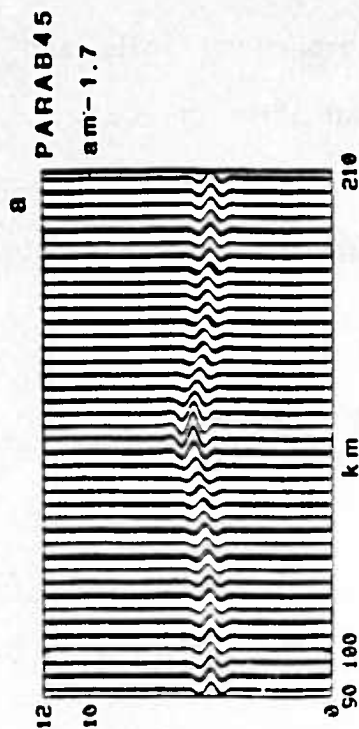


Figure 15 Single velocity heterogeneity with a radius of 25 km, a velocity contrast of 10%, the distance from the initial wavefront to the heterogeneity of 120 km, and the distance from the initial wavefront to the stations of 210 km. a) Parab45 seismograms Gaussian beam seismograms for b) $L_0=13 \text{ km}^{1/2}$, c) $L_0=10 \text{ km}^{1/2}$, d) $L_0=7 \text{ km}^{1/2}$, and e) $L_0=4 \text{ km}^{1/2}$.

In the final case, the inhomogeneity radius is reduced to 15 km with a velocity contrast of 10%, a distance from the initial wavefront to the heterogeneity of 30 km, and a distance from the initial wavefront to the stations of 210 km. The Parab45 result is shown in Figure 16a. The Gaussian beam results are shown in Figures 16b-e for several initial beam-widths. For this case the Gaussian beam amplitudes are less stable. The Gaussian beam solution with the smallest beam-width at the heterogeneity compares best with the Parab45 result. All the Gaussian beam solutions have a larger secondary arrival than the Parab45 solution. This may in part be due to dip filtering in the Parab45 formulation.

These preliminary results indicate that when using the Gaussian beam method in a heterogeneous medium, the critical beam-width be smaller than the heterogeneity scale. This is more restrictive than simply small wavelength with respect to heterogeneity scale, and is related to the approximation made in the final term in equations (5) or (7). This restriction is similar to that given for geometric optics by Kravtsov and Orlov [1980] in which the medium and wave parameters must be slowly varying over the first Fresnel zone.

Example H

In this example, SII waves in a 2-D basin structure are investigated. This problem has been studied by a number of investigators using

SINGLE INHOMOGENEITY

a-15 km A2
v-6.-.6 km/sec

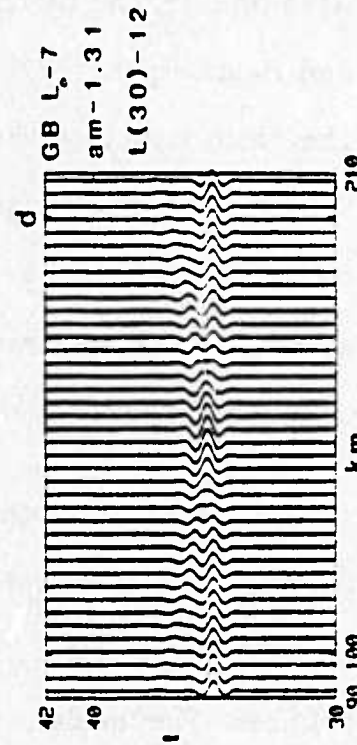
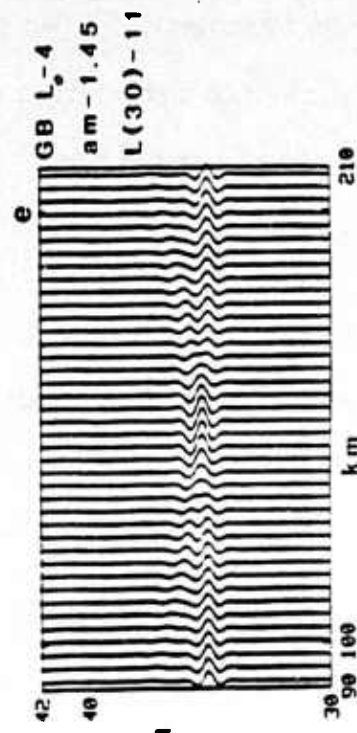
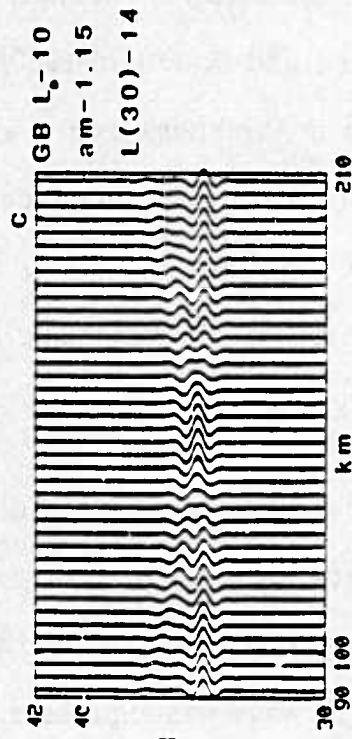
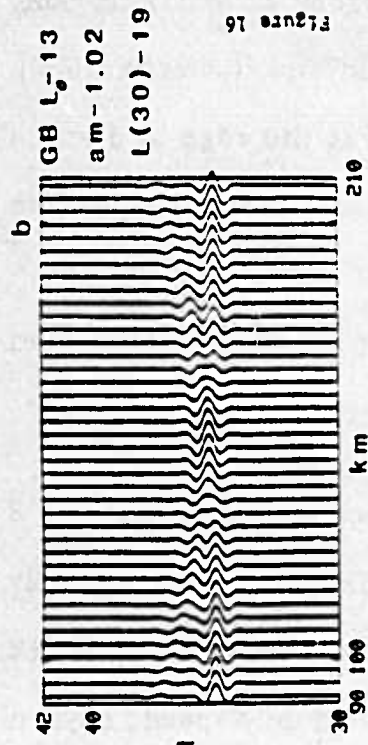
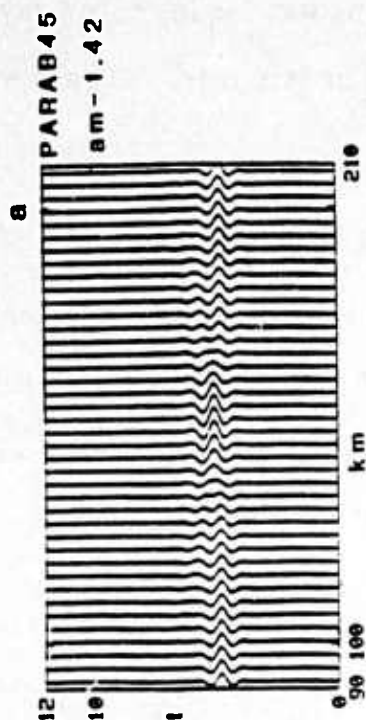


Figure 16 Single velocity heterogeneity with a radius of 15 km, a velocity contrast of 10%, the distance from the initial wavefront to the heterogeneity of 50 km, and the distance from the initial wavefront to the stations of 230 km. a) Parab45 seismograms Gaussian beam seismograms for b) $L_0=13 \text{ km}^{1/2}$, c) $L_0=10 \text{ km}^{1/2}$, d) $L_0=7 \text{ km}^{1/2}$, and e) $L_0=4 \text{ km}^{1/2}$

techniques including; finite difference, finite element, discrete wavenumber, and Glorified Optics(a ray method) [Boore et. al., 1971; Hong and Helmberger, 1977; Bard and Bouchon, 1980; Aki and Richards, 1980]. The basin structure varies in thickness from 1 km at the edge to 6 km in the center over a range of 25 km. The interface has a cosine shape. The material parameters in the basin are $v_1=7 \text{ km/sec}$ and $\rho_1=2.0 \text{ gm/cm}^3$ and the material parameters of the basement are $v_2=3.5 \text{ km/sec}$ and $\rho_2=3.3 \text{ gm/cm}^3$.

Figure 17 shows the ray diagrams for the direct wave and the first 2 bounces. The rays represent an upward traveling SH plane wave normally incident from below. The Gaussian beam synthesis was made up of 12 bounces. The incident plane wave was expressed by equally spaced rays in the basement. No two-point ray tracing was required; the rays were simply allowed to reverberate in the basin. For this case, 150 rays were used with a ray spacing of 1 km.

The damped cosine source wavelet used in the Gaussian beam synthesis (see eqn. 21) was adjusted to simulate an often used Ricker wavelet [see Bard and Bouchon, 1980]. In Figure 10, the solid line is the Ricker wavelet, and the dashed line is the damped cosine wavelet used in the Gaussian beam synthesis with $T=16.5 \text{ sec}$ and $\gamma=3.2$. The value of L_0 was set at $2.0 \text{ km}^{1/2}$.

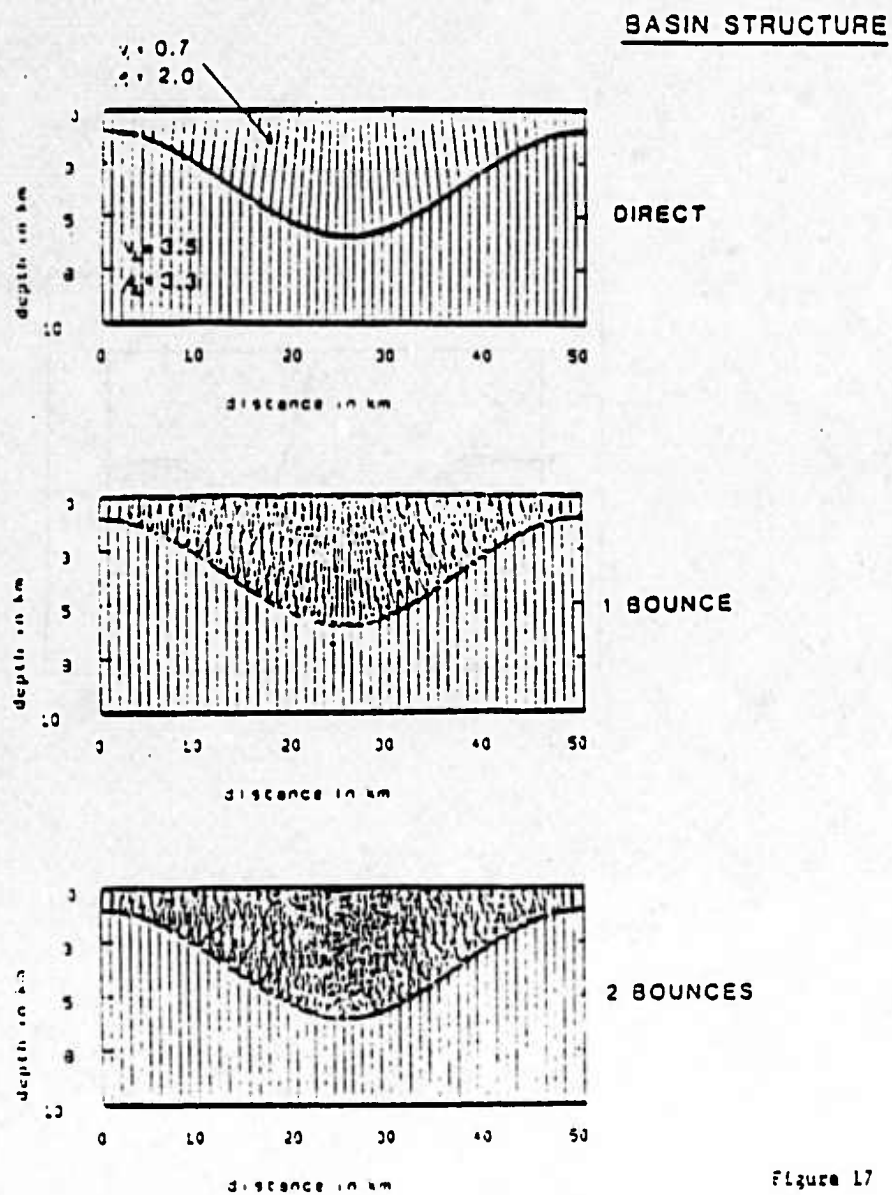


Figure 17

Figure 17 Ray diagrams for an SH plane wave incident on a 2-D basin structure. The direct wave and first 2 bounces are shown.

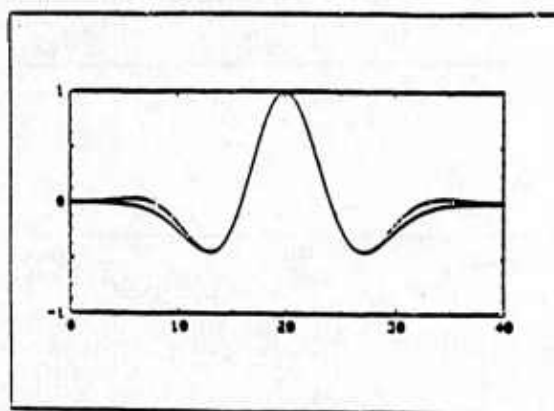


Figure 18

Figure 18 Comparison between a Ricker wavelet (solid line) used in other studies of the soft basin, and the damped cosine wavelet (dotted line) used for the Gaussian beam seismograms with $T=16.5$ sec and $\gamma=3.2$.

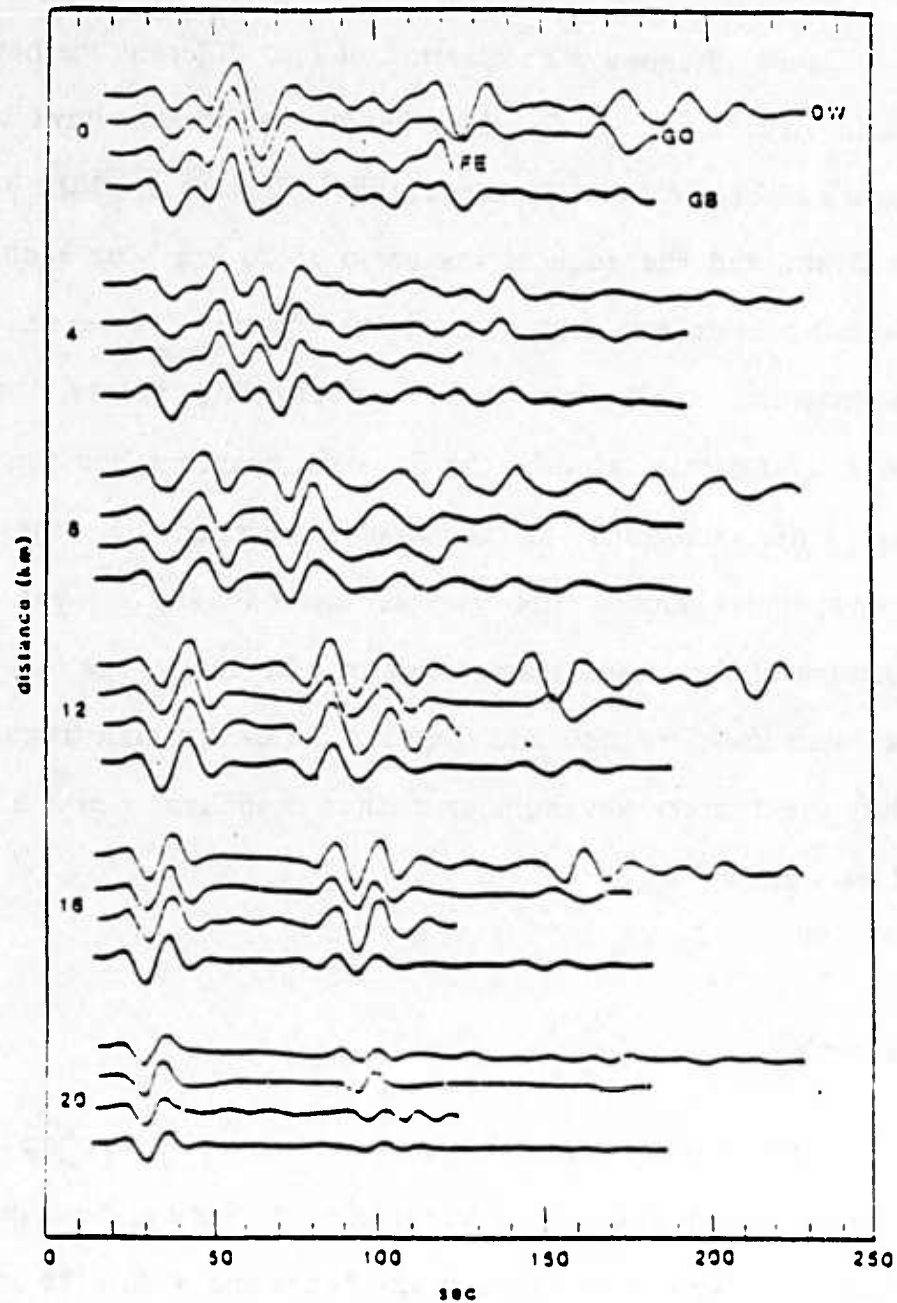


Figure 19

Figure 19 Comparison of seismograms computed by various methods for a soft basin: DW - discrete wavenumber, GO - Glorified Optics (a ray method), FE - finite element, GB - Gaussian beam. The center of the basin is marked as 0 km, and the edge as 20 km.

Figure 19 shows a comparison of four different methods for the 2-D basin structure. The Gaussian beam seismograms have been added to Figure 13.26 in Aki and Richards [1980]. The center of the basin is marked as 0 km, and the edge of the basin as 20 km. For each distance, the Gaussian beam seismogram is at the bottom. The second phase on the seismograms represents energy propagating across the basin. The seismograms calculated by the Gaussian beam method compare well with the other solutions at the earlier portions of the seismograms. Discrepancies among the various methods are evident in the latter portions of the seismograms. This may be due to the fact that both the Gaussian beam method and Glorified Optics are high-frequency methods, while the discrete wavenumber method is applicable over a broader range of frequencies.

Example 1

In this example, the Gaussian beam method is applied to the study of volcanic earthquakes under Mt. St. Helens. During the summer of 1981, a 9-station array was deployed on the flanks and within the crater of the Mt. St. Helens volcano. This was a joint effort undertaken by Oregon State University, Massachusetts Institute of Technology, and the U.S. Geological Survey [Fehler and Chouet, 1982]. Two of the more significant observations of the recorded high-frequency volcanic events were: 1) the

arrival times at the crater station were between 0.4 and 0.7 seconds earlier than those at the flank stations, and 2) the amplitudes at the crater station were about 10 times greater than those recorded at the flank stations.

Figure 20 shows the 2-D model of the velocity structure which was tentatively adopted for Mt. St. Helens. This model is based on evidence obtained by various workers on Mt. St. Helens [Malone, 1982] as well as on other similar types of volcanoes such as St. Augustine in Alaska [Kienle et al., 1979] and Showa-shinzan in Japan [Hayakawa, 1957]. In all these cases, the internal structure of a volcano is depicted as a high-velocity central body with a P velocity of about 4 km/sec covered by lower-velocity surficial layers.

It was first thought that the observed high amplitude and early arrival time could be explained by locating the earthquake foci at the bottom of the hypothesized high-velocity column. Figure 20 shows the ray diagram for a focus located 3.5 km below the summit. The high-velocity column tends to cause energy to diverge from the summit. Vertical component Gaussian beam seismograms for stations located across the volcano from 2.5 km to 7.5 km along the surface are shown in Figure 21 using frequencies of 5, 10, and 20 Hz. The amplitudes at the flank stations appear to be comparable to those at the crater, and the differences in arrival time are within 0.25 seconds.

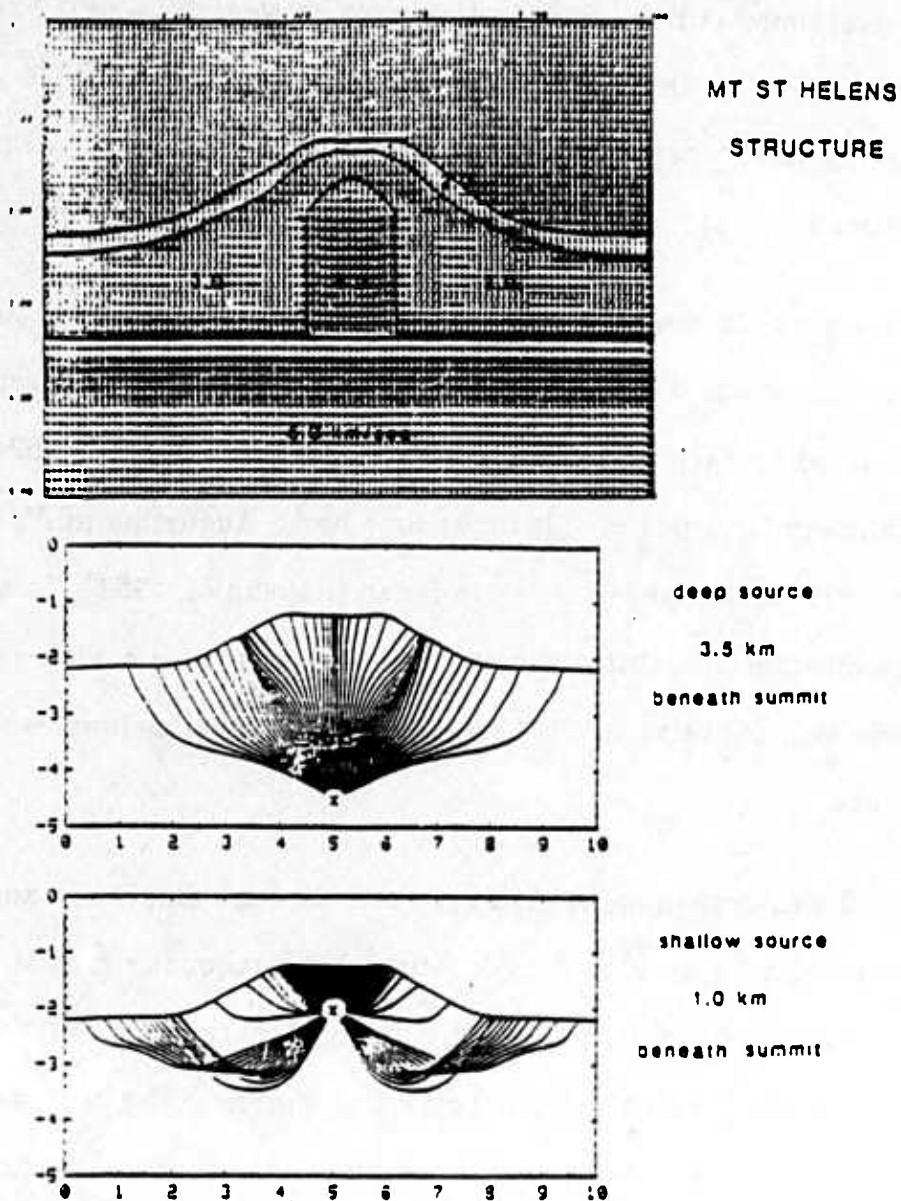


Figure 20

Figure 20 Velocity model used to depict the structure under Mt. St. Helens. Ray diagrams are shown below for a source located 3.5 km beneath the summit and a source located 1.0 km below the summit.

If the focal depth is moved upward to 1 km below the summit, the seismic signal changes dramatically. The ray diagram for a focus 1 km below the summit is shown in Figure 20. In this case, the ray paths going directly upward are dense and short, while those toward the flank of the volcano are sparse and take a detour. The resultant seismograms for 5, 10, and 20 Hz are shown in Figure 22. These show an amplitude at the crater station an order of magnitude greater than at the flank stations, and about 0.5 sec earlier arrival, in agreement with observation.

In order to check the accuracy of the computed seismograms, reciprocal seismograms were compared in Figure 23. In case A, a 2-D vertical point force is applied at a point 3.5 km beneath the summit, and the receiver is on the flank of the volcano. In case B, the vertical point force is on the flank of the volcano and the receiver is located 3.5 km beneath the summit. The free surface is neglected here. The vertical component seismograms are shown, for both cases, for frequencies of 5, 10, and 20 Hz. There is a satisfactory agreement in waveform for all three frequencies. The amplitudes agree to within 1% for 20 Hz, 6% for 10 Hz, and 12% for 5 Hz, showing better accuracy for higher frequencies, as expected.

Thus, the P-velocity structure depicted in Figure 20 and focal depths less than 1 km from the summit can explain the two significant observations. A similar observation was made for St. Augustine volcano by

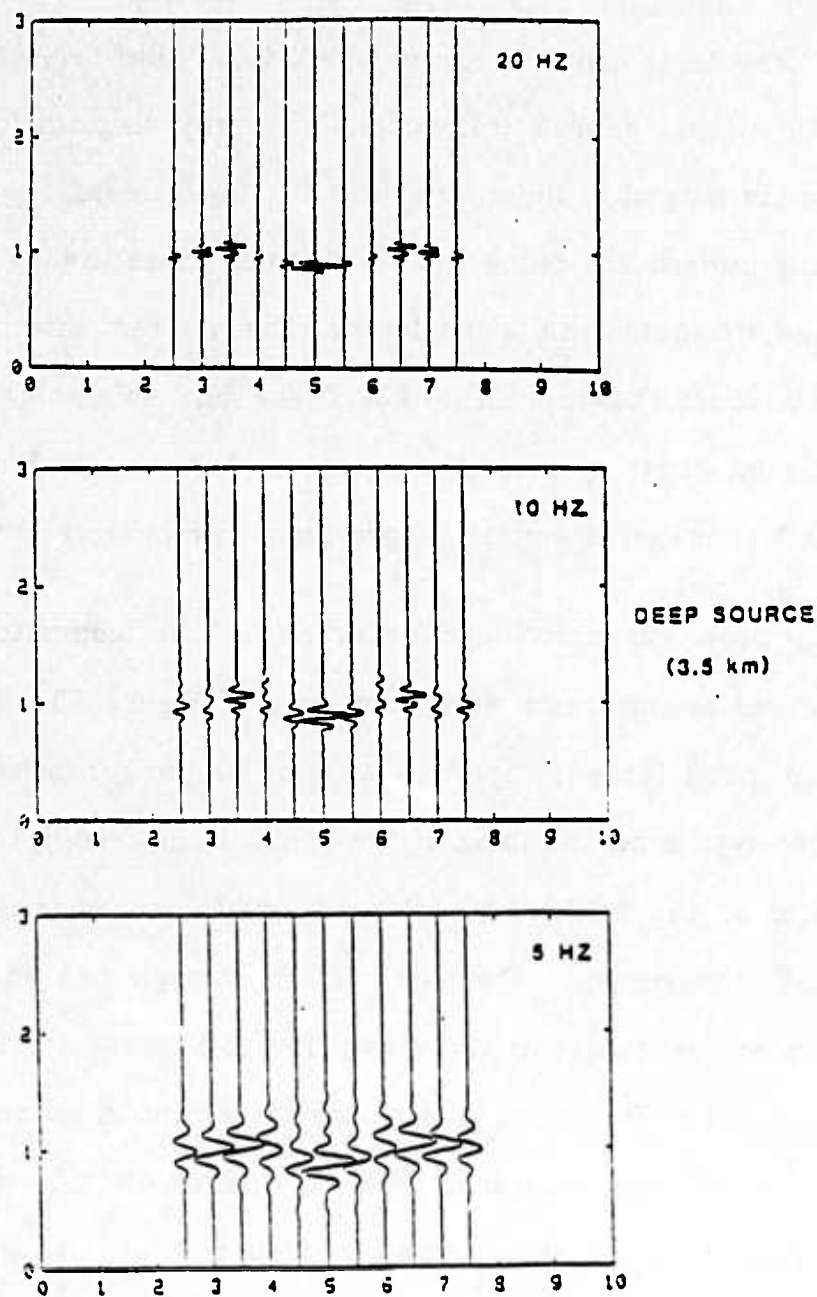


Figure 21

Figure 21 Gaussian beam seismograms for a source located 3.5 km below the summit using three frequencies; 5, 10, and 20 hz.

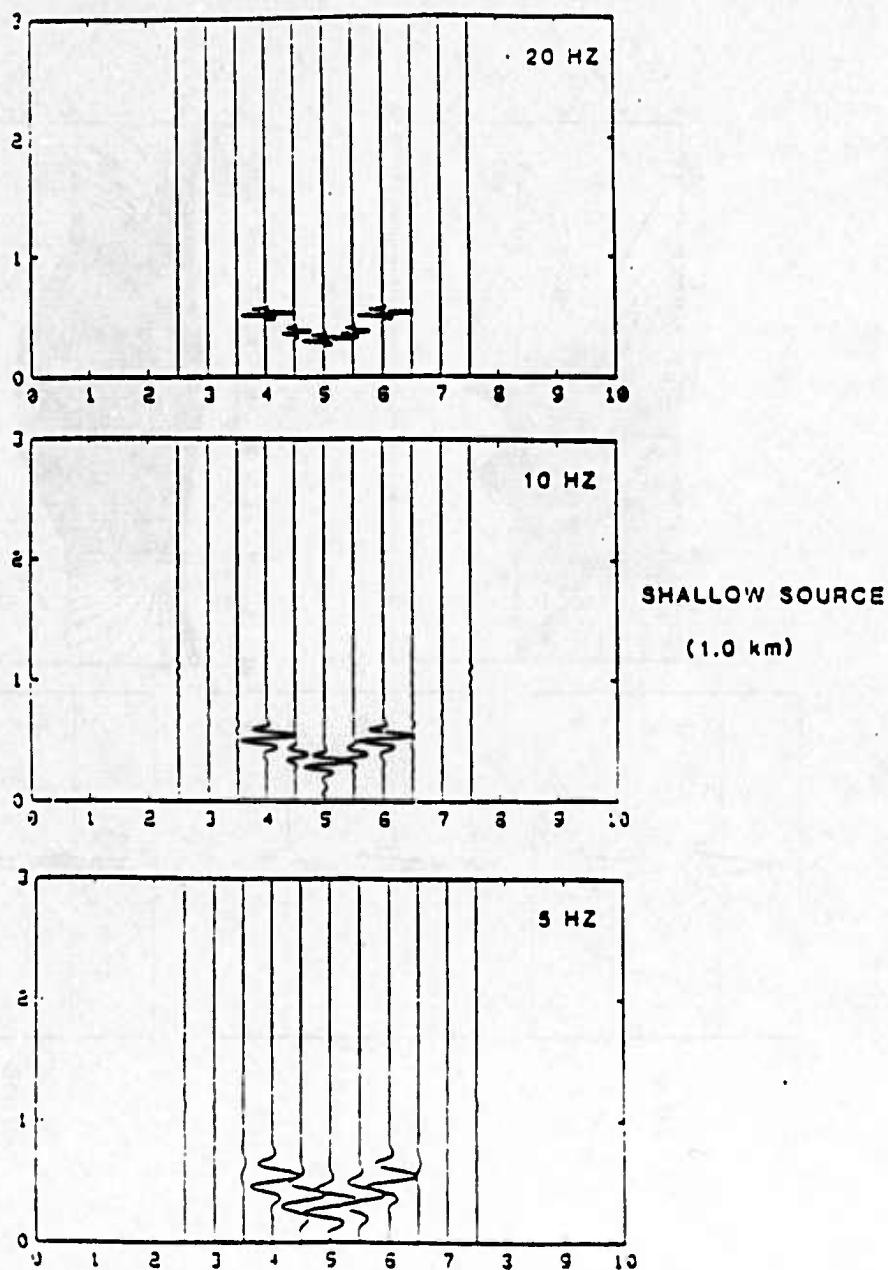


Figure 22

Figure 22 Gaussian beam seismograms for a source located 1.0 km below the summit using three frequencies; 5, 10, and 20 hz.

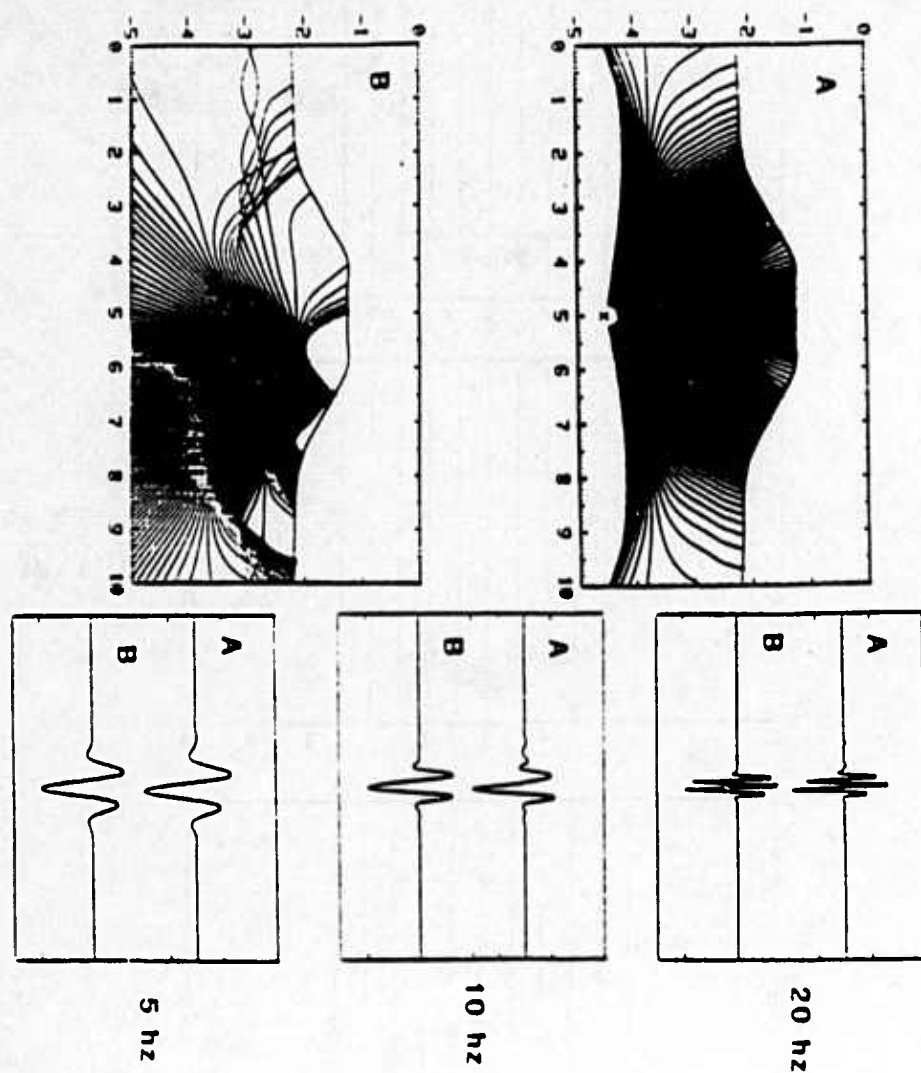


Figure 23

Figure 23 Ray diagrams and seismograms shown for reciprocal cases, A and B. The reciprocal seismograms are shown for three frequencies; 5, 10, and 20 Hz.

Lalla and Kienle [1982], who found a correlation between the amplitude ratio of summit to flank stations and the focal depth. For focal depths shallower than about 2 km, the summit amplitude was greater by an order of magnitude than the flank amplitude, and for deeper events, the opposite was observed.

In any case, the sensitivity of the seismic signal to the focal depth is strong, due to the heterogeneous structure under a volcano. This result also suggests a promising way of accurately locating volcanic earthquakes using variations in amplitude as well as in arrival time.

Example J

In this final example, the Gaussian beam method is used to study the influence of small-scale fluctuations on teleseismic P-waveforms. The models consist of several realizations of a lithosphere with randomly fluctuating velocities. The layer thickness is 120 km with a correlation length of 15 km. The average velocity is 8 km/sec and the root mean square fluctuation is 3%, roughly representing the results obtained for LASA [Aki, 1973; Capon, 1974].

Normally distributed random velocities were used to construct a 2-D velocity grid with a 15-km spacing between mesh points. This was then interpolated using bivariate splines to give continuous velocities as well as

first and second derivatives. The first realization for the randomly fluctuating velocities is shown in Figure 24. Above is shown a contour plot of the random velocities contoured at 0.1 km/sec interval. Below is shown a printer plot of the random velocities where each integer on the plot represents a 0.3 km/sec velocity interval and the integer "4" represents the velocity interval 7.7 to 8.0 km/sec.

Figure 25 shows, the ray diagram for a vertically incident plane wave perturbed by the random velocity layer. The receivers are located just above the random layer at $z=75$ km. The range of receivers is from 125 to 275 km, spaced at 5 km intervals. Figure 26 shows the resulting wavefield for two dominant frequencies, 2 Hz and 1 Hz. The amplitude and phase have both been significantly distorted by the random velocity layer. A multipath effect can also be seen in the range from 130 to 160 km. No significant P coda with a duration beyond about 3 seconds has developed as a result of the random velocity layer.

In order to check the accuracy of the method for this example, seismograms from reciprocal geometries have been compared using a 2-D point force as a source. The first comparison, shown in Figure 27, is for a source and receiver 300 km above and below the random velocity layer. In this geometry, the rays for the reciprocal cases sample the same region of the random layer. There is a good agreement in waveshape between reciprocal cases. Also, the amplitudes agree within 5%. For this case, the

random velocity layer

8 km/sec $\pm$ 3%, a=15 km iseed=0

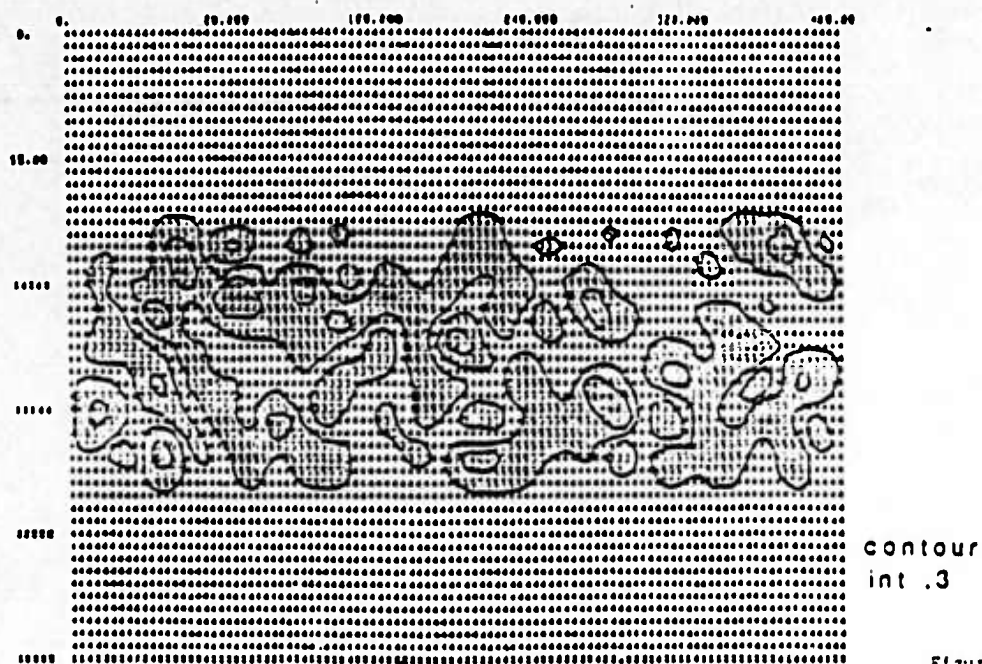
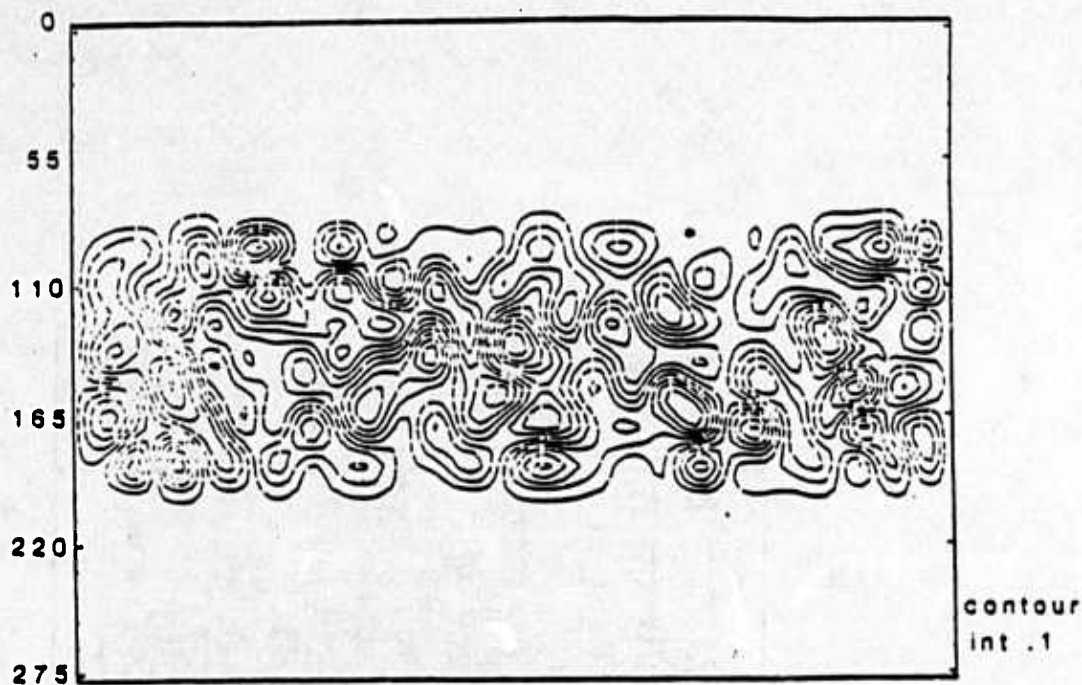


Figure 24

Figure 24 Contour plot and printer plot of a layer with randomly fluctuating velocities. The average velocity is 8 km/sec with a root mean square fluctuation of 3%. The correlation length is 15 km. The contour interval is .1 km/sec for the contour plot, and .3 km/sec for the printer plot. iseed=0.

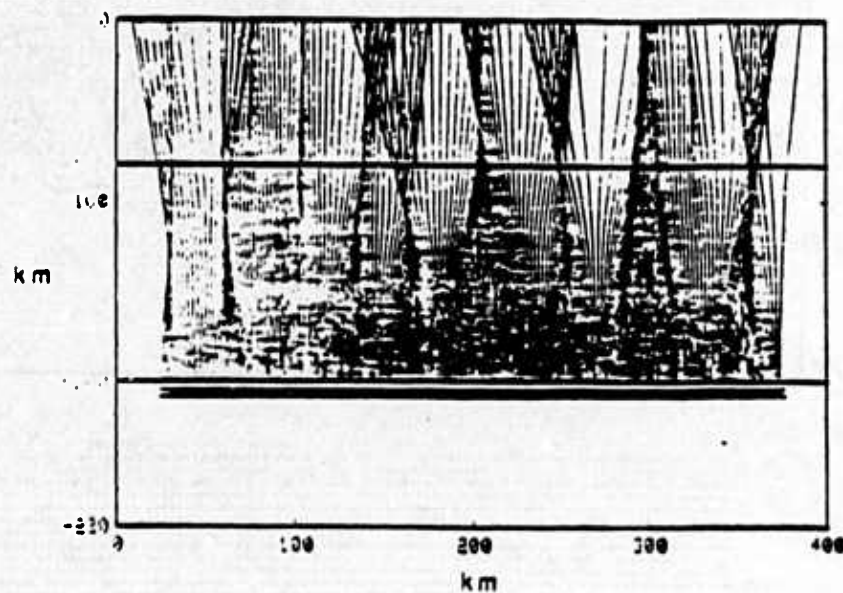


Figure 25

Figure 25 Ray diagram of vertically incident plane wave perturbed by the random velocity layer shown in Figure 24.

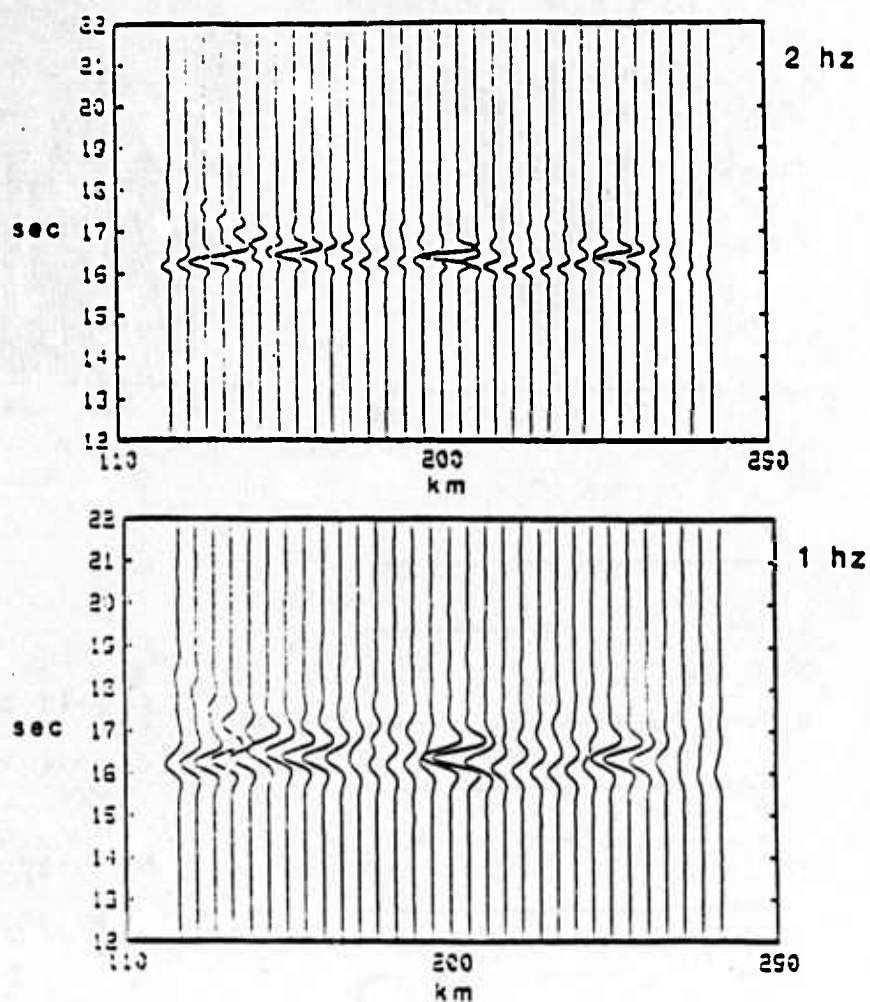


Figure 25

Figure 26 Wavefield of vertically incident plane wave propagated through random layer shown in Figure 24. Two frequencies are used, 1 and 2 hz.

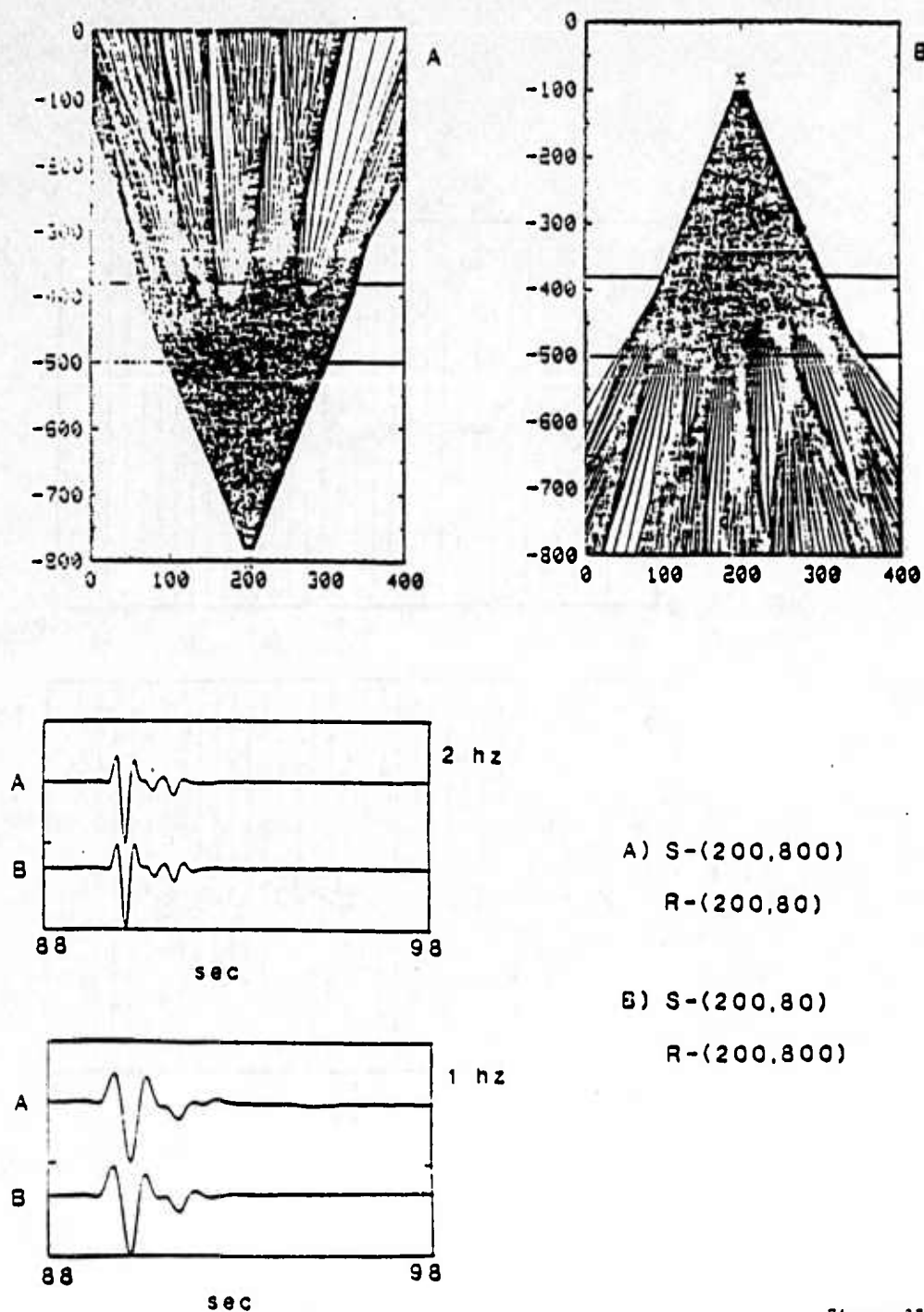


Figure 27

Figure 27 Ray diagrams and seismograms for the shown reciprocal cases A and B. For case A, the source is at (200,800) and the receiver is at (200,80). For case B, the source is at (200,80) and the receiver at (200,800). The random layer velocities are shown in Figure 24. The seismograms are computed for two frequencies, 1 and 2 Hz.

parameter, S_0 , was specified to put the minimum beam width (beam waist) just before the ray entered the random layer, and L_0 was specified as $12. \text{ km}^{1/2}$.

In addition to reciprocity checks, the calculation was repeated using the finite difference solution of the 45 degree equation [Claerbout, 1976]. Figure 28 shows a comparison between the parabolic result, Parab45, and the the Gaussian beam result for the 1 hz case. The results are similar in the significant details.

In Figure 29, a second realization for the random velocity layer is shown. The perturbed rays for an incident plane wave are shown in Figure 30. The wavefields observed on the surface of the random velocity layer for the two dominant frequencies of 2 Hz and 1 Hz are shown in Figure 31. This realization perturbs the wavefield less than the previous realization. The results have been checked using reciprocity and compared against the Parab45 solution and showed good agreement. For this realization, a comparison has been made between the ln-amplitude and the residual phase delay. This is shown in Figure 32. The positive correlation between $\ln(A)$ and the phase delay residual is in agreement with observations of teleseismic P waves at the Montana LASA by Aki [1973].

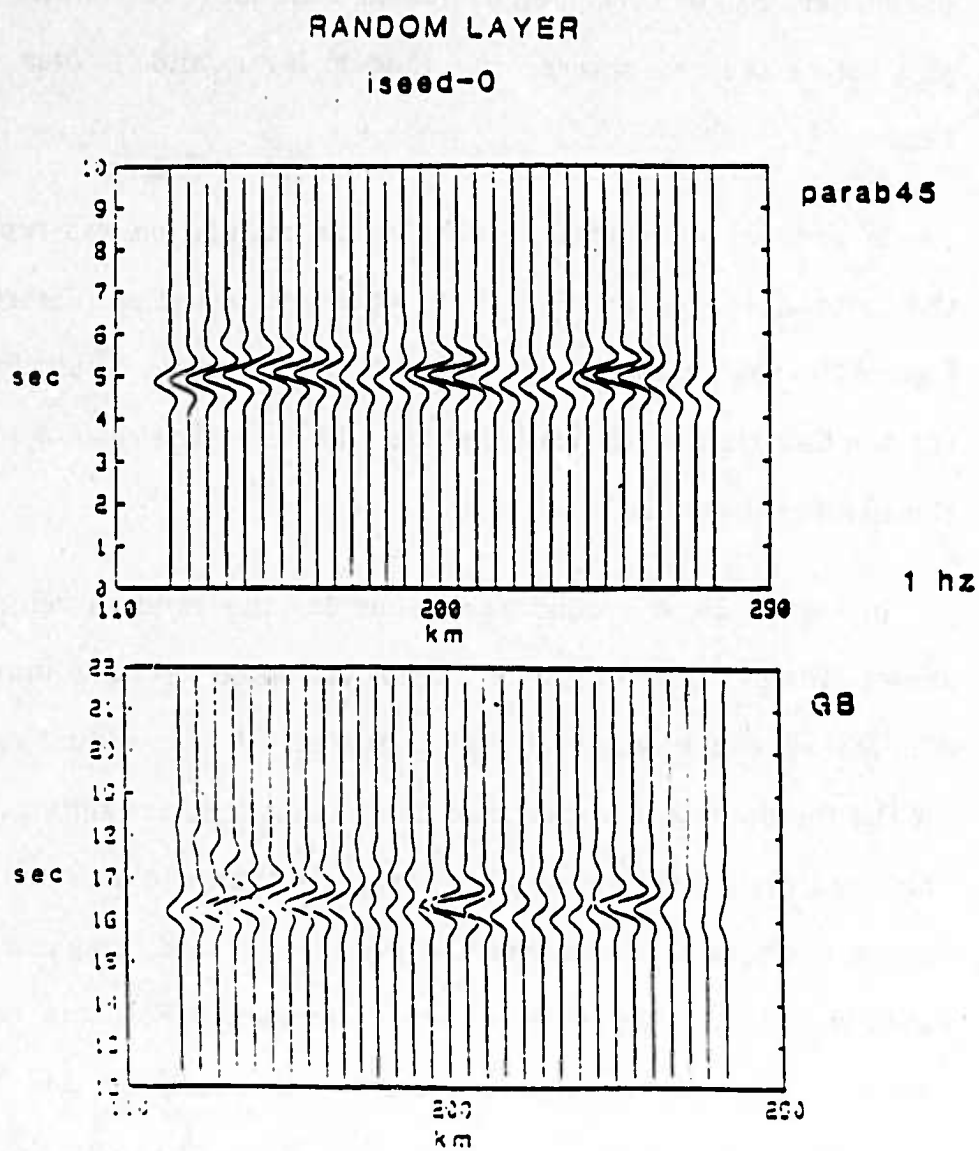


Figure 23

Figure 28 Comparison between the Parab45 solution and the Gaussian beam solution for the random layer shown in Figure 24.

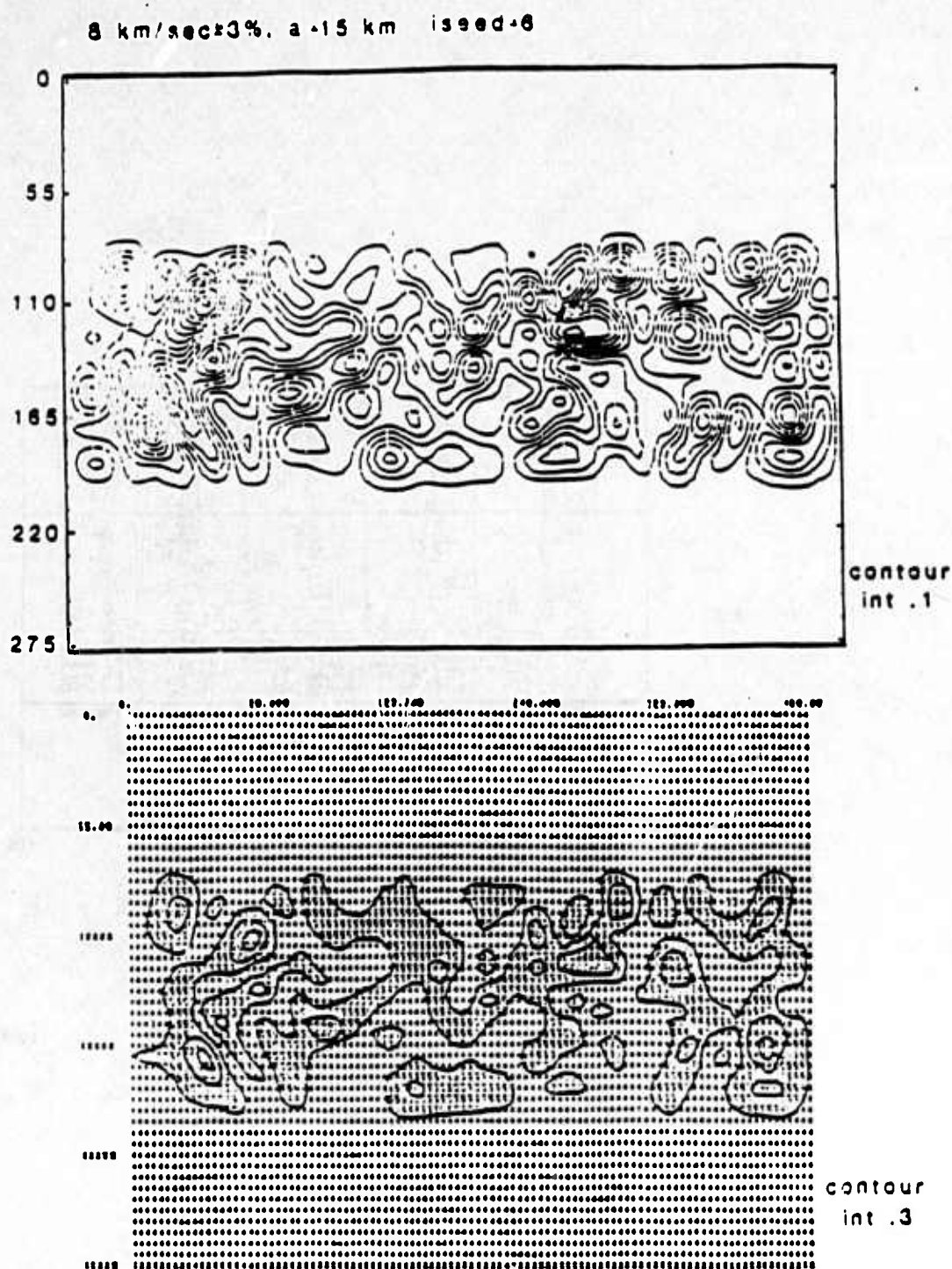


Figure 29

Figure 29 Contour plot and printer plot of a layer with randomly fluctuating velocities. The average velocity is 8 km/sec with a root mean square fluctuation of 3%. The correlation length is 15 km. The contour interval is 1 km/sec for the contour plot, and .3 km/sec for the printer plot. iseed = 6.

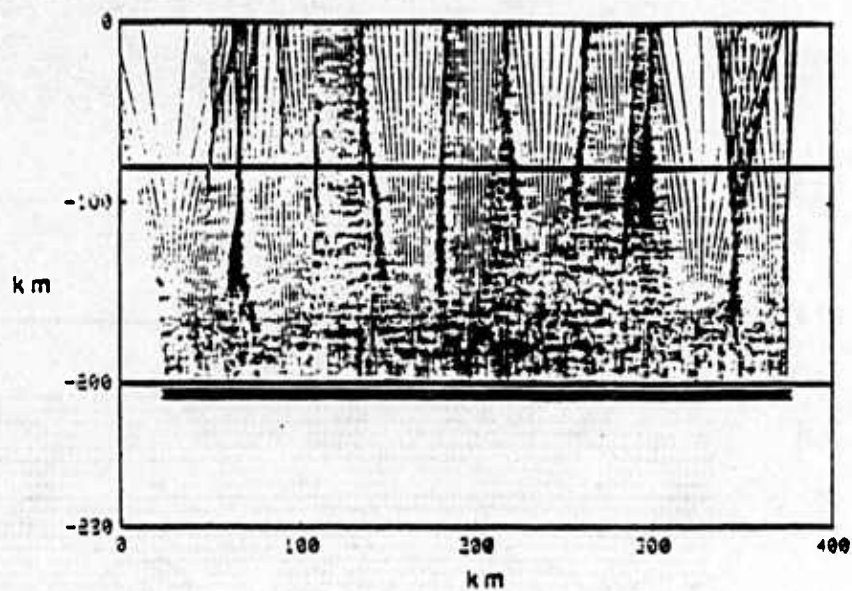


Figure 30

Figure 30 Ray diagram of vertically incident plane wave perturbed by the random velocity layer shown in Figure 29.

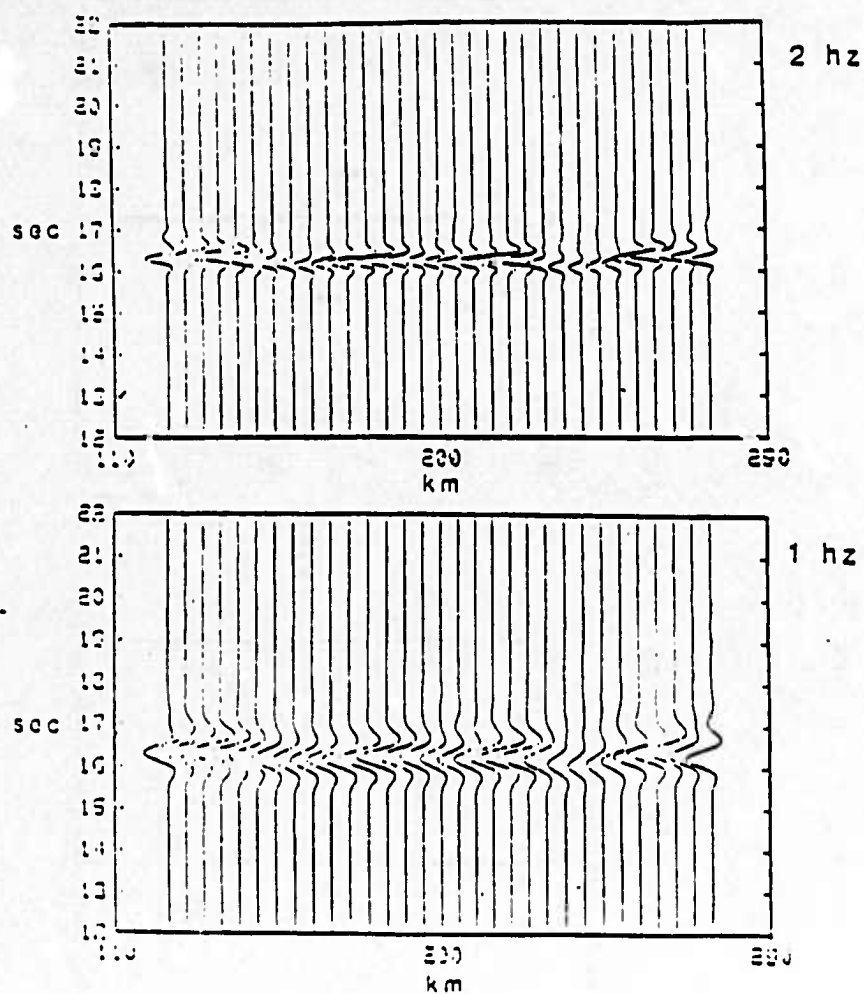


Figure 31

Figure 31 Wavefield of vertically incident plane wave propagated through random layer shown in Figure 29. Two frequencies are used, 1 and 2 Hz.

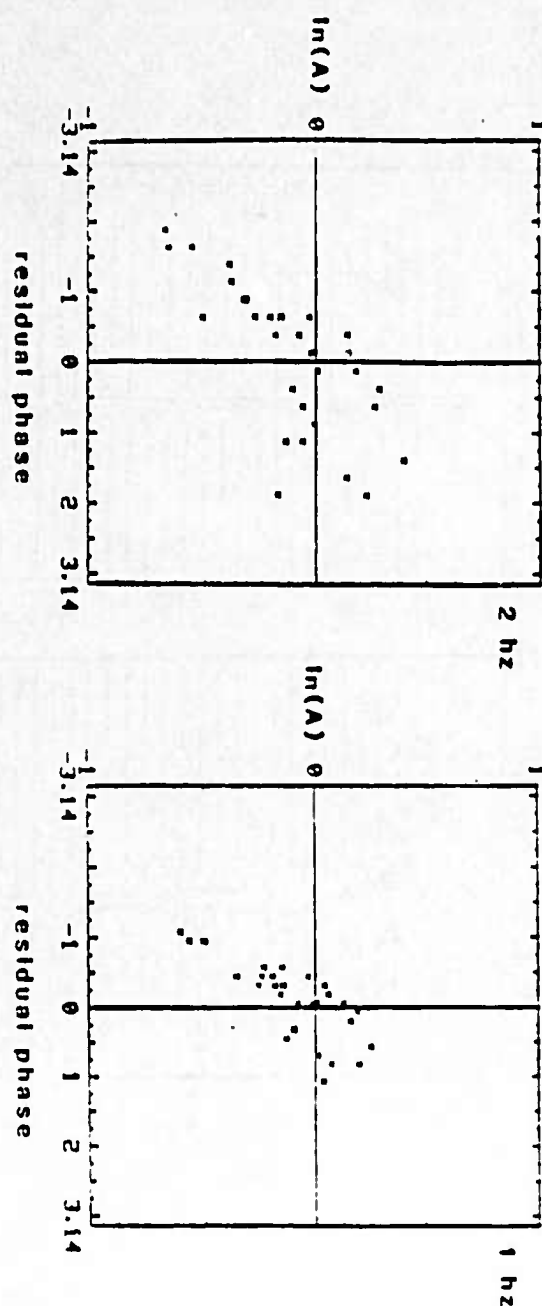


Figure 32

Figure 32 Plots of $\ln(\text{amplitude})$ versus residual phase delay for the wavefields shown in Figure 31.

2.2.4 CONCLUSIONS

There are several advantages in using the Gaussian beam method. First, the method is always finite at caustics. In addition, no prior knowledge of caustic locations is required as in other methods. Next, since the Gaussian beam method relies on local parabolic approximations, no global direction of propagation must be assumed as in the standard parabolic approximation. Finally, the Gaussian beam method is comparable in cost to ray methods, and possibly faster since no two-point ray tracing is done.

The Gaussian beam result depends on the choice of the initial beam-width. The expansion of a plane wave by Gaussian beams in a homogeneous medium requires initial beam-widths comparable or smaller than the critical initial beam width for a given station distance. This is similar to an expansion of an initial plane wave into point-like sources as viewed from the station. The approximate expansion of a line source into Gaussian beams in a homogeneous medium requires initial beam-widths comparable or larger than the critical initial beam-width. This is similar to an expansion of a line source into plane waves. The use of initial beam-widths near critical avoids end effects but doesn't simulate certain arrivals, such as head waves in the welded interface example.

The Gaussian beam method in a heterogeneous medium has been tested using two approaches. The first is the application of the reciprocal

theorem where the discrepancy between seismograms calculated for reciprocal cases is considered as a measure of error. The second is the application of the Gaussian beam method to cases, including the 2-D basin problem, which have solutions from other approximate methods. Preliminary results for a single heterogeneity using the finite difference solution to the parabolic equation and the Gaussian beam method gave best comparisons for critical beam-widths smaller than the heterogeneity scale.

The Gaussian beam method was then applied to the study of volcanic earthquakes under Mt. St. Helens. The observed differences in amplitude and arrival time between a station inside the crater and those at the flank can be explained by the combined effects of an anomalous velocity structure and a shallow focal depth. The method was also applied to study the influence of small-scale velocity fluctuations on teleseismic P waveforms. Significant amplitude and phase fluctuations due to a lithosphere with randomly fluctuating velocities were found, in agreement with observations by Aki [1973] at the Montana LASA.

2.2.5 APPENDIX 1

In this appendix, we give a derivation of the parabolic equation from the eikonal equation [see Gloge and Marcuse, 1989; and Marcuse, 1982]. By considering a variable endpoint in Fermat's integral, the eikonal or Hamilton-Jacobi equation can be obtained

$$|\nabla\tau|^2 = v^{-2} \quad (\text{A1.1})$$

The rays are the characteristics of this equation. In ray-centered coordinates, the eikonal equation can be written

$$\frac{\partial\tau}{\partial s} - \frac{h_1}{v} \left[1 - v^2 \left(\frac{\partial\tau}{\partial n} \right)^2 \right]^{1/2} = 0 \quad (\text{A1.2})$$

where h_1 is the scale factor in (2). The Hamiltonian is written as

$$H(n, p_n) = -\frac{h_1}{v} \left[1 - v^2 p_n^2 \right]^{1/2} \quad (\text{A1.3})$$

and $p_n = \frac{\partial\tau}{\partial n}$ is the slowness in the n direction and is identified as the generalized momentum. The Hamiltonian can be approximated to second order about the central ray [see Cerveny, 1981] as:

$$H(n, p_n) = -\frac{1}{v} + \frac{1}{2} v p_n^2 + \frac{1}{2} v^{-2} v_{,nn} n^2 \quad (\text{A1.4})$$

Following Marcuse [1982], a parabolic equation can be derived by treating all variables as operators. By analogy to quantum mechanics

$$n \rightarrow n, \quad p_n \rightarrow -i\kappa \frac{\partial}{\partial n}, \quad H \rightarrow i\kappa \frac{\partial}{\partial s} \quad (\text{A1.5})$$

where κ is a parameter to be determined. Applying the operator H to a wave function, ψ ,

$$i\kappa \frac{\partial}{\partial s} \psi = H\psi \quad (\text{A1.6})$$

Using equations (A1.4) and (A1.5), this is written

$$\frac{2i}{v\kappa} \psi_{,s} + \psi_{,nn} - \frac{1}{\kappa^2 v^3} v_{,nn} n^2 \psi + \frac{2}{v^2 \kappa^2} \psi = 0 \quad (\text{A1.7})$$

Writing ψ in terms of the envelope, $\psi = W e^{-i\omega(t - \int_0^s \frac{ds}{v(s)})}$, and assuming $v = \omega^{1/2} n$ where $v = O(1)$, gives

$$\frac{2i}{\kappa \omega v(s)} W_{,s} + W_{,nn} - \frac{1}{\kappa^2 \omega^2 v(s)^3} v_{,nn} v^2 W + \left(\frac{2}{\kappa^2 \omega v(s)^2} - \frac{2}{\kappa v(s)^2} \right) W = 0 \quad (\text{A1.8})$$

Letting $\kappa = \omega^{-1}$ gives

$$\frac{2i}{v(s)} W_{,s} + W_{,nn} - v(s)^{-3} v_{,nn} v^2 W = 0 \quad (\text{A1.9})$$

This is equivalent to the parabolic equation in (7) derived from asymptotic analysis. Note that in quantum mechanics $\kappa = \frac{h}{2\pi}$ where h is Planck's constant, but here $\kappa = v_0 \frac{\lambda_0}{2\pi}$. v_0 can be avoided here by writing the eikonal equation (A1.2) in terms of $\tau' = v_0 \tau$. Classical ray theory is valid when λ_0 is much smaller than the heterogeneity scale. The natural analogy between mechanics and ray theory thus provides a link between the eikonal equation and the parabolic equation. Further results from this analogy can be found in Marcuse [1982] and Marcuvitz [1980].

2.2.6 APPENDIX 2

In this appendix, a brief review of the properties of a Gaussian beam in a homogeneous medium is given. From the dynamic ray tracing equations, (10), and the initial conditions for a plane wave and line source, one obtains

$$\begin{aligned} q_1(s) &= 1 & p_1(s) &= 0 \\ q_2(s) &= s - s_0 & p_2(s) &= 1/\nu_0 \end{aligned} \quad (\text{A2.1})$$

For the source at $s_0=0$, with the beam waist at the source, $S_0=0$, and using (16), the complex spreading can be written

$$q(s) = s - i \frac{\omega}{2\nu} L_0^2 \quad (\text{A2.2})$$

where L_0 is the initial beam-width at the beam waist. The beam-width can be written from (17) as

$$L(s) = L_0 \left[1 + \left(\frac{\lambda s}{\pi L_0^2} \right)^2 \right]^{1/2} \quad (\text{A2.3})$$

The phase front curvature can be written from (18) as

$$K(s) = \frac{1}{s \left[1 + \left(\frac{\pi L_0^2}{\lambda s} \right)^2 \right]} \quad (\text{A2.4})$$

At the beam waist, $L(s)=L_0$ and $K(s)=0.0$. Thus at the beam waist, the minimum beam-width is achieved and the wavefront is planar. As $s \rightarrow \infty$ the angular spread or the far field diffraction angle can be defined as

$$\varphi = \lim_{|s| \rightarrow \infty} \left[\frac{L(s)}{s} \right] = \frac{\lambda}{\pi L_M}$$

This leads to the uncertainty relation

$$\varphi L_M = \frac{\lambda}{\pi} \quad (\text{A2.5})$$

For a small L_M there is a large far-field spread, whereas for a large L_M , there is a small far-field spread. The distance s_C at which the beam-width starts to significantly differ from the initial beam-width is

$$\varphi \approx \frac{L_M}{s_C} \rightarrow s_C = \frac{\pi L_M^2}{\lambda} \quad (\text{A2.6})$$

This is similar to the Fraunhofer range for an aperture of radius a which has the form $s \gg \frac{a^2}{\lambda}$. For $|s| \gg s_C$, $K(s) \approx \frac{1}{s}$, which is the curvature of the wavefront for a line source. Extrema in the curvature occur at $|s| = s_C$.

The critical value of L_M is the one which gives the smallest beam-width, $L(s)$, at the receiver distance, s . In a homogeneous medium, this places the receiver distance at $s = s_C$. Thus from (A2.6), $L_{opt} = \sqrt{\lambda s / \pi}$.

2.2.7 APPENDIX 3

In this appendix, the weight function $\psi(\delta)$ for a 2-D point force in an elastic medium is obtained by comparing the Gaussian beam synthesis with the exact plane-wave decomposition. For a general point force at the origin directed in the x_3 direction, the displacement field for an isotropic, homogeneous medium is given by

$$U_i = \frac{1}{4\pi\rho} \left[(3\gamma_i\gamma_3 - \delta_{i3}) \left(\frac{r}{i\omega\beta} + \frac{1}{\omega^2} \right) r^{-3} - (\gamma_i\gamma_3 - \delta_{i3})(\beta^2 r)^{-1} \right] e^{i\omega \frac{r}{\beta}} \quad (\text{A3.1})$$

$$- \frac{1}{4\pi\rho} \left[(3\gamma_i\gamma_3 - \delta_{i3}) \left(\frac{r}{i\omega\alpha} + \frac{1}{\omega^2} \right) r^{-3} - \gamma_i\gamma_3(\alpha^2 r)^{-1} \right] e^{i\omega \frac{r}{\alpha}}$$

[Aki and Richards, 1980], where a time dependence of $e^{-i\omega t}$ is assumed, and the direction cosines of $\vec{r}$ are given by $\gamma_i = x_i / r$. In order to express the above displacement as an integral superposition of plane waves, the Weyl integral is used to obtain

$$\frac{\partial}{\partial x_3} \frac{\partial}{\partial x_i} \left[\frac{e^{i\omega \frac{r}{c}}}{r} \right] = \frac{-i}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} k_i e^{ik_1 x_1 + ik_2 x_2 + ik_c x_3} dk_1 dk_2 \quad (\text{A3.2})$$

$$= \omega^2 \left[(3\gamma_i\gamma_3 - \delta_{i3}) \left(\frac{r}{i\omega c} + \frac{1}{\omega^2} \right) r^{-3} - \gamma_i\gamma_3(c^2 r)^{-1} \right] e^{i\omega \frac{r}{c}}$$

where $k_c = +\sqrt{\frac{\omega^2}{c^2} - k_1^2 - k_2^2}$, and c is either α or β . Then, for the 3-D point force

$$U_i = \frac{i}{8\pi^2\rho\omega^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} k_i e^{i(k_1 x_1 + k_2 x_2 + k_c x_3)} dk_1 dk_2 \quad (\text{A3.3})$$

$$+ \frac{-i}{8\pi^2 \rho \omega^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [k_i - \frac{\delta_{i3}}{k_\beta \beta^2}] e^{i(k_1 x_1 + k_2 x_2 + k_\beta x_3)} dk_1 dk_2$$

Now we will investigate the "P-wave" integral alone, and integrate with respect to x_2 to obtain the line force solution. Thus

$$U_{L2} = \int_{-\infty}^{\infty} U_i dx_2 = \frac{i}{4\pi \rho \omega^2} \int_{-\infty}^{\infty} k_i e^{i(k_1 x_1 + k_\alpha x_3)} dk_1 \quad (A3.4)$$

where $k_\alpha = k_\beta = +\sqrt{(\frac{\omega}{\alpha})^2 - k_1^2}$. Let ϑ be the angle with the x_3 axis, then

$k_1 = \frac{\omega}{\alpha} \sin \vartheta$ and $k_3 = \frac{\omega}{\alpha} \cos \vartheta$. Now, writing the integral over ϑ ,

$$U_{L1} = \int_{-\frac{\omega}{\alpha}}^{+\frac{\omega}{\alpha}} U_1 dx_2 = \int_{-\frac{\pi}{2}}^{+\frac{\pi}{2}} \frac{i}{4\pi \rho \alpha^2} \cos \vartheta \sin \vartheta e^{i \frac{\omega}{\alpha} (x_1 \sin \vartheta + x_3 \cos \vartheta)} d\vartheta \quad (A3.5)$$

$$U_{L3} = \int_{-\frac{\omega}{\alpha}}^{+\frac{\omega}{\alpha}} U_2 dx_2 = \int_{-\frac{\pi}{2}}^{+\frac{\pi}{2}} \frac{i}{4\pi \rho \alpha^2} \cos \vartheta \cos \vartheta e^{i \frac{\omega}{\alpha} (x_1 \sin \vartheta + x_3 \cos \vartheta)} d\vartheta$$

The Gaussian beam solution for a homogeneous medium is written

$$U_{GB}(\vec{x}, \omega) = \int_{-\frac{\pi}{2}}^{+\frac{\pi}{2}} \phi(\delta) u_\delta(s, n, \omega) d\delta \quad (A3.6)$$

$$U_{GB}(\vec{x}, \omega) = \int_{-\frac{\pi}{2}}^{+\frac{\pi}{2}} \phi'(\delta) \sqrt{\frac{\alpha_0 \rho_0 q_0}{\alpha \rho q}} e^{i \omega \frac{r}{\alpha} + i \frac{\omega}{2} \frac{p}{q} n^2} d\delta$$

where it has been assumed that $\phi(\delta)$ has the form, $\phi(\delta) = \phi'(\delta) \sqrt{\alpha_0 \rho_0 q_0}$.

For a homogeneous medium,

$$\begin{aligned}
 q &= r + \varepsilon, \quad q_0 = \varepsilon \\
 p &= v_0^{-1}, \quad p_0 = v_0^{-1} \\
 p/q &= \frac{1}{v_0[r + \varepsilon]}
 \end{aligned}$$

where $\varepsilon = S_0 - i \frac{\omega}{2v_0} L_M^2$. To compare with the plane wave solution, let $L_M \rightarrow \infty$

and let $\delta = \vartheta$. Then

$$U_{GB}(\vec{x}, \omega) = \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \phi'(\vartheta) e^{i \frac{\omega}{\alpha} (x_1 \sin \vartheta + x_3 \cos \vartheta)} d\vartheta \quad (\text{A3.7})$$

Comparison with (A3.5) gives

$$\phi'(\vartheta) = \frac{i \cos \vartheta}{4\pi \rho_0 \alpha_0^2}$$

and

$$\phi(\vartheta) = \frac{i \cos \vartheta \sqrt{\varepsilon}}{4\pi \rho_0^{1/2} \alpha_0^{3/2}} \quad (\text{A3.8})$$

where $\phi(\vartheta)$ is the approximate weight for the Gaussian beam representation of P-wave radiation from a 2-D point force, and ϑ is the angle between the direction of the point force and the beam direction.

2.2.8 REFERENCES

- Aki, K., Scattering of P waves under the Montana LASA, *J. Geophys. Res.*, 78, 1334-1346, 1973.
- Aki, K., and P.G. Richards, *Quantitative Seismology: Theory and Methods*, W.H. Freeman and Company, San Francisco, 1980.
- Archuleta, R.J., Analysis of near-source static and dynamic measurements from the 1979 Imperial Valley earthquake, *Bull. Seism. Soc. Am.*, 72, 1927-1956, 1982.
- Azbel, I.Y., L.A. Dmutrieva, and T.B. Yanovskaya, A method of calculating geometrical divergence for a medium inhomogeneous in three dimensions, *Computational Seismology*, 13, 113-127, 1980.
- Babich, V.M., and M.M. Popov, Propagation of concentrated sound beams in a three-dimensional inhomogeneous medium, *Soviet Physics-Acoustics*, 27, 459-462, 1981.
- Bard, P.-Y., and M. Bouchon, The seismic response of sediment-filled valleys. Part 1. The case of incident SH waves, *Bull. Seism. Soc. Am.*, 70, 1263-1286, 1980.
- Boore, D.M., K.L. Larner, and K. Aki, Comparison of two independent Methods for the solution of wave-scattering problems: response of a sedimentary basin to vertically incident SH waves, *J. Geophys. Res.*, 76, 558-569, 1971.
- Capon, J., Characterization of crust and upper mantle structure under LASA as a random medium, *Bull. Seism. Soc. Am.*, 64, 235-266, 1974.
- Cerveny, V., Seismic wave fields in structurally complicated media (ray and Gaussian beam approaches), *Lecture Notes*, Rijksuniversiteit Utrecht, Vening-Meinesz Laboratory, Utrecht, 1981.
- Cerveny, V., Expansion of a plane wave into Gaussian beams, *Studia Geophys. Geod.*, 26, 120-131, 1982.
- Cerveny, V., Synthetic body wave seismograms for Laterally varying layered structures by the Gaussian beam method, *Geophys. J. R. Astron. Soc.*, 73, 389-426, 1983.
- Cerveny, V., and F. Hron, The ray series method and dynamic ray tracing system for three-dimensional inhomogeneous media, *Bull. Seism. Soc.*

- Am., 70, 47-77, 1980.
- Cerveny, V., I.A. Molotkov, and I. Psencik, Ray Method in Seismology, Karlova Univerzita, Praha, 1977.
- Cerveny, V., M.M. Popov, and I. Psencik, Computation of wave fields in inhomogeneous media - Gaussian beam approach, Geophys. J.R. Astron. Soc., 70, 109-128, 1982.
- Cerveny, V., and I. Psencik, Ray amplitudes of seismic body waves in laterally inhomogeneous media, Geophys. J.R. Astron. Soc., 57, 91-106, 1979.
- Cerveny, V., and I. Psencik, Gaussian beams in two-dimensional elastic inhomogeneous media, Geophys. J.R. Astron. Soc., 72, 417-433, 1983a.
- Cerveny, V., and I. Psencik, Gaussian beams and paraxial ray approximation in three-dimensional elastic inhomogeneous media, J. Geophys., preprint, 1983b.
- Cerveny, V. and Psencik, I., Gaussian beams in elastic two-dimensional laterally varying layered structures, preprint, 1983c.
- Chapman, C.H., A new method for computing synthetic seismograms, Geophys. J.R. Astron. Soc., 54, 481-513, 1978.
- Chapman, C.H., and R. Drummond, Body-wave seismograms in inhomogeneous media using Maslov asymptotic theory, Bull. Seism. Soc. Am., 72, s277-s317, 1982.
- Clairbout, J.F., Fundamentals of Geophysical Data Processing, McGraw-Hill, New York, 1976.
- Cline, A.K., FITPACK - a software package for curve and surface fitting employing splines under tension, Dept. of Computer Science, University of Texas, Austin, 1981.
- De Boor, C., A practical guide to splines, Springer Verlag, 1980.
- Fehler, M., and B. Chouet, Operation of a digital seismic network on Mount St. Helens volcano and observations of long period seismic events that originate under the volcano, Geophys. Res. Lett., 9, 1017-1020, 1982.
- Fuis, G.S., W.D. Mooney, J.H. Healy, G.A. McMechan, and W.J. Lutter, The Imperial Valley earthquake of October 15, 1979: crustal structure of the Imperial Valley region, U.S. Geol. Surv. Profess. Paper 1254, 1982.

- Gloge, D., and D. Marcuse, Formal quantum theory of light rays, *J. Optical Soc. Am.*, 59, 1629-1631, 1969.
- Hayakawa, M., Prospecting of the underground structure of "Showa-shinzan" by various geophysical methods, particularly seismic survey, *Japan Jour. Geophys.*, 1, 13-20, 1957.
- Hong, T.L., and D.V. Helmberger, Glorified optics and wave propagation in non-planar structures, *Bull. Seism. Soc. Am.*, 68, 1313-1330, 1977.
- Kienle, J., D.J. Lalla, C.F. Pearson, and S.A. Barrett, Search for shallow magma accumulation at Augustine volcano, Final Report to U.S.D.O.E., contract No. AT(45-1)-2229, 1979.
- Kirpichnikova, N.J., Construction of solutions concentrated close to rays for the equations of elasticity theory in an inhomogeneous isotropic space, in *Mathematical Problems of Theory of Diffraction and Propagation of Waves*, Vol. 1, Nauka, Leningrad, 1971 (in Russian, English translation by American Mathematical Society, 1974)
- Klimes, L., and V. Cervený, Synthetic body wave seismograms for three-dimensional laterally varying media, preprint, 1983.
- Kravtsov, Y.A., Two new asymptotic methods in the theory of waves in laterally inhomogeneous media (review), *Soviet Physics-Acoustics*, 14, 1-17, 1968.
- Kravtsov, Y.A., and Y.I. Orlov, Limits of applicability of the method of geometric optics and related problems, *Sov. Phys. Usp.*, 23, 750-762, 1980.
- Lalla, D.J., and J. Kienle, Comparison of synthetic and observed seismograms for volcanic earthquakes at Augustine volcano, abstract, Pacific NW AGU meeting, 1982.
- Malone, S., Lecture Notes given at the Enrico Fermi summer school, Italy, 1982.
- Marcuse, D., *Light Transmission Optics*, Van Nostrand, New York, 1982.
- Marcuvitz, N., Quasi-particle view of wave propagation, *Proceedings of the IEEE*, 68, 1330-1395, 1980.
- McMechan, G.A., and W.D. Mooney, Asymptotic ray theory and synthetic seismograms for laterally varying structures: theory and application to the Imperial Valley California, *Bull. Seism. Soc. Am.*, 70, 2021-2035.

1980.

Popov, M.M., A new method of computation of wave fields in high-frequency approximation, in *Mathematical Problems of Theory of Propagation of Waves*, Vol. 11, 195-216, Nauka, Leningrad, 1981. (in Russian, translated to English in *Journal of Soviet Mathematics*, 1982)

Popov, M.M., A new method of computation of wave fields using Gaussian beams, *Wave Motion*, 4, 85-97, 1982.

Popov, M.M., and I. Psencik, Ray amplitudes in inhomogeneous media with curved interfaces, *Geofyzikalni Sbornik*, 24, 118-129, 1978.

Popov, M.M., and I. Psencik, Computation of ray amplitudes in inhomogeneous media with curved interfaces, *Studia Geophys. Geod.* 22, 248-258, 1978.

Pretlova, V., Bicubic spline smoothing of two-dimensional geophysical data, *Studia Geoph. Geod.*, 20, 168-177, 1976.

Psencik, I., Program for the computation of kinematic and dynamic parameters of seismic body waves in 2-D laterally inhomogeneous media with curved interfaces, in *Programs for the Interpretation of Seismic Observations*, Vol. 3, Nauka, Leningrad, 1983. (in Russian).

Tappert, F.D., The parabolic approximation method; in *Wave Propagation and Underwater Acoustics*, *Lecture Notes in Physics*, 70, 224-287, Springer-Verlag, Berlin, 1976.

2.3 TOPICS IN GAUSSIAN BEAM SYNTHESIS

2.3.1 INTRODUCTION

Many decompositions of wavefields are possible in terms of local, global, or intermediate building blocks. These building blocks include rays, modes, and beams. In this paper we investigate several topics in the approximate decomposition of wavefields in terms of beams. These topics include the specification of the beam parameter, c , and its interpretation in terms of a complex source point. Asymptotic and exact evaluations of the Gaussian beam integral for a line source are then investigated. Then, the problem connected with the decomposition of a point source in a velocity gradient is discussed, and reduces to the difference between an initial plane wave and an initial Snell wave. A comparison with the PP phase in the Imperial Valley structure is then made. Next, a slant stack of a wavefield is computed and compared with the beam decomposition for a large beam width, showing good agreement. A first order correction to the plane wave reflection coefficient typically used at interfaces in the Gaussian beam method is next described. Finally, the finite frequency scattering effects from a dented interface are computed using the Gaussian beam method and compared with the effects computed using the Aki-Larner method.

2.3.2 Choice of the complex parameter ϵ

Each Gaussian beam in the Gaussian beam superposition requires the choice of a complex parameter, ϵ . To be consistent within the superposition, one value of ϵ is usually chosen for all beams in the sum. The real part of ϵ shifts the beam waist away from the start point on the ray. The imaginary part of ϵ specifies the beam width at the beam waist. The appropriate range of values for ϵ was shown by Nowack and Aki [1984] to depend on the type of initial wavefield to be approximated by Gaussian beams. When decomposing an initial plane wave into Gaussian beams, useful results with a finite aperture occur for small $|\epsilon|$. In the limit as $|\epsilon| \rightarrow 0$, the Gaussian beam superposition reduces to a Kirchhoff integral with a paraxially approximated Green's function. When decomposing a point source into Gaussian beams, useful results occur for large $|\epsilon|$ (see Nowack and Aki [1984], example C). In the limit as $|\epsilon| \rightarrow \infty$, the Gaussian beam superposition reduces to a plane wave decomposition or a Weyl integral for a point or line source. A finite value of $|\epsilon|$ in the above cases has the result of smoothing over diffraction effects (see Nowack and Aki [1984], example B for the diffraction effects of a half plane and example D for the generation of a head wave). A finite $|\epsilon|$ compresses the energy to a vicinity of the stationary points. The choice of ϵ depends on what part of the wavefield one is interested in modelling. The critical value of $\text{Im}(\epsilon)$ was used by Nowack and Aki [1984] (or the optimal value in the terminology of Cerveny et. al. [1982]) as a soft cutoff between large and

small values of $|\epsilon|$ for a given source and receiver geometry. The above results are summarized in Figure 1. Nowack and Aki [1984] also found in their Example G that the beam width should be smaller than the heterogeneity scale. Moving the beam waist by changing the real part of ϵ may be useful particularly when a localized heterogeneity exists. The beam waist can then be moved to the region of the heterogeneity. For heterogeneities distributed along the whole path, the beam waist is typically put at the source or the receiver.

2.3.3 THE COMPLEX SOURCE POINT APPROACH TO GAUSSIAN BEAMS

In this section, Gaussian beams resulting from a complex source coordinate is described. The complex parameter, ϵ , for each Gaussian beam in the Gaussian beam synthesis can be considered as a complex distance from a source point in complex space to a start point along the ray in real space. Thus

$$\epsilon = d - ib \quad (1)$$

where d is the real distance between the source point, or the beam waist, and the start point and corresponds to S_0 in the notation of Cervený et. al [1984]. When ϵ is real, it corresponds to the radius of curvature of the wavefront at the start point. For a complex ϵ , d can be used to shift the beam waist along the ray. Putting the source point in complex space by

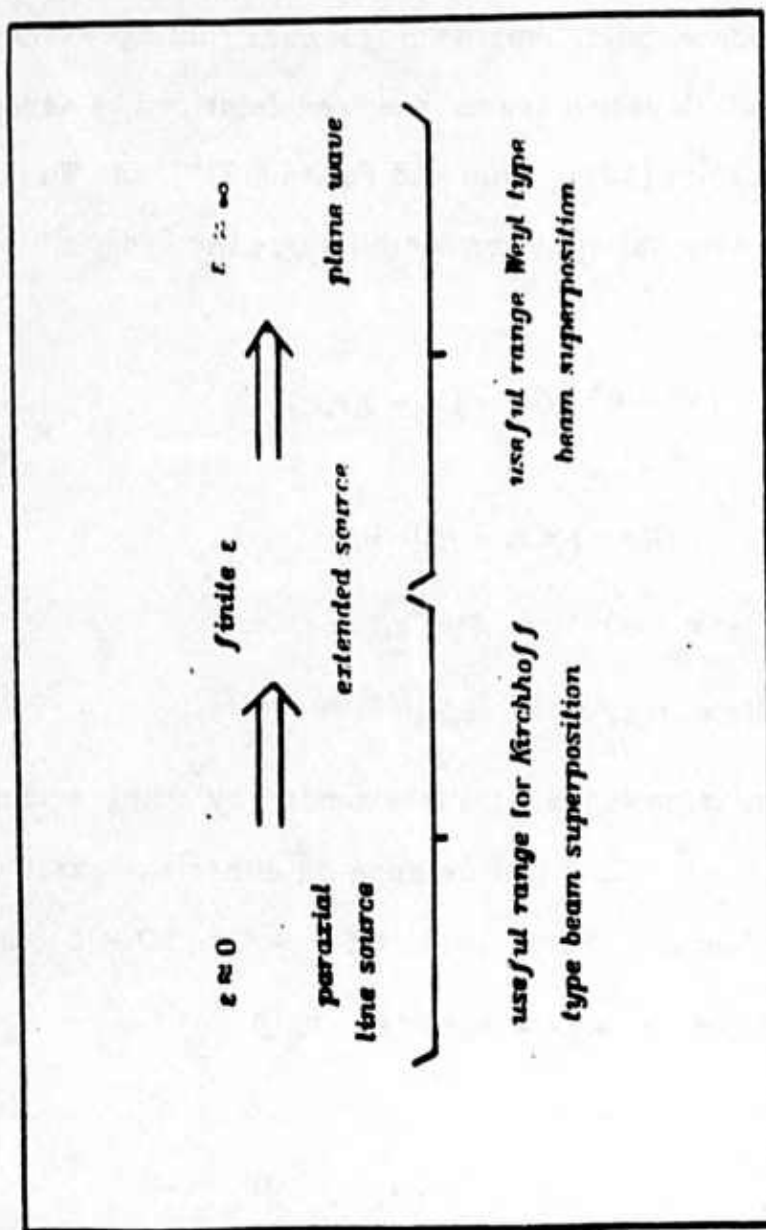


Figure 1 The range of the complex parameter, ϵ , for an individual Gaussian beam and their useful range in beam synthesis.

specifying $b \neq 0$ results in an evanescent wavefield which attenuates along the wavefront.

The complex source point approach to constructing evanescent wavefields, in particular Gaussian beams, has been described by Keller and Streifer [1969], Deschamps [1971], Shin and Felsen [1977], and Wu [1985]. Here, the 2-D Helmholtz equation is considered where the Green's function is specified from

$$(\nabla^2 + k^2) G(\mathbf{r}, \mathbf{r}') = -\delta(\mathbf{r}, \mathbf{r}')$$

and

$$G(\mathbf{r}, \mathbf{r}') = i/4 H_0^{(1)}(kD)$$

where $D = ((x - x')^2 + (z - z')^2)^{1/2}$. For $kD \gg 1$,

$$G(\mathbf{r}, \mathbf{r}') \approx 1/4 (2/\pi kD)^{1/2} e^{ikD - i\pi/4} \quad (2)$$

The Green's function is now analytically extended by using a complex source point. Thus $\mathbf{r}' \rightarrow \hat{\mathbf{r}}'$. This can be done on either the exact or an approximate Green's function. For example, let $\mathbf{r}' \rightarrow \hat{\mathbf{r}}' = (0, 0 + ib)$, then

$$D \rightarrow \hat{D} = [x^2 + (z - ib)^2]^{1/2} \quad \text{Re } \hat{D} > 0$$

In the paraxial region of the z axis

$$\hat{D} \approx (z - ib) \left[1 + 1/2 \left(\frac{x}{z - ib} \right)^2 \right]$$

or

$$\hat{D} = z - ib + \frac{zx^2}{2(z^2 + b^2)} + \frac{ibx^2}{2(z^2 + b^2)}$$

The analytically extended Green's function in the paraxial region is then

$$\bar{G}(r, r') \approx C/4 \frac{e^{ik(z-ib)}}{\sqrt{\pi k (z-ib)}} e^{\frac{ikxz^2}{2(z^2+b^2)}} e^{\frac{-kbz^2}{2(z^2+b^2)}} \quad (3)$$

This is a Gaussian beam where C is a normalizing constant. The beam curvature is

$$K(z) = \frac{1}{z(1 + b^2/z^2)}$$

and the beam width is

$$L(z) = (2b/k)^{1/2} (1 + z^2/b^2)^{1/2}$$

Comparing with equations (A2.3) and (A2.4) of Nowack and Aki [1984] gives

$$b = \frac{k}{2} L_M^2 = L_0^2$$

where L_M is the initial beam width at the beam waist and L_0 is a normalized initial beam width independent of frequency used by Nowack and Aki [1984]. The complex parameter can then be written in two ways

$$\varepsilon = S_0 - i \frac{k}{2} L_M^2 = d - ib$$

where in the second case ε refers to the complex distance from the source to the start point on the ray.

The fact that an analytically extended Green's function is a Gaussian beam has been used by a number of authors (see for example Felsen [1976], Ra et. al. [1973], Pott and Harris [1984]) to model the effects of beam scattering. The exact or approximate scattering results for a point

source is found and then extended to a complex source point to give the scattering results for a beam. Thus the scattering results for certain types of extended sources, Gaussian beams, can be gotten from the analogous point source problem without performing any integration. In addition the plane wave response can be gotten from the point source response by letting $b \rightarrow \infty$.

In the Gaussian beam synthesis method, a quite different approach is taken where approximate Gaussian beams are traced through the medium and superposed to simulate the scattering effects for a point source or an arbitrary extended source. The complex source point is expected to be useful in evaluating how appropriate the approximate Gaussian beams are in different complicated geologic media.

2.3.4 THE USE OF COMPLEX PHASE

Introducing a finite imaginary part to ϵ results in a complex phase for the component beams in the Gaussian beam synthesis as in eqn. (3). This imaginary phase part will act to modify the integrand of the Gaussian beam synthesis. In order to represent a point source by beam superposition, a value of b less than infinity will act to filter the plane wave components in a Gaussian manner.

The use of complex phase in wave superposition integrals has been used by several authors in order to obtain various computational advantages. In the discrete wavenumber method (see Lerner [1970], Bouchon and Aki [1977]) a complex frequency is used to reduce the effects of artificially induced periodic sources resulting from discretizing the wavenumber integral. Thus

$$e^{i\omega t} \rightarrow e^{i\omega_r t} e^{-\omega_i t}$$

This results in an exponentially damped time domain trace and smoothing and broadening in the frequency domain since

$$e^{-\omega_i \operatorname{sgn}(\omega) t} \leftrightarrow \frac{2\omega_i}{\omega_i^2 + \omega^2}$$

are a Fourier transform pair. Since $\omega = vk$, adding a small complex part to the frequency and requiring v to be real, adds a complex part to the wavenumber, k , also. This has the effect of moving singularities off the real axis. Thus, adding a small complex part to the frequency in this case can be used to gain numerical stability, economize on computation time, and reduce effects of artificial periodic sources when sampling the wavenumber integral. The integral is slightly distorted, but this is corrected after the synthesis by multiplying by $e^{i\omega_i t}$.

An interesting twist on this strategy was given recently by Madariaga and Papadimitriou [1985] to numerically compute the Gaussian beam impulse response having singularities at the geometric arrivals. A small

complex part is added to the phase time on the order of the time spacing of the impulse response to be computed. Thus

$$\psi \rightarrow \hat{\psi} + i\Delta t$$

The result of this is to filter the high frequencies and smooth the transient signal by convolving with a function which has a width on the order of the sample spacing Δt , since

$$e^{-|\omega|\Delta t} \leftrightarrow \frac{2\Delta t}{t^2 + \Delta t^2}$$

are also a Fourier transform pair. Thus singularities at geometric arrivals are slightly smoothed to give a stable sampled impulse response without the need to use an explicit source time function such as the Gabor wavelet as was used by Nowack and Aki [1984]. A particular source time function can then be incorporated in a convenient manner at a later point.

2.3.5 Asymptotic and exact solution of the Gaussian beam synthesis for a line source

In this section, the asymptotic and exact solutions to the Gaussian beam superposition for a line source are given. The geometry for a line source at the origin and a receiver at (x, z) is shown in figure (2a), where $R^2 = (x^2 + z^2) = (s^2 + r^2)$ and

$$s = R \cos(\varphi_r - \varphi) = z \cos \varphi + x \sin \varphi = v (xp_x + zp_z)$$

Figure 2

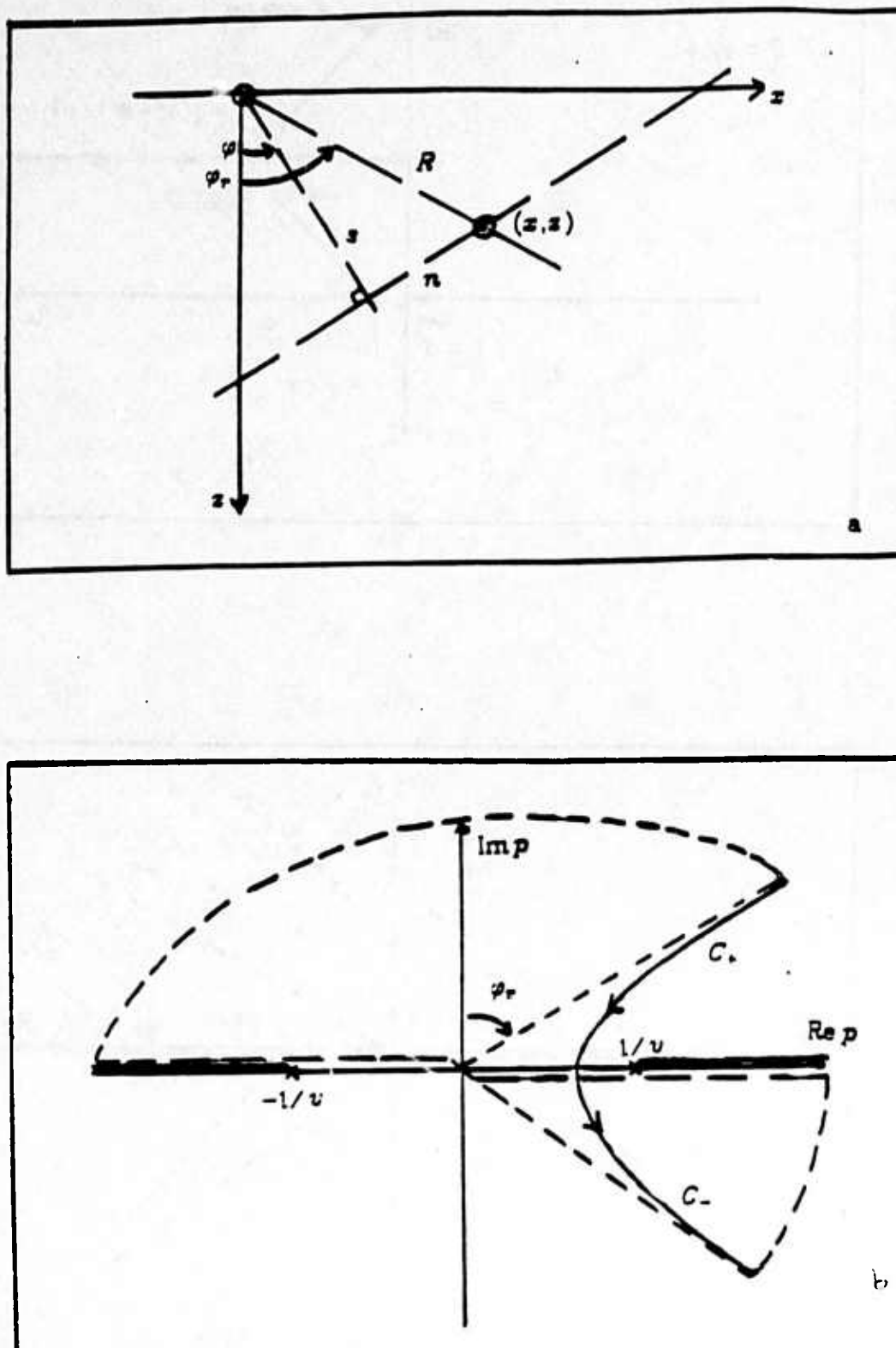
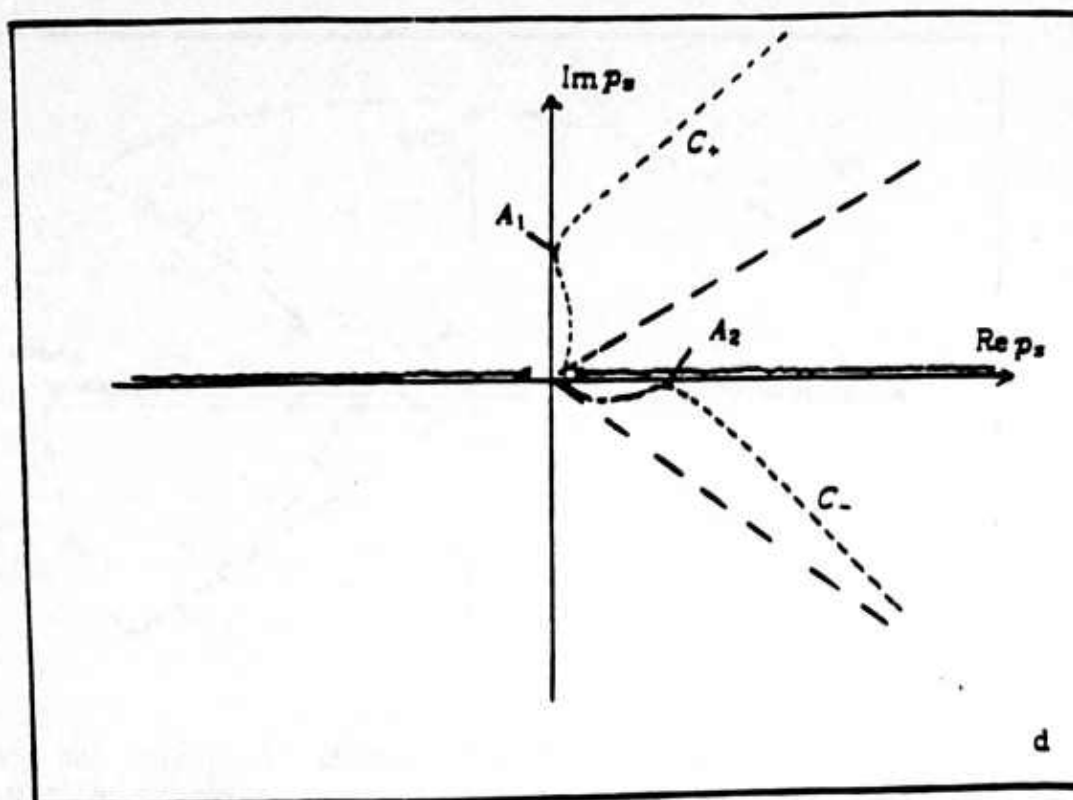
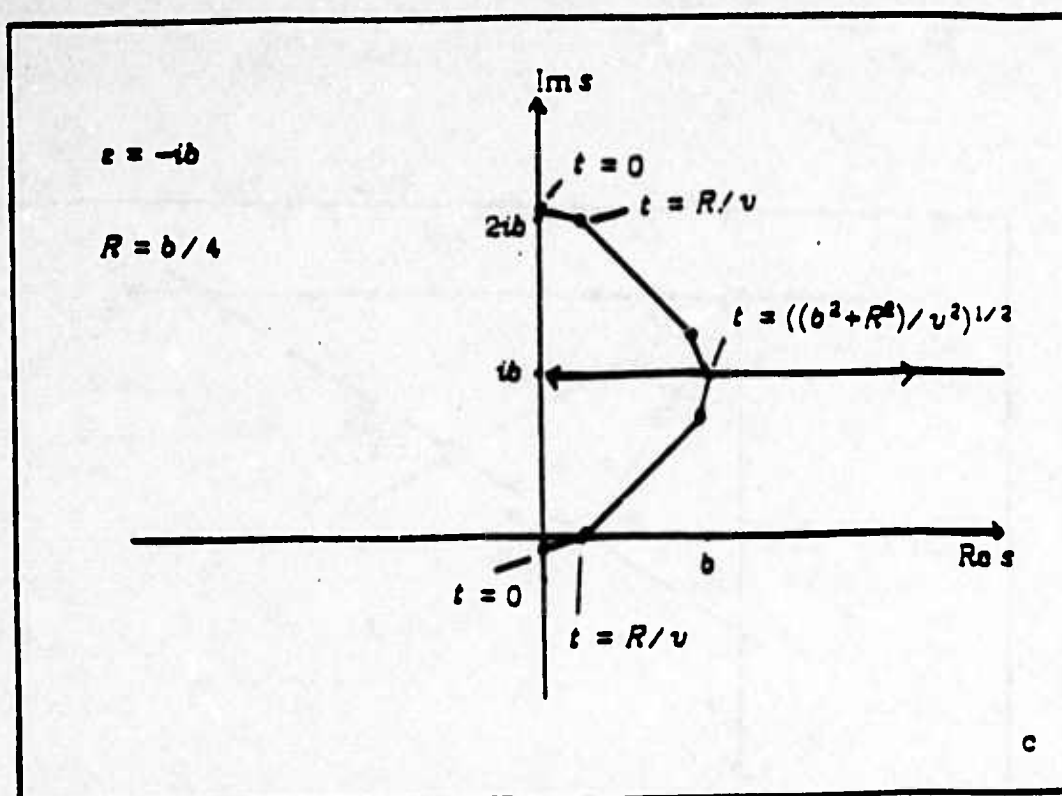
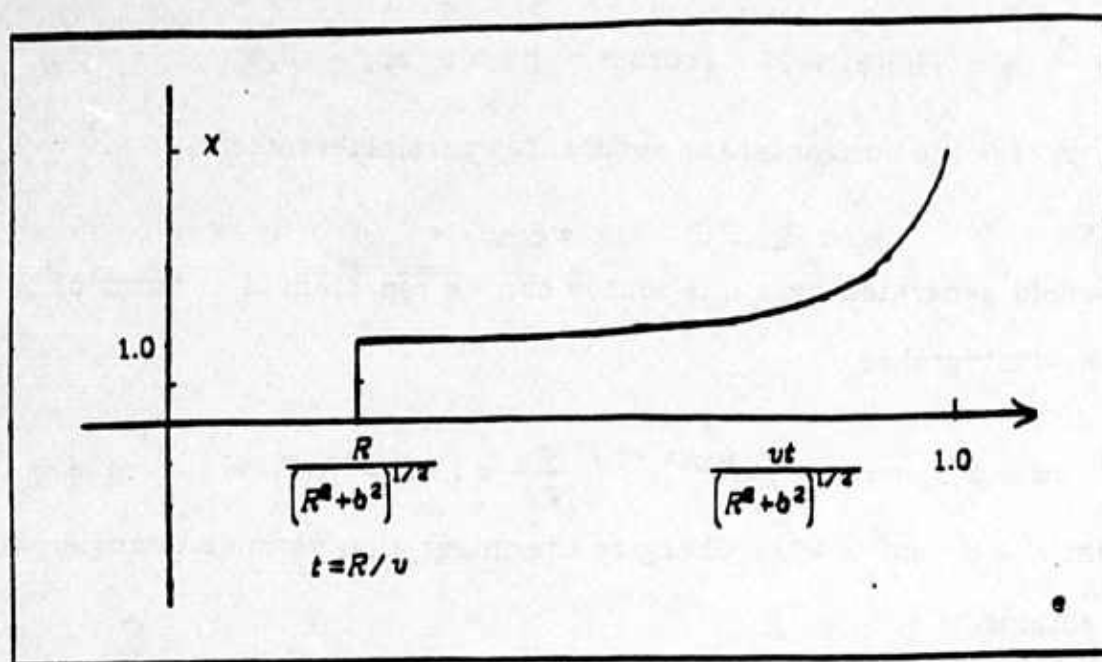


Figure 2 a) line source geometry. b) complex p_s plane with branch cuts defined for $\text{Re } p_s > 0$. original contour from $-\infty$ to ∞ , and Cagniard contour for exact line source solution. c) $s = s(t)$ for Gaussian beam solution in the complex s plane with $\epsilon = -ib$ and $R = b/4$. d) Cagniard contour for Gaussian beam integral for $\epsilon = -ib$, $R = 1 = b/4$, $v = 1$. e) The modification factor for the Gaussian beam solution compared to the exact solution.





$$n = R \sin(\varphi_r - \varphi) = x \cos \varphi - z \sin \varphi = v (x p_x - z p_z)$$

p_x and p_z are the horizontal and vertical ray parameters where

$$p_x = \sin \varphi / v \quad p_z = \cos \varphi / v$$

The wavefield generated by a line source can be represented in terms of the 2-D Weyl integral as

$$u(x, z, \omega) = i / 4\pi \int_{-\infty}^{+\infty} e^{i\omega(p_x x + p_z z)} \frac{dp_x}{p_x} = i / 4 H_0^{(1)}(\omega R / v) \quad (5a)$$

where $\text{Im } p_x \geq 0$ and $z \geq 0$. Changing the integration variable to angle, φ , one obtains

$$u(x, z, \omega) = i / 4\pi \int_L e^{i\omega/v(x \sin \varphi + z \cos \varphi)} d\varphi \quad (5b)$$

where the contour L goes from $-\pi/2 - i\infty$ to $+\pi/2 + i\infty$. The time domain solution due to a line source is given by

$$u(x, z, t) = 1/2\pi \frac{H(t - R/v)}{\sqrt{t^2 - R^2/v^2}} \quad (5c)$$

The Gaussian beam synthesis in terms of angle can be written

$$u_{gb}(x, z, \omega) = i / 4\pi \int_{-\pi/2}^{+\pi/2} \left[\frac{\epsilon}{\epsilon + s} \right]^{1/2} e^{i\omega/v \left[s + \frac{\pi^2}{2(\epsilon + s)} \right]} d\varphi \quad (6)$$

where ϵ is the complex beam parameter described in a previous section. Usually only homogeneous central rays are considered. As $\epsilon \rightarrow \infty$ and extending the contour to include the inhomogeneous waves, then this reduces to the exact solution given in (5b).

The steepest descent method is now applied to (6) in order to obtain the asymptotic solution for large ω , following Cerveny et. al. [1982]. Equation (6) can be written in the form

$$I(\omega) = \int_C F(\varphi) e^{i\omega f(\varphi)} d\varphi \quad \text{for } \omega \rightarrow \infty$$

where

$$F(\varphi) = i/4\pi \left| \frac{\varepsilon}{\varepsilon + s} \right|^{1/2} e^{1/2 \text{ARG} \frac{\varepsilon}{\varepsilon + s}}$$

and

$$f(\varphi) = i/v \left[s + \frac{\kappa^2}{2(\varepsilon + s)} \right]$$

This can be written as

$$f(\varphi) = i/v \left[R \cos \Delta\varphi + \frac{R^2 \sin^2 \Delta\varphi}{2(\varepsilon + R \cos \Delta\varphi)} \right]$$

where $\Delta\varphi = \varphi_r - \varphi$. The saddle point is defined by $f'(\varphi) = 0$, where

$$f'(\varphi) = \frac{i}{v} R \sin \Delta\varphi \left[1 - \frac{R \cos \Delta\varphi}{(\varepsilon + R \cos \Delta\varphi)} - \frac{R^2 \sin^2 \Delta\varphi}{2(\varepsilon + R \cos \Delta\varphi)^2} \right]$$

A saddle point occurs at $\varphi = \varphi_r$, the same as for the exact line source representation. There is an additional saddle point coming from equating the bracket to zero which will be discussed subsequently. At the saddle point, $\varphi = \varphi_r$, in the Gaussian beam case

$$f''(\varphi)|_{\varphi=\varphi_r} = \frac{-i\varepsilon R}{v(\varepsilon + R)}$$

The exponent can then be approximately written as

$$f(\varphi) \approx f(\varphi_r) + 1/2 f''(\varphi)|_{\varphi=\varphi_r} \Delta\varphi^2$$

or

$$f(\varphi) \approx iR/v - \frac{i\epsilon R}{2v(\epsilon + R)} (\varphi_r - \varphi)^2$$

Let $(\varphi_r - \varphi)^2 = A e^{i2\vartheta}$ then

$$f(\varphi) \approx iR/v - R/2v \left| \frac{\epsilon}{\epsilon + R} \right| A e^{i(2\vartheta + \pi/2 + \text{ARG}(\frac{\epsilon}{\epsilon + R}))}$$

or

$$f(\varphi) \approx iR/v - R/2v \left| \frac{\epsilon}{\epsilon + R} \right| A [\cos\beta + i\sin\beta]$$

where $\beta = 2\vartheta + \pi/2 + \text{ARG}\left(\frac{\epsilon}{\epsilon + R}\right)$. Now choose ϑ such that $\sin\beta = 0$

which gives the constant phase or steepest contour. Thus $\beta = n\pi$. This

results in $\vartheta = -\frac{\pi}{4} - \frac{1}{2}\text{ARG}\left(\frac{\epsilon}{\epsilon + R}\right)$ for steepest descent. The integral can

then be approximated as (see for example Aki and Richards [1980], box 6.3)

$$I(\omega) = \frac{2\pi}{\omega f''(\varphi_r)} \cdot 1/2 F(\varphi_r) e^{i\vartheta} e^{\omega f(\varphi_r)}$$

or

$$u_{\text{ph}}(x, z, \omega) \approx \frac{1}{4\pi} \left| \frac{2\pi v}{\omega R} \right|^{1/2} e^{i\omega R/v + i\pi/4} \frac{i}{4} H_0^{(1)}(\omega R/v) \quad (7)$$

which is also the asymptotic expansion for the exact solution. Thus for large ω , the Gaussian beam synthesis for the saddle point $\varphi = \varphi_r$ gives the correct response. This will be invalid when $\epsilon = 0$ for the Gaussian beam synthesis since then $f''(\varphi_r) = 0$. In order to investigate this behavior, an

exact transient evaluation of the Gaussian beam integral is done.

The Cagniard method is first used to evaluate the exact transient response (5c) from (5a). The substitution $s = -i\omega$ is made in (5a) where ω is real and positive. Equation (5a) can then be written

$$u(x, z, s) = i/4\pi \int_{-\infty}^{\infty} e^{-s(p_z x + p_z z)} \frac{dp_z}{p_z}$$

This allows the branch cuts to be changed to those in figure (2b) since now $\text{Re } p_z \geq 0$. Now let $t = p_z x + p_z z$, and deform the contour in the complex p -plane so that t is real. Thus

$$u(x, z, s) = i/4\pi \int_C e^{-st} \frac{dp_z}{p_z}$$

The Cagniard contour C_+ and C_- is shown in figure (2b) where

$$p_{z(+/-)} = \frac{xt}{R^2} (+/-) \frac{i|z|}{R^2} \left[t^2 - R^2/v^2 \right]^{1/2} \quad t > R/v$$

According to Jordan's Lemma the contributions on the large arcs in figure (2b) go to zero leaving the result

$$u(x, z, s) = i/4\pi \int_{C_+} \frac{e^{-st}}{p_z} dp_z + i/4\pi \int_{C_-} \frac{e^{-st}}{p_z} dp_z$$

and

$$u(x, z, s) = -i/4\pi \int_{R/v}^{\infty} \frac{e^{-st}}{p_{z+}} \left(\frac{dp_{z+}}{dt} \right) dt + i/4\pi \int_{R/v}^{\infty} \frac{e^{-st}}{p_{z-}} \left(\frac{dp_{z-}}{dt} \right) dt$$

where

$$\left(\frac{dp_z(+/-)}{dt} \right) = \frac{x}{R^2} (+/-) \frac{i|z|t}{R^2} \left[t^2 - R^2/v^2 \right]^{-1/2}$$

and

$$\left(\frac{dp_z(+/-)}{dt} \right) \frac{1}{p_z(+/-)} = \frac{(+/-)i}{\left[t^2 - R^2/v^2 \right]^{1/2}}$$

Then

$$u(x, z, \bar{s}) = 1/2\pi \int_{R/v}^{\infty} \frac{1}{\left[t^2 - R^2/v^2 \right]^{1/2}} e^{-\bar{s}t} dt$$

Using the fact that $L^{-1} \left[\int_{t_1}^{\infty} f(t) e^{-\bar{s}t} dt \right] = f(t) H(t - t_1)$, the exact time

domain response is given by (5c).

The Cagniard method is now used to obtain the transient response of the Gaussian beam integral in (6). First, the Gaussian beam integral is converted to p_z and extended from $-\infty$ to $+\infty$ to include the full range of ray parameter. The substitution $\bar{s} = -i\omega$ is then made, where $\bar{s}$ is real and positive. Thus

$$u_{gb}(x, z, \bar{s}) = i/4\pi \int_{-\infty}^{\infty} \left(\frac{\varepsilon}{\varepsilon + s} \right)^{1/2} e^{-\bar{s} \left[\frac{s}{v} + \frac{n^2}{2v(\varepsilon + s)} \right]} \frac{dp_z}{p_z}$$

where the branch cuts are defined by $\text{Re } p_z \geq 0$. Now the substitution is

made $t = \left[\frac{s}{v} + \frac{n^2}{2v(\varepsilon + s)} \right]$. Since $s^2 + n^2 = R^2$, then

$t = \left[\frac{s}{v} + \frac{R^2 - s^2}{2v(\varepsilon + s)} \right]$. We now want to deform the contour in the complex

p-plane so that t is real where $s = p_z x + p_z z$. First, s is solved in terms of t , giving

$$s = (vt - \epsilon) (+/-) \sqrt{v^2 t^2 + \epsilon^2 - R^2}$$

The two solutions are shown in figure (2c) for $\epsilon = -ib$ and $R = b/4$. For $\epsilon = -ib$, then the solution that reduces to $s = vt$ for $b \rightarrow \infty$ is

$$s(t) = vt + ib - iX \quad (R^2 + b^2) \geq v^2 t^2$$

where $X = \left\{ (R^2 + b^2) - v^2 t^2 \right\}^{1/2}$, and

$$s(t) = vt + ib + X \quad v^2 t^2 \geq (R^2 + b^2)$$

where $X = \left\{ v^2 t^2 - (R^2 + b^2) \right\}^{1/2}$. The factor $p = p(s)$ can be written as before as

$$p_z(+/-) = \frac{is}{vR^2} (+/-) \frac{i|z|}{vR^2} \left\{ s^2 - R^2 \right\}^{1/2} \quad |s| > R$$

and

$$p_z(+/-) \approx \frac{is}{vR^2} (+/-) \frac{i|z|s}{vR^2} \quad |s| \gg R$$

The contour C_+ and C_- are shown in figure (2d). The original contour is deformed into this contour in order to make t real. Then

$$u_{yb}(x, z, \xi) = i/4\pi \int_{C_+} \left(\frac{-ib}{-ib + s} \right)^{1/2} e^{-\pi} \frac{dp_z}{p_z} + i/4\pi \int_{C_-} \left(\frac{-ib}{-ib + s} \right)^{1/2} e^{-\pi} \frac{dp_z}{p_z}$$

Converting this to an integration in time

$$u_{yb}(x, z, \xi) = -i/4\pi \int_{r/v}^{\infty} \left(\frac{-ib}{-ib + s} \right)^{1/2} \left\{ \frac{1}{p_{z+}} \frac{dp_{z+}}{ds} \right\} \left\{ \frac{ds}{dt} \right\} dt$$

$$+ i/4\pi \int_{r/v}^{\infty} \left(\frac{-ib}{-ib + s} \right)^{1/2} \left(\frac{1}{p_z -} \frac{dp_z}{ds} \right) \left(\frac{ds}{dt} \right) dt$$

where $\left(\frac{1}{p_z(+/-)} \frac{dp_z(+/-)}{ds} \right) = \frac{(+/-)i}{[s^2 - R^2]^{1/2}}$

$$\left(\left(\frac{-ib}{-ib + s} \right) \frac{ds}{dt} \right) = \frac{v}{X} \sqrt{b} \left(X + ivt \right)^{1/2} \quad (R^2 + b^2) > v^2 t^2$$

$$\left(\left(\frac{-ib}{-ib + s} \right) \frac{ds}{dt} \right) = \frac{v}{X} \sqrt{-ib} \left(X + vt \right)^{1/2} \quad v^2 t^2 > (R^2 + b^2)$$

also

$$[s^2 - R^2]^{-1/2} = [2(X + ivt)(b - X)]^{-1/2} \quad (R^2 + b^2) > v^2 t^2$$

$$[s^2 - R^2]^{-1/2} = [2(X + vt)(ib + X)]^{-1/2} \quad v^2 t^2 > (R^2 + b^2)$$

The final response is

$$\left(\left(\frac{-ib}{-ib + s} \right)^{1/2} \frac{ds}{dt} [s^2 - R^2]^{-1/2} \right) = \frac{H(t - R/v) \chi}{2\pi [t^2 - R^2/v^2]^{1/2}}$$

where χ is a modification factor to the exact solution when using the Gaussian beam superposition, where

$$\chi = \frac{\sqrt{b} (v^2 t^2 - R^2)^{1/2}}{X [2(b - X)]^{1/2}} \quad (R^2 + b^2) > v^2 t^2$$

$$\chi = \frac{\sqrt{-ib} (v^2 t^2 - R^2)^{1/2}}{X [2(ib + X)]^{1/2}} \quad v^2 t^2 > (R^2 + b^2)$$

This modification factor is shown in figure (1e). The modification factor has a singularity at $v^2 t^2 = (R^2 + b^2)$. After this time the modification factor becomes complex, but is bounded. This additional singularity moves to infinity as $b \rightarrow \infty$. To check the existence of this singularity, we note

that the function $f'(\varphi) = 0$ given previously has an additional zero when the bracket term is zero, giving $s = R \cos \varphi = ib (+/-) (R^2 + b^2)^{1/2}$. For the plus sign and using the formula for $s(t)$, this corresponds to $t = (R^2 + b^2)^{1/2}/v$, which is what is shown here. For a critical value of b , this secondary singularity occurs at 1.41 times the geometric arrival. In the p -plane shown in figure (2d) this corresponds to the two points, A_1 and A_2 . The problem of this secondary singularity occurs when extending the original contour from $-\infty$ to $+\infty$ to include the evanescent beam waves. It would be difficult to define central rays for these. The usual Gaussian beam formulation only includes central rays with real takeoff angles, and the additional singularity won't in general be included in the usual Gaussian beam synthesis. The usual Gaussian beam integral thus does not correspond to the exact transient solution for a line source. It approaches the exact solution as $\varepsilon \rightarrow \infty$. For large ω and real takeoff angles, the Gaussian beam integral is asymptotically equivalent to the exact integral representation as long as $\varepsilon \neq 0$. In the general case, the Gaussian beam response is good near the geometric arrival but deteriorates for larger times.

2.3.6 Initial point in a medium with a velocity gradient

In this section, the initial conditions for the variables $p(s)$ and $q(s)$ that solve the dynamic ray equations are discussed. The variable $q(s)$ has the interpretation of a spreading and $\frac{1}{v} \frac{p}{q}$ has the interpretation of a wavefront curvature. Gaussian beams can be formed as a linear combination of two independent sets of solutions to the dynamic ray equations with a complex weighting factor. Thus

$$q(s) = q_2(s) + \epsilon q_1(s)$$

$$p(s) = p_2(s) + \epsilon p_1(s)$$

and $\epsilon = d - ib$ is described in an earlier section. With b strictly greater than zero, then a beam is defined in which $q(s) \neq 0$. This uses the fact that the determinant of the matrix formed by the two solutions, $\begin{bmatrix} q_1 & q_2 \\ p_1 & p_2 \end{bmatrix}$, is an invariant of the motion, and is sometimes called the Lagrange invariant. Cerveny et. al. [1982] uses two independent solutions with initial conditions, $\begin{bmatrix} 1 & 0 \\ 0 & 1/v_0 \end{bmatrix}$. The initial conditions, $q_1(0) = 1$ and $p_1(0) = 0$, correspond to an initial planar wavefront with, $p_1/q_1 = 0$. The initial conditions, $q_2(0) = 0$ and $p_2(0) = 1/v_0$ correspond to an initial point source, with $p_2/q_2 \rightarrow \infty$. The source region is usually assumed to be locally homogeneous.

Madariaga [1984] constructed a beam solution in terms of a small perturbation of a constant Snell wave having a constant horizontal ray parameter. For a source imbedded in a vertically varying medium, he found that the (q_1, p_1) solution must be modified in order to specify a constant Snell wave, thus

$$q_1(0) = 1 \quad p_1(0) = \frac{-p_z^2 v_{,z}(0)}{\cos \varphi_0} \quad (8)$$

where φ_0 is the initial takeoff angle of the ray from the vertical and $p_z = \sin \varphi_0 / v_0$ is the horizontal ray parameter which in a vertically varying medium is constant along a ray. The (q_2, p_2) solution is the same in each case. For a homogeneous source region, where $v_{,z}(0) = 0$, Madariaga's constant Snell wave solution is the same as the previous plane wave solution. When $\varepsilon \rightarrow \infty$, Madariaga's beam superposition in a vertically varying media reduces to the standard WKBJ result.

Madariaga's initial conditions in a vertically varying medium is shown by imbedding a thin homogeneous layer within the gradient as in figure (3a). A source is put in the constant velocity layer. Boundary conditions are then matched on the top and bottom interfaces and the layer thickness shrunk to zero. Boundary conditions for p and q at an interface are given by Cervený and Hron [1979] derived using the phase matching principle. The 2-D results are given in a compact form by Cervený [1981] as

$$\begin{bmatrix} \tilde{q} \\ \tilde{p} \end{bmatrix} = (+/-) \frac{\cos \tilde{\vartheta}}{\cos \vartheta} \begin{bmatrix} 1 & 0 \\ C + I_n & \frac{\cos^2 \vartheta}{\cos^2 \tilde{\vartheta}} \end{bmatrix} \begin{bmatrix} q \\ p \end{bmatrix}$$

where (q, p) are the incident values and $(\tilde{q}, \tilde{p})$ are the transmitted or reflected values. The upper sign is for the transmitted case and the lower sign for the reflected case. The angles are defined in figure (3b). The effects due to interface curvature, represented by C , and the effects due to velocity heterogeneity, represented by I_n , are additive, where

$$C = G \left\{ \frac{\cos \vartheta}{v} (-/+) \frac{\cos \tilde{\vartheta}}{\tilde{v}} \right\} \frac{1}{\cos^2 \tilde{\vartheta}}$$

and

$$I_n = \frac{-\sin \vartheta}{v \cos^2 \tilde{\vartheta}} \left\{ 2 \left(\frac{v \cdot \vec{m}}{v} \cos \vartheta (-/+) \frac{\tilde{v} \cdot \vec{m}}{\tilde{v}} \cos \tilde{\vartheta} \right) + \frac{\sin \vartheta}{v} (v_{,l} - \tilde{v}_{,l}) \right\}$$

where G is the local interface curvature at the point of incidence, where G is positive if the interface is convex as seen by the incident ray. $\vec{l}$ is the local tangent to the ray and $\vec{m}$ is perpendicular to it as seen in figure (3b).

In the case of a ray striking the lower horizontal interface as in figure (3b), then

$$\tilde{v}_{,m} = -\tilde{v}_{,z} \sin \vartheta \quad \tilde{v}_{,l} = \tilde{v}_{,z} \cos \vartheta$$

Also, since $v = \tilde{v}$ at the interface, then $\vartheta = \tilde{\vartheta}$ by Snell's law. The interface curvature term is zero, $C = 0$, and the velocity inhomogeneity

term is $I_n = \frac{-v_{,z} \sin^2 \vartheta}{v^2 \cos \vartheta}$ then

$$\tilde{q} = q \quad \tilde{p} = p - \frac{v_z \sin^2 \vartheta}{v^2 \cos \vartheta} q$$

Shrinking the layer thickness to zero results in the modified initial conditions for (q_1, p_1) as in eqn (8). For a vertical ray, these conditions reduce to Cerveny's initial conditions. For a horizontal grazing ray, the boundary conditions at the interface break down.

Part of the discrepancy in the two initial conditions for (q_1, p_1) is that an initial plane wave imbedded in a medium with a vertical velocity gradient is not a wave with a constant horizontal ray parameter. Madariaga's initial conditions for (q_1, p_1) construct an initial constant Snell wave. A constant Snell wave has a caustic or focus at its bottoming point. This results in a problem when using Madariaga's beam superposition for an initial ray that is exactly horizontal for which both the (q_2, p_2) solution and the (q_1, p_1) solution are singular and thus no longer independent. In the WKB method, the contributions from horizontal rays are typically windowed out (see Sinton and Frazier [1982]).

By using Cerveny's initial conditions, two independent solutions are always guaranteed. The disadvantage of this approach is that the (q_1, p_1) may no longer correspond to a constant Snell wave in the global coordinates. Also, since the weight function in the beam superposition is derived by matching to a known solution in a homogeneous source region, then in an arbitrary source region, the weight function will only be

approximate. Beydoun and Ben Menahem [1985] used Cerveny's initial conditions in various gradients and compared with exact analytic solutions. They showed that good agreement occurred provided that several high frequency conditions were satisfied. In addition, the Gaussian beam superposition was found to have more restrictions than the ray theoretical solution. This may be in part due to the approximate weight function used in the Gaussian beam superposition.

Several examples are now done using Cerveny's initial conditions, C.i.c., and Madariaga's initial conditions, M.i.c., in a vertical gradient. In the first example a surface source is used. The medium has a linear gradient with velocities from 1.5 km/sec to 5.6 km/sec from 0 to 10 km. A Gabor wavelet with $f = 4\text{hz}$ and $\gamma = 3$ was used (see Nowack and Aki [1984] eqn. (22)). Results in the time-distance domain are shown in figure (4) with the C. i.c. shown on the left and the M. i.c. shown on the right. Two initial beam widths are used, $L_0 = 5\text{ km}^{1/2}$ and $L_0 = 15\text{ km}^{1/2}$, where $L_0 = 5\text{ km}^{1/2}$ is near an average critical value. The C. i.c. for $L_0 = 15\text{ km}^{1/2}$ has noise predominantly in front of the geometric arrival. The M. i.c. for $L_0 = 15\text{ km}^{1/2}$ has noise later in time than the geometric arrival.

In figure (5), the contributions from each beam is shown for a station at 12 km. The same model as in figure (4) is used. The rays are numbered from 1 to 20, from shallow to steeply dipping at the source. The ray number-phase time plots are shown for both initial conditions and initial

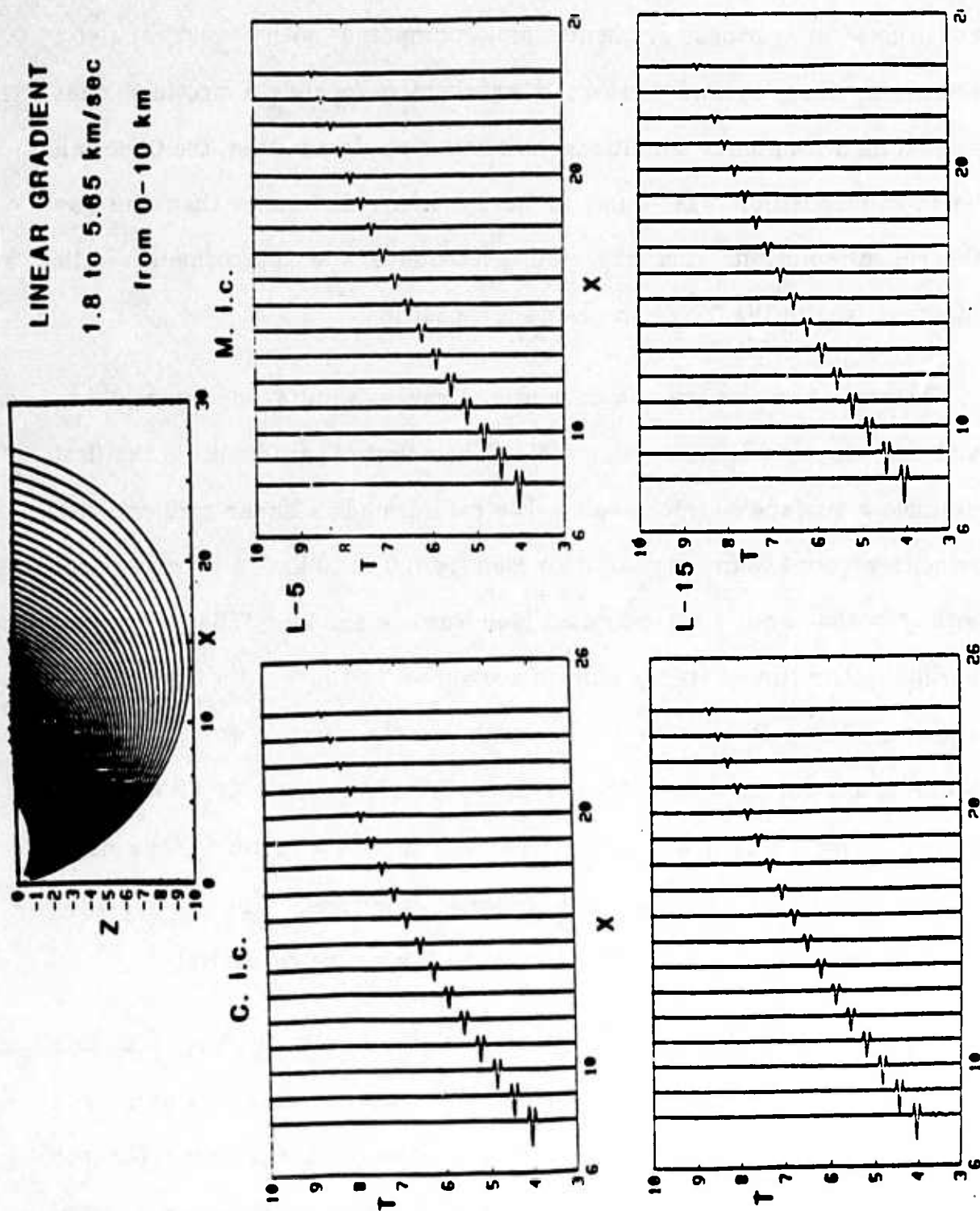


Figure 4 Time distance wavefields for a surface source in a linear gradient for two initial conditions, C. i. c. and M. i. c., and two initial beam widths, $L_0 = 5, 15 \text{ km}^{1/2}$

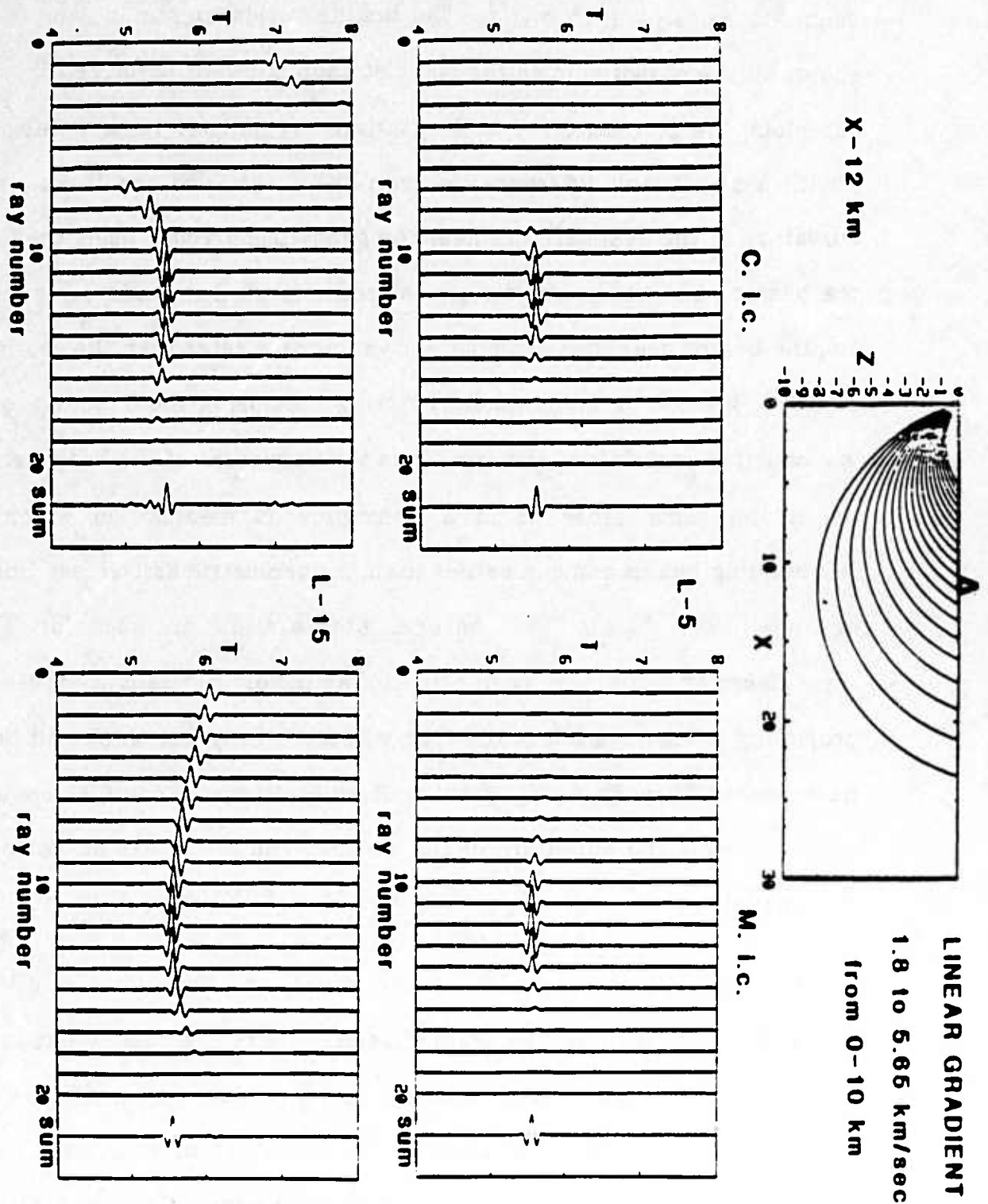


Figure 5 Ray number-phase time plots for a surface source in linear gradient with a station at 12 km. Two initial conditions, C. i.c. and M. i.c., and two initial beam widths, $L_0 = 5.15 \text{ km}^{1/2}$ are shown.

beam widths, $L_0 = 5, 15 \text{ km}^{1/2}$. The resultant seismogram in each case is shown by the scaled sum on the right of each sub-plot in figure (5). In all sub-plots, the geometric ray to the station, ray number 12, is a stationary point. A significant difference between the C. i.c. and the M. i.c. is the curvature of the beam arrivals near the geometric arrival. Using the C. i.c., the beams near the geometric arrival come in earlier, while using the M. i.c., the beams near the geometric arrival come in later than the geometric arrival. This can be understood by noting that using the C. i.c. no q_1 or q_2 caustics occur along the ray. Thus the curvature of the beams should be in the same sense as in a homogeneous medium, in which the surrounding beams come in earlier than the geometric arrival (see Nowack and Aki [1984] figure (5b)). Several later arrivals are seen for C. i.c., $L_0 = 15 \text{ km}^{1/2}$, but this is due to an artificial q_1 caustic caused by projecting these rays above the free surface. Using the M. i.c., all beams have passed through a q_1 caustic at the bottoming point of each ray which changes the curvature of the beams. The C. i.c. has no bottoming point caustic.

Another difference between the two initial conditions, C. i.c. and M. i.c., is the phase of the individual beam wavelets. There is a phase delay for the C. i.c. and a phase advance for the M. i.c.. In each case, the phase of the individual beam wavelets depends on the initial beam width used. In all cases, the resulting sum is approximately zero phase. Thus in each case, the individual beam phase shifts is offset by the curvature of the beams

away from the geometric arrival. This occurs in such a way as to give a seismogram with no resulting phase shift. This results from using a weight function in the superposition which was derived by matching to a known solution.

In figure (6), the time-distance wavefields are shown for a source at 4. km depth in a linear gradient. The velocity is 1.8 km/sec to 5.6 km/sec from 0. to 20. km in depth. A Gabor wavelet with $f = 4 \text{ Hz}$ and $\gamma = 4$ For the M. i.c. there is now a small second arrival for distances greater than 16 km, using both initial beam widths, 4 and 20 $\text{km}^{1/2}$. Also, the amplitude is slightly larger in the 14-16 km range for the M. i.c. case. In order to investigate this, ray number-phase time plots are constructed for several station locations.

In figure (7), the ray number-phase time plots are shown for a station range of 22 km, for both initial conditions and initial beam widths. The rays are numbered from 1 to 20 from shallow to steeply dipping at the source, as shown in the above ray diagram. The geometric arrival is specified by ray number 17. For the C. i.c., the beams near the geometric beam arrive earlier in time. The noise in front of the geometric arrival in the scaled sum for $L_0 = 20 \text{ km}^{1/2}$ can be reduced by using more beams in the superposition. For the M. i.c., the beams nearby the geometric arrival, ray number 17, come in later in time. But, another stationary point occurs at ray number 9. This is a maximum stationary point, and has lower

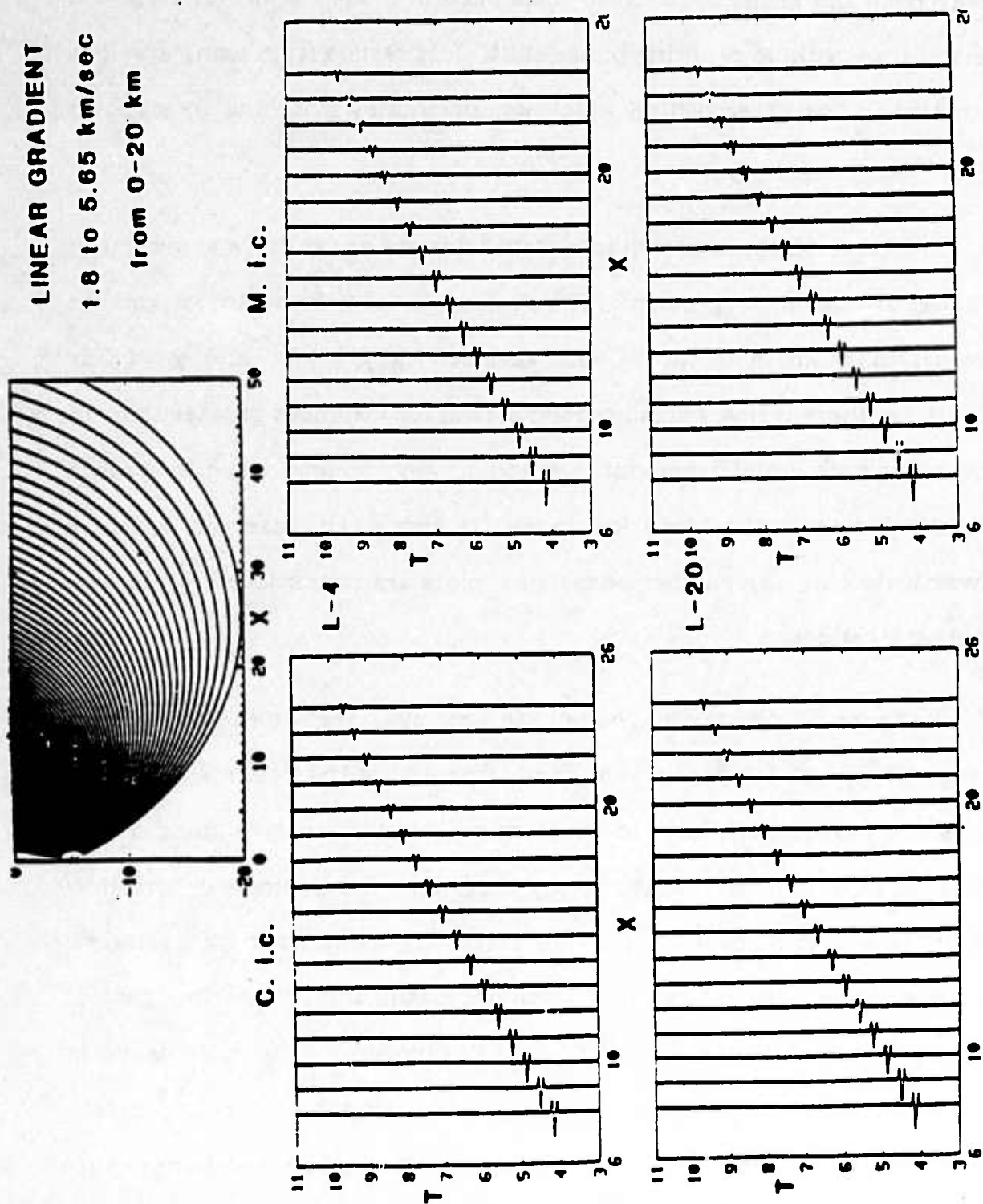


Figure 6 Time distance wavefields for a source at 4 km depth in a linear gradient for two initial conditions, C. i.c. and M. i.c., and two initial beam widths, $L_0 = 4.20 \text{ km}^{1/2}$.

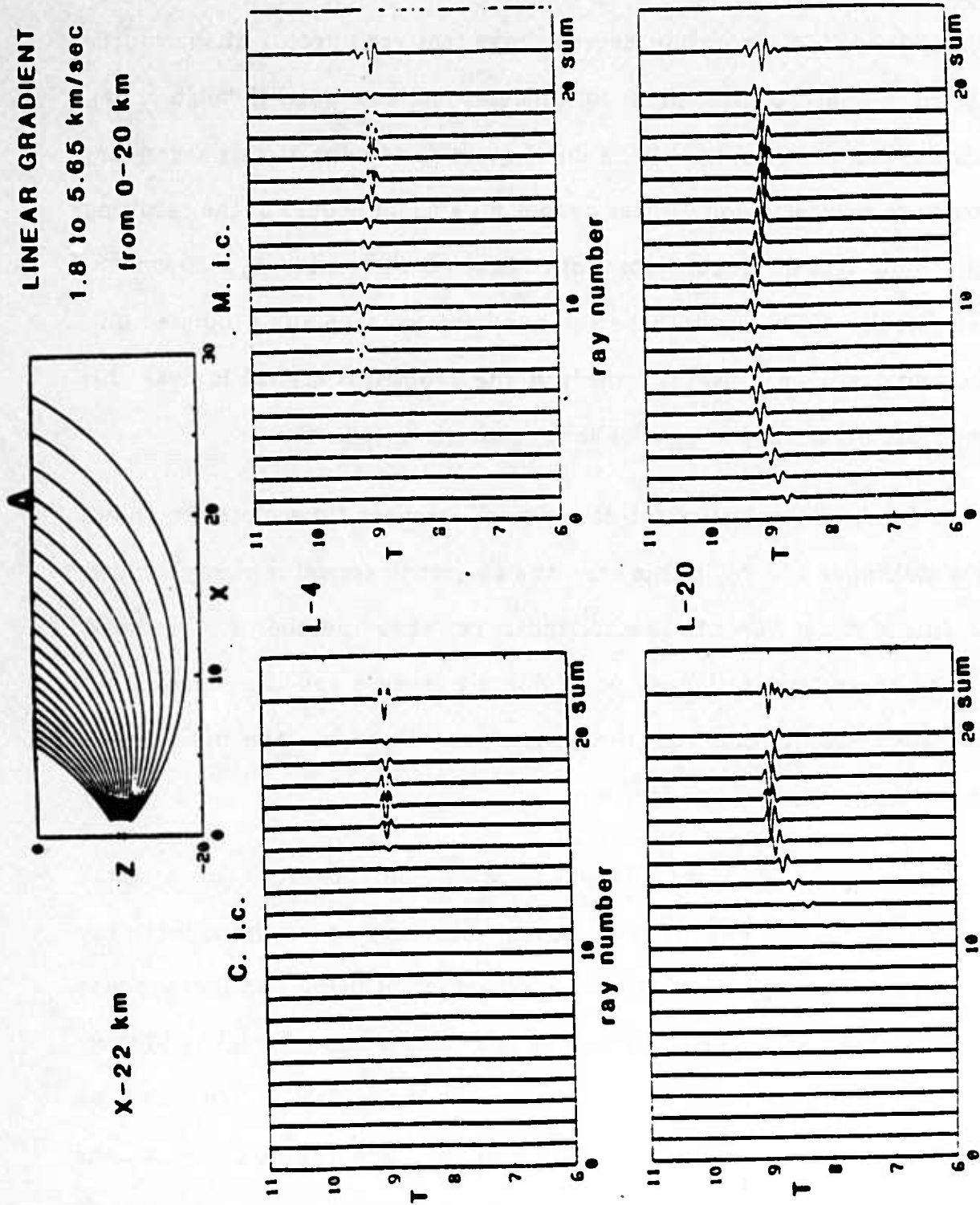


Figure 7 ray number-phase time plots for a source at 4 km depth in a linear gradient with a station at 22 km. Two initial conditions, C.I.C. and M.I.C., and two initial beam widths, $L_0 = 4, 20 \text{ km}^{1/2}$ are shown.

amplitudes. It corresponds to the ray that takes off horizontally at the source. There is a separation between rays that are upgoing at the source and those that go through a bottoming point and pass through a q_1 caustic. Since the contributions don't quite go to zero at this secondary stationary point and don't phase cancel, an artifact occurs in the resulting sum. This artifact occurs for both $L_0 = 4 \text{ km}^{1/2}$ and $L_0 = 20 \text{ km}^{1/2}$. Again, in the WKBJ result these ray numbers are typically windowed out. This would not be a useful remedy if the geometric arrival is near this horizontal ray as in the Imperial Valley example to follow.

In figure (8), a similar set of ray number-phase time plots are shown for a station at 15 km. In this case, the geometric arrival is at ray number 13. This is much closer to the horizontal ray at ray number 9. For the M. i.e., the secondary stationary point doesn't cause a spurious arrival, but now interferes directly with the main pulse. Modifying the initial beam width doesn't alleviate the problem.

Madariaga [1984] considered a beam superposition as a small perturbation of a WKBJ superposition with respect to horizontal ray parameter. This approach loses the advantage of using two independent solutions for (p,q) for a horizontal ray. Madariaga's initial conditions cause a spurious phase for the rays near horizontal which can't be eliminated by changing the initial beam width. Cerveny's initial conditions are guaranteed to give two independent solution for (p,q) , but doesn't

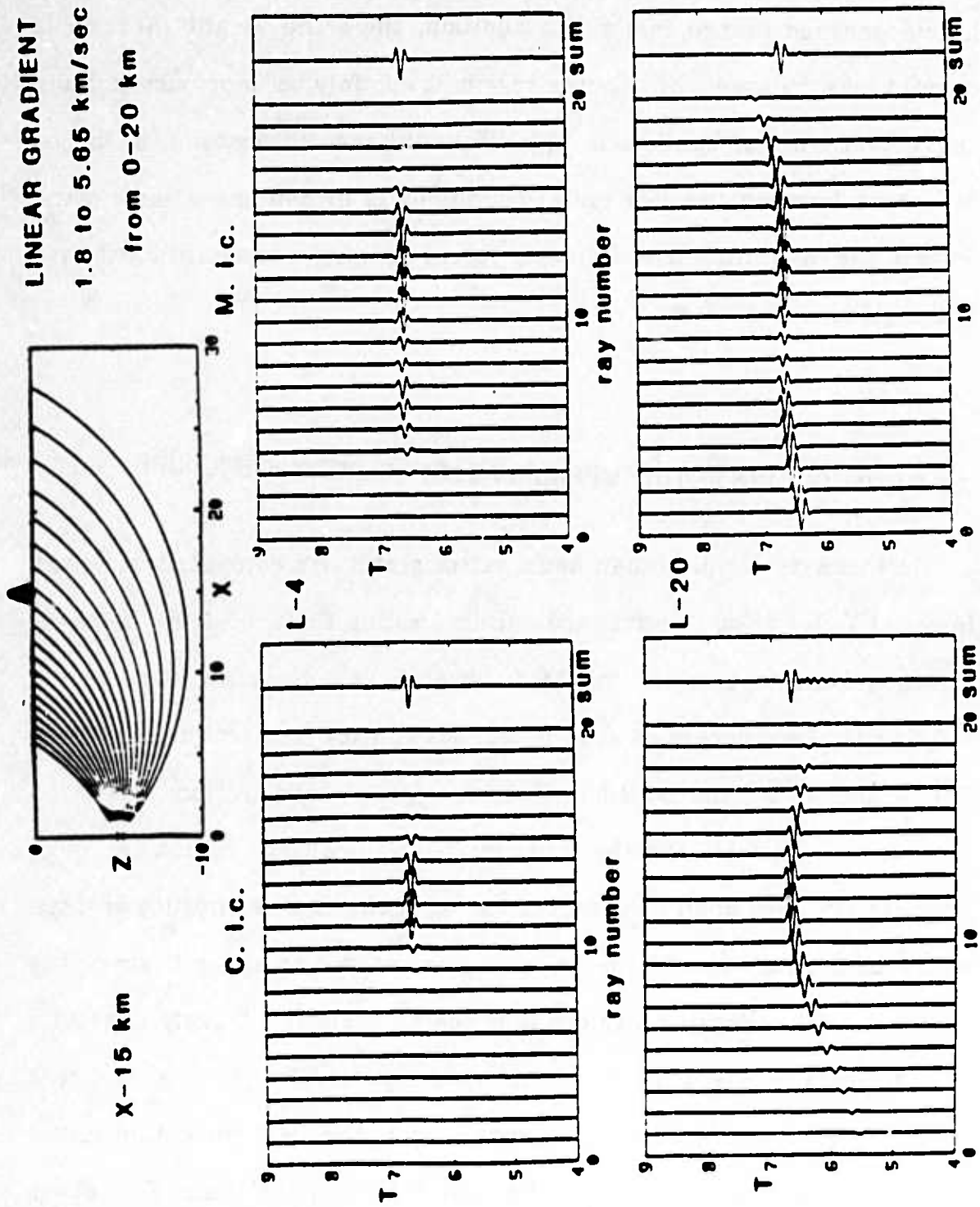


Figure 8 ray number-phase time plots for a source at 4 km depth in a linear gradient with a station at 15 km. Two initial conditions, C. l. c. and M. l. c., and two initial beam widths, $L_0 = 4.20 \text{ km}^{1/2}$ are shown.

correspond to a small perturbation of the WKBJ solution in an inhomogeneous source region. In addition, since the weight function is derived for a homogeneous source region, it will only be approximate using the Cerveny initial conditions in an inhomogeneous region. The major difference between the two initial conditions is in defining a local plane wave in the ray centered coordinates versus defining a constant Snell wave in the global coordinates.

2.3.7 SEISMOGRAMS IN THE IMPERIAL VALLEY VELOCITY STRUCTURE

In this example, Gaussian beam seismograms are computed using the Imperial Valley velocity structure. An interesting feature of the recorded strong ground motion for the 1979 Imperial Valley earthquake was the large vertical accelerations at some but not all sites (see Archuleta [1982]). For example, at the station E06, the peak acceleration was 1705 cm/sec^2 . Several explanations for this localized region of high acceleration have been considered, including the source rupture process, and a path effect due to the velocity gradient in the upper 5 km of the Imperial Valley. Here we explore only the path effect. A likely candidate for the large acceleration resulting from a path effect is the PP phase which was a prominent arrival recorded during the 1979 refraction survey of the Imperial Valley (see McMechan and Mooney [1980], and Fuis et. al. [1982]). Because of the focusing effect due to the multiple bounce, the PP

phase has a localized high amplitude over a limited distance range. Archuleta [1982] found that the travel times of the observed high accelerations were consistent with a PP phase originating from a source point at 4.0 km depth and 5.5 km north of the hypocenter, although he now attributes the main cause of localized high acceleration to a source effect.

The Imperial Valley structure derived from the refraction survey is shown in figure (9). The ray diagram for the PP phase is also shown. In figure (10), seismograms computed using the Gaussian beam method are compared to seismograms computed using the Reflectivity method, without attenuation, by R. Archuleta. The Gaussian beam method uses the Cerveny initial conditions described in the previous section. A significant portion of the energy comes from nearly horizontal rays at the source. The source is modelled using a point double couple radiation pattern with a dip of 75 degrees, a 0.0 degree slip angle, and the seismic section 13 degrees off strike. A Gabor wavelet with a frequency of 6 hz was used in the Gaussian beam seismograms. To obtain the Gaussian beam result for a point source, the GB results were multiplied by a ray method correction given by Cerveny et. al. [1982]. The value of L_0 was set at $2.0 \text{ km}^{1/2}$ resulting in average beam widths at the receivers of about 2 km. Both seismic sections show the predominant high amplitudes of the PP phase in the distance range of 15-18 km, with the Gaussian beam result being slightly closer. Significant S wave energy is seen in the Reflectivity section. The Reflectivity section shown in Archuleta [1982] included the effects of

PP PHASE IN IMPERIAL VALLEY STRUCTURE

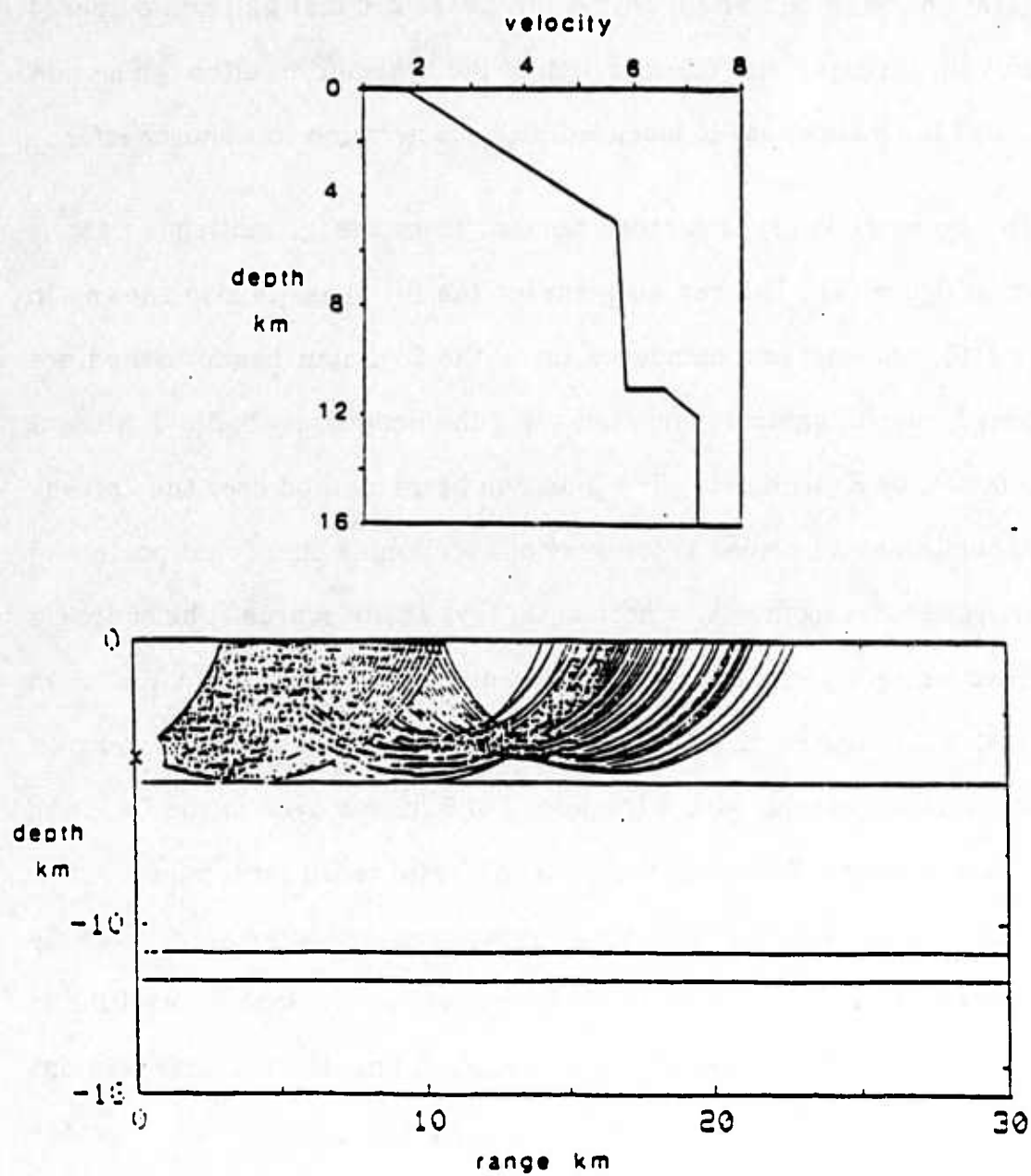


Figure 9 Imperial Valley velocity structure and ray diagram for the PP phase.

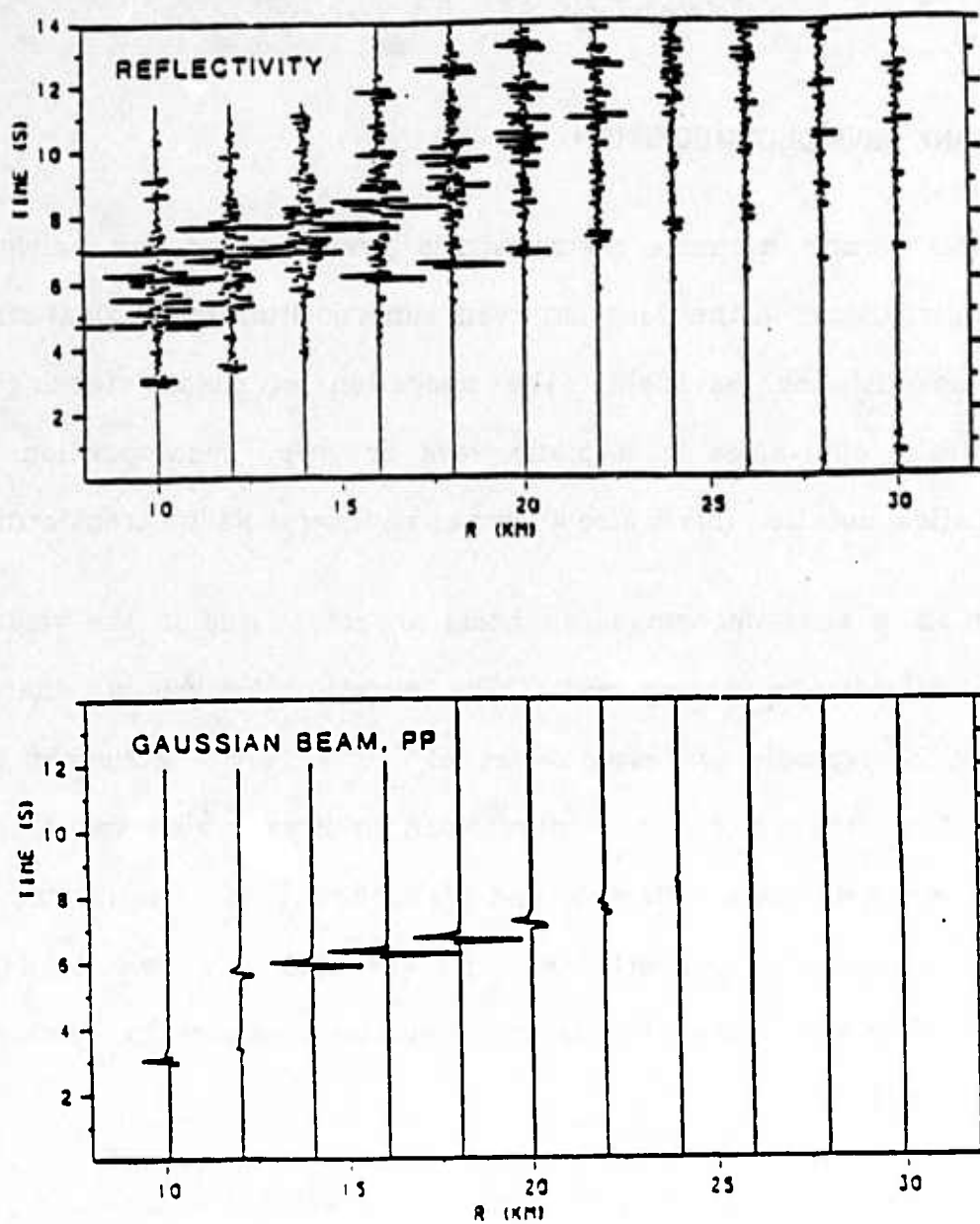


Figure 10 Reflectivity and Gaussian beam seismic sections computed using a 4 km depth source in the Imperial Valley structure.

attenuation which diminished the S waves, and further enhanced the PP arrival.

2.3.8 PLANE WAVE DECOMPOSITION

In this section, a simple comparison is given between the individual beam contributions in the Gaussian beam superposition and a slant stack of a time-distance wavefield. The operation of slant stacking is kinematically equivalent to a plane wave or τ - p decomposition. In mathematical notation this is also known as an inverse Radon transform.

The plane wave decomposition holds a central role in the velocity analysis in vertically varying media. The τ -method for velocity analysis was used for example by Bessonova et. al. [1974, 1976]. McMechan and Ottolini [1980] showed that τ - p curves can be directly observed in slant stacked wavefield data. Clayton and McMechan [1981] performed an iterative migration to convert the τ - p wavefield to a velocity depth estimate. This was applied to the upper mantle structure by Walck and Clayton [1984].

Using the τ - p wavefield has the advantage that triplications are unwrapped and thus made more easily interpretable. Also, in a vertically varying medium, each τ - p trace corresponding to a ray parameter, p , satisfies a 1-D wave equation. The relationship between the τ - p wavefield

and seismogram synthesis has been investigated by Chapman [1978,1981] and Treitel et.al. [1982]. For a line source in a vertically varying medium, Chapman [1981] gives the Radon transform formulas for wavefield decomposition and synthesis as

$$\tilde{u}(\tau, p) = \int_{-\infty}^{\infty} u(\tau + px, x) dx \quad (9a)$$

and

$$u(t, x) = \int_{-\infty}^{\infty} \frac{-1}{2\pi} \tilde{u}^+(t - px, p) dp \quad (9b)$$

where $\tilde{u}^+$ is the time derivative, Hilbert transform of $\tilde{u}$. These are reversible linear transformations which map a line in one domain to a point in the other domain. Each trace of $\tilde{u}(\tau, p)$ represents an individual plane wave, or in a vertically varying medium a constant Snell wave since in general they won't be planar.

An intermediate step in the Gaussian beam method is the evaluation of each beam contribution along a specific ray arriving at a given station. Each beam satisfies a local wave equation along each central ray. The final GB solution using the slowness method is a superposition of the beam contributions. Thus

$$u(t, x) = \int_{-1/v}^{1/v} u_{gb}(s, n, t) dp \quad (10)$$

where u_{gb} is an individual beam contribution and the integral is over the ray number measured at the source. In the Gaussian beam method this

integral is usually only over real angles. Using large initial beam widths in a vertically varying medium, then each beam will approach an initial plane wave contribution as in equation (9b).

A simple comparison is now done between a slant stacked wavefield and the individual beam contributions in equation (9b). The velocity model consists of a 15 km homogeneous layer with a velocity of 5.6 km/sec over a linear gradient from 5.6 to 8.0 km/sec. A Gabor wavelet is used with $f = 4 \text{ Hz}$ and $\gamma = 3$. The ray diagram is shown in figure (11a). A reduced time wavefield is shown in figure (11b). In order to use the same format as the Gaussian beam decomposition, each ray number is simply counted from 1 to 18 from shallow to steeply dipping at the source as in figure (12a). A station distance of 120 km is used. Equation (9a) is then applied to wavefield in figure (11b) resulting in a $\tilde{u}(\tau, p)$ wavefield. This plane wave decomposition is shown in figure (12b) in the ray number format for a station at 120 km. The slight phase shift shown on the first several traces may be the result of using a finite x aperture in the slant stack. Filtering this plot with ray number corresponds to a beam forming operation in which energy from certain directions is enhanced with respect to other directions.

Figure (12c) shows the ray number-phase time plot for the Gaussian beam synthesis using a large initial beam width of $L_0 = 128 \text{ km}^{1/2}$. For this initial beam width, some amplitude filtering is still seen for the first

LAYER OVER GRADIENT

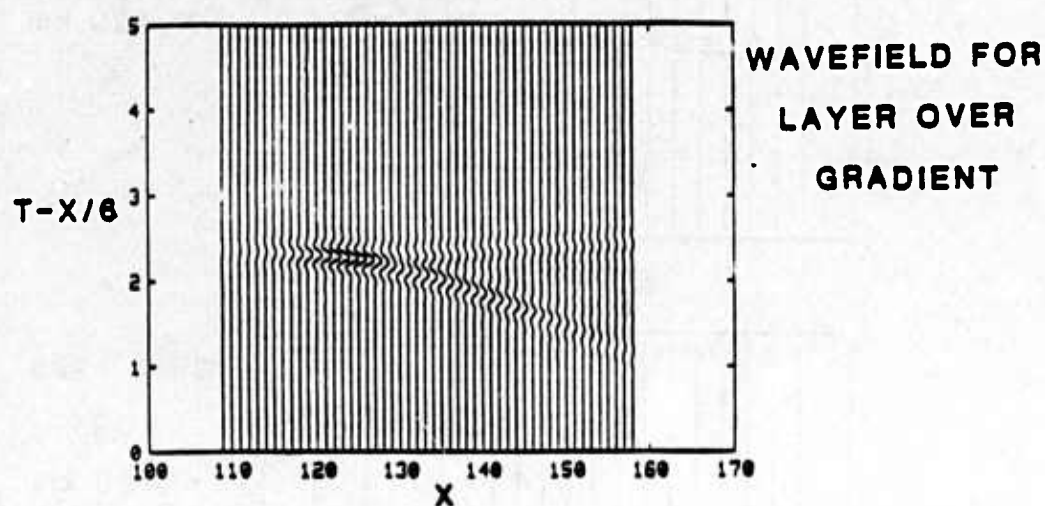


Figure 11: a) ray diagram, and b) reduced time-distance wavefield for a layer over a gradient model.

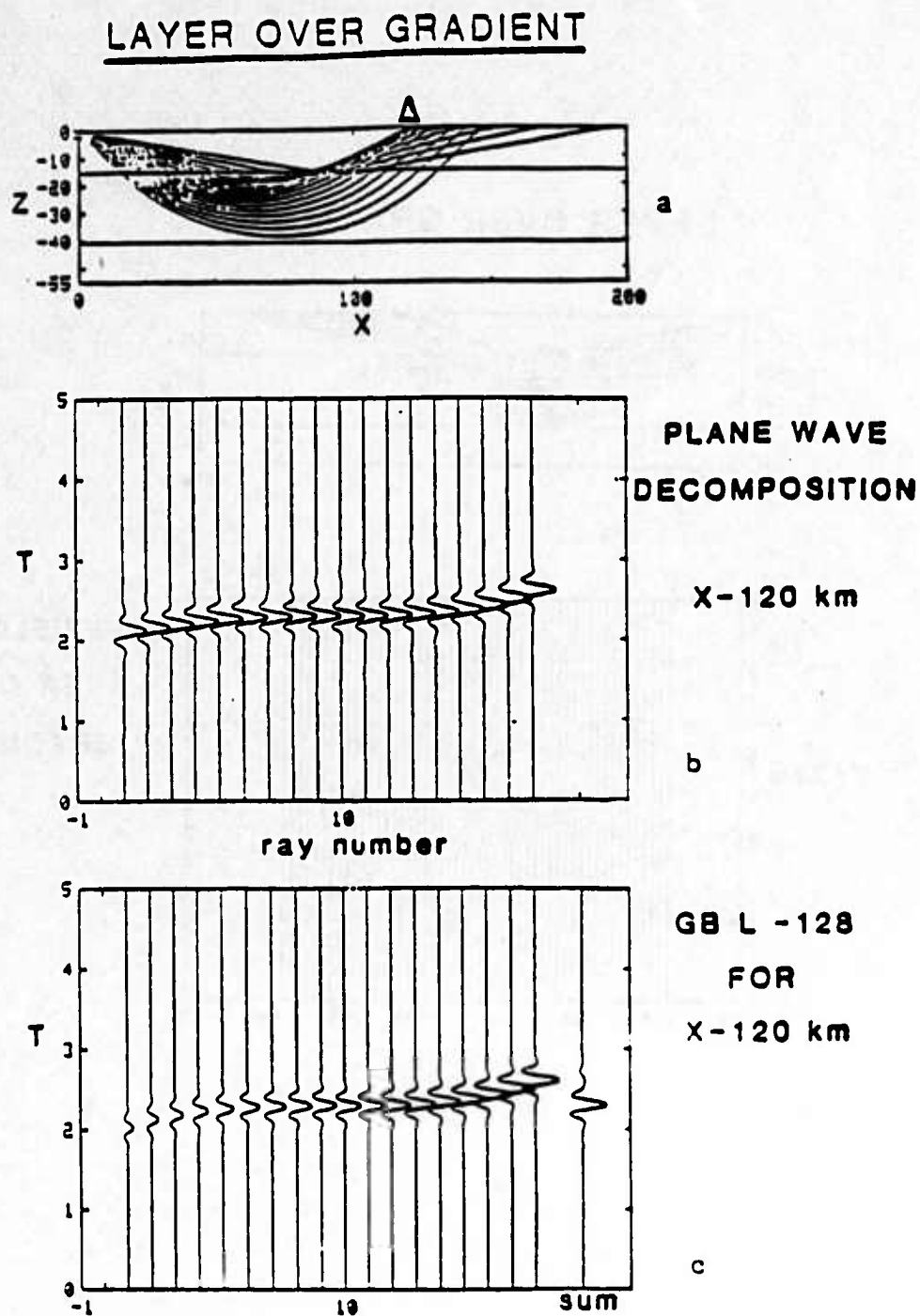


Figure 12 a) ray diagram for ray number-phase time plots for a layer over a gradient model with a station at 120 km. b) slant stack of wavefield in figure (11b). c) Gaussian beam decomposition for $L_0 = 120 \text{ km}^{1/2}$.

several ray number traces. The resulting scaled sum for a range of 120 km is shown on the right of figure (12c) corresponding to the synthesis operation in eqn (10). Comparing figures (12b) and (12c) shows that the locus of maximum amplitudes map out the same curve for the slant stack and the Gaussian beam decomposition. It is interesting that the individual wavelets in figures (12b) and (12c) have similar phase shifts. This results from the approximations in using the Gabor wavelet in the Gaussian beam method, in particular for the Hilbert transform.

Since the slant stacking operation is a linear beam forming operation, an important extension of this is to investigate the variations of the ray parameter-phase time wavefields in laterally varying media. This has been investigated for several structures by Treitel et. al [1982]. Because of its usefulness in vertically varying media, plane wave or beam decompositions are anticipated to be important data representations in more complicated media. This representation is straightforward to generate as a by product of many seismogram synthesis routines, in particular, as ray centered beam responses in the Gaussian beam method.

2.3.9 Q1 CAUSTICS

Previous results have indicated that in decomposing a point source into beams that stable results occur for large initial beam widths. Nowack and Aki [1984] suggested that the critical value of L_0 be used to separate

large from small initial beam widths, where for $S_0 = 0$, $L_{cr} = |q_2/q_1|^{1/2}$. When initial beam widths smaller than critical are used, then phase interference and amplitude decay only weakly restrict the number of beams contributing at a given station. A finite aperture will then result in a reduced amplitude of the resulting seismogram.

A q_1 caustic is defined where $q_1 = 0$ in the wavefield. For a station at a q_1 caustic, then $L_{cr} \rightarrow \infty$. Thus the initial beam width at the source resulting in the narrowest beam width at the receiver approaches infinity, i.e. a plane wave. However, in this case using a slightly smaller initial beam width should still give a satisfactory amplitude, since the aperture of rays will still be large compared to the focused energy at the q_1 caustic. Typically an average value of L_0 is used for the suite of rays in the vicinity of the station. This would then smooth over the beam solutions near the station. Thus at a q_1 caustic, the finite aperture should not pose a problem for a suitably large but finite L_0 . However, a finite L_0 will average over nearby rays and lose some spatial resolution.

In surface wave ray tracing, Yomogida and Aki [1985] found a q_1 caustic along the equator for a high latitude source when using a Mercator mapping from spherical to cartesian coordinates. They found that as long as L_0 was large enough to avoid truncation errors from the aperture, then the amplitude was accurate to 1%. But, an increasing amount of averaging of surrounding beams, without the aide of phase interference, is seen in

figure (4) of Yomogida and Aki [1985]. A slight increase in the beam width at the q_1 caustic to be several times the ray spacing should stabilize the solution. Another possible way of avoiding the problems associated with a q_1 caustic is to use the real part of ε , S_0 , to shift the beam waist along the ray. But, this will also introduce an additional free parameter.

2.3.10 GAUSSIAN BEAMS INCIDENT ON A PLANAR INTERFACE

In this section, the asymptotic analysis of a beam reflected on a planar interface is given, following the development of Ra et. al. [1973]. The saddle point contribution to the reflected field is given by

$$u(x_s, z_s, \omega) \approx \frac{R(\theta)}{\sqrt{k_1 D_r}} e^{ik_1 D_r - i\pi/4} \quad (11)$$

where $k_1 D_r \gg 1$, D_r is the distance from source to receiver, and $R(\theta)$ is a plane wave reflection coefficient. As shown in a previous section, the reflected beam contribution can be gotten by using a complex source point. Using the image source point as in figure (13), then

$$\begin{aligned} x' &\rightarrow ib \sin \alpha \\ z' &\rightarrow z_0 - ib \cos \alpha \end{aligned}$$

and the angle θ is defined from $\sin \theta = \frac{x_s - x'}{D_r}$. When $b \rightarrow 0$, then

$D_r \rightarrow \rho_r$ and $\theta \rightarrow \vartheta_r$ in figure (13).

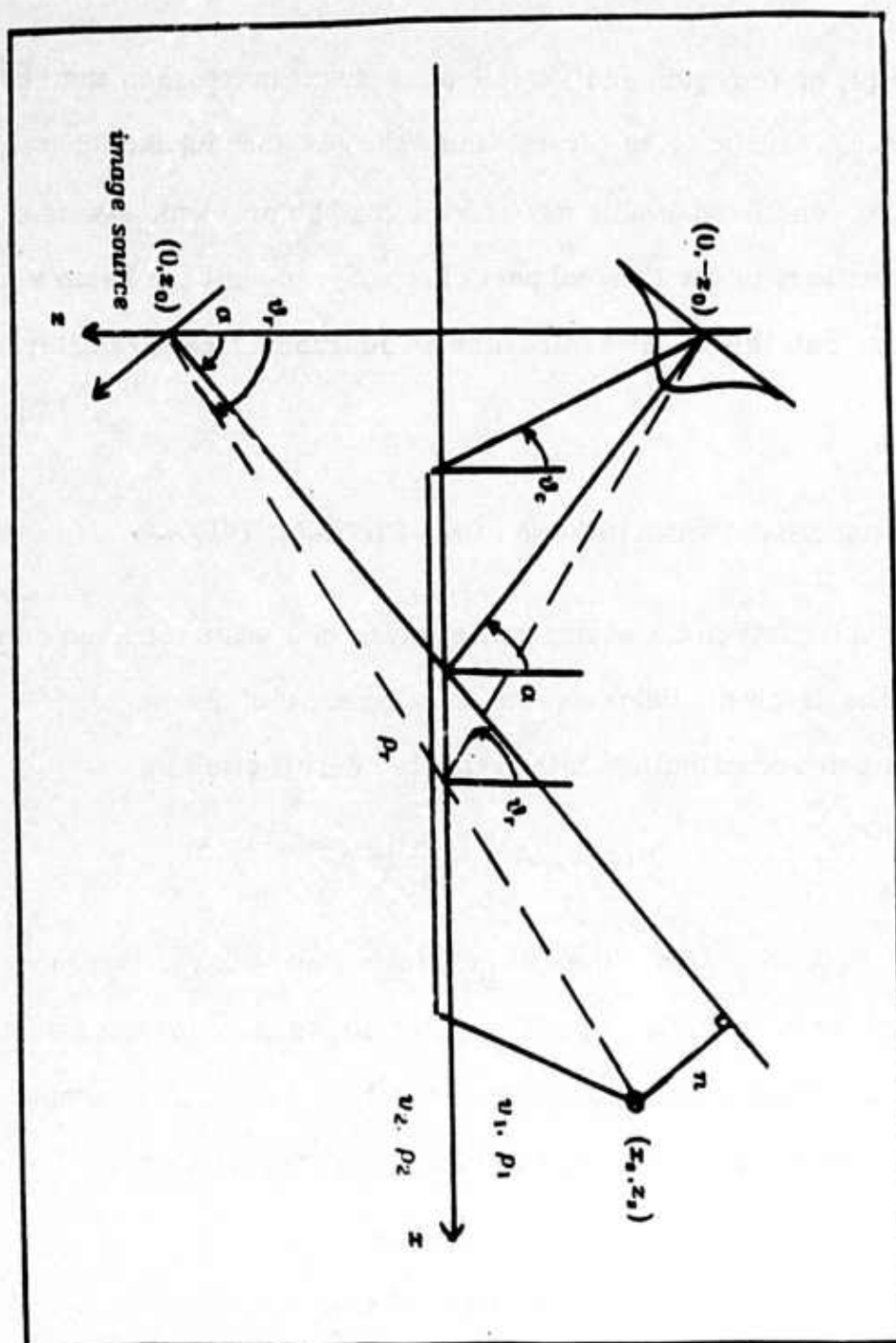


Figure 13 geometry for beam incident on a plane interface. The angle, α , corresponds to the reflection of the beam axis. The angle, θ_r , is the geometric reflection from the source to the station. θ_c is the critical angle.

For $\vartheta > \vartheta_c = \sin^{-1}(v_1/v_2)$, then in addition to the reflected field there will also be a diffracted head wave field. In the vicinity of ϑ_c , a uniform approximation must be used. Going to a Gaussian beam by using a complex source point requires using the appropriate complex angles in these formulas. Here only the reflected field will be investigated. In the Gaussian beam method as currently implemented, the real reflection coefficient for the beam axis is used, and all ray tracing is done along the beam axis. Thus in the reflection coefficient, Θ is replaced by α .

The reflection coefficient in the acoustic case has a slightly different form than that for the SH case. The reflection coefficient in the acoustic case is given by

$$R_a(\vartheta) = \frac{\rho_2 v_2 \cos \vartheta - \rho_1 v_1 \left[1 - \frac{v_2^2}{v_1^2} \sin^2 \vartheta \right]^{1/2}}{\rho_2 v_2 \cos \vartheta + \rho_1 v_1 \left[1 - \frac{v_2^2}{v_1^2} \sin^2 \vartheta \right]^{1/2}}$$

For the SH case, the reflection coefficient is given by

$$R_{SH}(\vartheta) = \frac{\rho_1 v_1 \cos \vartheta - \rho_2 v_2 \left[1 - \frac{v_2^2}{v_1^2} \sin^2 \vartheta \right]^{1/2}}{\rho_1 v_1 \cos \vartheta + \rho_2 v_2 \left[1 - \frac{v_2^2}{v_1^2} \sin^2 \vartheta \right]^{1/2}}$$

This difference in reflection coefficient is related to the different type of impedance in the acoustic and the SH cases (see for example Aki and Richards [1980], Box 5.4, or Aki [1970], lecture notes pp. 60-61)

It is assumed that the beam is well defined with $k_1 b \gg 1$. In addition, the paraxial region is considered with $\vartheta_r \approx \alpha$. The reflection coefficient is then approximated as

$$\begin{aligned} R(\theta) &\approx R(\alpha) + R'(\vartheta)|_{\vartheta=\alpha} \Delta\vartheta \\ R(\theta) &\approx R(\alpha) \left\{ 1 + \frac{R'(\vartheta)|_{\vartheta=\alpha}}{R(\vartheta)} \Delta\vartheta \right\} \\ R(\theta) &\approx R(\alpha) e^\nu \end{aligned}$$

where ν is a complex phase correction which has the effect of relocating the maximum of the reflected field (see Ra et. al. [1973], Felsen [1984]).

In the acoustic case, then

$$\frac{R'(\vartheta)|_{\vartheta=\alpha}}{R(\alpha)} \approx \frac{2 \sin \alpha \chi_a}{\left[1 - \frac{v_2^2}{v_1^2} \sin^2 \alpha \right]^{1/2}}$$

$$\text{where } \chi_a = \frac{\frac{\rho_2 v_2}{\rho_1 v_1} \left(\frac{v_2^2}{v_1^2} - 1 \right)}{\left[\frac{\rho_2^2 v_2^2}{\rho_1^2 v_1^2} \cos^2 \alpha + \frac{v_2^2}{v_1^2} \sin^2 \alpha - 1 \right]} \quad \text{For } \rho_1 = \rho_2, \text{ then } \chi_a = \frac{v_2}{v_1} \text{ in}$$

the acoustic approximation. In a similar manner, for the SH case

$$\frac{R'(\vartheta)|_{\vartheta=\alpha}}{R(\alpha)} \approx \frac{2 \sin \alpha \chi_{SH}}{\left[1 - \frac{v_2^2}{v_1^2} \sin^2 \alpha \right]^{1/2}}$$

$$\text{where } \chi_{SH} = \frac{\frac{\rho_1 v_1}{\rho_2 v_2} \left(\frac{v_2^2}{v_1^2} - 1 \right)}{\left[\frac{\rho_1^2 v_1^2}{\rho_2^2 v_2^2} \cos^2 \alpha + \frac{v_2^2}{v_1^2} \sin^2 \alpha - 1 \right]} \quad \text{The complex beam shift is}$$

then

$$\nu = \frac{R'(\vartheta)|_{\vartheta=\alpha}}{R(\alpha)} \Delta\vartheta_r \approx \frac{2\sin\alpha \chi}{\left[1 - \frac{v_2^2}{v_1^2} \sin^2\alpha\right]^{1/2}} \Delta\vartheta_r$$

For a real source point

$$\Delta\vartheta_r = \sin^{-1}\left(\frac{n}{\rho_r}\right) \approx \frac{n}{\rho_r}$$

For a complex source point, $\Delta\vartheta_r \rightarrow \hat{\Delta\vartheta}_r$ and $\rho_r \rightarrow D_r \approx (\rho_r - ib)$, then

$$\hat{\Delta\vartheta}_r \approx \frac{n}{\rho_r - ib} = \frac{\rho_r}{\rho_r - ib} \sin\Delta\vartheta_r$$

The formula for the complex beam shift correction, ν , is then given by

$$\nu = 2\sin\alpha \chi \left\{ \frac{\rho_r + ib}{\left[1 - \frac{v_2^2}{v_1^2} \sin^2\alpha\right]^{1/2}} \right\} \frac{\rho_r \sin\Delta\vartheta_r}{(\rho_r^2 + b^2)}$$

As $b \rightarrow 0$, this is a real perturbation. For $b \rightarrow \infty$, then the beam approaches a plane wave and ν goes to 0. The Gaussian beam result is modified by a first order correction to the beam axis reflection coefficient at an interface by a factor, e^ν .

In order to see how the factor e^ν shifts the maximum of the reflected beam amplitude, the maximum of the factor $e^{\frac{-n^2}{L^2} + nB}$ is found where

$$L^2 = \frac{2b}{k} \left(1 + \frac{\rho_r^2}{b^2} \right) \text{ and } B = \frac{2\sin\alpha \chi}{(\rho_r^2 + b^2)} \operatorname{Re} \left\{ \frac{\rho_r + ib}{\left[1 - \frac{v_2^2}{v_1^2} \sin^2\alpha\right]^{1/2}} \right\}. \text{ Then}$$

$$n_{\max} = +L^2 B = \frac{2 \sin \alpha \chi}{kb} \operatorname{Re} \frac{\rho_r + ib}{\left[1 - \frac{v_2^2}{v_1^2} \sin^2 \alpha\right]^{1/2}}$$

or

$$n_{\max} = \frac{2 \sin \alpha \chi}{k} \left(\frac{\rho_r}{b} \right) \frac{1}{\left[1 - \frac{v_2^2}{v_1^2} \sin^2 \alpha\right]^{1/2}} \quad \alpha < \vartheta_c$$

$$n_{\max} = \frac{2 \sin \alpha \chi}{k} \frac{1}{\left[\frac{v_2^2}{v_1^2} \sin^2 \alpha - 1\right]^{1/2}} \quad \alpha > \vartheta_c$$

This is analogous to the formulas given in Ra et. al. [1973]. Note that the beam shift goes to zero for $k \rightarrow \infty$. This is sometimes used to justify neglecting this factor in high frequency beam propagation.

As a simple example, consider a surface source in a 35 km thick crust with $\beta_1 = 3.5 \text{ km/sec}$ and $\rho_1 = 2.74 \text{ g/cm}^3$ and $\beta_2 = 4.6 \text{ km/sec}$ and $\rho_2 = 3.3 \text{ g/cm}^3$. The critical angle, ϑ_c , is 49.54 degrees and a critical distance of 82. km. Let angle, α of the beam axis be 20. degrees with a range of 25.47 km. The angle, ϑ_r , corresponding to a range of 40 km is 29.74 degrees. ρ_r is 80.62 km. The critical value of $b = L_0^2 = |q_2/q_1| = \rho_r$ is used. For $f = 5 \text{ hz}$, $n_{\max} = -0.9 \text{ km}$, or about an eighth of a wavelength. For larger b , this value would be correspondingly less in the pre-critical range.

2.3.11 FINITE FREQUENCY EFFECTS OF AN INDENTATION ON AN INTERFACE

In this final section, the scattering effects due to an indentation on an interface is investigated. This was first studied by Larner [1970] for a dented Moho under the Montana large aperture seismic array (LASA). Larner [1970] used the discrete wavenumber technique to compute the amplitude and phase perturbations of a single frequency signal impinging at different angles on various interfaces indentations. He found a significant frequency variation in the scattered signal. Bard and Bouchon [1980] extended this to the analysis of transient signals.

The model geometry is shown in figure (14a). W and h are the width and height of the indentation. The observation plane is 25 km above the interface. The upper half space has $v_1 = 3.0 \text{ km/sec}$ and $\rho_1 = 2.8 \text{ g/cm}^3$, and the lower half space has $v_2 = 4.0 \text{ km/sec}$ and $\rho_2 = 3.3 \text{ g/cm}^3$. An SH wave is incident from below with an angle of 32 degrees from the vertical. A Gabor wavelet is assumed with $\gamma = 4$ and several central frequencies. The peak amplitude of the time domain trace is then picked along the observation plane.

The results using the Gaussian beam method are compared with the time domain implementation of the Aki-Larner discrete wavenumber technique (see Bard and Bouchon [1980]). The peak amplitudes along the observation plane are shown in figure (14). Figures (14b-e) show the amplitude patterns for an increasing indentation height h from 0 km to 5

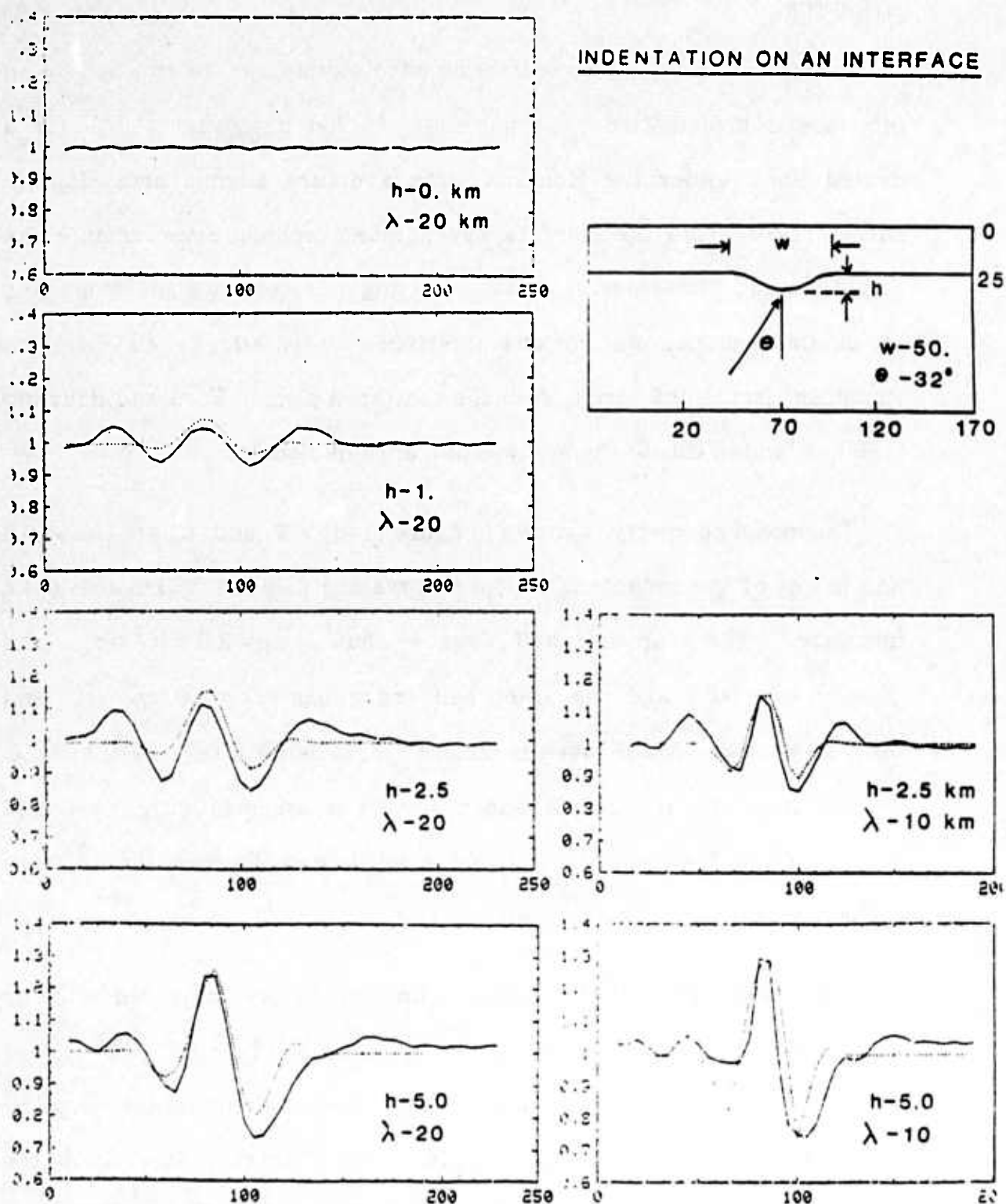


Figure 14 Amplitude patterns for a plane wave incident on an indented interface. a) model geometry. b-e) $\lambda = 20$ km incident wavelet for indentation height, $h = 0, 1, 2.5, 5.0$ km. f-g) $\lambda = 10$ km incident wavelet for indentation height, $h = 2.5, 5.0$ km.

km for a pulse with a 20 km incident wavelength. The solid line is the amplitude pattern computed using the Gaussian beam method and the dotted line is the amplitude computed using the Aki-Larner discrete wavenumber method. The center of the indentation is located at 70 km. As the indentation height increases, the amplitude pattern increases with an asymmetrical appearance. The Gaussian beam solution consistently shows larger side lobes than the discrete wavenumber solution. The difference may be due to the fact that the Gaussian beam method is a high frequency method and the Aki-Larner method is a low frequency method. The differences in the solutions is probably an artifact of using the Gaussian beam method for the low frequencies of this example. However, the width of the main peak of the amplitude pattern is comparable for the two methods. In figures (14f-g), the amplitude anomaly for the two methods is shown for indentation heights of 2.5 and 5.0 km with an incident wavelength of 10 km. The main amplitude peak is narrower for this case than for the 20 km wavelength case and is also higher in amplitude. There is thus a frequency dependence of the scattering anomaly. This was also observed by Larner [1970]. A simple use of ray theory for an example such as this would never reveal the frequency dependent nature of the scattering. The high but still finite frequency methods like the Gaussian beam method can be used to bring out some of these finite frequency effects.

2.3.12 CONCLUSIONS

Several topics in Gaussian beam synthesis have been investigated in this paper including the choice of the beam parameter, ϵ , and its relation to a complex source point. The asymptotic and exact evaluations to the Gaussian beam integral shows an additional singularity which goes to infinity as $\epsilon \rightarrow \infty$. But, this typically won't be computed in the usual Gaussian beam synthesis. An initial point source in a vertically varying gradient distinguishes between an initial Snell wave, which is singular for a horizontal ray, and an initial plane wave. In a homogeneous source region these are the same, but in an inhomogeneous source region each has different advantages. A first order correction to the plane wave reflection coefficients are then described. This results in a beam shift which goes to zero for high frequency beam propagation. A simple example of a 35 km layer shows a modest beam shift for a 5 hz signal. Finally, the finite frequency scattering effects for an indented interface are computed using the Gaussian beam method and shows reasonable agreement with the effects computed using the Aki-Larner method.

Further work using the Gaussian beam method include investigation of the real part of the complex beam parameter, ϵ , which can be used to shift the beam waist presumably to where the strong heterogeneity is located. For small, continuously distributed heterogeneities, a multiple beam Kirchhoff scheme would be one possible approach. This would involve

decomposing the wavefield at regular intervals. Finally, the development of a practical implementation of the full 3-D Gaussian beam formalism will be important in synthetic seismogram modelling.

2.3.13 REFERENCES

- Aki, K., and P.G. Richards, *Quantitative Seismology: Theory and Methods*, W.H. Freeman and Company, San Francisco, 1980.
- Archuleta, R.J., Analysis of near-source static and dynamic measurements from the 1979 Imperial Valley earthquake, *Bull. Seism. Soc. Am.*, 72, 1927-1956, 1982.
- Bard, P.-Y., and M. Bouchon, The seismic response of sediment-filled valleys. Part 1. The case of incident SH waves, *Bull. Seism. Soc. Am.*, 70, 1263-1286, 1980.
- Ben-Menahem, A., and W.B. Beydoun, Range of validity of seismic ray and beam methods in general inhomogeneous media, Part I: General theory, preprint, 1985.
- Bessonova, E.N., V.M. Fishman, V.Z. Ryabayi, and G.A. Sitnikova, The Tau method for inversion of travel times - I. deep seismic sounding data, *Geophys. J.*, 36, 377-398, 1974.
- Bessonova, E.N., V.M. Fishman, M.G. Shirman, G.A. Sitnikova, and L.R. Johnson, The Tau method for inversion of travel times - II. earthquake data, *Geophys. J.*, 46, 87-108, 1976.
- Beydoun, W.B., and A. Ben-Menahem, Range of validity of seismic ray and beam methods in general inhomogeneous media, Part II: a canonical problem, preprint, 1985.
- Bouchon, M., and K. Aki, Discrete wavenumber representation of seismic-source wavefields, *Bull. Seism. Soc. Am.*, 67, 259-277, 1977.
- Cerveny, V., Seismic wave fields in structurally complicated media (ray and Gaussian beam approaches), Lecture Notes, Rijksuniversiteit Utrecht, Vening-Meinesz Laboratory, Utrecht, 1981.
- Cerveny, V., Dynamic ray tracing in 2-D media and dynamic ray tracing across curved interfaces, Stanford Exploration Project no. 28, Stanford Univ., California, 1981.
- Cerveny, V., Expansion of a plane wave into Gaussian beams, *Studia Geophys. Geod.*, 26, 120-131, 1982.
- Cerveny, V., Synthetic body wave seismograms for Laterally varying layered structures by the Gaussian beam method, *Geophys. J.R. Astron. Soc.*,

73, 389-426, 1983.

Cerveny, V., The application of ray tracing to the propagation of shear waves in complex media, in Seismic Exploration, eds. Treitel, S. and K. Helbig, Vol. Seismic Shear Waves, eds. Dohr, G., Geophysical Press, in press, 1985.

Cerveny, V., and F. Hron, The ray series method and dynamic ray tracing system for three-dimensional inhomogeneous media., Bull. Seism. Soc. Am., 70, 47-77, 1980.

Cerveny, V., M.M. Popov, and I. Psencik, Computation of wave fields in inhomogeneous media - Gaussian beam approach, Geophys. J.R. Astron. Soc., 70, 109-128, 1982.

Chapman, C.H., A new method for computing synthetic seismograms, Geophys. J.R. Astron. Soc., 54, 481-513, 1978.

Chapman, C.H., Generalized Radon transform and slant stacks, Geophys. J.R. astr. Soc., 66, 445-453, 1981.

Clayton, R.W., and G.A. McMechan, Inversion of refraction data by wavefield continuation, Geophysics, 46, 860-868, 1981.

Deschamps, G.A., Gaussian beams as a bundle of complex rays, Electronics Lett., 7, 684,685, 1971.

Felsen, L.B., Complex-source-point solutions of the field equations and their relation to the propagation and scattering of Gaussian beams, Symp. Matemat., XVIII, 40-56, Istituto Nazionale di Alta Matematica, Academic Press, London, 1976.

Felsen, L.B., Geometrical theory of diffraction, evanescent waves, complex rays and Gaussian beams, Geophys. J.R. astr. Soc., 79, 77-88, 1984.

Fuis, G.S., W.D. Mooney, J.H. Healy, G.A. McMechan, and W.J. Lutter, The Imperial Valley earthquake of October 15, 1979. crustal structure of the Imperial Valley region, U.S. Geol. Surv. Profess. Paper 1254, 1982.

Keller, J.B., and W. Streifer, Complex rays with an application to Gaussian beams, J. Opt. Soc. Am., 61, 40, 1969.

Larner K.L., Near-receiver scattering of teleseismic body waves in layered crust mantle models having irregular interfaces, Ph.D. thesis MIT, Cambridge, 1970.

- Madariaga, R., Gaussian beam synthetic seismograms in a vertically varying media, *Geophys. J.R. astr. Soc.*, 79, 589-612, 1984.
- Madariaga, R., and P. Papadimitriou, Gaussian beam modelling of upper mantle phases, preprint, 1985.
- McMechan, G.A., and W.D. Mooney, Asymptotic ray theory and synthetic seismograms for laterally varying structures: theory and application to the Imperial Valley, California, *Bull. Seism. Soc. Am.*, 70, 2021-2035, 1980.
- McMechan, G.A., and R. Ottolini, Direct observation of a p-tau curve in a slant stacked wavefield, *Bull. Seism. Soc. Am.*, 70, 775-789, 1980.
- Nowack, R., and K. Aki, The 2-D Gaussian beam synthetic method: testing and application, *J. Geophys. Res.*, 89, 7797-7819, 1984.
- Pott, J., and J.G. Harris, Scattering of an acoustic Gaussian beam from a fluid-solid interface, *J. Acoust. Soc. Am.*, 76, 1829-1838, 1984.
- Ra, J.W., H.L. Bertoni, and L.B. Felsen, Reflection and transmission of beams at a dielectric interface, *SIAM J. appl. Math.*, 24, 396-413, 1973.
- Shin S.Y., and L.B. Felsen, Gaussian beam modes by multipoles with complex source points, *J. Opt. Soc. Am.*, 67, 699-700, 1977.
- Shin S.Y., and L.B. Felsen, Lateral shifts of totally reflected Gaussian beams, *Radio Science*, 12, 551-564, 1977.
- Sinton, J.B., and L.N. Frazier, A method for the computation of finite frequency body wave synthetic seismograms in laterally varying media, *Geophys. J.R. astr. Soc.*, 71, 37-55, 1982.
- Treitel, S., P.R. Gutowski, and D.E. Wagner, Plane-wave decomposition of seismograms, *Geophysics*, 47, 1375-1401, 1982.
- Walck, M.C., and R.W. Clayton, Analysis of upper mantle structure using wavefield continuation of P waves, *Bull. Seism. Soc. Am.*, 70, 1263-1286, 1980.
- Wu, Ru-shan, Analytic extension of Green's function, Gaussian beams and complex rays in smoothly inhomogeneous media, preprint, 1985.
- Yomogida, K., and K. Aki, Total waveform synthesis of surface waves in laterally heterogeneous Earth by Gaussian beam method, accepted to *J. Geophys. Res.*, 1985.

2.4 ITERATIVE INVERSION FOR VELOCITY USING WAVEFORM DATA

2.4.1 INTRODUCTION

In this paper, an iterative inversion method for earth structure is investigated. Previous inversions for velocity using array data have mostly used travel times in constructing the tomography problem (see Aki [1977], Aki et.al. [1977]). Recently iterative methods, such as the algebraic reconstruction method, have been applied to travel time inversions (see McMechan [1983], Clayton and Comer [1983], Humphreys et.al. [1984]). Several of these iterative algorithms are described by Censor [1981]. The iterative methods have advantages over generalized inversion methods when dealing with large sparse matrices, but can have convergence problems when small eigenvalues are present (Ivansson [1983]). Still, there is much more information in seismic data than just first arrival travel times, including amplitudes and waveforms. Figure 1 shows a sketch of a heterogeneous region with surrounding sources and receivers. A simple transmitted waveform is shown.

A number of studies have incorporated amplitude and waveform data in seismic inversion and interpretation. For example, spatial variations of phase time and amplitude from teleseismic body waves recorded at the Montana LASA were interpreted by Larner [1970] in terms of a dented Moho. Haddon and Husebye [1979] and Thomson and Gubbins [1982] did joint interpretations of travel time and amplitude to infer the structure beneath NORSAR. Waveform inversion using refraction data assuming a

vertically varying structure has been done by Shaw [1983]. Brown [1984] did a comparison of travel time inversions and waveform inversions using synthetic data also in a vertically varying structure. Migration of reflection seismic data is a form of inversion (see Schneider [1978]), and it has been recently shown to be kinematically similar to a linearized Born inversion (see Tarantola [1984b], Miller et.al. [1984]). Direct inversions based on the linearized model have been investigated by Cohen and Bleistein [1979] and Raz [1981]. There has also been some progress in generalizing exact 1-D inversion methods to higher dimensions (see Newton [1983]).

A recent formulation by Tarantola [1984a] uses an iterative linearized approach and is described in the next section. A linear sensitivity operator must be derived, and this can be done economically by using reciprocity of the Green's function. Several descent algorithms are described which avoid a large matrix inversion. A fast and reasonably accurate forward modelling scheme is required, and here we make use of the Gaussian beam method in a slowly varying medium. Data errors and a priori information are incorporated using covariance operators. Different types of linearizations are possible including a linearization in terms of the field, the Born approximation, and a linearization in terms of the log of the field, the Rytov approximation. Their relative merits are discussed. Finally several numerical examples are performed using a linearization in terms of the field and transmitted waveforms in order to test the method.

2.4.2 LINEARIZED INVERSION

In this section a review of the linearized inverse approach is given following Madden [unpublished lecture notes], Tarantola [1981a, 1981b, 1984c], and Lailly [1983]. In the linearized approach, first a linear sensitivity or Frechet derivative operator must be derived and then used in an iterative procedure for the model parameters. In addition, several descent algorithms are described which don't require the inversion of a large matrix.

The forward problem can be written in the form

$$L(p(\vec{r}, t)) = S(\vec{r}, t) \quad (1)$$

where L is a model dependent differential operator, $p(\vec{r}, t)$ is the field variable, and $S(\vec{r}, t)$ is the source term. For example, for the scalar wave equation

$$L = \left[\frac{1}{v(\vec{r})^2} \frac{\partial^2}{\partial t^2} - \nabla^2 \right] \quad (2)$$

The solution of (1) can be written

$$p(\vec{r}, t) = \int dV(\vec{r}') g(\vec{r}', t; \vec{r}) * S(\vec{r}', t) + \text{boundary terms}$$

where $g(\vec{r}', t; \vec{r})$ is the Green's function from $\vec{r}$ to $\vec{r}'$, and $*$ is a time convolution. This can be derived from the bilinear identity [see Lanczos, 1961]. For the scalar wave equation the Kirchhoff integral is the boundary term. In the following, homogeneous boundary conditions will be assumed

giving zero for the boundary term. Finally, for a self adjoint operator L with homogeneous boundary conditions, the Green's function is reciprocal with respect to source and receiver location. Thus

$$g(\vec{r}, t; \vec{r}_g) = g(\vec{r}_g, t; \vec{r})$$

In order to obtain a linear sensitivity operator, a perturbed problem is constructed

$$\left[L + \delta L \right] \left[p + \delta p \right] = S + \delta S \quad (3)$$

or

$$L(\delta p) = -\delta L \left[p + \delta p \right] + \delta S$$

Assuming that $\delta L \delta p$ is small is equivalent to the Born approximation which requires that the perturbed field δp be much smaller than the unperturbed primary field p . To first order, (3) can be written

$$L(\delta p) = -\delta L \left[p \right] + \delta S \quad (4)$$

where $-\delta L(p)$ is an equivalent source term for the medium perturbations and δS is the term for the source perturbations. Here we will investigate the "inverse medium problem" for a given source, thus $\delta S = 0$. The solution for the perturbed field can then be written as a space integration over secondary equivalent sources

$$\delta p(\vec{r}_g, t; \vec{r}_s) = -\int dV(\vec{r}) g(\vec{r}, t; \vec{r}_g) * \delta L \left[p(\vec{r}, t; \vec{r}_s) \right] \quad (5)$$

where δL includes the model perturbations, and p is the incident primary field. Equation (5) can be written as

$$\delta p(\vec{r}_g, t; \vec{r}_s) = \frac{\partial p}{\partial L} \delta L = F \delta L \quad (6)$$

where F is a linear sensitivity or Frechet derivative operator, and the operator δL has imbedded within it the model variations. Assuming that the primary field can be written as $p(\vec{r}, t; \vec{r}_s) = g(\vec{r}, t; \vec{r}_g)$, then

$$\delta p(\vec{r}_g, t; \vec{r}_s) = - \int dV(\vec{r}) g(\vec{r}, t; \vec{r}_g) \circ \delta L \left[g(\vec{r}, t; \vec{r}_s) \right]$$

Thus the linear sensitivity operator can be constructed by computing $g(\vec{r}, t; \vec{r}_s)$ which propagates the field from the source $\vec{r}_s$ to each model point $\vec{r}$, and $g(\vec{r}, t; \vec{r}_g)$ which backpropagates the field back from the geophones into the model at $\vec{r}$. This requires the computation of $N_g + N_s$ forward problems evaluated at the interior points of the model where N_g is the number of geophones and N_s is the number of sources

For a given L , the perturbed operator, δL , can be separated into operations on the forward field, $p(\vec{r}, t; \vec{r}_s)$, and perturbations of the model parameters. For example for the scalar wave equation

$$\delta p(\vec{r}_g, t; \vec{r}_s) = \int dV(\vec{r}) \left\{ \frac{2}{v^3(\vec{r})} \ddot{p}(\vec{r}, t; \vec{r}_s) \circ g(\vec{r}, t; \vec{r}_g) \right\} \delta v(\vec{r}) \quad (7)$$

where the term in the brackets is the linear sensitivity operator, $F = \frac{\partial p}{\partial v}$.

$\delta v(\vec{r})$ is the velocity model perturbation at $\vec{r}$, and $\ddot{p}(\vec{r}, t; \vec{r}_s)$ is the second time derivative of the primary field computed from the source point to the point $\vec{r}$ in the medium. The linear sensitivity operators for the elastic wave equation with respect to the field are given by Tarantola

[1984c]. The linear sensitivity operators for the acoustic wave equation with respect to the field are given by Tarantola [1984a], and used in a simple 1-D example in Appendix 3.

In the following, the linearized problem (5) will be imbedded within an iterative procedure. The effect of ignoring the $\delta L \delta p$ term in the linearization will result in no multiple interactions with the perturbations within δL , or no "cross talk" between perturbations during each iteration. Multiple interactions from previous iterations are included via the Green's functions. In principle, an iterative strategy should be able handle large velocity contrasts. This approach requires the calculation of Green's functions in inhomogeneous media inexpensively and reasonably accurately. Various forward modelling approaches could be considered including ray theoretical methods, the discrete wavenumber method, or the finite difference or finite element methods. Any inversion method is only as good as the forward modelling scheme on which it is based.

The forward problem as a function of model parameters is in general nonlinear function and can be written

$$\vec{p} = f(\vec{m}) \quad (8)$$

where $\vec{p}$ is the field and $\vec{m}$ is the model parameters. In general, $\vec{p}_{obs} \neq f(\vec{m}_{prior})$, where $\vec{p}_{obs}$ is the observed field and $\vec{m}_{prior}$ is the a priori model based on previous information. The objective is to find the combined vector $\begin{bmatrix} \vec{p} \\ \vec{m} \end{bmatrix}$ that satisfies $\vec{p} = f(\vec{m})$ and is closest to

$[\vec{p}_{obs}, \vec{m}_{prior}]$ in some sense. This distance could be defined in various ways depending on the noise structure of the problem, including the L_2 norm, the L_1 norm, or the L_∞ norm. Here the weighted L_2 norm will be considered with its associated induced inner product. This results in a least squares formulation. Thus we want to minimize the functional [see Tarantola and Valette, 1982, Tarantola, 1984a]

$$S(\vec{m}) = \frac{1}{2} \left\{ |\vec{p}_{obs} - f(\vec{m})|^2 + |\vec{m} - \vec{m}_{prior}|^2 \right\} \quad (9)$$

where the $1/2$ has been introduced for later convenience. The linear sensitivity operator, F_k at $\vec{m}_k$ can then be defined as

$$f(\vec{m}_k + \delta\vec{m}) = f(\vec{m}_k) + F_k \delta\vec{m}_k + O(|\delta\vec{m}_k|^2)$$

The negative of the gradient of $S(\vec{m}_k)$ gives the local direction of maximum descent at $\vec{m}_k$. To first order this can be written

$$\vec{\gamma}_k = -\nabla S(\vec{m}_k) = \left\{ F_k^* \delta\vec{p} - (\vec{m}_k - \vec{m}_{prior}) \right\} \quad (10)$$

where $\delta\vec{p}_k = \vec{p}_{obs} - f(\vec{m}_k)$, and F_k^* is the adjoint of the linear sensitivity operator, F_k . The adjoint is defined from the bilinear identity

$$\langle \delta\vec{p}, F \delta\vec{m} \rangle_{\delta p} = \langle F^* \delta\vec{p}, \vec{m} \rangle_{\delta m} = \text{boundary term} \quad (11)$$

where $\langle, \rangle$ is a defined inner product. The boundary term is assumed to be zero using homogeneous boundary conditions. Thus, F is a linear mapping from model to data space, and the adjoint, F^* , is a linear

mapping from data space back to model space. Inverse operators are in general difficult to construct, but adjoint operators through the bilinear identity are straightforward to obtain. In addition, adjoint operators have very useful properties which aide in the construction of inverse and generalized inverse operators.

Consider a linear problem, $F\vec{m} = \vec{p}_{obs}$, with noisy observations but where the nullspace of F is zero, $N(F) = \vec{0}$. The nullspace, $N(F)$, represents the subspace of the model space that satisfies $F\vec{m} = \vec{0}$. The standard least squares solution results by using the adjoint theorem, $R^\perp(F) = N(F^*)$, where R is the range of $F\vec{m}$ in data space. Thus mapping $F\vec{m}$ and $\vec{p}_{obs}$ through F^* annihilates the component of $\vec{p}_{obs}$ in $R(F)$ resulting in a consistent set of equations. Then, $F^*F\vec{m} = F^*\vec{p}_{obs}$ or $\vec{m} = (F^*F)^{-1}F^*\vec{p}_{obs}$. For $N(F) = \vec{0}$, (F^*F) is full rank and an inverse exists. The standard least squares solution thus results from two operations on the data, $\vec{p}_{obs}$. First, the data is operated on by the adjoint operator, F^* , which projects $\vec{p}_{obs}$ from data space to model space. This blurred image in model space is then filtered with the operator $(F^*F)^{-1}$. When $N(F) \neq \vec{0}$, a generalized inverse can be computed by using additional adjoint theorems. This is equivalent to the Lanzcos formulation which results in a minimum-norm least-squares solution [Aki and Richards, 1980]. An alternative is to stabilize the (F^*F) operator using a maximum likelihood procedure which reduces to minimizing a functional like (9) with appropriate data and model covariance operators defined. Any a priori

knowledge about the component of the model in the null space can be incorporated into the final solution. This might include smoothness of the solution.

The maximum likelihood algorithm of Tarantola and Valette [1982] is a linear iterative procedure for solving the nonlinear inverse problem, (8), by minimizing the functional (9). It can be written

$$\vec{m}_{k+1} = \vec{m}_k + [H_k]^{-1} \left\{ F_k^* \delta \vec{p} - (\vec{m}_k - \vec{m}_{prior}) \right\} \quad (12)$$

where the term in the brackets is the negative gradient of the functional $S(\vec{m})$, and

$$H_k = [I + F_k^* F_k] \quad (13)$$

Covariance operators can be introduced by a suitable definition of the data and model space scalar products [see Tarantola, 1984c]. Thus

$$\begin{aligned} \langle \vec{m}_1, \vec{m}_2 \rangle_m &= \vec{m}_1^T C_m^{-1} \vec{m}_2 \\ \langle \vec{p}_1, \vec{p}_2 \rangle_p &= \vec{p}_1^T C_p^{-1} \vec{p}_2 \end{aligned}$$

where $\vec{m}^T$ is the transpose of $\vec{m}$, C_m is the model covariance operator, and C_p is the data covariance operator. From this definition of the scalar products, the adjoint operator, F_k^* , of F_k is related to the transpose by

$$F_k^* = C_m F_k^T C_p^{-1} \quad (14)$$

The adjoint is equal to the transpose when $C_m = I$ and $C_p = I$. For a complex operator F , the transpose is replaced by the conjugate transpose operator.

A major problem with the iterative scheme (12) is that in discretized form it requires the inversion of a possibly large matrix. Since the goal is to iteratively solve the original problem, (8), only an approximate solution of the the imbedded linearized problem may be required. Thus a simple iterative scheme could be written

$$\bar{m}_{k+1} = \bar{m}_k + \alpha_k W_k \bar{\gamma}_k \quad (15)$$

where $\bar{\gamma}_k$ is the negative gradient of the functional $S(\bar{m})$, W_k is called a preconditioner, and α_k is a scalar that approximately solves a 1-D line search of $S(\bar{m})$ in the direction $W_k \bar{\gamma}_k$. α_k is thus chosen to minimize $S(\bar{m}_k + \alpha_k W_k \bar{\gamma}_k)$ with respect to α_k giving

$$\alpha_k = \frac{\langle W_k \bar{\gamma}_k, \bar{\gamma}_k \rangle}{\langle W_k \bar{\gamma}_k, H_k W_k \bar{\gamma}_k \rangle} \quad (16)$$

where H_k is given in (13). For the maximum likelihood algorithm, (11), $W_k = H_k^{-1}$ giving $\alpha_k = 1$. This would then be similar to a modified Newton's method for solving the nonlinear problem, (8), dependent on the choice of the covariance operators used.

Various simplifying choices for W_k could be considered. Choosing $W_k = I$ results in the steepest descent method. Thus

$$\bar{m}_{k+1} = \bar{m}_k + \alpha_k \bar{\gamma}_k \quad (17)$$

where

$$\alpha_k = \frac{\langle \tilde{\gamma}_k, \tilde{\gamma}_k \rangle}{\langle \tilde{\gamma}_k, H_k \tilde{\gamma}_k \rangle}$$

with H_k given in (13). For a quadratic $S(m)$, $\tilde{\gamma}_{k+1}$ and $\tilde{\gamma}_k$ will be perpendicular. One thus moves in the direction $\tilde{\gamma}_k$ until tangent to a level curve of $S(\tilde{m})$. At this point, $\tilde{\gamma}_{k+1}$ is chosen perpendicular to the level curve in a direction of maximum descent.

Another possible choice for W_k is [see Tarantola, 1984a]

$$W_k = \left[\text{DLAG}(H_k) \right]^{-1} \quad (18)$$

This is similar to a modified Jacobi method with suitably defined covariance operators.

In any of the preconditioning strategies, the single step convergence properties for quadratic functionals, $S(\tilde{m})$, are governed by the difference in the smallest and largest eigenvalues of $W_k H_k$ with H_k given in (13). From Luenberger(1984),

$$S(\tilde{m}_{k+1}) \leq \left[\frac{\lambda_{\max} - \lambda_{\min}}{\lambda_{\max} + \lambda_{\min}} \right]^2 S(\tilde{m}_k)$$

Thus the closer W_k is to H_k^{-1} , the better the single step convergence. A general overall strategy is to construct a preconditioning scheme that is easy to compute and possesses a favorable eigenvalue structure at each step.

Instead of using the gradient directions in the steepest descent approach, improved convergence for very little extra effort can be gotten by using so called conjugate gradient directions defined with respect to a new inner product. $\langle x, x \rangle = x^T H_k x$ where H_k is given in (13). The Fletcher-Reeves implementation of the conjugate gradient method to nonquadratic problems is briefly outlined. First, given $\bar{m}_{prior}$ and $\bar{\gamma}_0 = -\nabla S(\bar{m}_{prior})$, set $\bar{d}_0 = \bar{\gamma}_0$. Now let $\bar{m}_{k+1} = \bar{m}_k + \alpha_k \bar{d}_k$ where α_k minimizes $S(\bar{m}_k + \alpha_k \bar{d}_k)$. The next direction of descent, $\bar{d}_{k+1}$ is chosen to be H_k -orthogonal to the previous direction $\bar{d}_k$. This is done by setting $\bar{d}_{k+1} = \bar{\gamma}_{k+1} + \beta_k \bar{d}_k$, where $\beta_k = \langle \bar{\gamma}_{k+1}, \bar{\gamma}_{k+1} \rangle / \langle \bar{\gamma}_k, \bar{\gamma}_k \rangle$. $\bar{d}_{k+1}$ is thus a direction modified from the steepest descent direction. In a similar fashion to steepest descent, $\bar{d}_{k+1}$ and $\bar{d}_k$ are orthogonal but with respect to H_k . The conjugate gradient method has the important property that the estimate $\bar{m}_k$ minimizes a quadratic functional $S(m)$ over a subspace spanned by all the previous directions $(\bar{d}_0, \dots, \bar{d}_{k-1})$ [see Luenberger, 1984].

All of the above methods require the the linear sensitivity operator F_k and the adjoint of this operator F_k^* . For example for the scalar wave equation (2)

$$F_k = \left\{ \frac{2}{v^3(r')} p(r', t, r_s) q(r', t, r_g) \right\} = \frac{\partial p}{\partial v} \quad (19)$$

where velocity v is the model parameter. Thus to first order $\delta p = \frac{\partial p}{\partial v} \delta v$

Using the bilinear identity, (11), with $C_m = I$ and $C_p = I$ and homogeneous boundary conditions, then $F^* = F^T$ and

$$\begin{aligned} \langle \delta p, F \delta v \rangle_p &= \sum_{r_s} \sum_{r_g} \int dt \delta p F(\delta v) \\ &= \sum_{r_s} \sum_{r_g} \int dt \delta p(r_g, t; r_s) \int dV(r') \left\{ \frac{2}{v^3(r')} \ddot{p}(r', t; r_s) * g(r', t; r_g) \right\} \delta v(r') \end{aligned}$$

Then,

$$\begin{aligned} \langle F^* \delta p, \delta v \rangle_v &= \int dV(r') F^*(\delta p) \delta v \\ &= \int dV(r') \left\{ \sum_{r_s} \sum_{r_g} \int dt \left(\frac{2}{v^3(r')} \ddot{p}(r', t; r_s) * g(r', t; r_g) \right) \delta p(r_g, t; r_s) \right\} \delta v(r') \end{aligned}$$

Thus F_k^* is just F_k with summations now over geophone and source locations and an integration in time.

The final adjoint operator includes the data and model covariance operators. Assuming uncorrelated errors over source and receiver locations and time, then

$$C_p(\vec{r}_g, t; \vec{r}_s, \vec{r}_g, t; \vec{r}_s) = \sigma_{gs}^2 \delta_{gg} \delta_{ss} \delta_{tt} \quad (20)$$

where σ_{gs} represents the estimated error in the trace corresponding to the g^{th} receiver and the s^{th} source. For the model covariance operator, a spatial Gaussian correlation is a commonly used choice [see Aki and Richards, 1980; Tarantola, 1984a]. Thus

$$C_v(\vec{r}, \vec{r}') = (2\pi)^{-3/2} \frac{\sigma_v^2}{L^3} \exp \left\{ -\frac{1}{2} \frac{(\vec{r} - \vec{r}')^2}{L^2} \right\} \quad (21)$$

where v is the model parameter $\tilde{m}$, L is the medium correlation length for the Gaussian correlation function, and σ_v represents the estimated departures of $\tilde{m}(\vec{r})$ from $\tilde{m}_{prior}(\vec{r})$. As L goes to zero, then $C_v(\vec{r}, \vec{r}') = \sigma_v^2 \delta(\vec{r} - \vec{r}')$. The model covariance operator will constrain the iterated model parameters as well as act like a spatial smoothing filter. The complete adjoint can be written, $F_k^* = C_v F_k^T C_p^{-1}$, as in (14).

For the scalar wave equation, the preconditioned descent method can be written

$$\tilde{v}_{k+1} = \tilde{v}_k + \alpha_k W_k \left[C_v F^T \delta \vec{p}' - (\tilde{v}_k - \tilde{v}_{prior}) \right] \quad (22)$$

where

$$\delta \vec{p}' = C_p^{-1} \delta \vec{p}$$

are the weighted data residuals. Now let

$$\delta \tilde{v} = C_v F_k^T \delta \vec{p}' \quad (23)$$

$$\delta \tilde{v} = C_v \sum_{\vec{r}_g} \sum_{\vec{r}_s} \int dt \left(\frac{2}{v^3(\vec{r})} p(\vec{r}, t; \vec{r}_s) * g(\vec{r}, t; \vec{r}_g) \right) \delta p'(\vec{r}_g, t; \vec{r}_s) \quad (24)$$

This can be rewritten as

$$\delta \tilde{v}(\vec{r}) = C_v \sum_{\vec{r}_s} \sum_{\vec{r}_g} \int dt \frac{2}{v^3(\vec{r})} p(\vec{r}, t; \vec{r}_s) \left\{ g(\vec{r}, -t; \vec{r}_g) * \delta p'(\vec{r}_g, t; \vec{r}_s) \right\} \quad (25)$$

where the term in the brackets corresponds to the propagation of the weighted residuals from the geophones into the model backwards in time. This is then compared with the computed primary field from the sources to

the model point. A summation over all source and receiver locations is then done and an integration over time. Finally, the a priori model covariance operator is applied. This is reminiscent of an imaging principle in which the reflector (velocity perturbation) exists where the downgoing and upgoing waves coincide in time (Claerbout, 1976, Tarantola, 1984b).

In order to implement this algorithm, $p(\vec{r}, t; \vec{r}_g)$ and $g(\vec{r}, t; \vec{r}_g)$ must be computed where $\vec{r}$ is a point in the model (see Figure 2). Since for an inhomogeneous medium an inexpensive forward modelling scheme is required, one possibility is that paraxial ray theory or the Gaussian beam method be used to calculate p and g (see Cervený et al., 1982, Nowack and Aki, 1984). An advantage of using these methods is that no two-point ray tracing is required. In addition, the Gaussian beam method produces a smoothed field with no unphysical singularities in amplitude resulting from caustics which may have adverse effects on an inversion.

Using the ray approximation in 3-D,

$$g(\vec{r}, t; \vec{r}_g) \approx A_{rr_g} \delta(t - T_{rr_g})$$

$$\ddot{p}(\vec{r}, t; \vec{r}_g) \approx A_{rr_g} \ddot{S}(t - T_{rr_g})$$

where A_{rr_g} and T_{rr_g} are the amplitude and travel time computed by ray theoretical methods and $\ddot{S}(t)$ is the second derivative of the source time pulse. For the 2-D case, an additional $\pi/4$ phase shift in the far field must be included. Equation (24) is then

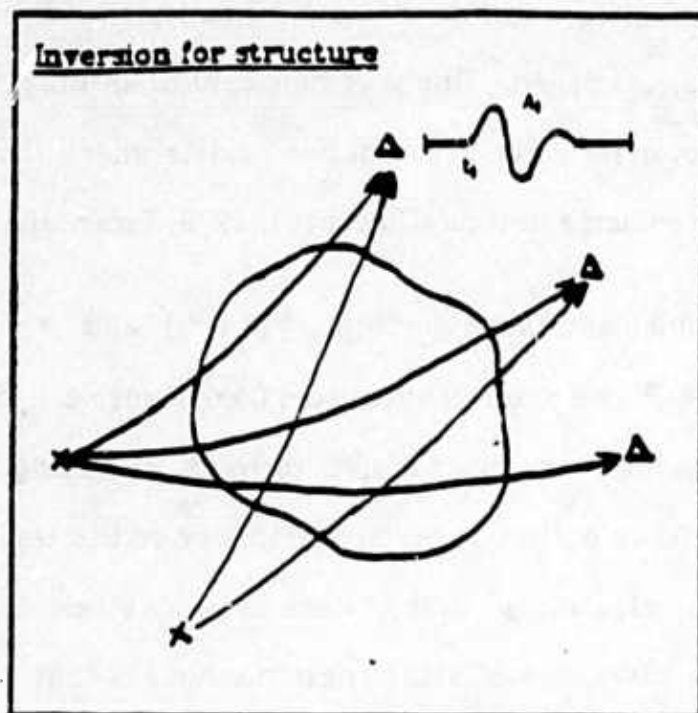


Figure 1

Figure 1 Sketch of heterogeneous body with surrounding sources and receivers. A simple transmitted waveform is shown.

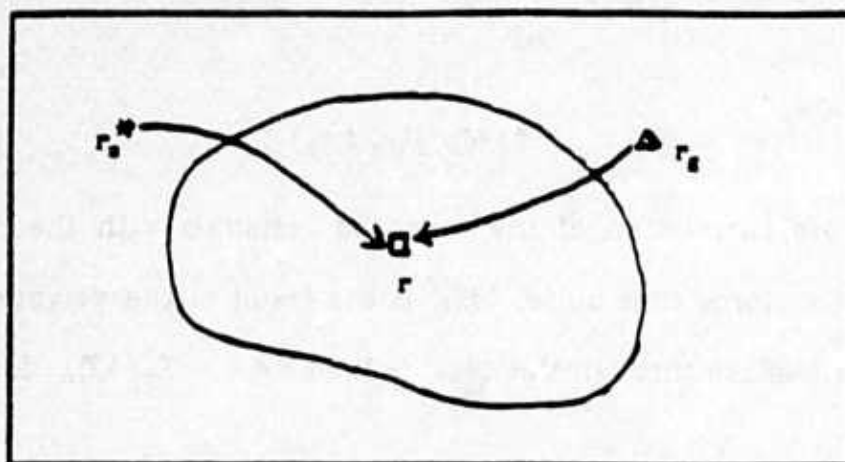


Figure 2

Figure 2 The computation of the linear sensitivity operator requires a forward problem for each source and receiver evaluated at the interior points of the model.

$$\delta v_k = C_v F_k^T \delta p' = C_v \sum_{r_s, r_g} \frac{2}{v^3(\vec{r})} A_{rr_s} A_{rr_g} \int dt \ddot{S}(t - T_{rr_s}) \delta p'(r_g, t + T_{rr_s}, r_s) \quad (26)$$

This can be rewritten as

$$\delta v_k = C_v \sum_{r_s, r_g} \frac{2}{v^3(\vec{r})} A_{rr_s} A_{rr_g} \delta \tilde{p}_k(r_g, t = T_{rr_s} + T_{rr_g}, r_s) \quad (27)$$

where

$$\delta \tilde{p}_k = \ddot{S}(-t) * \delta p'_k(r_g, t; r_s) \quad (28)$$

$\delta \tilde{p}$ is the cross correlation of the weighted residuals with the second derivative of the source time pulse. δv_k is the result of the weighted sum over the phase surface through the data defined by $t = T_{rr_s} + T_{rr_g}$ for each model point $\vec{r}$.

When the a priori starting model is close enough to the true velocity model so that the Born approximation is strictly valid, then $F_k^* = F_0^*$ for all iterations. In addition, if the a priori model is homogeneous, then

$$\delta v_k = C_v F_0^T \delta p' = C_v \sum_{r_s, r_g} \frac{1}{v^3 8\pi^2 R_{rr_s} R_{rr_g}} \delta \tilde{p}(r_g, t = \frac{R_{rr_s} + R_{rr_g}}{v}, r_s) \quad (29)$$

where $\delta \tilde{p}$ is defined above, R_{rr_s} is the distance between source and model point, and R_{rr_g} is the distance between the geophone and the model point. This is similar to a Kirchhoff migration and a cross correlation with the source time pulse [Schneider, 1978; Tarantola, 1984b]. Thus assuming the Born approximation is strictly valid, then a velocity inversion using a descent approach is equivalent to an iterative sequence of Kirchhoff migrations.

A complete iterative step involves a preconditioner, W_k , and a scalar α_k , which solves a 1-D line search. W_k is assumed to be some approximation to $(I + F_k^* F_k)^{-1}$. Approximating F_k by a matrix, then

$$F = \{B_{lj}\}$$

where

$$B_{lj} = \left\{ \frac{2}{v^3(r_j)} \ddot{p}(r_j, t; r_s) \mathfrak{Y}(r_j, t; r_g) \right\}$$

with $l = 1, N_p$ where N_p is the number of data values $= N_t \cdot N_s \cdot N_g$, and $j = 1, N_v$ where N_v is the number of velocity model values. Assuming

$F_k^* = \frac{\sigma_v^2}{\sigma_p^2} F_k^T$ with $C_p = \sigma_p^2 I$ and $C_v = \sigma_v^2 I$, then

$$(H_k)_{ij} = (I + F_k^* F_k)_{ij} = \left\{ \delta_{ij} + \frac{\sigma_v^2}{\sigma_p^2} \sum_{m=1}^{N_p} B_{mi} B_{mj} \right\}$$

Thus

$$(I + F_k^* F_k)_{ij} = \delta_{ij} + \sum_{r_s, r_g} \int dt \frac{4\sigma_v^2}{\sigma_p^2 v^3(r_s) v^3(r_g)} \left\{ \ddot{p}(r_i, t; r_s) \mathfrak{Y}(r_i, t; r_g) \right\} \left\{ \ddot{p}(r_j, t; r_s) \mathfrak{Y}(r_j, t; r_g) \right\} \quad (30)$$

In the ray approximation this can be written

$$(I + F_k^* F_k)_{ij} = \delta_{ij} + \frac{4\sigma_v^2}{\sigma_p^2 v^3(r_i) v^3(r_j)} \sum_{r_s, r_g} A_{r_i, r_s} A_{r_i, r_g} A_{r_j, r_s} A_{r_j, r_g} \tilde{S}(t = T_{r_i, r_s} + T_{r_i, r_g} - T_{r_j, r_s} - T_{r_j, r_g}) \quad (31)$$

where

$$\tilde{S}(t) = \ddot{S}(-t) \bullet \ddot{S}(t)$$

With a high frequency source, this matrix will be sparse. For a diagonally dominant operator, H_k , then an improved single step convergence can be gotten using $W_k = (\text{DIAG}(H_k))^{-1}$, where

$$\text{DIAG}(H_k)_{ii} = \left\{ 1 + \frac{4\sigma_v^2 \tilde{S}(0)}{\sigma_p^2 v^6(r_i)} \sum_{r_s} \sum_{r_g} A_{r_i r_s}^2 A_{r_i r_g}^2 \right\} \quad (32)$$

Thus a simple preconditioned iterative scheme can be written

$$v_{k+1}(r_i) = v_k(r_i) + \alpha_k \left\{ 1 + \frac{4\sigma_v^2 \tilde{S}(0)}{\sigma_p^2 v^6(r_i)} \sum_{r_s} \sum_{r_g} A_{r_i r_s}^2 A_{r_i r_g}^2 \right\}^{-1} \quad (33)$$

$$C_v \sum_{r_s} \sum_{r_g} \frac{2}{v^3(r)} A_{r_i r_s} A_{r_i r_g} \int dt \tilde{S}(t - T_{r_i r_s}) \delta p'(r_g, t + T_{r_i r_s}, r_s)$$

When the a priori model is homogeneous, the starting model is sufficiently close to the true model so that the Born approximation is strictly valid, and with $\sigma_v \rightarrow \infty$, then

$$v_{k+1}(r_i) = v_k(r_i) + \alpha_k 8\pi^2 v^3(r_i) \left\{ \frac{C_v}{\sigma_v^2} \right\} \left\{ \sum_{r_s} \sum_{r_g} \left\{ \frac{1}{R_{r_i r_s}^2} \frac{1}{R_{r_i r_g}^2} \right\} \right\}^{-1} \quad (34)$$

$$\sum_{r_s} \sum_{r_g} \left\{ \frac{1}{R_{r_i r_s}^2} \frac{1}{R_{r_i r_g}^2} \right\} \frac{\sigma_p^2}{\tilde{S}(0)} \delta \tilde{p}(r_g, t = \frac{R_{r_i r_s} + R_{r_i r_g}}{v}, r_s)$$

where $\left\{ C_v / \sigma_v^2 \right\}$ includes the spatial filtering part of the Covariance operator, $\delta \tilde{p}$ is given in equation (28) and is normalized by $\tilde{S}(0) / \sigma_p^2$, and α_k is gotten from equation (16). The factor $\left\{ \sum_{r_s} \sum_{r_g} \left\{ \frac{1}{R_{r_i r_s}^2} \frac{1}{R_{r_i r_g}^2} \right\} \right\}^{-1}$

approximately corrects for geometric spreading. A simple variant of this formula will be used in the examples to follow.

2.4.3 ALTERNATIVE LINEARIZATIONS

The formulation in the previous section was applied to the linearization of the field variables. General limitations to this approach are summarized in appendix 1. A simple example of the limitations in using a linearization in terms of the field can be seen as follows. Consider a field in an unperturbed homogeneous medium (see Brown, [1984]).

$$u_0(r, t) = \frac{\delta(t - R/v_0)}{4\pi R}$$

where $R = |r - r_0|$. In a slightly perturbed homogeneous model with velocity $v_1 = v_0 + \delta v$, the field is

$$u_1(r, t) = \frac{\delta(t - R/v_1)}{4\pi R}$$

The perturbed minus unperturbed field is then approximated by the linearization

$$u_1 - u_0 = \delta u \approx \frac{\partial u}{\partial v} \delta v \quad (34)$$

or

$$\frac{\delta(t - R/v_1)}{4\pi R} - \frac{\delta(t - R/v_0)}{4\pi R} \approx \frac{\partial}{\partial v} \left[\frac{\delta(t - R/v_0)}{4\pi R} \right] \delta v$$

which reduces to

$$\delta(t - R/v_1) - \delta(t - R/v_0) \approx \delta'(t - R/v_0) \left(\frac{R}{v_0^2} \right) \delta v$$

Thus the actual differential seismogram is approximated by a constant times the derivative of a delta function. As seen from Figure 3, depending on the frequency content this will be a poor approximation when

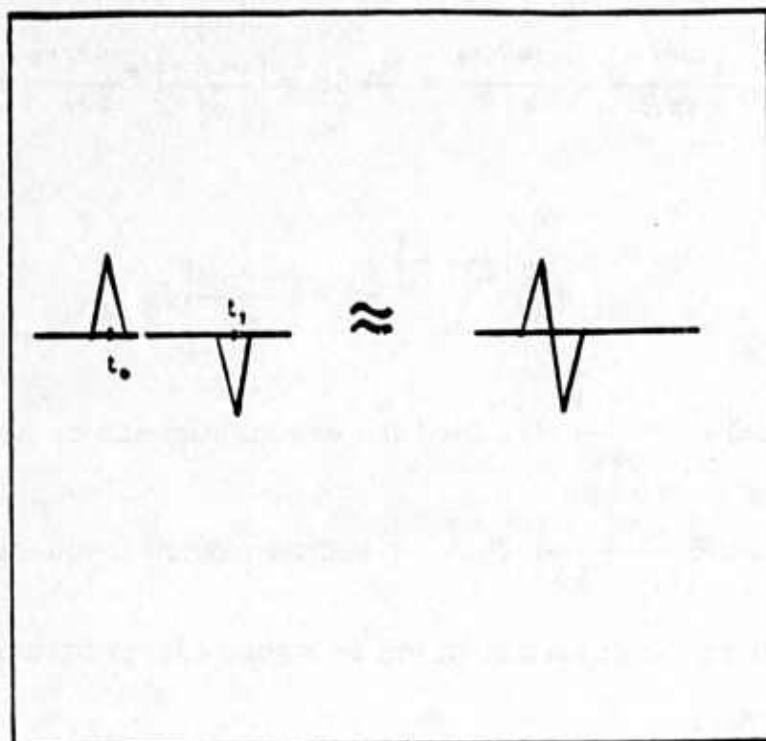


Figure 3

Figure 3 A linearization in terms of the field approximates the differential seismogram on the left with the doublet on the right. This will be a good approximation only for low enough frequencies.

$\delta t = \left[\frac{R}{v_1} - \frac{R}{v_0} \right]$ is large. For a single harmonic frequency in a homogeneous medium, then $u = \frac{e^{i\omega R/v}}{4\pi R}$. The linearization in terms of the field can then be written

$$\frac{e^{i\omega R/v_1}}{4\pi R} - \frac{e^{i\omega R/v_0}}{4\pi R} \approx \frac{\partial u}{\partial v} \delta v = \left(\frac{-i\omega R}{v_0^2} \right) \frac{e^{i\omega R/v_0}}{4\pi R} \delta v$$

or

$$e^{i\omega R \left(\frac{1}{v_1} - \frac{1}{v_0} \right)} - 1 \approx \left(\frac{-i\omega R}{v_0^2} \right) \delta v$$

for small $\omega R \left(\frac{1}{v_1} - \frac{1}{v_0} \right) \ll 1$, then the exponential can be approximated as

$$e^{i\omega \delta t} \approx 1 + i\omega R \left(\frac{1}{v_1} - \frac{1}{v_0} \right) \delta v$$

Thus for sufficiently low frequencies, such that

$\omega \delta t \ll 1$, then the linearization will be a good approximation to the actual differential field.

An alternative linearization could be performed in terms of the log of the field variables. This has been successfully applied ^{to} electromagnetic problems by Madden (personal communication). By perturbing the velocity in a homogeneous medium, we again consider the extreme case of

$ka \gg 1$. With $\ln(u_0) = \ln \left(\frac{e^{i\omega R/v_0}}{4\pi R} \right)$, then

$$\ln(u_1) - \ln(u_0) \approx \frac{\partial(\ln(u))}{\partial(\ln v)} \delta(\ln v)$$

or

$$i\omega R\left(\frac{1}{v_1} - \frac{1}{v_0}\right) \approx i\omega R\left(\frac{-1}{v_0^2}\right)(v_0) \delta(\ln v)$$

$$\left(\frac{1}{v_1} - \frac{1}{v_0}\right) \approx \frac{-1}{v_0^2}(v_1 - v_0)$$

Thus for this simple example of a homogeneous medium perturbation there is no dependence on frequency for the complex phase linearization. The region of validity for the log field parameterization is more involved in the general case where there is an amplitude perturbation as well. In any event, this type of linearization has been ~~used~~ successfully applied in travel time inversions (see Aki et al. [1977], Clayton and Comer [1983]), where only the imaginary part of the complex phase is used.

The linear sensitivity operator in terms of the log field is now derived and compared with the Born linearization. The log field linearization is additive in the phase, and is equivalent to Rytov's method of smooth perturbations (see Tatarskii [1971], section 45; Ishimaru [1978]). Rytov's approximation is compared in Appendix 2 with other asymptotic, multiplicative approximations including geometric optics and the parabolic equation method. We start with the Helmholtz equation

$$\nabla^2 u + \frac{\omega^2}{v^2(\vec{x})} u = 0$$

and let $\psi = \ln(u(\vec{x}, \omega))$, then the following nonlinear Riccati equation is obtained

$$\nabla^2 \psi + (\nabla \psi)^2 + \frac{\omega^2}{v(\vec{x})^2} = 0$$

Now a perturbed problem is derived with $\psi = \psi_0 + \epsilon \psi_1$ and $v = v_0(\vec{x}) + \epsilon v_1(\vec{x})$, assuming $\epsilon \ll 1$, $\psi_0 \approx \psi_1$, and $v_0 \approx v_1$. Then

$$\begin{aligned} & \left\{ \nabla^2 \psi_0 + (\nabla \psi_0)^2 + \frac{\omega^2}{v_0^2(\vec{x})} \right\} \\ & + \epsilon \left\{ \nabla^2 \psi_1 + 2 \nabla \psi_0 \nabla \psi_1 - \frac{2 \omega^2 v_1(\vec{x})}{v_0^3(\vec{x})} \right\} \\ & + \epsilon^2 \left\{ (\nabla \psi_1)^2 + \frac{3 \omega^2 v_1^2(\vec{x})}{v_0^4(\vec{x})} \right\} = 0 \end{aligned}$$

The first bracket is equal to zero by solving the forward problem in the unperturbed medium $v_0(\vec{x})$. Dropping the third bracket in ϵ^2 constitutes the Rytov approximation which is now linear in ψ_1 and v_1 . Letting $\delta v(\vec{x}) = \epsilon v_1$ and $\delta \psi = \epsilon \psi_1$ gives

$$\nabla^2(\delta \psi) + 2 \nabla \psi_0 \nabla(\delta \psi) = \frac{2 \omega^2 \delta v(\vec{x})}{v_0^3(\vec{x})}$$

This is a linear equation for $\delta \psi$ which can be simplified by the substitution, $\delta \psi = \bar{\psi} e^{-\psi_0}$

$$\nabla^2 \bar{\psi} + \frac{\omega^2}{v_0^2(\vec{x})} \bar{\psi} = \frac{2 \omega^2 \delta v}{v_0^3} e^{\psi_0}$$

which has a solution

$$\bar{\psi}(r_g, r_s) = - \int dV(r) g(r_g, r') e^{i\phi_0(r', r_s)} \frac{2\omega^2 \delta v}{v_0^3}$$

or

$$\delta\psi(r_g, r_s) = - \int dV(r) g(r_g, r') e^{i\phi_0(r', r_s) - i\phi_0(r_g, r_s)} \frac{2\omega^2 \delta v}{v_0^3} \quad (35)$$

This gives the linear sensitivity operator in terms of the log field variables in the form $\delta\psi = \frac{\partial\psi}{\partial v} \delta v$.

To see how this compares to the Born approximation or a linearization in terms of the field, equation (35) is exponentiated

$$e^{\delta\psi} = e^{i\psi - i\phi_0} = e^{-\int dV(r)(\dots)}$$

For small $-\int dV(r)(\dots)$

$$e^{\delta\psi} \approx e^{i\psi} - \int dV(r) p(r', r_s) g(r_g, r') \frac{2\omega^2 \delta v(r')}{v_0^3(r')}$$

which is equivalent to the frequency transformed equation (7). Thus equation (7) can be derived from (35) assuming that the exponentiated argument is small. It appears that the Rytov approximation includes some multiple scattering effects not included in the Born approximation. However, the Rytov approximation was originally derived assuming $\delta\psi \ll 1$. Thus, there has been an ongoing controversy as to whether the Rytov is in fact any better than the Born approximation (see Tatarskii [1971], Barabanenko et. al. [1971], Aki [1973], Brown [1966, 1967], Taylor [1967], Yura et. al. [1983]). In fact, Aki and Richards [1980] derive their

expressions for the log amplitude and phase from the exponentiated Born series (EBS).

However, asymptotic equivalence may not correspond to a practical equivalence between the two approaches as shown in the previous simple example. Keller [1969] showed that the Born and the Rytov approximations are equally accurate in their dependence on the small velocity parameter, ϵ , but show quite different asymptotic properties as a function of range. Mueller et al [1979] noted that both the Born and the Rytov approximations required that some norm of the velocity perturbations over the volume be small compared to unity. In addition, the Born approximation required small ka , where a is the scale of the heterogeneity, in agreement with the range of validity for the single scattering formulation given in Appendix 1. Thus low spatial wavenumber, band limited velocity perturbations appears to favor the Rytov approximation. From the context of an inverse problem, it is natural to estimate the low spatial wavenumbers of the velocity first via a multiple scale type of procedure. For small scale residual heterogeneities, the Born approximation to the field may have an advantage.

Experiments in diffraction tomography comparing the Born and the Rytov formulations have been conducted by Kaveh et. al. [1982,1983] who showed that when a large forward scattering component exists, the linearized reconstructed images based on the Born approximation were

more in error than the Rytov, but the Rytov required the determination of the phase of the scattered waves. There are numerical difficulties in obtaining the log of the field since the phase of the complex log must be unwrapped. Still, numerical algorithms for this exist (see Tribolet [1978]). Also, algorithms based on the smoothness of the phase at adjacent spatial points have had some initial success (see Kaveh et.al. [1982], Kaveh et.al. [1984]).

Another consideration as to whether to use a linearization based on the Born (field) or the Rytov (log of the field) approximations is the noise structure of the problem (see Tatarskii [1971]). If the data is expressed as a sum of signals plus independent Gaussian noise, N , then the linearized inversion should be applied to the waveform or to the real and imaginary Fourier components. If the noise structure is multiplicative and proportional to the signal, then $\log N$ will behave in a Gaussian manner around $\log u$. In this case, the log should be taken (see Aki and Richards [1980], pp 639). Detailed work on the effects of additive and multiplicative noise for moment tensor inversion of surface wave data was done by Patton and Aki [1979].

If the noise is Gaussian and additive in the field, then the fluctuations in the amplitude should follow a Rayleigh distribution. If the noise is Gaussian and additive in the complex phase, then the fluctuations in the amplitude should follow a lognormal distribution. Experiments of light

propagation in the atmosphere show closer agreement with the lognormal distribution (see Tatarskii [1971]).

In order to check the linearization in terms of the log field for a large scale heterogeneity, a homogeneous velocity perturbation, δv , is again considered. Equation (35) can then be written

$$\delta\psi(r_g, r_s) = \frac{-2\omega^2\delta v}{v_0^3} \int dV(r') \frac{e^{ikR_{g,r'}}}{4\pi R_{g,r'}} \frac{e^{ikR_{r,s}}}{4\pi R_{r,s}} 4\pi R_{r,s} e^{-ikR_{g,r,s}}$$

Using the Fresnel approximation with z axis oriented from the source to receiver

$$\delta\psi(r_g, r_s) = \frac{-\omega^2\delta v z}{2\pi v_0^3} \int_0^z \frac{dz'}{(z-z')z'} \int_{-\infty}^{\infty} dx' e^{\frac{ikx'^2}{2z'(z-z')}} \int_{-\infty}^{\infty} dy' e^{\frac{iky'^2}{2z'(z-z')}}$$

Using the relation $\int_{-\infty}^{\infty} e^{-ax^2} = (\pi/a)^{1/2}$, then

$$\delta\psi(r_g, r_s) = \frac{-\omega^2\delta v z}{2\pi v_0^3} \int_0^z \frac{dz'}{z'(z-z')} \frac{2\pi z'(z-z')}{-ikz}$$

or

$$\delta\psi(r_g, r_s) = \frac{-i\omega z \delta v}{v_0^2}$$

Now let $\delta\psi = i\omega\delta t = i\omega z \left(\frac{1}{v_1} - \frac{1}{v_0} \right) \approx \frac{-i\omega z \delta v}{v_0^2}$. Thus, the linearization in terms of log variables will retrieve a global velocity change in one iteration without the apparent low frequency requirement of the Born approximation of equation (7).

There are several basic linearizations that can be used in inversion studies, including those based on the additive nature of the field, Born, or the log field, Rytov. Travel time inversions are more similar to the Rytov approximation except that only the imaginary part of the complex phase is utilized and usually only first arrival times. Keller [1969] noted that any advantages that the Rytov approximation may have over the Born approximation may be lost when the waveform contains more than one wave. This would occur when there is high wavenumber velocity fluctuations causing significant backscatter. This may be the reason for the success of simple migration algorithms which map small scale reflectors and which are kinematically equivalent to Born inversion (see Miller et al. [1984]). Another consideration as to whether to use the log field data, or the amplitude and phase, is the numerical difficulties in unwrapping the phase when taking the complex log, although work on this has been done (Tribolet [1978], Kaveh [1982, 1984]).

One strategy for inversion of material parameters is to use a multiple scale procedure in which the large scale ¹fluctuation are inverted for first, and compensated for. This might include several Rytov iterations followed by a series of Born iterations. A compromise approach would be to use an iterative travel time inversion first to adjust for the major phase differences and then use the iterative Born inversion on the waveforms to incorporate more of the data and fine tune the model. In the following numerical examples, the Born linearization is used on residual waveform data assuming small velocity perturbations.

2.4.4 EXAMPLES

In this section, several examples are given using a linearization in terms of the field as described previously. A simple descent method is used

$$v_{k+1} = v_k + \alpha W_k (F_k^* \delta p_k) \quad (36)$$

where $F_k^* = C_v F_k^T C_p$, W_k is a preconditioner, and α is a scaling factor. For an initial homogeneous model close to the true model, then $F_k^* \approx F_0^*$. Operating on the residuals by the adjoint is then given in (29) as

$$\delta v_k = F_k^* \delta p_k \approx \frac{C_v}{v^3 8\pi^2} \sum_{r_s} \sum_{r_g} \left[\frac{1}{R_{rr_s} R_{rr_g}} \right] \delta \tilde{p}(r_g, l = \frac{R_{rr_s} + R_{rr_g}}{v}, r_s)$$

where $\delta \tilde{p}$ is the cross correlation of the second derivative of the source time function with the weighted data residuals, C_v is the model covariance operator and includes a spatial smoothing function. Here, a Gaussian smoothing function as in (21) is used. Operating on the residuals by the adjoint is kinematically similar to a Kirchhoff migration. If the preconditioner, W_k , is chosen as I , then the velocity inversion in (36) with $F_k^* = F_0^*$ is equivalent to a sequence of Kirchhoff migrations. An alternative choice for the preconditioner, W_k , is the inverse of the diagonal of $(I + F^* F)$ as in (32). The part of this preconditioner that approximately corrects for geometric spreading is

$$W_k = \left\{ \sum_{r_s} \sum_{r_g} \left[\frac{1}{R_{rr_s}^2 R_{rr_g}^2} \right] \right\}^{-1} \quad (37)$$

This simple preconditioner will be used in addition to the identity in the next example.

The geometry of the first example is shown in Figure 4. There are nine sources marked by x's and nineteen receivers marked by triangles. The axes are in km to give specific units and the initial velocity is 8 km/sec. For this example, a single smooth heterogeneity with a radius of 10 km is reduced in velocity by .01 km/sec. This is shown by the circle in Figure 4. The interior box in Figure 4 shows the region where the velocity is to be inferred.

Figure 5 shows the seismograms computed using the Gaussian beam method for the single smooth heterogeneity shown in Figure 4 for three of the nine source locations. The source wavelet is the Gabor wavelet

$$f(t) = e^{-(2\pi f_0 t / \gamma)^2} \cos(2\pi f_0 t + \varphi_0)$$

where $f_0 = 4 \text{ Hz}$, $\gamma = 3$, $\varphi_0 = 0$. This results in a 2 km wavelength. The correction given by Cerveny et.al. [1982] is used to approximately compute the 3-D response. In Figure 5, a slightly larger amplitude can be seen on some seismograms for different source locations. For example, for $S = (30., 140.)$, the amplitude increases due to focusing at a range of 180 km and a time of 24 sec.

The residual, perturbed minus unperturbed, seismograms are shown in Figure 6 for three source locations. The amplitudes of the residuals are

single smooth heterogeneity

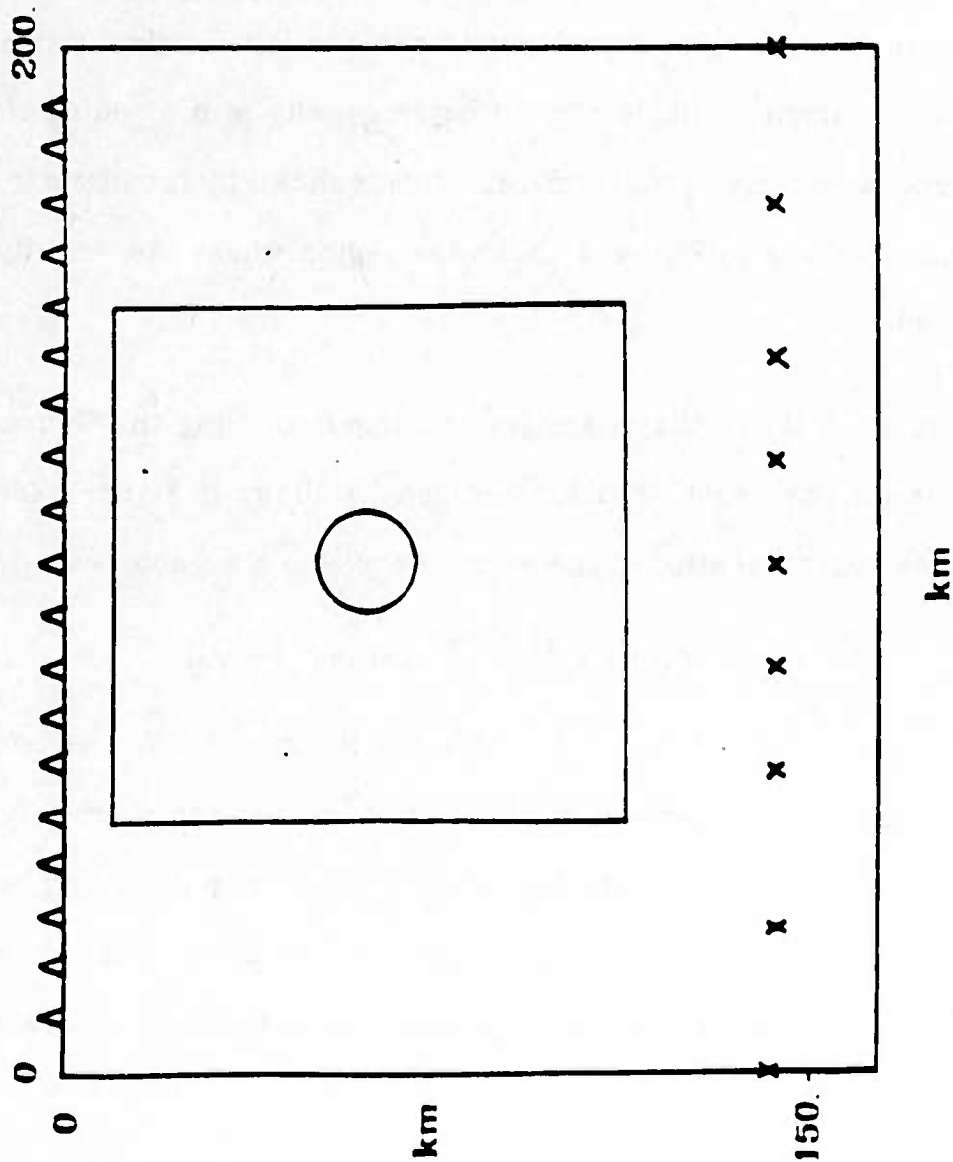


Figure 4 Model geometry for the single smooth heterogeneity numerical example.

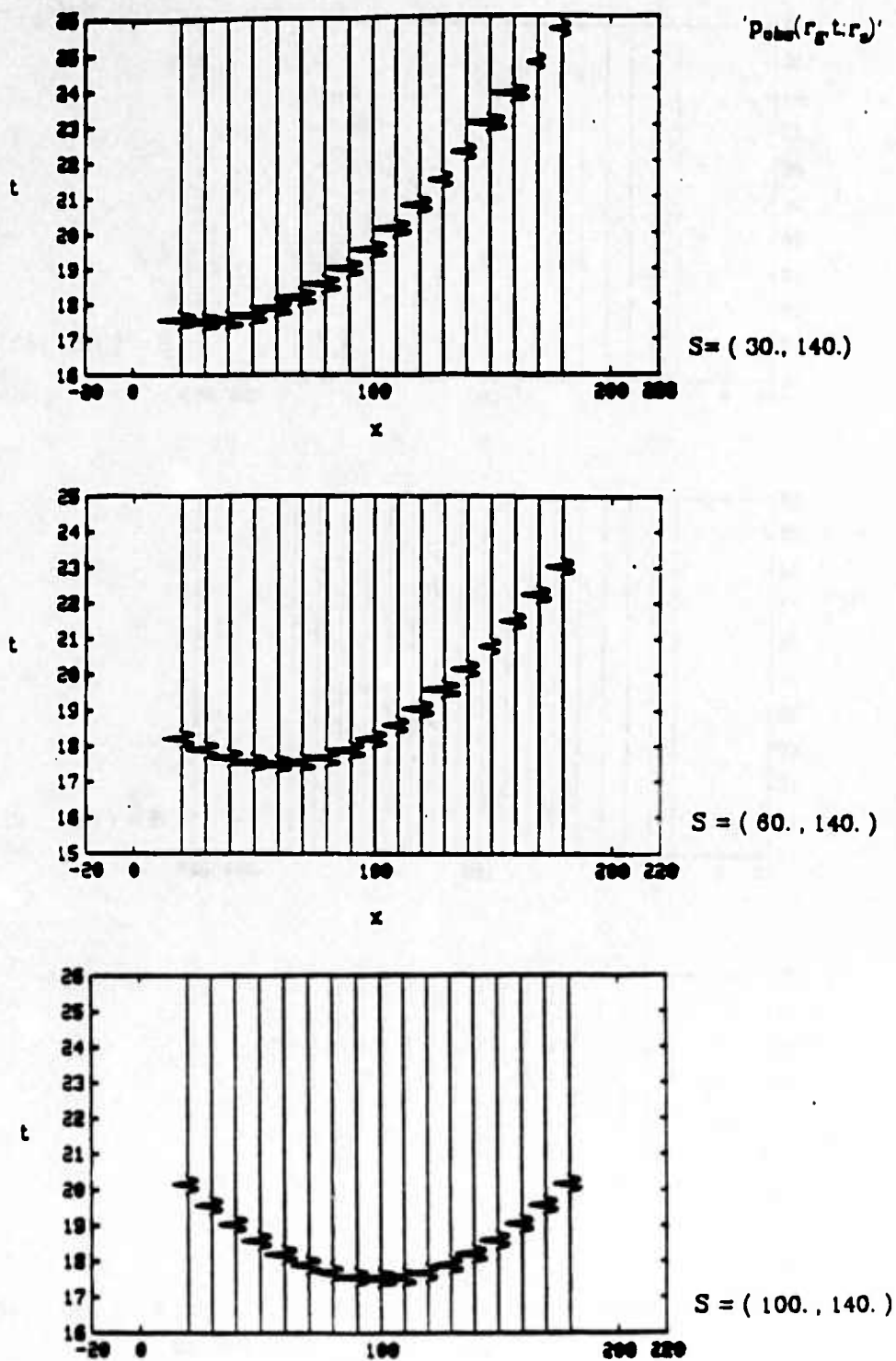


Figure 5 Gaussian beam seismograms for three source locations for a single smooth heterogeneity.

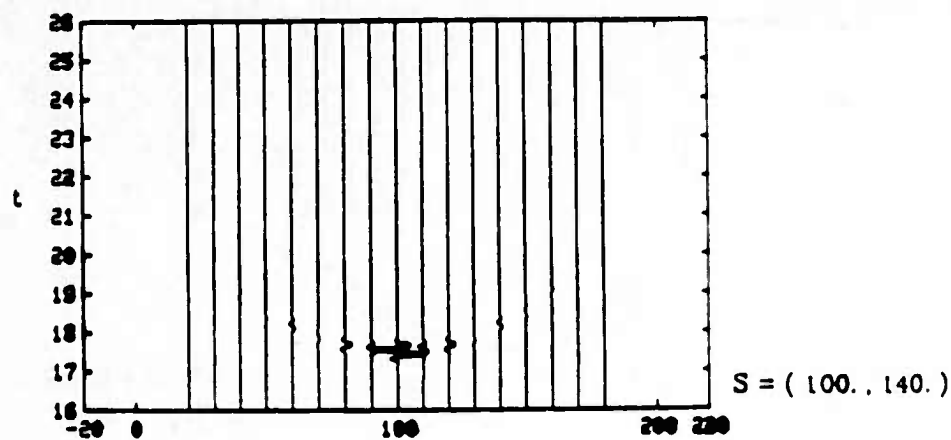
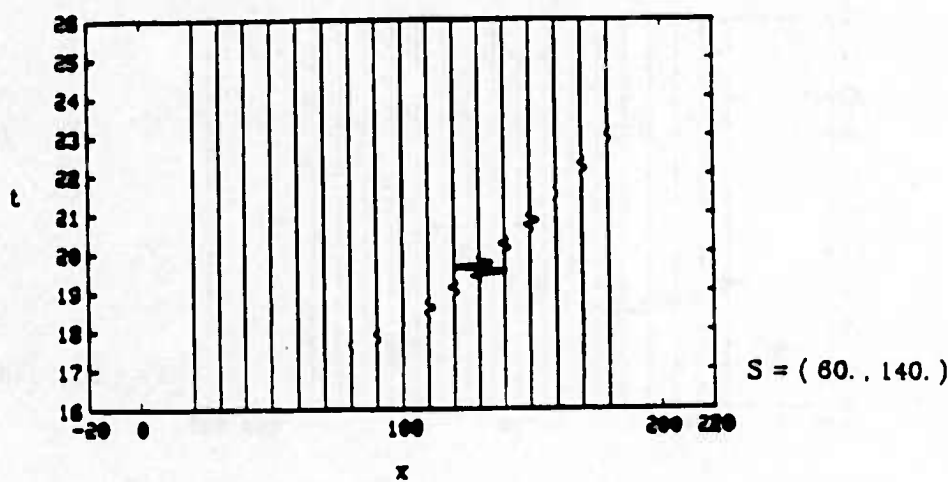
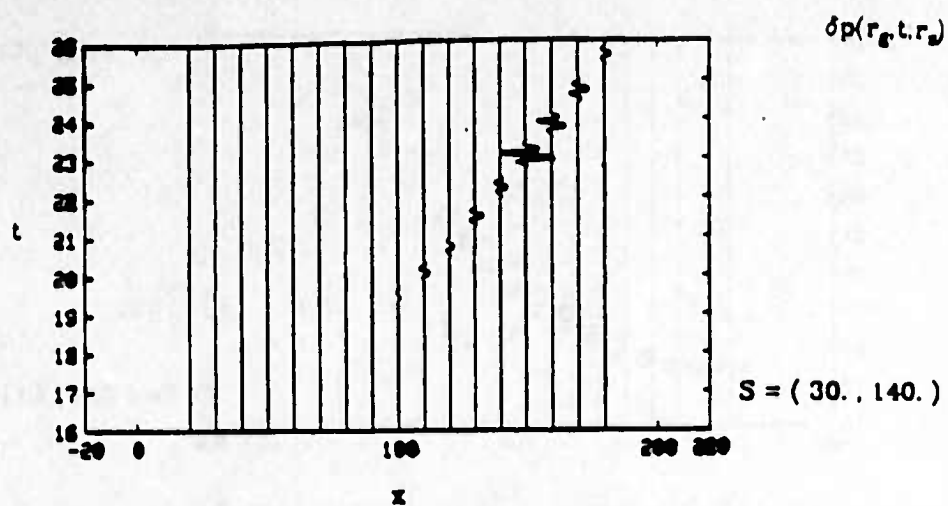


Figure 6 Residual seismograms for three source locations for a single smooth heterogeneity.

about ten times smaller than the amplitudes of the unperturbed seismograms. This satisfies the Born requirement that the scattered field be small compared to the primary field. The wavelet cross correlation with the data results in a symmetric wavelet centered on the arrival time, and the second derivative approximately results in a factor, $-\omega_0^2$, assuming a Gabor wavelet.

Figure 7 shows the result of back propagating the residuals into model space using the adjoint operator. This would be the first iteration in a descent procedure with $W_k = I$ and $C_v = \sigma_v \delta(\vec{r} - \vec{r}')$. A grid of 100 by 100 points has been evaluated in model space. The result in Figure 7 is scaled from -1. to +1.. The contours are at (-.75, -.25, .25, .75) with the -.75 contour in the center. The parallelogram and the small plus represent the location of the true heterogeneity. The data residuals appear to be streaked along lines from the receivers to the stations in model space. The response to the simple transpose operator for a single heterogeneity thus has a streaked appearance in model space.

There are additional operations that can be included in an individual velocity inversion step. This includes a more involved preconditioner as in equation (37) which approximately corrects for geometric spreading, and is an approximation to filtering the backprojection. Figure 8 shows the result of using this preconditioner. It has the effect of reducing some of the streaks in the backprojection. Again, the true heterogeneity is

single heterogeneity

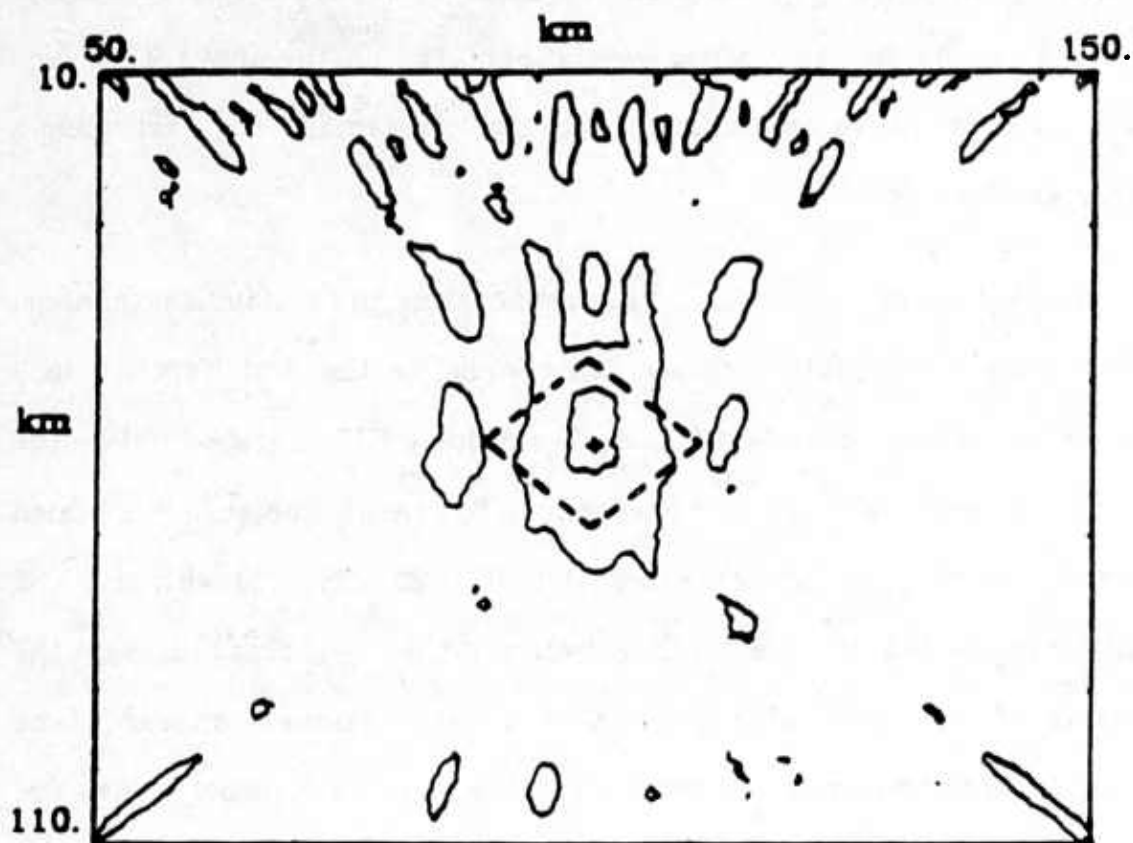


Figure 7

Figure 7 The backpropagation of the residuals into model space for a single smooth heterogeneity.

single heterogeneity

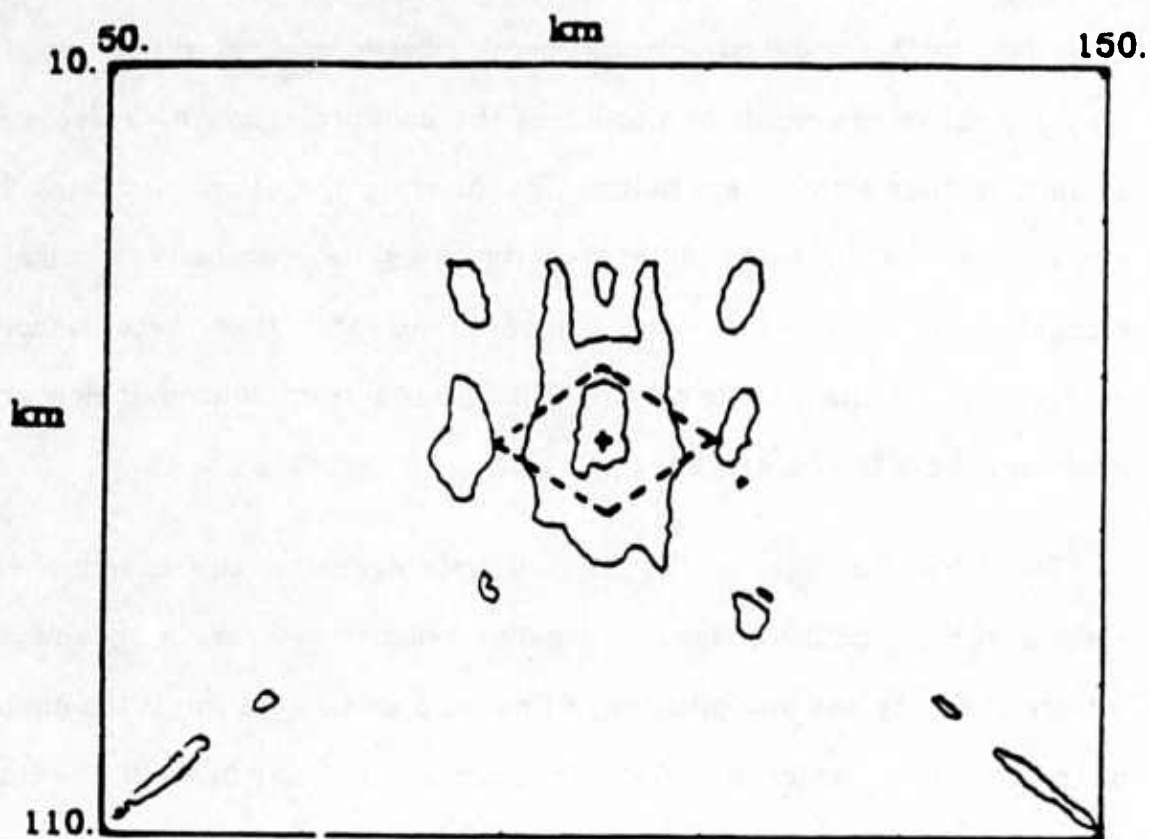


Figure 8

Figure 8 The first iteration of a velocity inversion for a single smooth heterogeneity with W_k given by eqn. (37) and $C_v = \sigma_v \delta(\vec{r} - \vec{r}')$.

represented by the parallelogram.

Another operation that can be applied to the backprojection is spatial smoothing by the model covariance operator based on a priori information. Figure 9 shows the result of smoothing the backprojection by a Gaussian smoothing filter with a 5 km radius. This filtering operation eliminates the streaks and localizes the heterogeneity. This heterogeneity is slightly elongated in the z direction compared to the true heterogeneity represented by the parallelogram. This results from sources below and receivers above the heterogeneous region.

To check the sign of the derived heterogeneity, the example was redone with a positive instead of negative velocity heterogeneity, and the inferred velocity was also positive. A final step in an iteration is the choice of the scalar, α , which specifies how far in a particular descent direction one moves. A particular value for the scaling factor can be computed from equation (16), although other values can be used to over or under relax the problem. The scaling factor gotten to match the single heterogeneity case will be used to relax each iteration step in the next examples.

In the next example, three smooth heterogeneities are used, each with a velocity heterogeneity of -01 km/sec lower than the background. The geometry is shown in Figure 10. The sources are below the heterogeneity and the stations above. The result of backprojecting the residuals into model space and applying a 4 km Gaussian covariance operator is shown in

single heterogeneity

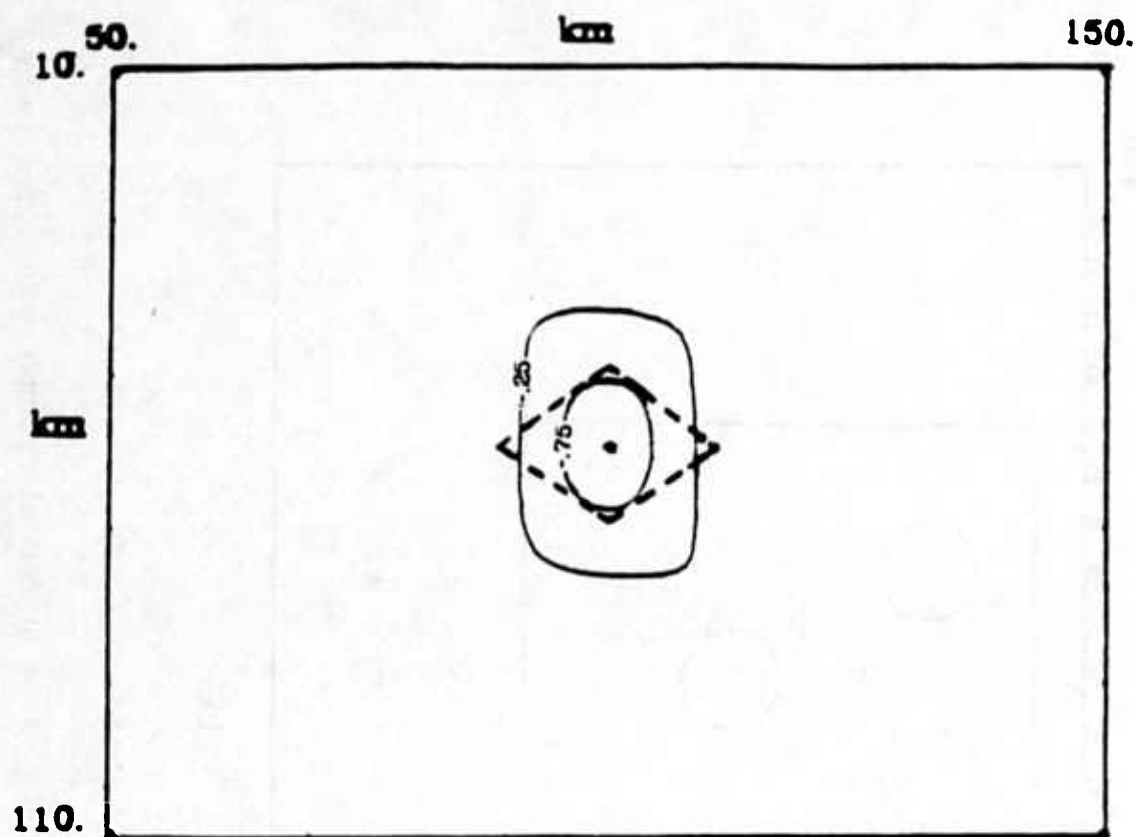


Figure 9

Figure 9 The first iteration of a velocity inversion for a single smooth heterogeneity with $W_k = 1$ and C_v given by a Gaussian function with a 5 km radius.

three heterogeneities

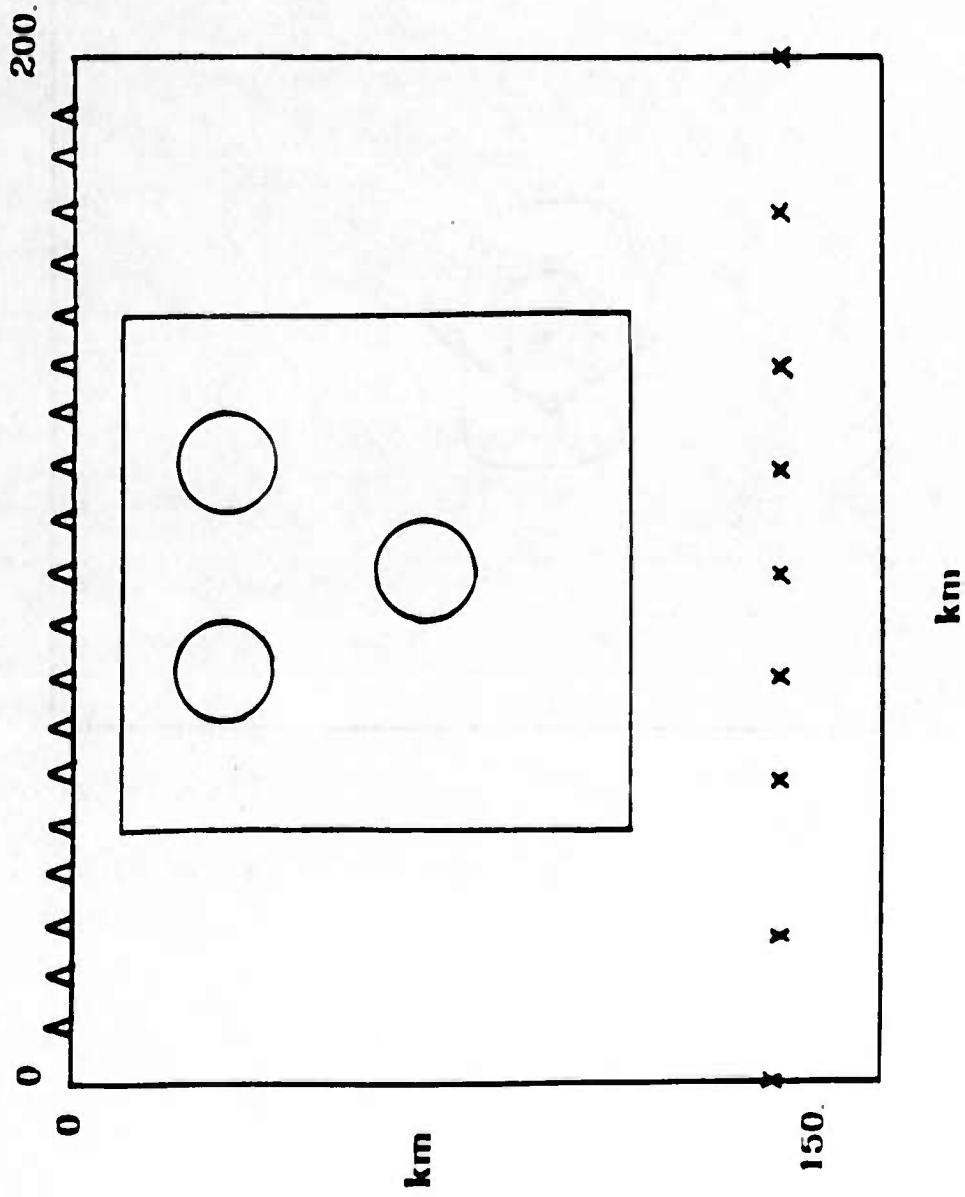


Figure 10 Model geometry for a numerical example with three smooth heterogeneities

Figure 11. The contour interval goes from -1. to +1., and the contours are at (-.75, -.25, .25, .75). The true locations of the heterogeneities are represented by the parallelograms. Using the scaling factor gotten from the single heterogeneity case, then the lower heterogeneity has a velocity perturbation of -.0109 which is approximately accounted for in this iteration. The upper two heterogeneities have about half that perturbation.

Figure 12 shows the second iteration result of backprojecting and smoothing with a 4 km Gaussian model covariance operator. The upper two heterogeneities can be seen along with side lobes. Using the scaling factor gotten from the single heterogeneity case, the upper two heterogeneities have a velocity perturbation of -.0052 km/sec. When combined with the perturbation from the first iteration, then all three heterogeneities have been accounted for in the inversion.

In the next example, a ring heterogeneity is considered. The model geometry is shown in Figure 13. This is an idealization of the structure derived by Thurber [1984] under the Kilauea volcano using a travel time inversion. He found a high velocity core of the volcano surrounding an interior lower velocity region, which was inferred to be a magma complex. For this numerical example, just the ring is reduced in velocity by -.01 km/sec. Figure 14 shows the result of smoothing the backprojection with a Gaussian model covariance operator with a 4 km radius. The contours go

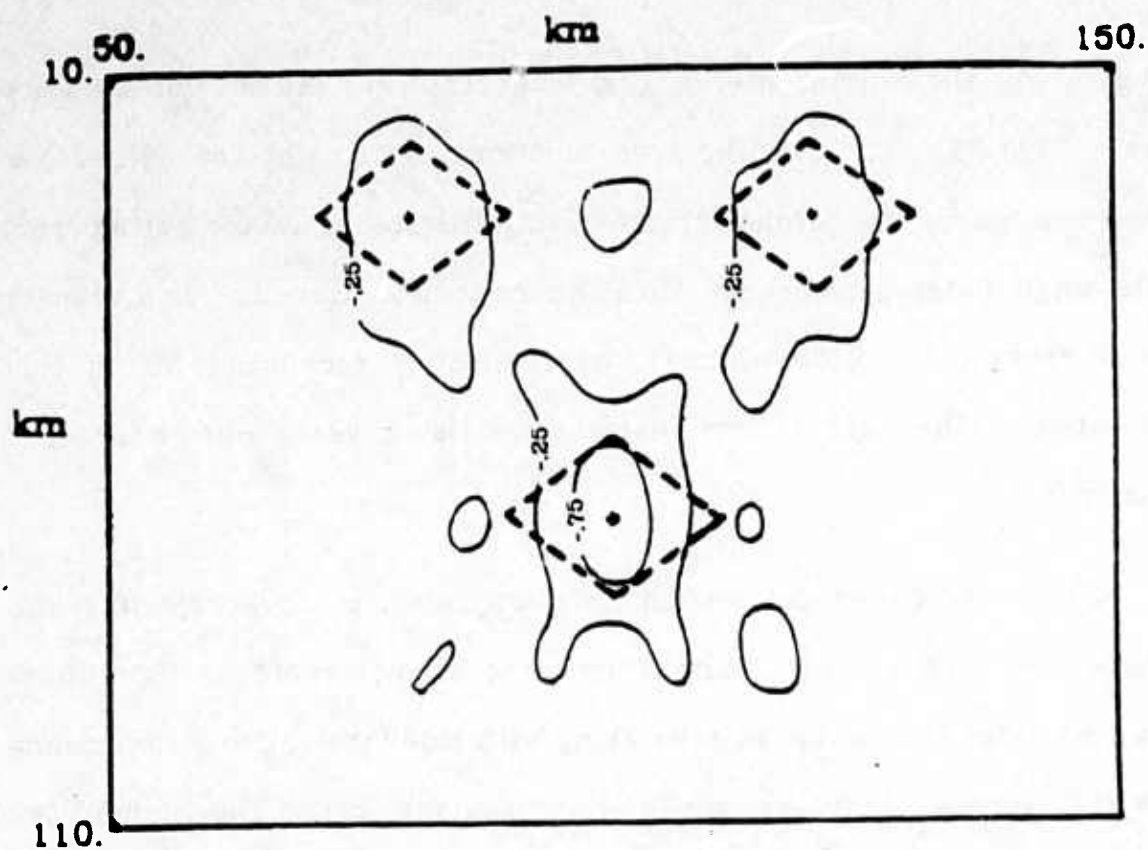


Figure 11

Figure 11 The first iteration in a velocity inversion for three smooth heterogeneities with $W_k = 1$ and C_v given spatially by a Gaussian function with a 4 km radius.

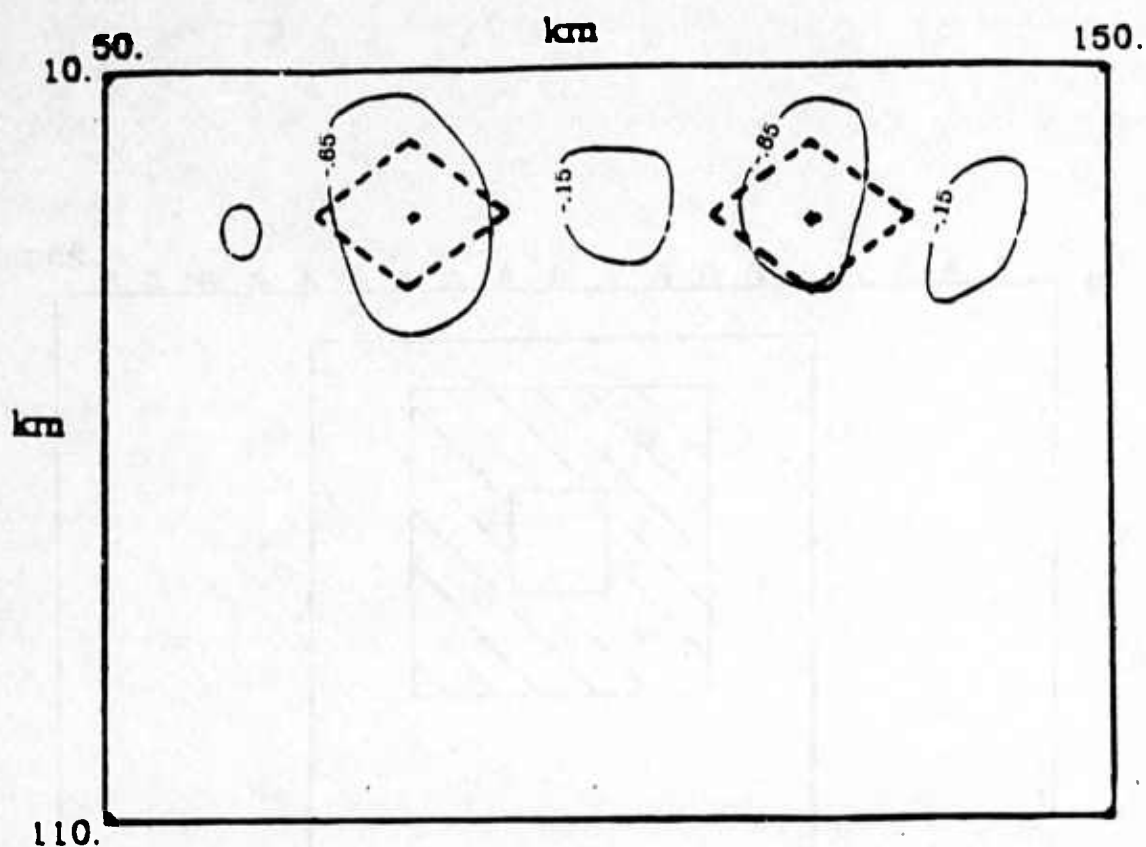


Figure 12

Figure 12 The second iteration of a velocity inversion for three smooth heterogeneities with W_k and C_v given spatially by a Gaussian function with a 4 km radius.

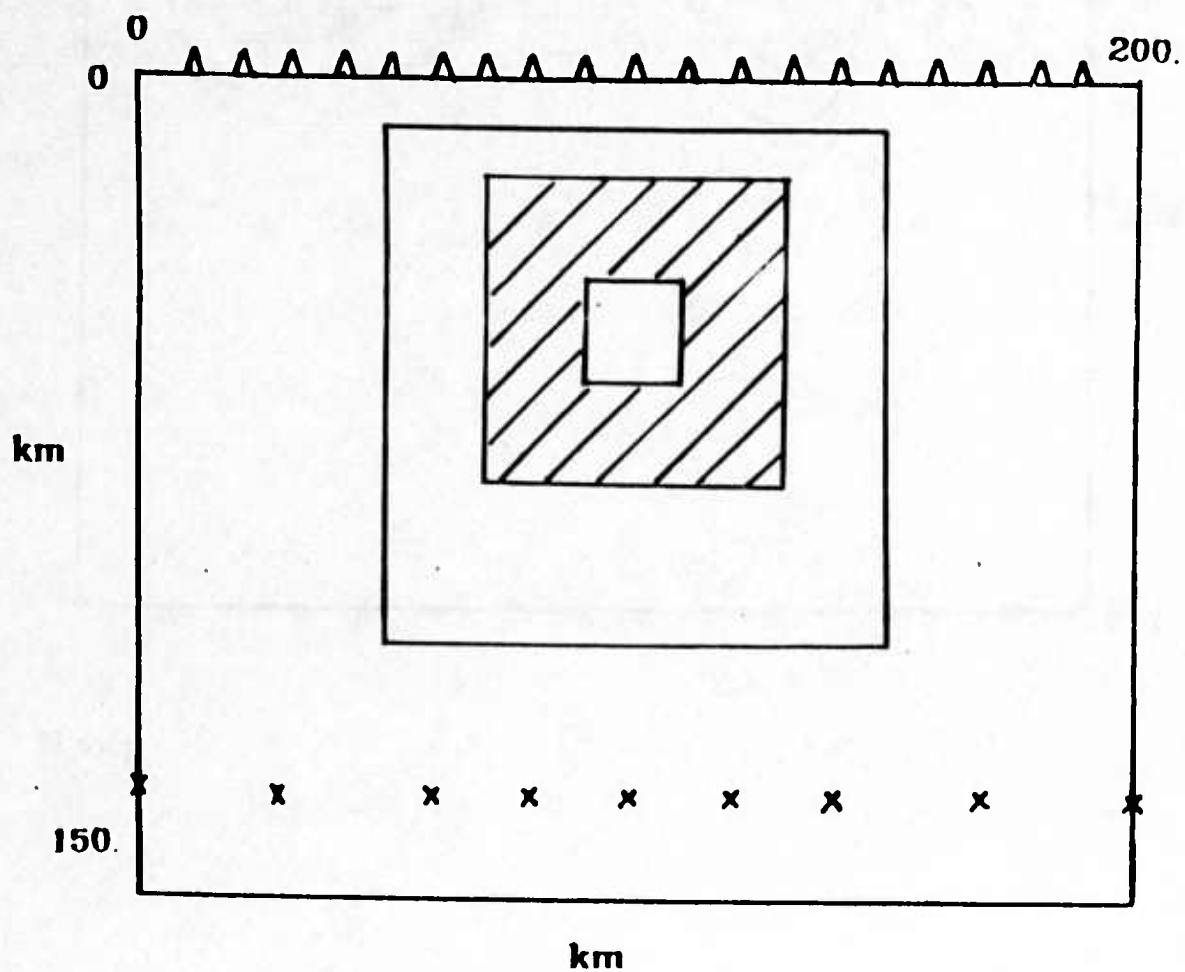


Figure 13 Model geometry for a numerical example for a ring heterogeneity.

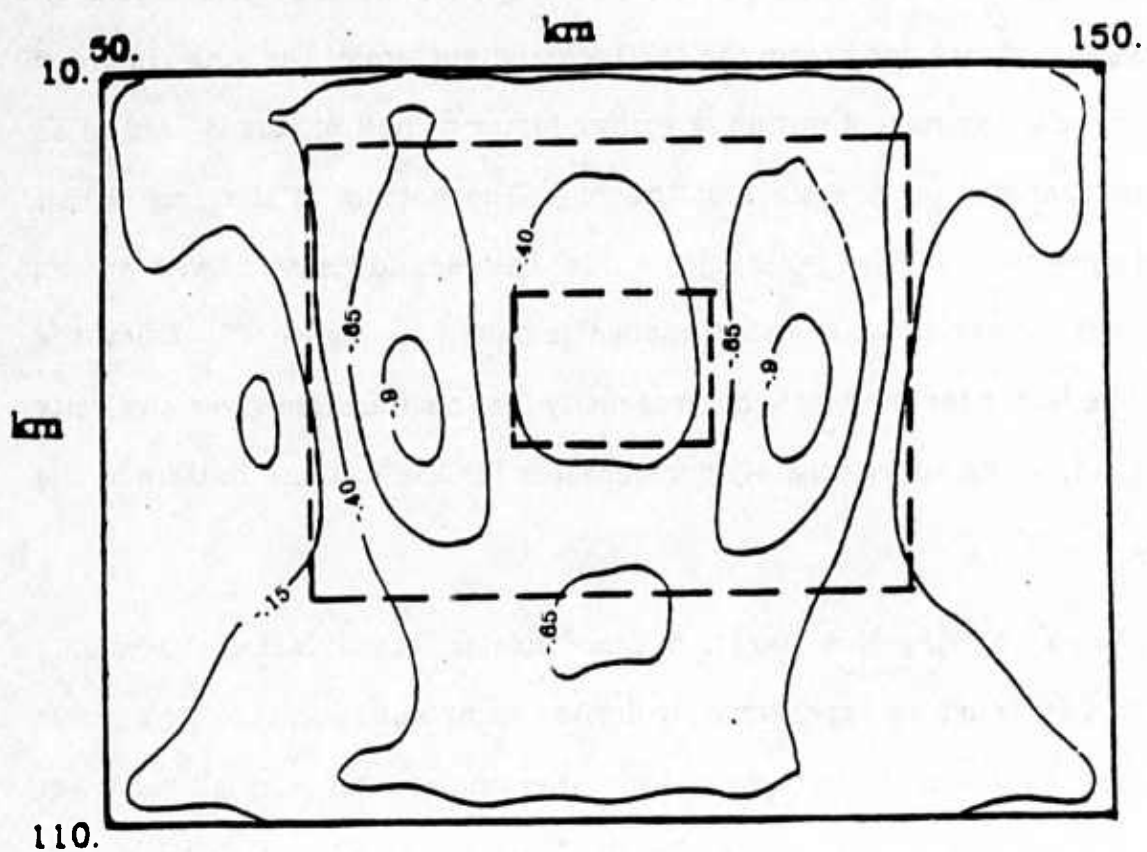


Figure 14 The first iteration of a velocity inversion for a ring heterogeneity with $W_k = 1$ and C_0 given spatially by a Gaussian with a 4 km radius.

from -1. to +1. with a .25 contour interval. The velocity perturbation forms a ring with the largest perturbations on the sides of the ring. Using a scaling factor from the single heterogeneity case gives an amplitude of -.02 on the sides of the ring, or twice as large as the true heterogeneity. Smaller perturbations form the the top and bottom of the ring. In order to examine the next iteration, a scaling factor of half of this is used so as to compensate for the sides of the ring. The bottom of the ring is half compensated for. The backprojection for the second iteration with a 4 km Gaussian covariance operator applied is shown in Figure 15. Using the scaling factor for the single heterogeneity for this iteration gives a velocity of -.011, which approximately compensates for the top and bottom of the ring.

When applying this iteration procedure to actual data, the scaling factor, α , must be approximately derived using equation (16), or an over relaxed version of it. In these numerical examples, only a rough value was used to see how convergence proceeds. Nonetheless, choice of this scaling factor is essential part of the problem.

Finally, a simple comparison of these results with travel time inversions using the same model geometries is done. The model is divided into 10 by 10 blocks or 100 model parameters. 171 travel times are used, not all of which give new information about the model. A damped inversion is used. The result for a single heterogeneity is shown in Figure 16a. Using

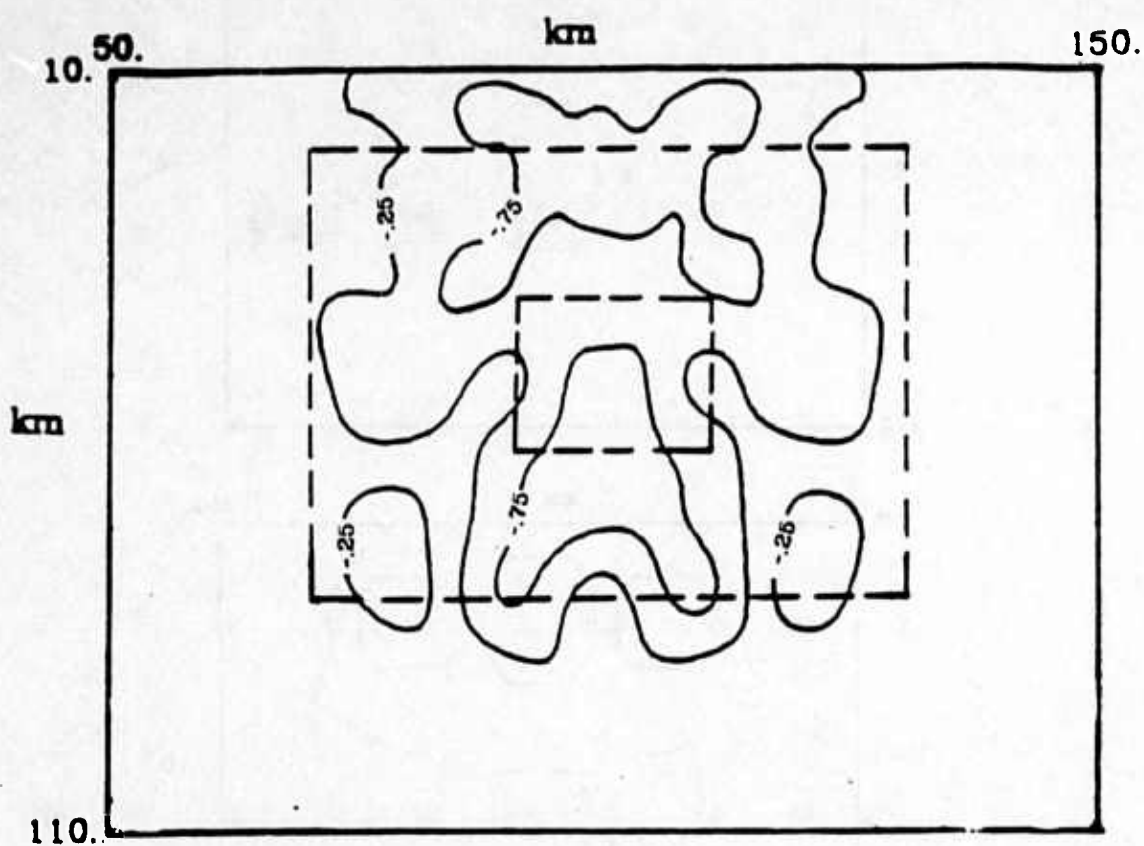


Figure 15

Figure 15 The second iteration for a velocity inversion for a ring heterogeneity with W_k and C_v given spatially by a Gaussian with a 4 km radius.

Figure 16

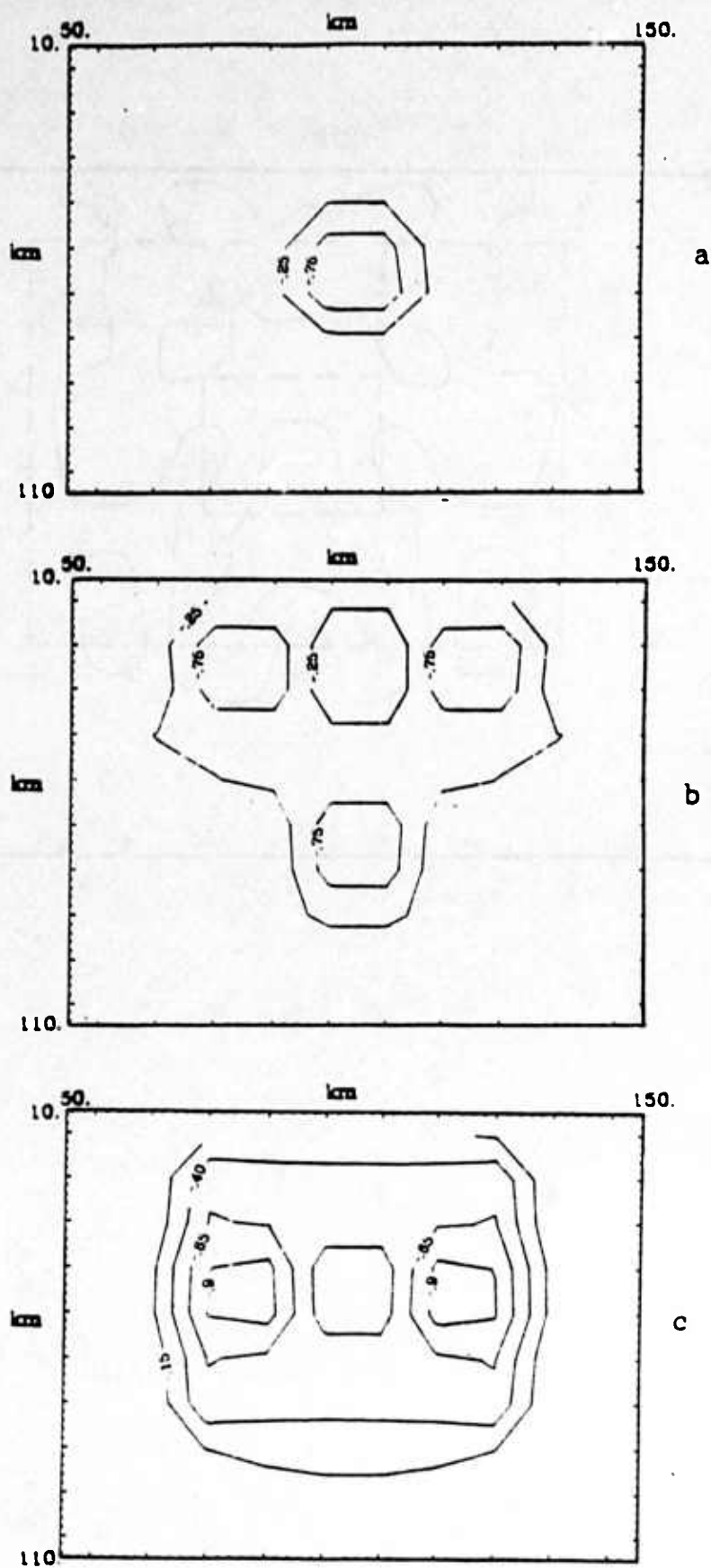


Figure 16 Travel time inversion results for a) single heterogeneity, b) three heterogeneities, c) ring heterogeneity.

a damping of .001, a velocity of -0.0102 km/sec is gotten. This result compares with the waveform inversion shown in Figure 8. The travel time inversion for three heterogeneities is shown in Figure 16b. This is again similar to the two iteration waveform result. Finally, the travel time inversion result for the ring structure is shown in Figure 16c. In this case the waveform result may do slightly better. As discussed in a previous section, for large scale heterogeneities, a complex phase linearization may be more robust. Travel time inversions are in this class, in which only the imaginary part of the phase is used with first arrivals. However, for small velocity perturbations the results are comparable. In actual cases, waveform inversions allow for a more complete use of the data when the phase perturbations are small, or have been previously compensated for.

2.4.5 CONCLUSIONS

In this paper, a linearized inversion procedure for material parameters has been investigated. In this procedure a linear sensitivity operator must be derived and can be computed economically by using reciprocity of the Green's function. Data errors and a prior model information are included via covariance operators. Large matrix inversions are avoided by using descent algorithms. A fast forward modelling scheme is required and here the Gaussian beam method for a laterally varying medium is used. This allows for inhomogeneous initial models.

Different linearizations are possible including a linearization in terms of the field, the Born approximation, and a linearization in terms of the log field, the Rytov approximation. Travel time inversions are in the class of linearizations for the log field, where only the imaginary part of the complex phase is used with first arrivals. Log field linearizations may be more robust than field linearizations for large scale heterogeneities where forward scattering predominates, but phase unwrapping may be difficult numerically. The linearization in terms of the field for small perturbations from a homogeneous background is kinematically equivalent to a sequence of Kirchhoff migrations. Field linearizations are expected to be useful for small scale heterogeneities which result in scattering effects that are additive in the field. Several numerical examples are performed using a field linearization which identify the trial structures. A comparison is then done with travel time inversion results.

2.4.6 APPENDIX 1

In this appendix a brief review is given for the conditions of validity of the Born approximation in a continuous fluctuating medium characterized by a spatial Gaussian correlation function where

$$N(\tau) = \frac{E[\mu(\tau)\mu(\tau+\tau')]}{E[\mu^2]} = e^{-|\tau|^2/a^2}$$

where $\mu(\tau) = -\delta v(\tau)/v(\tau)$, and $E[\dots]$ represents the expectation over an ensemble of random media, and a is the correlation length. If the process is ergodic then this expectation can be replaced by a spatial average over one realization. The fractional loss of energy to the primary wave for a medium with velocities with a Gaussian correlation function is (see Aki and Richards [1980])

$$\frac{\Delta I}{I} = \sqrt{\pi} E[\mu^2] k a k L (1 - e^{-k^2 a^2})$$

For the Born approximation to be valid then $\frac{\Delta I}{I} \ll 1$, or

$$\sqrt{\pi} E[\mu^2] k a k L \ll 1 \quad k a \gg 1$$

$$\sqrt{\pi} E[\mu^2] k^3 a^3 k L \ll 1 \quad k a \ll 1$$

where L is the propagation distance and a is the correlation distance. Thus, $k a \gg 1$ refers to large scale scatterers and $k a \ll 1$ refers to small scale scatterers with respect to the wavelength. $k a \ll 1$ corresponds to Rayleigh scattering. For large scale scatterers, $k a \gg 1$, then the mean

square velocity variation must be small for the strict validity of the Born approximation. For small scale scatterers, relatively large mean square velocity fluctuations can be accommodated by the Born approximation. Thus

$$ka = 10. \quad E[\mu^2]kL \ll .056$$

$$ka = 2.0 \quad E[\mu^2]kL \ll .280$$

$$ka = 0.5 \quad E[\mu^2]kL \ll 4.50$$

$$ka = 0.1 \quad E[\mu^2]kL \ll 564.$$

For a given propagation distance, L , linearizations based on the Born approximation will favor small scale fluctuations.

2.4.7 APPENDIX 2

In this appendix a comparison is done between perturbation solutions which to first order are additive in the complex phase or multiplicative in the field. These include the first order geometric optics approximation, the parabolic equation method, and the Rytov approximation. These types of approximations work best in a smoothly varying media with predominantly forward scattering. First, the geometric optics approximation is outlined and compared with the standard parabolic approximation. These are then compared to the Rytov approximation.

Starting with the Helmholtz equation, the first order geometric optics approximation has the following general trial solution

$$u = e^{\bar{k}\varphi_0 + \varphi_1} \quad (\text{A.2.1})$$

where $\bar{k}$ is a large parameter and $\varphi_0 \approx \varphi_1$. In this approximation $\bar{k}$ is usually identified with an average wavenumber and large $\bar{k}$ implies a small wavelength. By substituting this into the Helmholtz equation then

$$\begin{aligned} & \bar{k}^2 \left\{ \nabla \varphi_0 \nabla \varphi_0 + k^2 / \bar{k}^2 \right\} \\ & + \bar{k} \left\{ \nabla^2 \varphi_0 + 2 \nabla \varphi_0 \nabla \varphi_1 \right\} \\ & + \left\{ \nabla^2 \varphi_1 + \nabla \varphi_1 \nabla \varphi_1 \right\} = 0 \end{aligned} \quad (\text{A.2.2})$$

where $k = \omega / v(\vec{x})$. The particular trial solution usually used for geometric optics is $u = A(\vec{x}) e^{i\bar{k}S(\vec{x})}$ where $\bar{k} = \omega / \langle v \rangle$. Thus $\varphi_0 = iS(\vec{x})$

and $\varphi_1 = \ln A(\vec{x})$ and (A.2.2) can then be written as

$$\begin{aligned} & \bar{k}^2 \left\{ -(\nabla S)^2 + k^2 / \bar{k}^2 \right\} \\ & + 2i\bar{k} \left\{ \frac{1}{2} \nabla^2 S + \nabla S \nabla \ln A \right\} \\ & + \left\{ \nabla^2 A / A \right\} = 0 \end{aligned} \quad (\text{A.2.3})$$

For large $\bar{k}$, the first geometric optics solution involves dropping the $\left\{ \nabla^2 A / A \right\}$ term. The Eikonal equation results from equating the first bracket to zero and the transport equation results from equating the second bracket to zero. The transport equation relates the imaginary part of the phase to the real part of the phase.

The standard parabolic equation method is derived by using for (A2.2.1) the particular trial solution $u = Ue^{i\bar{k}z}$ where U is assumed to be a slowly varying function, then $\varphi_0 = i\bar{k}z$ and $\varphi_1 = \ln U(\vec{x})$. (A.2.2) can then be written

$$\begin{aligned} & \bar{k}^2 \left\{ -1 + k^2 / \bar{k}^2 \right\} \\ & + 2i\bar{k} \left\{ 0 + \frac{U_{,z}}{U} \right\} \\ & + \left\{ \frac{U_{,xx} + U_{,yy} + U_{,zz}}{U} \right\} = 0 \end{aligned} \quad (\text{A.2.4})$$

For large $\bar{k}$, the Eikonal equation give $k = \bar{k}$, and the transport equation

gives $U_{,x} = 0$. The parabolic approximation results from neglecting the $U_{,xx}/U$ in the final bracket but retaining the transverse part of the Laplacian, $\nabla_{\text{pert}}^2 = U_{,yy} + U_{,xx}$. This is not entirely consistent within the perturbation scheme used but it gives a one way operator in x ,

$$2ikU_{,x} + \nabla_{\text{pert}}^2 U + \left\{k^2 - \bar{k}^2\right\}U = 0 \quad (\text{A.2.5})$$

The transverse Laplacian is called the diffraction term and the term $\left\{k^2 - \bar{k}^2\right\}$ is called the thin lens term and locally adjusts for the refractive index.

The Rytov approximation results by using the trial solution

$$u = e^{\psi_0 + \epsilon\psi_1} \quad (\text{A.2.6})$$

where ϵ is a small parameter, and $\psi_0 \approx \psi_1$. The velocity is also expanded as $v(\vec{x}) = v_0(\vec{x}) + \epsilon v_1(\vec{x})$, where $v_0 \approx v_1$. Substituting this into the Helmholtz equation gives

$$\begin{aligned} & \left\{ \nabla^2 \psi_0 + (\nabla \psi_0)^2 + \omega^2 / v_0^2(\vec{x}) \right\} \\ & + \epsilon \left\{ \nabla^2 \psi_1 + 2\nabla \psi_0 \nabla \psi_1 - 2\omega^2 v_1(\vec{x}) / v_0^3(\vec{x}) \right\} \\ & + \epsilon^2 \left\{ (\nabla \psi_1)^2 + 3\omega^2 v_1^2(\vec{x}) / v_0^4(\vec{x}) \right\} = 0 \end{aligned} \quad (\text{A.2.7})$$

where the first term in the brackets is a solution of the Helmholtz equation with $v = v_0(\vec{x})$ and $u = e^{\psi_0}$. In the first Rytov approximation, the ϵ^2

term is dropped. Equating the second term to zero gives

$$\nabla^2 \psi_1 + 2\nabla \psi_0 \nabla \psi_1 = 2v_1 \omega^2 / v_0^3 \quad (\text{A.2.8})$$

In order to compare with the previous approximations let $\psi_0 = \bar{k} \varphi_0$ and $\psi = \varphi_1 / \varepsilon = \bar{k} \varphi_1$, then (A.2.7) can be written

$$\begin{aligned} & \left\{ \bar{k} \nabla^2 \varphi_0 + \bar{k}^2 (\nabla \varphi_0)^2 + \bar{k}^2 \left(\frac{k_0^2}{\bar{k}^2} \right) \right\} \\ & + \left\{ \nabla^2 \varphi_1 + \bar{k} 2 \nabla \varphi_0 \nabla \varphi_1 + \bar{k} \left(\frac{k_0^2}{\bar{k}^2} \frac{2v_1}{v_0} \right) \right\} \\ & + \left\{ (\nabla \varphi_1)^2 + \left(\frac{k_0^2}{\bar{k}^2} \frac{3v_1^2}{v_0^2} \right) \right\} = 0 \end{aligned} \quad (\text{A.2.9})$$

where $\bar{k} = \omega / \langle v_0(\vec{x}) \rangle$, $k_0 = \omega / v_0(\vec{x})$, and $\bar{k} \approx k_0$. The Rytov approximation again retains the first two brackets, where the first bracket is assumed known. The $\bar{k}^2$ terms relates to the Eikonal equation in equation (A.2.2), only now related to the velocity model $v_0(\vec{x})$. The $\bar{k}$ terms correspond to the transport equation in (A.2.2) plus a correction term to the velocity linear in $v_1(\vec{x})$. Finally, the term $\nabla^2 \varphi_1$ is retained, but not the term $(\nabla \varphi_1)^2$ and the higher order terms containing the velocity correction $v_1(\vec{x})$. This is similar to the parabolic approximation where again only a part of the third bracket in (A.2.2) is retained. In that case the transverse Laplacian perpendicular to the predominant direction of the incident wave is retained. For example if we let v_0 be a homogeneous velocity, and $\psi_0 = i\bar{k}x$, then by dropping terms in ε^2 , (A.2.7) can be written in terms of $U = e^{i\psi_1}$

$$2ikU_z + \left\{ \nabla^2 U - \frac{(\nabla U)^2}{U} \right\} - \frac{2\varepsilon v_1 \omega^2}{v_0^3} U = 0$$

This is similar to the parabolic equation given in (A.2.5) with the thin lens term approximated to first order in $\varepsilon v_1(\vec{x})$ and the term $\nabla_{\text{perp}}^2 U$ replaced by $\nabla^2 U - (\nabla U)^2/U$. Using the Transverse Laplacian results in a simpler equation but the Rytov equation above is consistent to order ε .

2.4.8 APPENDIX 3

In this appendix, the acoustic wave equation is investigated in a simple 1-D example, where

$$L = \left\{ \frac{1}{K(\tau)} \frac{\partial^2}{\partial t^2} - \nabla \left(\frac{1}{\rho(\tau)} \nabla \right) \right\}$$

The transposed linear sensitivity operators for a perturbation in density and bulk modulus from the field are given in Tarantola [1984a] where

$$\begin{aligned} \delta K_k' &= U_k^T \delta p = \\ &= \frac{1}{K^2(\tau)} \sum_{r_s} \sum_{r_g} \int dt \left\{ \ddot{p}(\tau, t; r_s) \cdot g(\tau, t; r_g) \right\} \delta p(r_g, t; r_s) \end{aligned}$$

$$\begin{aligned} \delta \rho_k' &= V_k^T \delta p = \\ &= \frac{1}{\rho^2(\tau)} \sum_{r_s} \sum_{r_g} \int dt \left\{ \nabla p(\tau, t; r_g) \cdot \nabla g(\tau, t; r_s) \right\} \delta p(r_g, t; r_s) \end{aligned}$$

or

$$\delta K_k' = \frac{1}{K^2(\tau)} \int dt \sum_{r_s} \ddot{p}(\tau, t; r_s) \left\{ \sum_{r_g} g(\tau, -t; r_g) \cdot \delta p(r_g, t; r_s) \right\} \quad (\text{A.3.1})$$

$$\delta \rho_k' = \frac{1}{\rho^2(\tau)} \int dt \sum_{r_s} \nabla p(\tau, t; r_s) \left\{ \sum_{r_g} \nabla g(\tau, -t; r_g) \cdot \delta p(r_g, t; r_s) \right\} \quad (\text{A.3.2})$$

In the 1-D case, a coincident source and receiver are considered where $r_g = r_s = z = 0$. Small perturbations of a homogeneous starting model to the true model will be assumed. The observed reflection seismogram is

$$P_{\text{obs}}(r_g, t; r_s) = P_{\text{obs}}(t)$$

and the predicted seismogram is

$$P_k(\tau_g, t; \tau_s) = P_k(t)$$

The differential seismogram is then

$$\delta p_k(\tau_g, t; \tau_s) = P_{obs}(t) - P_k(t)$$

The operation in (A.3.1) can then be written

$$\delta K_k' = \frac{1}{K_k^2(\tau)} \int dt \ddot{p}_k(\tau, t; \tau_s) \langle g(z, t; 0) \delta p_k(t) \rangle$$

where $\langle \dots \rangle$ denotes a cross correlation. The analogous operation on $\delta \rho_k'$ in (A.3.2) can be written

$$\delta \rho_k' = \frac{1}{\rho_k^2(z)} \int dt \frac{\partial}{\partial z} p_k(z, t; 0) \frac{\partial}{\partial z} \langle g(z, t; 0) \delta p_k(t) \rangle$$

With an initial homogeneous model with velocity v then

$$g(z, t; 0) = \delta(t - z/v)$$

and

$$\langle g(z, t; 0) \delta p_k(t) \rangle = \delta p_k(t + z/v)$$

Then

$$\delta K_k'(z) = \frac{1}{K_k^2(z)} \int dt \frac{\partial^2}{\partial t^2} p_k(z, t; 0) \delta p_k(t + z/v)$$

$$\delta \rho_k'(z) = \frac{1}{\rho_k^2(z)} \int dt \frac{\partial}{\partial z} p_k(z, t; 0) \frac{\partial}{\partial z} \delta p_k(t + z/v)$$

Now with $p_k(z, t; 0) = \delta(t - z/v)$ then

$$\delta K_k' = \frac{1}{K_k^2(z)} \delta \ddot{p}_k(t = 2z/v)$$

$$\delta \rho_k(z) = \frac{1}{v^2 \rho_k^2(z)} \delta \ddot{p}_k(t = 2z/v)$$

where $t = 2z/v$ places the model perturbation at $z = vt/2$. Since $v = \sqrt{K/\rho}$ and $\eta = \rho v = \text{impedance}$, then

$$\frac{\delta v}{v} = \left\{ \frac{\delta K}{K} - \frac{\delta \rho}{\rho} \right\}$$

$$\frac{\delta \eta}{\eta} = \left\{ \frac{\delta K}{K} + \frac{\delta \rho}{\rho} \right\}$$

Including the model covariances this can be written assuming,

$$C_K \approx \sigma_k^2 \approx K_k^2(z)$$

$$\delta K_k(z) = C_K \delta K_k(z) \approx \frac{\sigma_k^2}{K_k^2(z)} \delta \ddot{p}(2z/v) \approx \delta \ddot{p}(2z/v)$$

and with $\sigma_\rho^2 \approx \rho_k^2(z)$.

$$\delta \rho_k = C_\rho \delta \rho_k \approx \frac{\sigma_\rho^2}{v^2 \rho_k^2(z)} \delta \ddot{p}(2z/v) \approx \frac{1}{v^2} \delta \ddot{p}(2z/v)$$

Perturbations in velocity and impedance can then be written

$$\frac{\delta v}{v} = \frac{1}{2} \left[\frac{1}{K} - \frac{1}{v^2 \rho} \right] \delta \ddot{p}(2z/v) = 0$$

and

$$\frac{\delta \eta}{\eta} = \frac{1}{2} \left[\frac{1}{K} + \frac{1}{K} \right] \delta \ddot{p}(2z/v) = \frac{1}{K} \delta \ddot{p}(2z/v)$$

Thus for the 1-D reflection seismogram modelled using the acoustic wave equation, there is no velocity perturbation, but there is an impedance perturbation. This is as expected from the 1-D exact inversion results.

2.4.9 REFERENCES

- Aki, K., Three dimensional seismic velocity anomalies in the lithosphere. Method and summary of results, *Journal of Geophysics*, 43, 235-242, 1977.
- Aki, K., Scattering of P waves under the Montana Lasa, *J. Geophys. Res.*, 78, 1334-1346, 1973.
- Aki, K., A. Christoffersson, and E.S. Husebye, Determination of the three-dimensional seismic structure of the lithosphere, *J. Geophys. Res.*, 82, 277-296, 1977.
- Aki, K., and P.G. Richards, *Quantitative Seismology: Theory and Methods*, W.H. Freeman and Company, San Francisco, 1980.
- Barabanenkov, Yu. N., Yu. A. Kravtsov, S.M. Rytov, and V.I. Tamarskii, Status of the theory of propagation of waves in a randomly inhomogeneous medium, *Soviet Physics Usp.*, 13, 551-575, 1971.
- Brown, G.B., Linearized travel time, intensity, and waveform inversion in the ocean sound channel - a comparison, *J. Acoust. Soc. Am.*, 75, 1451-1461, 1984.
- Brown, W.P., Validity of the Rytov approximation in optical propagation calculations, *J. Opt. Soc. Am.*, 56, 1045-1052, 1966.
- Brown, W.P., Validity of the Rytov approximation, *J. Opt. Soc. Am.*, 57, 1539-1543, 1967.
- Censor, Y., Row-action methods for huge sparse systems and their applications, *SIAM Review*, 23, 444-466, 1981.
- Cervený, V., M.M. Popov, and I. Psencik, Computation of wave fields in inhomogeneous media - Gaussian beam approach, *Geophys. J.R. Astron. Soc.*, 70, 109-128, 1982.
- Clayton, R. and R. Comer, A tomographic analysis of mantle heterogeneities from body wave travel times, *EOS, Trans. Am. Geophys. Union*, 64, pp 776, 1983.
- Claerbout, J.F., *Fundamentals of geophysical data processing*, McGraw Hill, New York, 1976.

- Cohen, J.K., and N. Bleistein, Velocity inversion procedure of acoustic waves, *Geophysics*, 36, 467-481, 1979.
- Haddon, R.A.W. and E.S. Husebye, Joint interpretation of p-wave time and amplitude anomalies in terms of lithospheric heterogeneities, *Geophys. J.R. Astron. Soc.*, 55, 19-43, 1978.
- Humphreys, E., R. Clayton, and B.H. Hager, A tomographic image of the mantle structure beneath southern California, preprint, 1984.
- Ishimaru, A., Wave propagation and scattering in random media, v. 1,2, Academic Press, New York, 1978.
- Ivansson, S., Remark on an earlier proposed iterative tomographic algorithm, *Geophys. J.R. astr. Soc.*, 75, 855-860, 1983.
- Kaveh, M., M. Soumekh, and R.K. Mueller, A comparison of Born and Rytov approximations in acoustic tomography, in *Acoustical Imaging*, Vol. 11, J. Powers, Ed., New York, Plenum, 1981.
- Kaveh, M., M. Soumekh, Z.Q. Lu, R.K. Mueller, and J.F. Greenleaf, Further results on diffraction tomography using Rytov's approximation, in *Acoustical Imaging*, Vol. 12, New York, Plenum, 1982.
- Kaveh, M., M. Soumekh, and J.F. Greenleaf, Signal processing for diffraction tomography, *IEEE trans. on Sonics and Ultrasonics*, Vol. SU-31, 230-239, 1984.
- Keller, J.B., Accuracy and validity of the Born and Rytov approximations, *J. Opt. Soc. Am.*, 59, 1003-1004, 1969.
- Lailly, P., The seismic inverse problem as a sequence of before stack migrations, in *Conference on inverse scattering: theory and applications*, edited by Bednar et al., SIAM, Philadelphia, pp 206-220, 1983.
- Lakshminarayanan, A.V., and A. Lent, Methods of least squares and SIRT in reconstruction, *J. Theor. Biol.*, 76, 267-295, 1979.
- Larner, K.L., Near-receiver scattering of teleseismic body waves in layered crust-mantle models having irregular interfaces, Ph.D. thesis, MIT, 1970.
- Luenberger, D.G., Introduction to linear and nonlinear programming, Addison-Wesley, 1984.

- McMechan, G.A., Seismic tomography in boreholes, *Geophys. J.R. astr. Soc.*, 74, 601-612, 1983.
- Miller, D., M. Oristaglio, and G. Beylkin, A new formalism and an old heuristic for seismic migration, SEG extended abstract, 1984.
- Mueller, R.K., M. Kaveh, and G. Wade, Reconstructive tomography and application to ultrasonics, *Proc. IEEE*, 67, 567-587, 1979.
- Mueller, R.K., M. Kaveh, and R.D. Iverson, A new approach to acoustic tomography using diffraction techniques, in *Acoustic Holography*, Vol 8, edited by A.F. Metherell, Plenum, New York, 1980.
- Newton, R.G., The Marchenko and Gelfand-Levitan methods in the inverse scattering problem in one and three dimensions, in *Conference on Inverse Scattering: Theory and Applications*, edited by Bednar et.al., SIAM, Philadelphia, 1-74, 1983.
- Nowack, R. and K. Aki, The 2-D Gaussian beam synthetic method: testing and application, *J. Geophys. Res.*, 89, 7797-7819, 1984.
- Patton, H. and K. Aki, Bias in the estimate of seismic moment tensor by the linear inversion method, *Geophys. J. R. astr. Soc.*, 59, 479-495, 1979.
- Raz, S., Three-dimensional velocity profile inversion from finite offset scattered data, *Geophysics*, 46, 837-842, 1981.
- Schneider, W.S., Integral formulation for migration in two and three dimensions, *Geophysics*, 43, 49-76, 1978.
- Shaw, P., Waveform inversion of explosion data, Ph.D. thesis, U. of Calif. San Diego, 1983.
- Tarantola, A., Inversion of seismic reflection data in the acoustic approximation, *Geophysics*, 49, 1259-1266, 1984a.
- Tarantola, A., Iterative migration = inversion 1: linearized problem for an homogeneous reference model, preprint, 1984b.
- Tarantola, A., The principles of seismic imagery, preprint, 1984c.
- Tarantola, A. and B. Valette, Generalized nonlinear inverse problems solved using the least-squares criterion, *Rev. of Geophys. and Space Phys.*, 20, 219-232, 1982.
- Tatarskii, V.I., The effects of the turbulent atmosphere on wave

propagation, (translated from Russian by Israel Program for Scientific Translations Ltd., 1971)

Taylor, L.S., On Rytov's method, *Radio Science*, 2, 437-441, 1967.

Thomson, C.J. and D. Gubbins, Three-dimensional lithosphere modelling at NORSAR: linearity of the method and amplitude variations of the anomalies, *Geophys. J.R. Astron. Soc.*, 71, 1-36, 1982.

Thurber, C., Seismic detection of the summit magma complex of Kilauea volcano, Hawaii, *Science*, 223, 165-167, 1984.

Tribolet, J.M., A new phase unwrapping algorithm, *IEEE Trans. Acoust, Speech and Signal Processing*, Vol. ASSP-25, 170-177, 1977.

Yura, H.T., C.C. Sung, S.F. Clifford, and R.J. Hill, Second-order Rytov approximation, *J. Opt. Soc. Am.*, 73, 500-502, 1983.

3. DIRECT MEASUREMENT OF THE MANTLE ATTENUATION OPERATOR FROM BROADBAND P AND S WAVEFORMS

3.1 INTRODUCTION

Compelling evidence for the existence of frequency dependent Q has come primarily from comparisons of amplitude ratios of body waves recorded by instruments with widely separated and, often, narrow frequency bands (e.g., Gutenberg, 1958; Solomon, 1972; Der and McElfresh, 1977; Sipkin and Jordan, 1979; Lundquist and Cormier, 1980). It is now widely accepted that the incongruity posed by attenuation models derived for different frequency bands can not be resolved without forsaking the assumption of frequency independent attenuation (e.g., see the review by Cormier, 1982). Until recently, studies of attenuation over a broad and continuous range of frequencies, say, from several Hz to tens of seconds were few (among them Uleg and Berckhemer, 1984; Der et al, 1985). A major impediment to broadband studies has been that conventional seismograph systems, such as that of the World-Wide Standardized Seismograph Network, were usually designed to filter out data in the midband range of seismic frequencies (about 0.1-1.0 Hz). With recent improvements in digital instrumentation and increasing availability of broadband recordings, it becomes imperative that the behavior of intrinsic attenuation in the earth be understood across a broad and continuous spectrum of frequencies which encompasses the midband range. Clearly our ability to infer physical properties and processes in the earth can only be as complete as our understanding of geophysical parameters such as attenuation. More importantly, a correct description of body-wave attenuation is critical if studies of earth structure and source mechanism are to be conducted properly. The choice of an attenuation operator in the computation of synthetic seismograms or for the correction of propagation effects in data can inordinately influence the interpretation of the rupture process of earthquakes (Hanks, 1981).

A new method of measuring the frequency dependence of Q is presented.

The method consists of matching details of displacement and velocity records of broadband P- and SH-type body waves with synthetic body waves. The frequency dependent effects of a finite source are incorporated in the synthetics by using a source theory that provides a realistic and causal description of rupture and healing for simple earthquakes. The effects of attenuation are incorporated by using dispersive operators conforming to the Kramers-Krönig relations. The conditions imposed by causal rupture and causal attenuation predict phase spectra that can behave in such radically dissimilar fashion that the tradeoff between source-time function and attenuation is strongly minimized. It is possible to observe any antithetical behavior in the phase spectra only with waveforms that contain information over a broad and continuous frequency range. At the same time, the modeling technique is consistent with assumptions about source behavior that are commonly used by attenuation studies that model the asymptotic behavior of amplitude spectra at high and low frequencies. First, some general characteristics of broadband displacement and velocity waveforms are described, followed by a description of the technique used to generate theoretical seismograms. Afterwards, the method is applied to the analysis of P and

SH waveforms generated by a deep earthquake that occurred in the Sea of Okhotsk. Finally, there is a discussion of the implications of the final range of Q_α and Q_β models on the mechanisms of attenuation in the earth.

3.2 CHARACTERISTICS OF BROADBAND WAVEFORMS

The data analysed in this paper are taken from digitally recording stations of the GDSN and RSTN networks and the GRF array. Operating characteristics of the seismographs may be found in Ganse and Hutt (1982), Engdahl et al. (1982) and Harjes and Seidl (1978). Except for that of GRF array, most of the raw data are not directly proportional to either displacement or velocity over a wide and continuous range of frequencies. The most common seismograph configuration splits an incoming signal into long- and short-period components. Although an intermediate-period component is recorded by the RSTN and many GDSN stations, the bandwidth of the data needs to be augmented. In order to obtain broadband records of displacement and velocity, we use the method described by Harvey and Choy (1982) which recombines data recorded on separate recording channels by a multichannel instrument deconvolution. It is the high dynamic range of the seismographs and the digital format of the data which enable spectral information over a broad frequency range to be easily retrieved. The accuracy of the multichannel deconvolution has been checked by a comparison of the deconvolved P waves obtained from the short- and long-period channels of the SRO station at GRFO with the P wave recorded directly as broadband velocity at the same site by a GRF station (Choy, 1981). The accuracy of the multichannel deconvolution has been further checked by comparing the processed waveform with the waveforms recorded on an intermediate-period component at many stations. The P waves in this paper that are obtained by this processing generally have a flat displacement or velocity response from 0.02 to 5 Hz; the S waves generally have a flat response to displacement or velocity from .01 to between 2.0-5.0 Hz, depending on the recording instrument.

Details of the rupture process which are suppressed in narrowband data are often obvious with broadband data. Some of the advantages of broadband data are illustrated in Figures 3.2-3.4. Figure 3.2 shows the broadband P waves that were obtained for a deep event that occurred 20 April 1985 in the Sea of Okhotsk. The global distribution of the digitally recording stations about the epicenter is shown in Figure 3.1. That the earthquake ruptured in a complex fashion is immediately obvious from the pattern of double pulses that can be seen in the displacement records at all stations. The coherence of the complex rupture history across the suite of stations is even more evident in the velocity records. The doublet in displacement becomes the first two peaks and the ensuing trough in the velocity records. The strong coherence of these features is not surprising. Velocity, being directly related to energy radiated through the focal sphere, is more sensitive to the dynamics of the source than is displacement. This obvious complexity of rupture, however, would be virtually undetectable in conventional narrowband form. Figure 3.3, for instance, shows

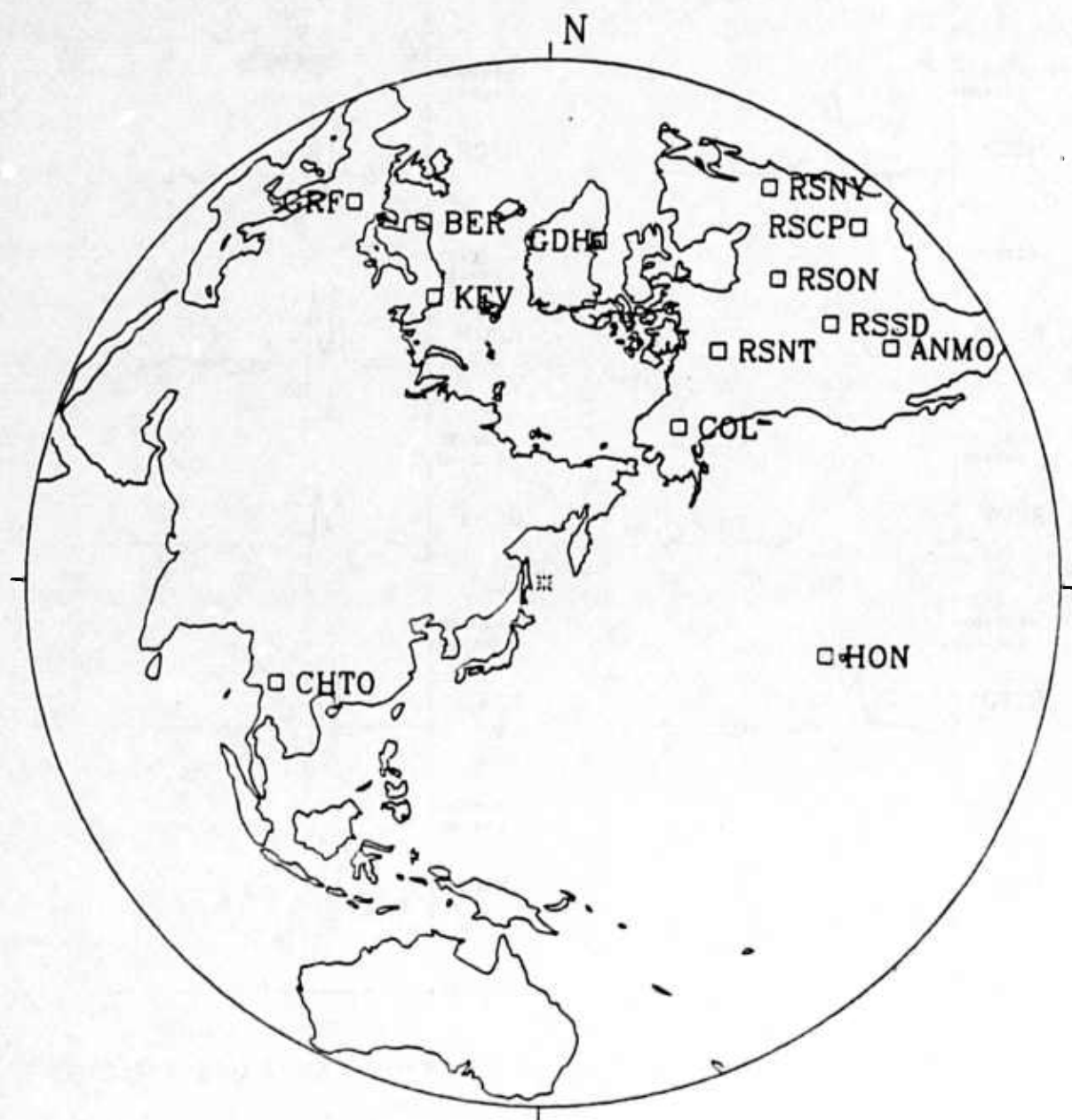


Figure 3.1. A polar projection showing stations of the GDSN, RSTN and GRF array within 90° of the epicenter of the deep (h 582 km) Sea of Okhotsk earthquake that occurred on 20 April 1984.

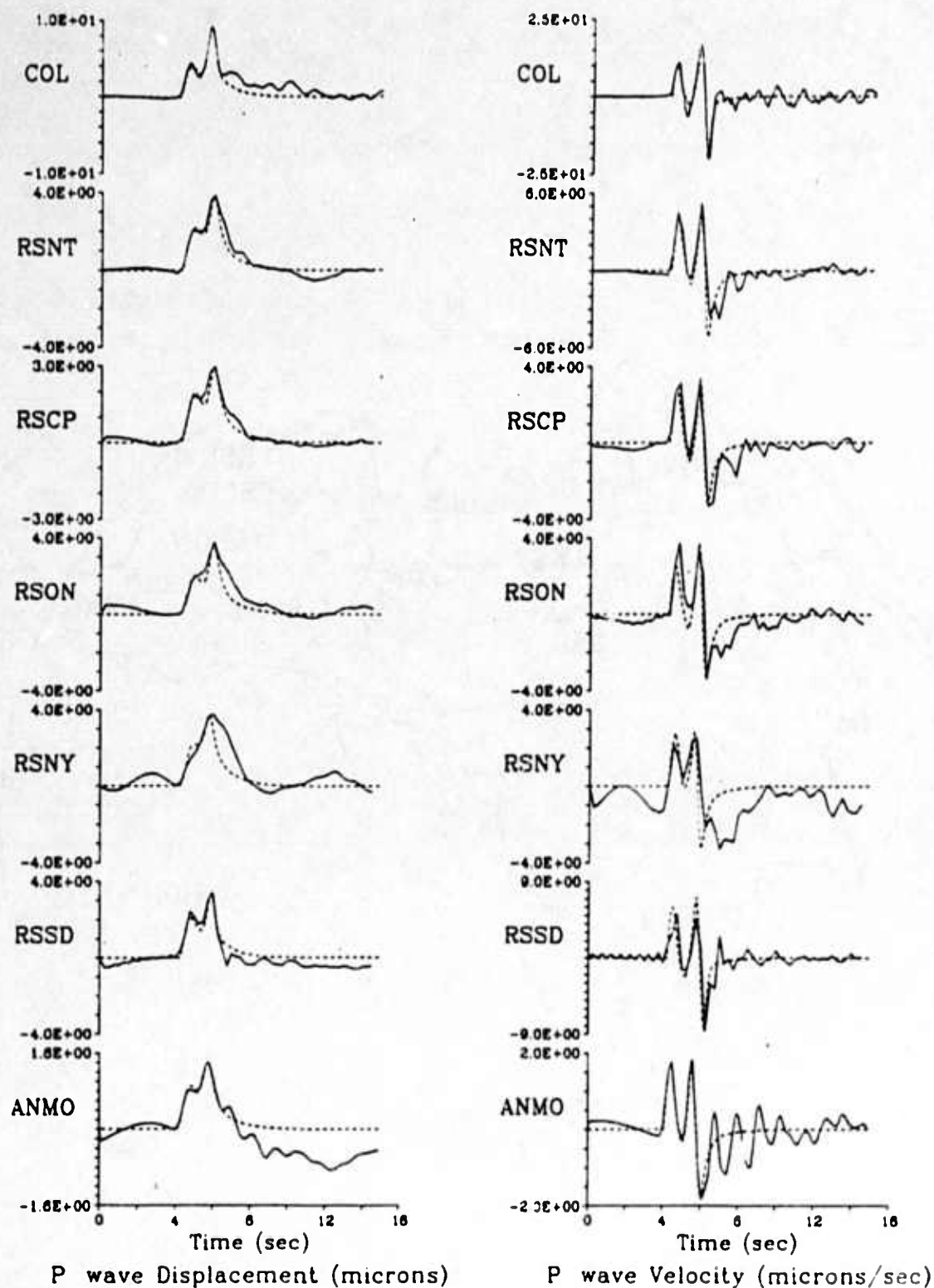


Figure 3.2. Broadband P-waves from the GDSN, RSTN and GRF array for the deep Sea of Okhotsk earthquake. Polarities of all first motions have been plotted as positive. Data are plotted as solid lines; synthetics are plotted as dashed lines. In some displacement records (right column), a double pulse is very evident. In other displacement records, the complexity is seen as a strong inflection in displacement. In all the velocity records (left column), the complexity of rupture is seen as the first two peaks and the subsequent trough.

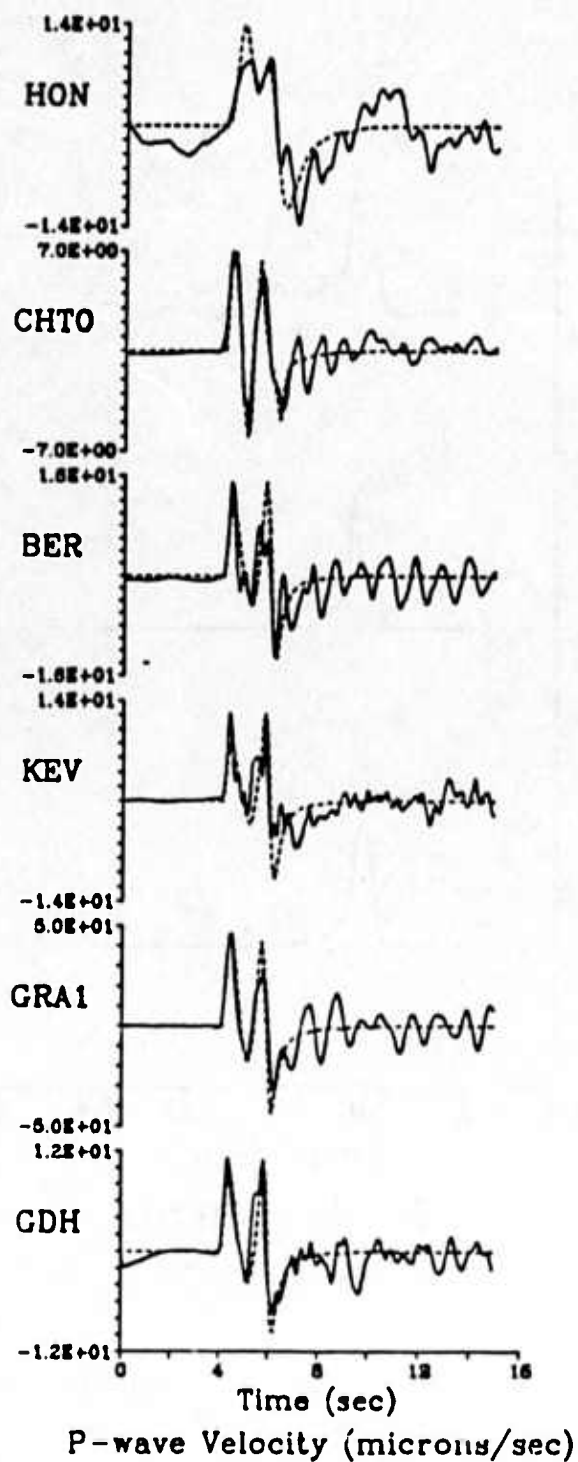
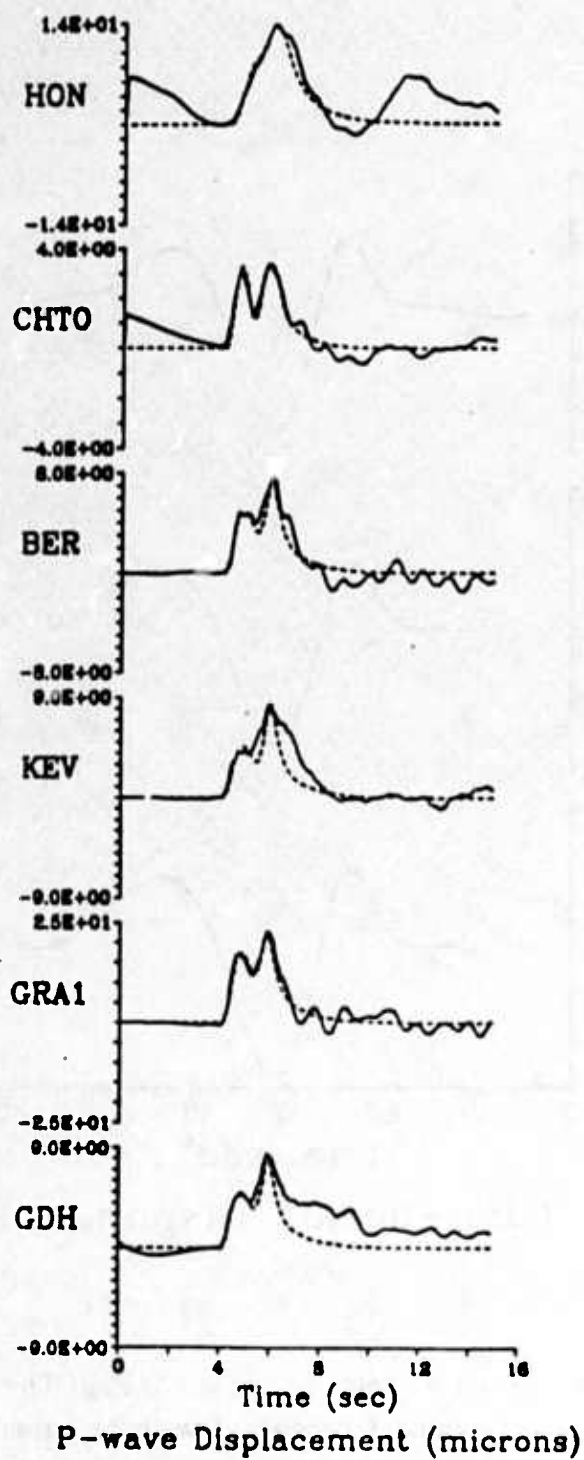


Figure 3.2 (continued)

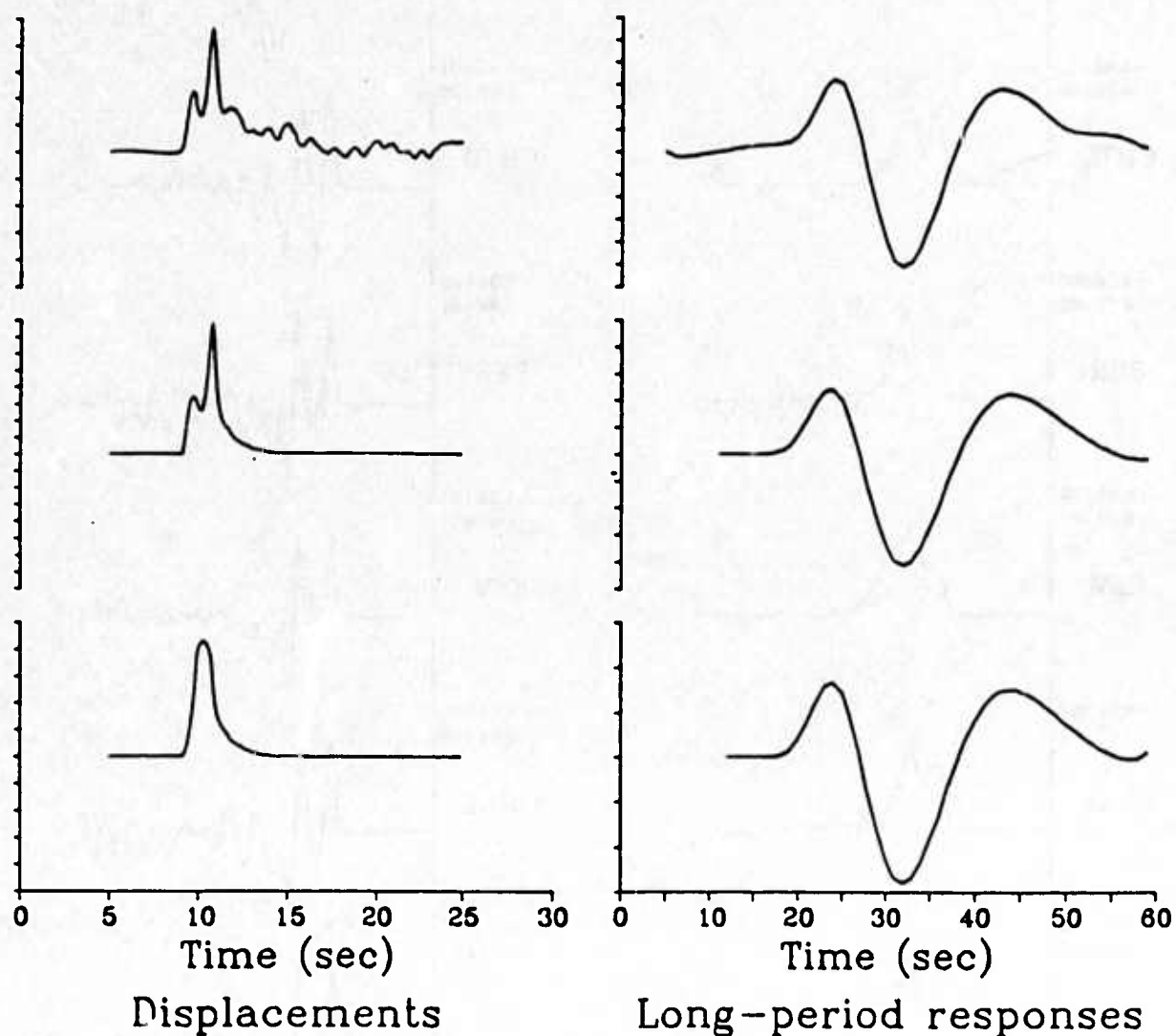


Figure 3.3. Three different displacement pulses are shown at the left. (*Top*) The actual P wave at COL. (*Middle*) The synthetic displacement convolved with an attenuation operator. (*Bottom*) An arbitrarily different source operator convolved with the same attenuation operator. This source operator has the same total duration as the source operator used in the middle example, but has no rupture complexity. The corresponding responses to convolution with a long-period seismograph response are shown at the right.

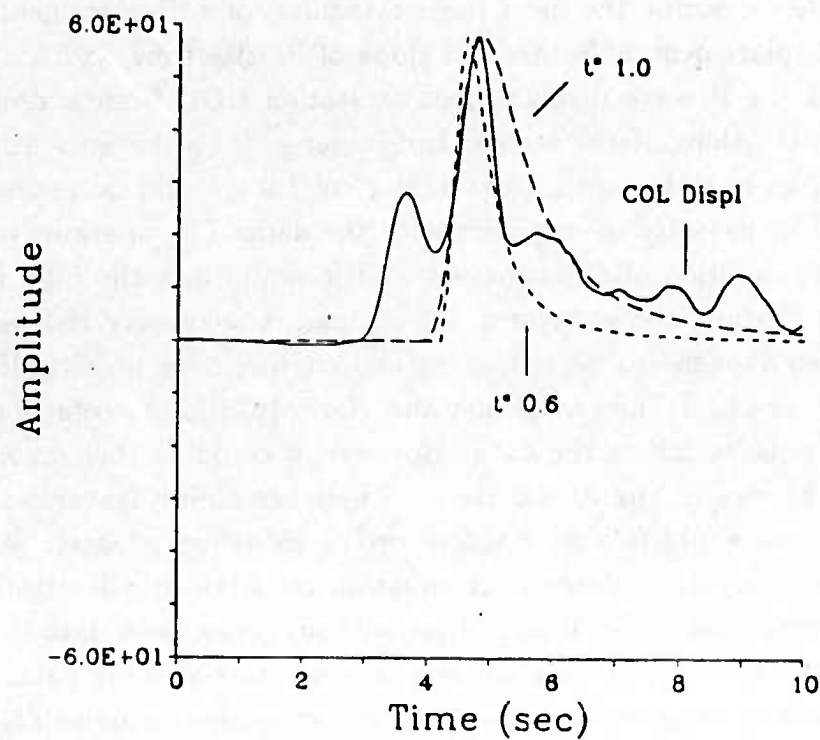


Figure 3.4. Comparison of actual P-wave displacement pulse (solid line) with frequency-independent attenuation operators. With a t^* of 1.0 s (long dashes), it is possible to match the slope of the rise time in the second subevent in the actual data. But such attenuation would be unable to match the pulse duration and it would lack the high frequency content that is shaping the pulse. A constant t^* of 0.6 s could match the pulse width of the subevent, but would be unable to match either the initial slope or the decay history of the data.

three different displacement records and the result of convolving them with a long-period seismograph response. The actual broadband displacement from COL shows the double pulse complexity. The second trace shows the best-fitting synthetic record using the source and attenuation operators derived later in this paper for COL. The last trace shows a much different displacement record obtained by convolving the attenuation operator with an arbitrarily different source operator. This source operator has the same total pulse duration as the previous operator but assumes a single simple source using the model of Sato and Hirasawa (1973). The corresponding long-period seismograms are apparently indistinguishable from each other (Figure 3.3b). The effect of the instrument response has been to suppress major differences between the source operators beyond visual perception.

Conversely, it is obvious with broadband data that taking a Q-model derived from low-frequency data and extending it in frequency-independent fashion to high frequencies cannot adequately describe the most basic attributes of a displacement pulse. Two basic features of a displacement pulse are the slope of its rise time and its total pulse width. Figure 3.4 shows the P wave displacement at station COL from a deep earthquake. In order to match the slope of the second and larger of the subevents using an attenuation operator alone, an operator with a constant t^* of 1.0 s would be necessary. But such an operator would be patently incongruent with the data. The operator would be unable to match the short duration of the subevent and it would lack the high frequency spectral content that is shaping the waveform. If, instead, a frequency independent $t^*=0.6$ s is used (which also happens to be the operator that would be predicted by the frequency-independent Q model of SL8 [Anderson and Hart, 1978]), the resulting waveform would have the same pulse width as the data. However, it could neither match the initial slope nor the decay history of the actual data. These remaining features can only be fit by incorporating more sophisticated features in the modeling process. As described later, these will be frequency dependence in attenuation and a realistic description for the rupture process. It is important to emphasize that without broadband data it would be difficult to evaluate the fit between various attenuation operators and the data. The intermediate and high frequencies which would otherwise unmistakably distinguish the effect of various attenuation operators on a source function are often suppressed upon convolution with the response of a long-period seismograph.

Finally, a distinct advantage of time-domain analysis over a purely spectral approach can be seen by comparing the P wave displacement pulse with its amplitude spectrum (Figure 3.5). The subevents which are evident in the time domain manifest themselves only as a pronounced modulation in the spectrum between approximately 0.1-1.0 Hz. In spectral form, it would be difficult to obtain any intuition about the rupture process. Also, the spectral modulation might tend to bias estimates of attenuation based on the spectral decay unless one averages the spectra of many events with different rupture histories from the same region.

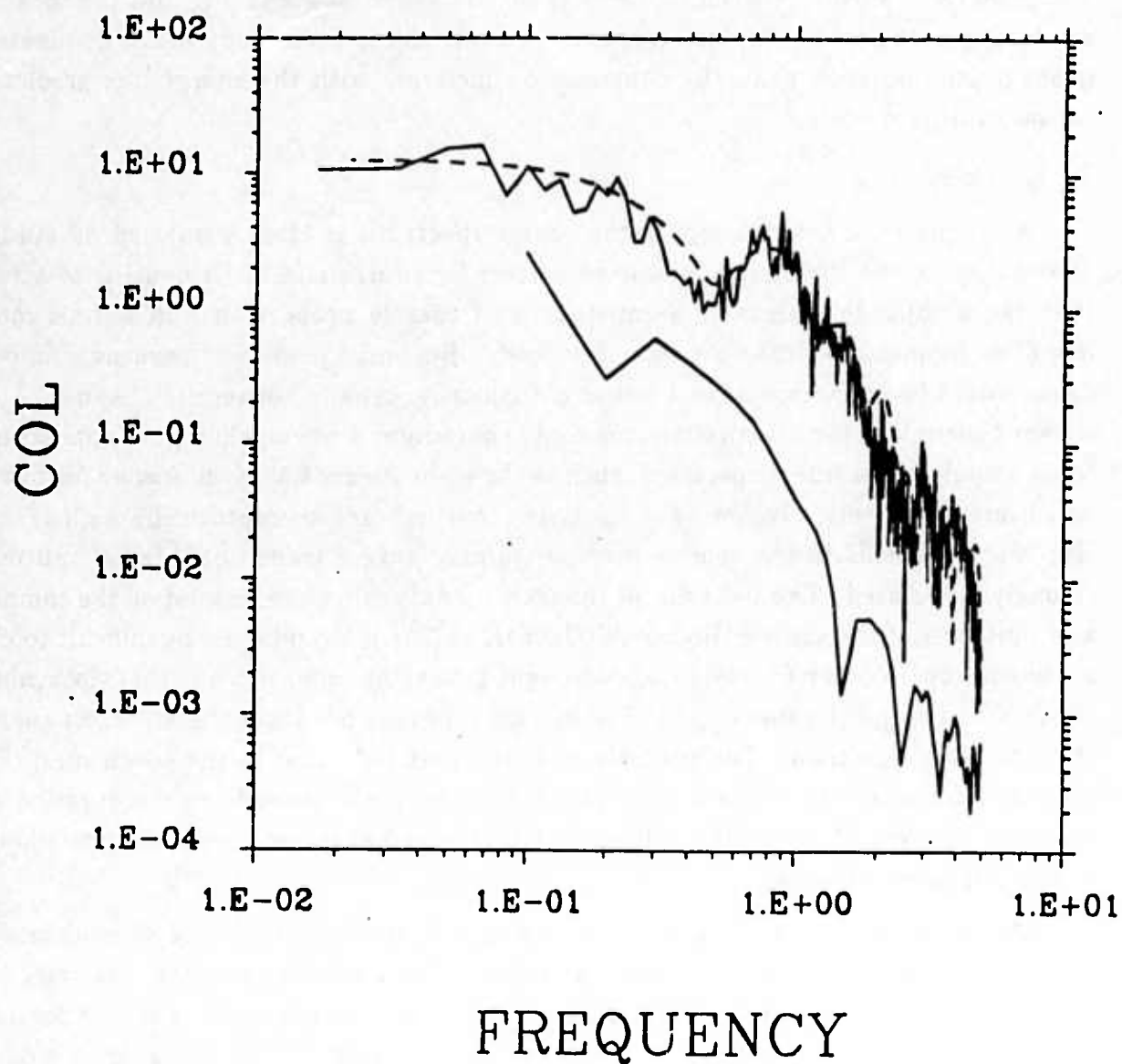


Figure 3.5. Amplitude spectrum of the P-wave displacement at COL (upper solid line). The noise spectrum at COL using a 10 s window preceding the arrival of P is shown by the lower solid line. Most stations had very good signal-to-noise ratio for the continuous range of frequencies between 0.02-5.0 Hz. The amplitude spectrum of the synthetic is shown by the dashed line.

3.3 COMPUTATION OF BODY WAVES

At teleseismic distances, the cone of take-off angles for the seismic rays describing the earth's response between source and receiver is very narrow, allowing source effects to be decoupled from propagation effects. In order to model body waves, our experimental procedure is to find those source and propagation operators which, when convolved with each other, match the observed data the best. In this section we first describe our method for computing source and propagation operators. Then we describe the constraints which minimize the tradeoff between these operators. It turns out that the constraints are strong because the dispersive property of intrinsic attenuation, which dominates the propagation operator, generally contrasts distinctively with the energy flux predicted by causal rupture models.

Source Operator

A simple parameterization of the source spectrum is often employed by studies of attenuation in the frequency domain to correct for source effects. It consists of assuming that the amplitude maintains a constant level that is proportional to seismic moment from low frequencies up to a corner frequency. Beyond the corner frequency, amplitude decay would be proportional to a power of frequency, usually between ω^{-2} and ω^{-3} . Even without correcting for attenuation this simple paradigm is obviously a poor approximation to an actual displacement spectrum, such as shown in Figure 3.5, for all frequencies but those which are asymptotically low ($f \ll 0.1$ Hz) or which are asymptotically high ($f \gg 1.0$ Hz). For intermediate frequencies the spectra may have a transitional falloff and may be strongly modulated. The behavior of the transitional region can depend on the complexity and finiteness of the source (Boore and Joyner, 1978). It would even be difficult to choose an unequivocal corner frequency. Boatwright (1984) has pointed out that, in a multiple rupture, the corner frequency of the dominant subevent produces the strongest change in the amplitude spectrum. For example, the strongest inflection in the spectrum of Figure 3.5 occurs at about 1 Hz. The inverse of this frequency corresponds to the duration of the stronger subevent of the double pulse at COL (Figure 3.4) rather than to the total duration of the displacement pulse.

Our experience with broadband data suggests that the majority of moderate-sized deep earthquakes rupture in a complex fashion. With broadband data, systematic variations in waveform about the focal sphere can be seen and related to source complexity for earthquakes as small as $m_b 5.5$ (Choy and Boatwright, 1981). In order to model the broadband information it is essential that a suitably realistic model of rupture be used. We follow the method of Choy and Boatwright (1981) to generate our source operators. An important aspect of the procedure is that constraints on the rupture process are obtained primarily through modeling the *velocity* pulse shapes. Velocity, being directly related to energy radiated through the focal sphere, naturally emphasizes energy near and above the corner frequency of an event. For each earthquake all the velocity pulse shapes are exam-

ined as a function of takeoff angle and azimuth about the fault plane in order to identify features which are coherent or which vary systematically from station to station. Source operators are then obtained by modeling the rupture process with the quasidynamic models of Boatwright (1980). These models describe rupture growth as self-similar and rupture healing as related causally to the stopping of rupture growth. At asymptotically high frequencies, the spectral falloff of these models varies between ω^{-2} and ω^{-3} , depending on the healing of the rupture. Thus, at high frequencies these models conform to asymptotic corrections used in many spectral amplitude approaches.

Propagation operator

Propagation effects include geometrical spreading, attenuation and interactions with discontinuities in the earth. We first discuss the attenuation operator which is the dominant propagation effect in the coherent portion of the waveforms. Then we discuss interactions with earth structure, which generally produce minor and late-arriving features in the waveforms.

Attenuation operators can be computed analytically by any of several methods of seismogram synthesis, such as the full wave, reflectivity and WKBJ methods (e.g., Cormier and Richards, 1976; Kennett, 1975; Chapman, 1978). If attenuation is both dispersive and frequency dependent, the computation becomes expensive and time-consuming. Our approach, therefore, is to first compute dispersive operators conforming to the Kramers-Krönig relations. The operators that fit our data best are then related to Q models that generate the same time-functions for the same propagation paths in the earth.

Assuming asymptotic ray theory, linear visco-elasticity, and no scattering losses, the complex spectrum of the Green's function for a body wave propagating from a reference point $\mathbf{x}_0$ on the rupture surface to a receiver at $\mathbf{x}$ is given by

$$g(\mathbf{x}, \omega) = G_o(\mathbf{x}, \mathbf{x}_0)A(\omega). \quad (3.1)$$

The function G_o is a product of a function of the densities and wave velocities at $\mathbf{x}_0$ and $\mathbf{x}$, a radiation pattern appropriate for the wave type (P or S) and the take-off angles for the ray from $\mathbf{x}_0$ to $\mathbf{x}$, and the geometric spreading from $\mathbf{x}_0$ to $\mathbf{x}$. $A(\omega)$ is defined by

$$A(\omega) = e^{i\omega\hat{T}(\omega)} \quad (3.2)$$

and includes the attenuation and phase shifts due to propagation and dispersion. The complex travel time $\hat{T}(\omega)$ in (3.2) is simply the integral of the complex slowness along the ray path:

$$\hat{T}(\omega) = \int \frac{ds}{\hat{c}(\omega)} \quad (3.3)$$

The complex, dispersive velocity $\hat{c}(\omega)$ in (3.3) is given by $\hat{c}(\omega) = \sqrt{\frac{\hat{M}(\omega)}{\rho}}$, where $\hat{M}(\omega)$ is the complex elastic modulus determined by the parameters of a linear, visco-elastic

reology. By invoking causality, a relation between $Re \hat{T}(\omega)$ and $Im \hat{T}(\omega)$ can be found, which is equivalent to the Kramers-Krönig relations for the complex index of refraction of electromagnetic waves or to those for the complex velocities of seismic waves (Futterman, 1962). For linear visco-elasticity, the fastest arriving frequency component will be the infinite frequency (Gurtin and Herrera, 1965). Thus causality can be satisfied by requiring that the time domain operator $A(t) = 0$, until $t = Re \hat{T}(\infty) = \int \frac{ds}{Re \hat{c}(\infty)}$. This requirement can be shown to lead to the relation

$$[Re \hat{T}(\omega) - Re \hat{T}(\infty)] = H[Im \hat{T}(\omega)], \quad (3.4)$$

where H is the Hilbert transform.

With the definition, $t^*(f) = 2 \int \frac{ds}{Im \hat{c}(\omega)}$, (3.4) can be written as

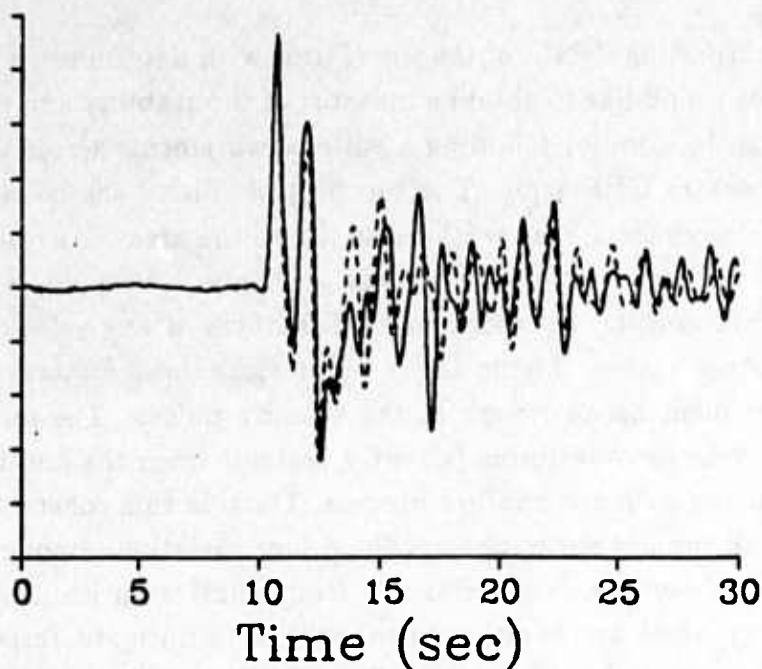
$$Re \hat{T}(\omega) = Re \hat{T}(\infty) + \frac{H[t^*(f)]}{2} \quad (3.5)$$

(If $Q \gg 1$, where $Q = \left| \frac{Re \hat{M}(\omega)}{Im \hat{M}(\omega)} \right|$, then $t^* = \int \frac{Q^{-1}}{Re \hat{c}(\omega)} ds$, and $Re \hat{c}(\omega)$ can be replaced by a velocity c_0 at some reference frequency in the band of interest.) Using (3.4) and the definition of t^* , the complex spectrum of the attenuation operator (3.2) can be rewritten as

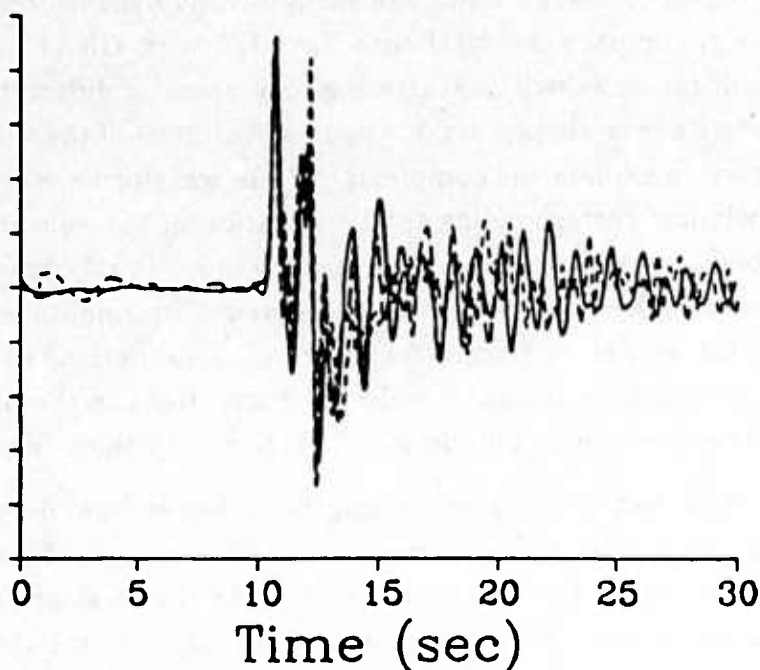
$$A(\omega) = e^{-\frac{\omega t^*(f)}{2}} e^{i\omega \left\{ Re \hat{T}(\infty) + \frac{H[t^*(f)]}{2} \right\}} \quad (3.6)$$

For a generally frequency dependent t^* , the shape in the time domain of $A(t)$ can be found by calculating the Hilbert transform of $t^*(f)$ and inverse Fourier transforming (3.6). $\hat{T}(\infty)$ need not be known and can be incorporated into a reference time needed to bring the apparent start of the pulse into the finite time window determined by the folding frequency of a fast Fourier transform. The frequency dependence of t^* in (3.6), however, cannot be completely arbitrary. Some restrictions on the allowable frequency dependence of $t^*(f)$ are imposed by linear visco-elasticity. In particular, since $t^* = \int \frac{Q^{-1}}{c_0} ds$, the frequency dependence of t^* must be consistent with $Q = C f^n$, where $|n| \leq 1$ (Minster and Anderson, 1981). Also the width and intensity of an absorption band at any given depth must give a positive, non-zero relaxed modulus at the zeroth frequency limit (Minster, 1978).

For linear elasticity and causal propagation, the form of the attenuation operator will be a minimum-delay pulse that arrives after the reference arrival time. Taking the attenuation operator to be minimum phase implies that stochastic processes such as scattering and multipathing are negligible. This is appropriate and even desirable as we are only modeling the features of waveforms that are coherent over the focal sphere. To insure that we



P-velocity at GRF



P-velocity at N. Eur.

Figure 3.6. (*Top*) The broadband velocity records from stations A1 (solid line) and C1 (dashed line) of the GRF array. The stations are about 85 km apart. The coherent portion of the waveforms are easily identified as the first two peaks and the subsequent trough. (*Bottom*) The velocity waveforms from northern European stations: the A1 station of the GRF array (solid line), KEV (short dashes), and BER (long dashes). The coherent component of the waveforms is also easily identifiable on this teleseismic scale.

are not overinterpreting details of the waveforms with deterministic models of rupture and propagation, we would like to obtain a measure of the stability and reliability of broadband pulses. This can be done by following a suite of waveforms across the broadband stations of the large aperture GRF array. The top part of Figure 3.6 shows the typical variation of velocity pulse shapes for a deep earthquake across the array. Any differences among waveforms would be expected to show up most glaringly in velocity, which emphasizes rapid changes in displacement. Yet the dominant features of any velocity waveform are easily found in the other traces. In the GRF waveforms, these features are the initial double peak and subsequent broad trough in the velocity pulses. The duration of the coherent portion of the velocity waveforms (about 4 seconds from the onset of P) corresponds to the duration of the dynamic rupture process. Outside this coherent window the codas of all the traces fall rapidly out of phase. The minor variations among waveforms within the time window of coherence arise primarily from relative enrichment or depletion of high frequency energy which serves either to enhance or to mitigate, respectively, the sharpness and amplitude of the peaks and troughs. The overall stability of the velocity pulses persists on a teleseismic scale. The bottom part of Figure 3.6, for instance, shows the pulse shapes for stations located in Europe. Differences among the velocity waveforms are confined to bursts of high frequency energy which are easily distinguishable from the coherent part of the waveform (e.g, compare the synthetics for KEV with GRA1 in Figure 3.2). Although differential amplification as well as scattering may arise for different propagation paths as frequency increases above about 1 Hz, it appears that most of the energy up to about 5 Hz, which is necessary to explain the complexity of the waveforms, will still arrive coherently within a time window corresponding to the duration of the rupture. In an examination of stability of body waves across the GRF array, Choy (1981) observed that coherence in phase persisted to higher frequencies than coherence in amplitude. This observation is consistent with the theory of Chernov scattering. Fluctuations in the *logarithm* of the amplitude spectrum are on the same order as fluctuations in the phase spectrum. Thus, fluctuations in the absolute amplitude will be larger than those in phase.

As this study is restricted to examining body waves from deep earthquakes, crustal corrections of direct P and SH body waves need only be considered for the receiver end of the propagation path. The crustal response near the receiver could be one source of incoherence among stations, as it is not likely that different stations would have identical crustal responses. However, incorporating the crustal response was generally found to have little effect on body waves during the interval of coherent rupture. This is not surprising as the nature of body waves from deep earthquakes minimizes the effect of crustal response. At teleseismic distances the angles of incidence of P and SH waves are very steep. Because we consider only the vertical component of P and the transverse component of SH, the crustal reverberations predicted by standard earth models are both negligible in amplitude and arrive much later in the wavetrain than the duration of the rupture. It would require local velocity inversions of rather large contrast ($> 25\%$) to produce significant ($> 10^{-7}$)

amplitudes. Such structures would also have to be shallow in order to produce arrivals in the coda that immediately follows the P waves. The variability of the P wave coda could be explained in part by ad hoc low-velocity layers at the surface or within the crust. However, the response of these ad hoc models would still be too small to alter the coherent part of the waveforms significantly. Furthermore, a portion of the coda might not even be related to near-receiver structure. Owens et al. (1983) have inferred the structure beneath stations of the RSTN network. Many of these stations have low-velocity sedimentary layers at the surface. Incorporating the crustal responses into the synthetics did not improve our fits. In fact, the behavior of the theoretical response was occasionally opposite to that observed in the data, suggesting that some other propagation effect can affect the coda. Fortunately, however, with the redundancy of data in North America the coherent part of the waveform can be identified easily and an occasional mismatch does not affect the overall inversion for the source. The velocity pulses consistently exhibit the double peak followed by a trough. More generally, the predominant coherence of displacement and velocity pulses over teleseismic distances indicates that anomalous velocity conditions do not exist at most of the stations we used. As it is not the purpose of this paper to detail crustal structure, the synthetics are presented without convolution with a crustal response.

3.4 MODELING CONSTRAINTS

In our analysis, we derive the best source and attenuation operators that, when convolved with each other, provide the optimum fit to both displacement and velocity data for a suite of pulse shapes about the focal sphere. In order to estimate the error bounds due to the nonuniqueness of this convolution, it is necessary to examine the tradeoff between source and attenuation operators.

Midband and high frequencies

For midband and high frequencies ($f > 0.1$ Hz), the tradeoff is not large because conditions imposed by intrinsic attenuation and causal rupture predict waveforms that can behave in drastically dissimilar fashion. By explicitly considering broadband data, these differences are distinguishable.

Intrinsic attenuation operators conforming to the Kramers-Krönig relations are minimum phase waveforms. That is, they have the minimum possible phase shift for their amplitude spectrum (Robinson, 1967). In the time domain, this describes an impulse response in which the energy buildup occurs as close as possible at the beginning of the waveform. Examples of such operators are shown in the first two columns of Figure 3.7. One of the operators employs a frequency independent t_0' of 0.6 s. The other operator uses the frequency dependent t_0' model that is derived later in this paper. The top trace shows the displacement for each operator. The second trace shows the corresponding velocity trace. Each positive pulse in displacement becomes a positive and negative pulse in velocity. The last trace shows the square of the velocity traces. As energy flux is directly related to veloc-

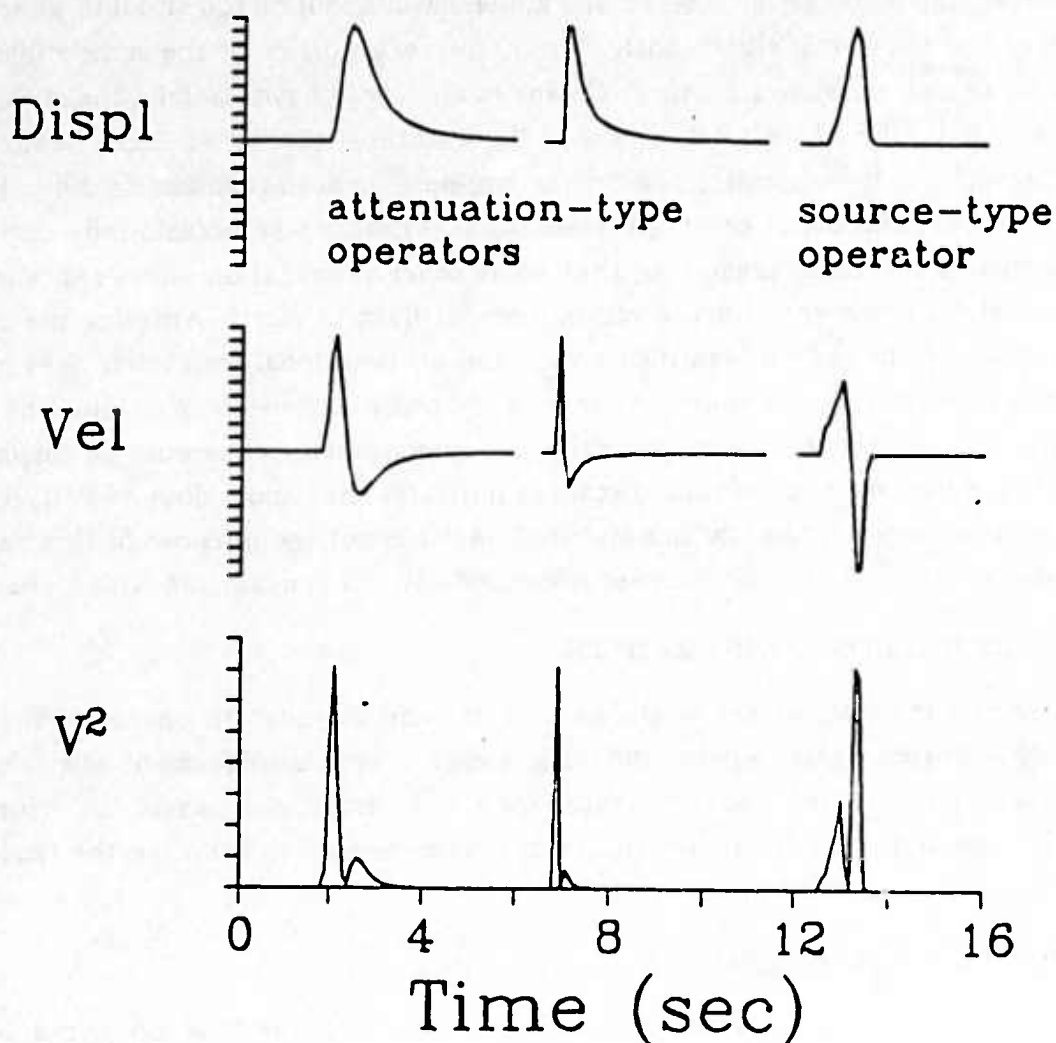


Figure 3.7. The displacement, velocity and velocity-squared records of typical attenuation and source operators. The first column shows an attenuation operator using a frequency-independent $t^* = 0.6$ s. The center column uses the frequency-dependent t^* (typical of a high-Q path (lower solid line in Figure 8)). The third column shows the behavior of a typical source operator. Because operators describing intrinsic attenuation are minimum phase operators, the amplitude of the first pulse of their v^2 records always exceeds the amplitude of the second pulse. The behavior of source operators is often drastically dissimilar. In this example, the amplitude of the second pulse greatly exceeds that of the first pulse.

ity squared, the latter waveforms demonstrate the two major attributes of minimum phase operators. First, the preponderance of energy arrives in the first pulse of velocity-squared. The amplitude of the first velocity-squared pulse is much greater than the amplitude of the second pulse. Secondly, the pulse width of the first pulse in velocity-squared is always less than the width of the second pulse.

Depending on the rupture geometry it is possible for the energy flux of a causal rupture process to vary considerably between widely separated points on the focal sphere. The greatest differences in an asymmetric planar rupture, for instance, would be seen among pulse shapes that are radiated in the direction of and behind the direction of slip. However, by analysing only events which meet rather simple criteria, some strong generalizations can be made about the behavior of energy flux. By using only events with dip-slip focal mechanisms, the takeoff angles for rays recorded at teleseismic distances are then all relatively steep with respect to either fault plane. Furthermore, the teleseismic rays subtend a small cone in which the variation of energy flux, although observable, does not change rapidly with takeoff angle. A majority of earthquakes then are found to rupture such that a substantial amount of energy is radiated as the rupture decelerates and heals. An example is shown in the last column of Figure 3.7. The rupture model used to compute the synthetic is the one that is derived for the dominant subevent of a deep earthquake that is examined later in this paper. The displacement pulse has a rise time and decay time which in velocity becomes two pulses. The first pulse, called the rupture phase, describes rupture growth. The second pulse, called the healing phase, describes healing of the rupture front. The velocity-squared trace is directly proportional to radiated energy flux. In contradistinction to the attenuation operators, this source operator has a velocity-squared function in which it is the amplitude of the second pulse which greatly exceeds that of the first pulse. The pulse widths of the rupture and healing phases are also of the same order. The gross dissimilarity in how energy is radiated over time is what permits us to reduce the tradeoff between source and attenuation. (However, we emphasize that an earthquake with a dip-slip mechanism does not guarantee that the healing phase is dominant. It is possible for some rupture processes to mimic the behavior of a minimum phase operator by radiating more energy during the rupture phase than during the healing phase. Those events in which source and attenuation effects are not easily distinguished are obviously not useful for deriving attenuation operators.)

Asymptotically low and high frequencies

Synthetic seismograms for broadband P waves in which the dominant corner frequency is greater than 0.1 Hz are not generally sensitive to low frequencies ($f < 0.1$ Hz). Our constraints for low frequencies are based on model ABM (Anderson and Given, 1982) which was derived from data that explicitly considered normal mode and long-period body wave data. The depth-corrected value of t_s^* at 0.01 Hz is about 0.8 s for all the teleseismic distances under consideration. For $f = 0.1$ Hz, the t_s^* of model ABM is about 0.6 s

However, neither our synthetics nor the ABM model preclude values up to 0.8 s. Our synthetics slightly favor a value of 0.7 s.

The spectral amplitudes of asymptotically high frequencies ($f \gg 1.0$ Hz) are two or more orders of magnitude smaller than the amplitude of the dominant corner frequency. Consequently, constraints in the high frequency range are easier to obtain by a comparison of synthetic and observed spectra rather than in the time domain. Our synthetic amplitude spectra are usually in excellent agreement with data down to 5.0 Hz (e.g., Figure 3.5). As our source models have an asymptotic falloff that conforms to the asymptotic corrections used in many purely spectral approaches, we expect our error bounds for $t^*(f)$ at high frequencies to be as good as that which would be obtained by spectral approaches.

3.5 P WAVE ANALYSIS

Waveforms for analysis are from a deep focus earthquake that occurred on 20 April 1985 at a depth of 582 km. For the focal mechanism we use the solution from first motions of P waves which is in agreement with the body-wave moment-tensor solution given by the NEIS (National Earthquake Information Service). We have also verified the solution using SH and sSH polarities and relative amplitudes. The other source parameters derived for this earthquake by modeling the broadband P waves are given in Table 3.1. The earthquake consists of two rupture events, separated by 1.0 s. The relative locations of their points of nucleation are not sufficiently well-constrained that a preferred fault plane can be determined. The shape of the source functions is not affected, however, by the choice of a fault plane because the individual events are approximately circular ruptures. As the areas of the rupture events probably overlap, it can be inferred that the initial nucleation was momentarily stopped by a fracture barrier. The failure of this barrier then resulted in the second rupture event.

In computing our synthetics the same source model was used for every station. The $t^*(f)$ operator was allowed to vary slightly in order to accommodate the reality of regional variations in attenuation. The synthetic seismograms obtained by modeling the displacement and velocity records are shown in Figure 3.2. Figure 3.5 shows a typical fit between our synthetic and the observed data in the frequency domain. As expected, a good fit in the time domain transforms to a good fit in the frequency domain. It must be remembered, however, that in order to arrive at the good spectral fit across a continuous frequency range from 0.02 to 5 Hz we first had to model in the time domain. The amplitude spectrum by itself, devoid of any phase information, is normally insufficient to give any intuition about the dynamic behavior of the rupture.

In a radially homogeneous earth and our frequency band of interest, $t_s^*(f)$ is a weak function of distance for teleseismic body waves. Regional variations must exist, however, as we still required two types of operators to synthesize the waveforms (Figure 3.8). These operators are appropriate for rays that propagate through apparently lower and higher

Parameter	Value
Depth	582 km
Strike	236°
Dip	12°
Rake	-90°
Moment	3.8×10^{15} dyn cm
Rupture radius 1	2.3 km
Rupture radius 2	3.5 km

Table 3.1 Source Parameters of the Sea of Okhotsk Earthquake of April 20, 1984 (Origin Time 0631:10:6; 50.12°N, 148.75°E)

The earthquake consisted of two subevents that had approximately circular geometries. The second subevent occurred 1.0 s after the initial event. The relative location of the second event is not well constrained, but its rupture area probably overlaps the rupture area of the first subevent.

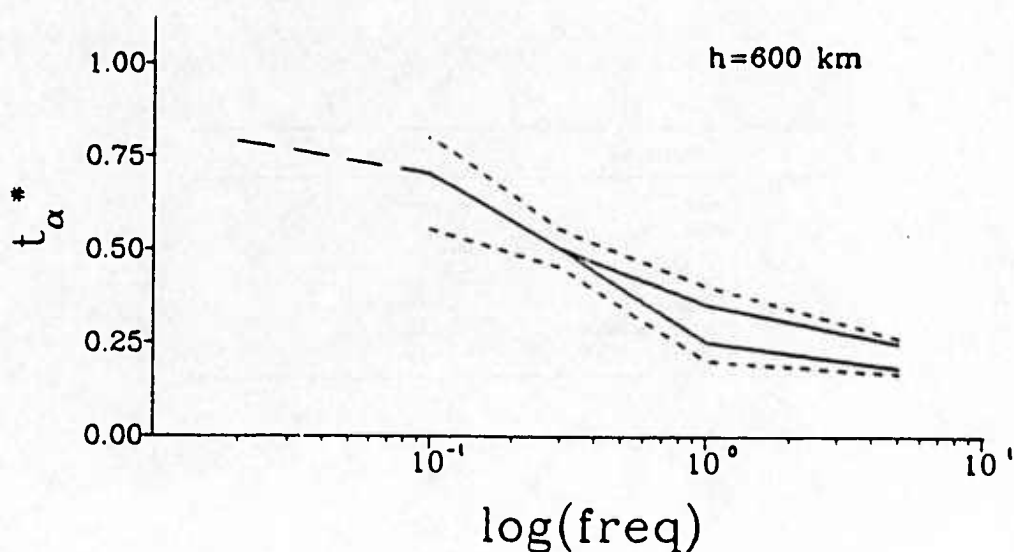


Figure 3.8. The frequency dependence of t_{α}^* for a source at depth 600 km. Two types of $t_{\alpha}^*(f)$ curves (solid lines) were needed to fit the observed waveforms. The two curves converge at about 0.3 Hz. The long-dashed line between 0.01-0.1 Hz indicates assumed values, as our synthetics are not sensitive to this frequency range. At 0.01 Hz, we assume the value, after depth-correction, of t_{α}^* given by model ABM (Anderson and Given, 1982). The short-dashed lines bracket the uncertainty in our attenuation operators.

attenuating zones in the earth. Also in Figure 3.8 are dashed lines bracketing the uncertainty in our $t_a^*(f)$ model. The long-dashed part of the $t_a^*(f)$ curve indicates the frequency range for which the data are not sensitive. Accordingly, we take the value of t_a^* at .01 Hz from model ABM. At 0.1 Hz, our preferred value is 0.7 s. The convergence of our low- and high-Q t_a^* operators indicates that we can resolve no strong regional differences for frequencies less than about 0.3 Hz. This does not mean that regional differences do not persist to lower frequencies. It merely indicates that our data set cannot resolve such differences within the indicated error bounds. Der and Lees (1985), for example, analysed different data that suggest that regional differences in attenuation persist down to 0.01 Hz.

Waveforms at stations BER, RSSD, CHTO and COL could be synthesized by using $t^*(f)$ typical of the high-Q path. All other stations except HON have waveforms that could be synthesized by the $t_a^*(f)$ for the low-Q path. The HON waveform required a nearly constant $t_a^* = 0.6$ s for all frequencies greater than 0.1 Hz. This anomalous curve is not included in Figure 3.8. The shape of these t_a^* curves are similar in principle to that of Der et al. (1982, 1985), but our results do not preclude downward concavity nor some complexity (such as an inflection) in the $t_a^*(f)$ curve in the midband range.

3.6 S WAVE ANALYSIS

In order to obtain the attenuation operator for S waves, we find the minimum phase operator that, when convolved with the source model derived from the P wave analysis, best fits the displacement and velocity pulse shapes. For a finite source, the major theoretical differences between P and S wave spectra involve a corner frequency shift and the rate of spectral falloff beyond the corner frequency. For the earthquake we are modeling, the corner frequency of the dominant subevent is shifted by a factor of 0.7 and the spectral falloff is approximately $\omega^{-1.5}$. A typical fit between our synthetic and observed S-wave spectra is shown in Figure 3.9. Noise, primarily signal-generated, usually exceeds signal for frequencies above about 2.0 Hz. In the time domain, the observed and synthetic seismograms of displacement, velocity and velocity-squared (v^2) are shown in Figures 3.10 and 3.11 for S and ScS body waves, respectively.

In an earth where attenuation is depth and frequency dependent, $t_b^*(f)$ may vary noticeably with distance. A typical $t_b^*(f)$ curve, appropriate for a distance of 50° , is shown in Figure 3.12. This curve was derived from analysis of the S wave recorded at RSNT. We now describe the constraints provided by RSNT. The derivation of $t_b^*(f)$ for other distances follows a similar procedure. Our broadband S waves (Figure 3.10) are not sensitive to very low frequencies ($f < 0.1$ Hz). The lack of constraint at low frequencies is indicated by the long-dashed portion of the $t_b^*(f)$ curve. At 0.01 Hz, we assume the value of the ABM model (Anderson and Given, 1982) which, after correction for a source at depth, is $t_b^* = 3.7$ s. However, in this low frequency range, the synthetic waveforms are not sensitive to values of t_b^* as high as 4.0 s, which is a value suggested by using, after correction for depth, the Q_{ScS} of Sipkin and Jordan (1979). Constraints near 0.1 Hz up to about 2.0 Hz

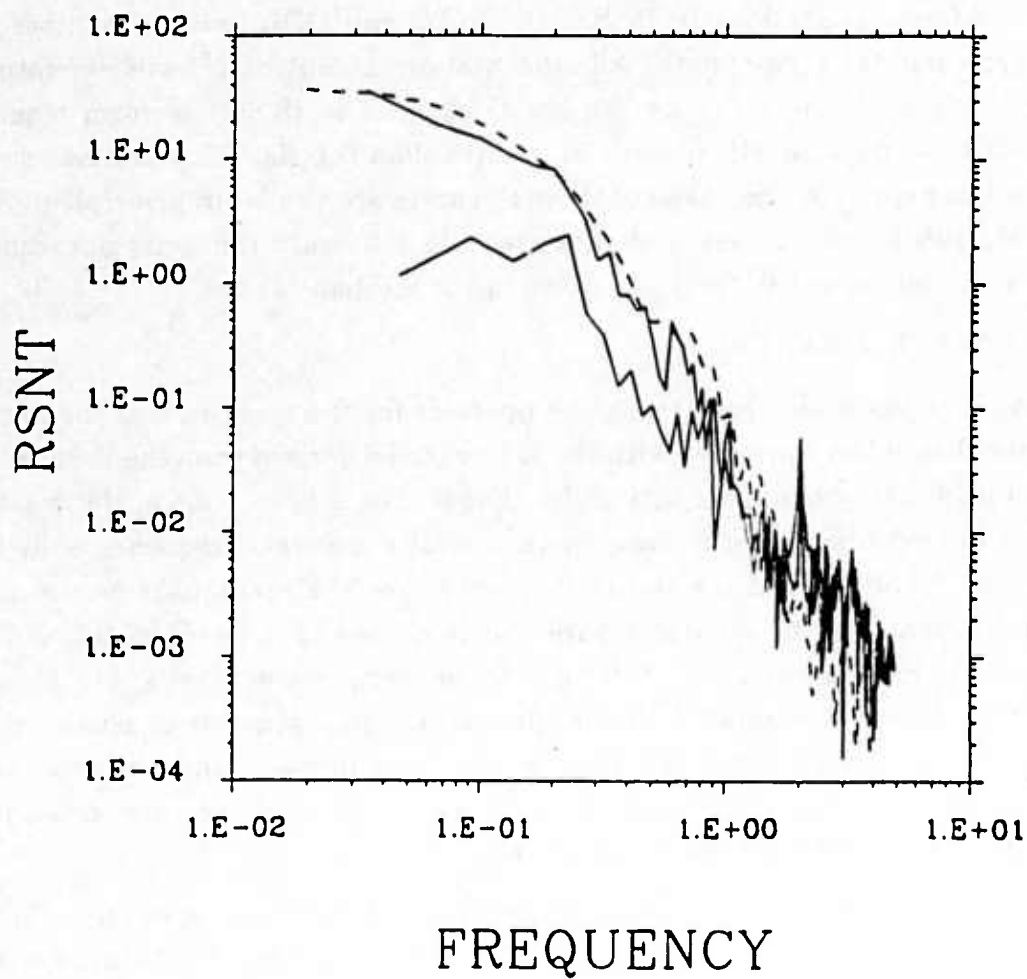


Figure 3.9. Amplitude spectrum of the S-wave displacement at RSNT (upper solid line). The noise spectrum using a 20 s window preceding the arrival of S is shown by the lower solid line. The amplitude spectrum of the synthetic is shown by the dashed line.

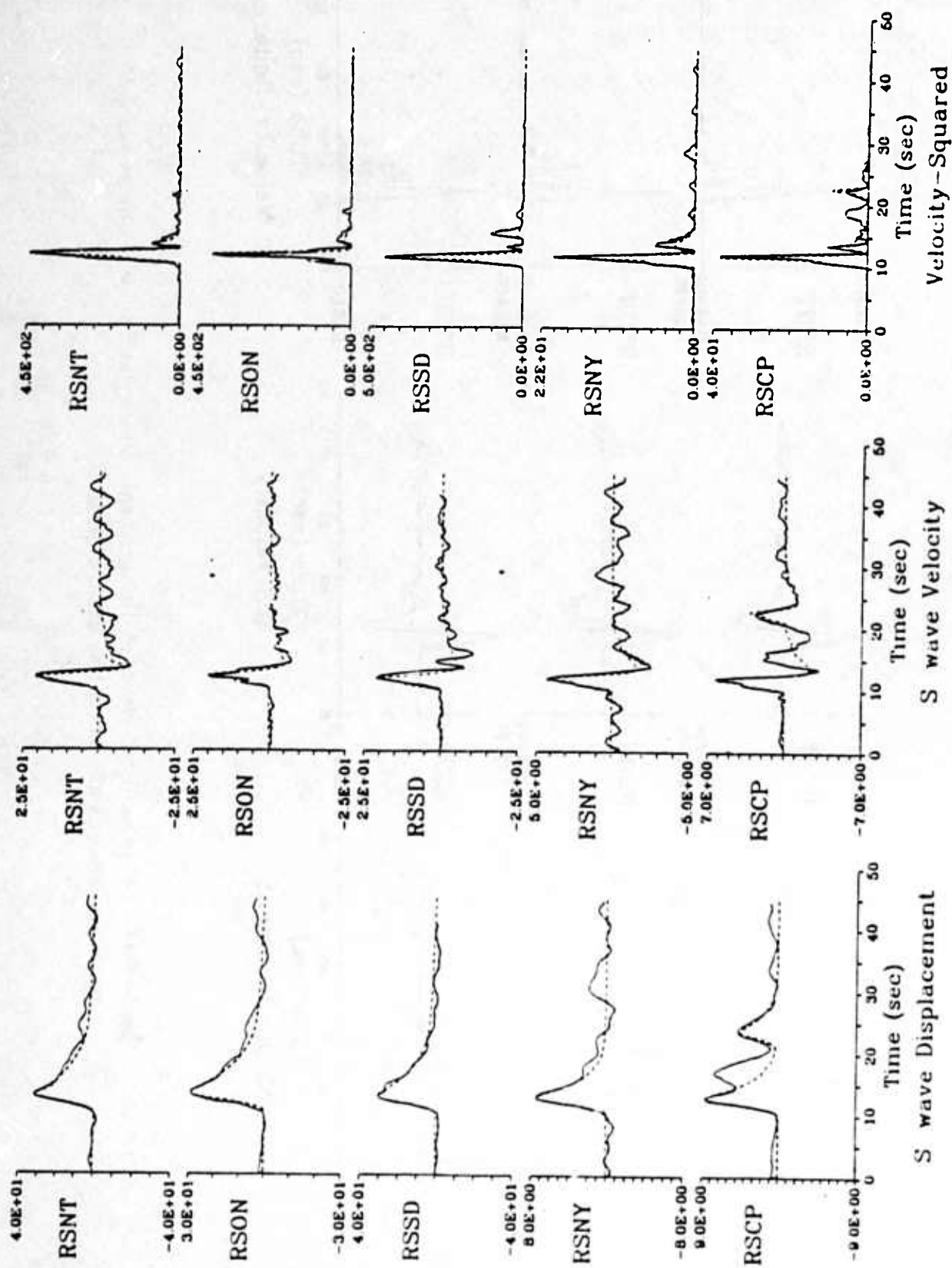


Figure 3.10. Broadband S waves of displacement, velocity and velocity-squared (solid lines). Synthetic records are plotted as dashed lines. Distances from the epicenter are 50° for RSNT, 66° for RSON, 67° for RSSD, and 78° for RSNY. At RSCP ($\Delta 82^\circ$) the waveform contains both S and ScS. An intermediate phase at RSCP was not modelled.

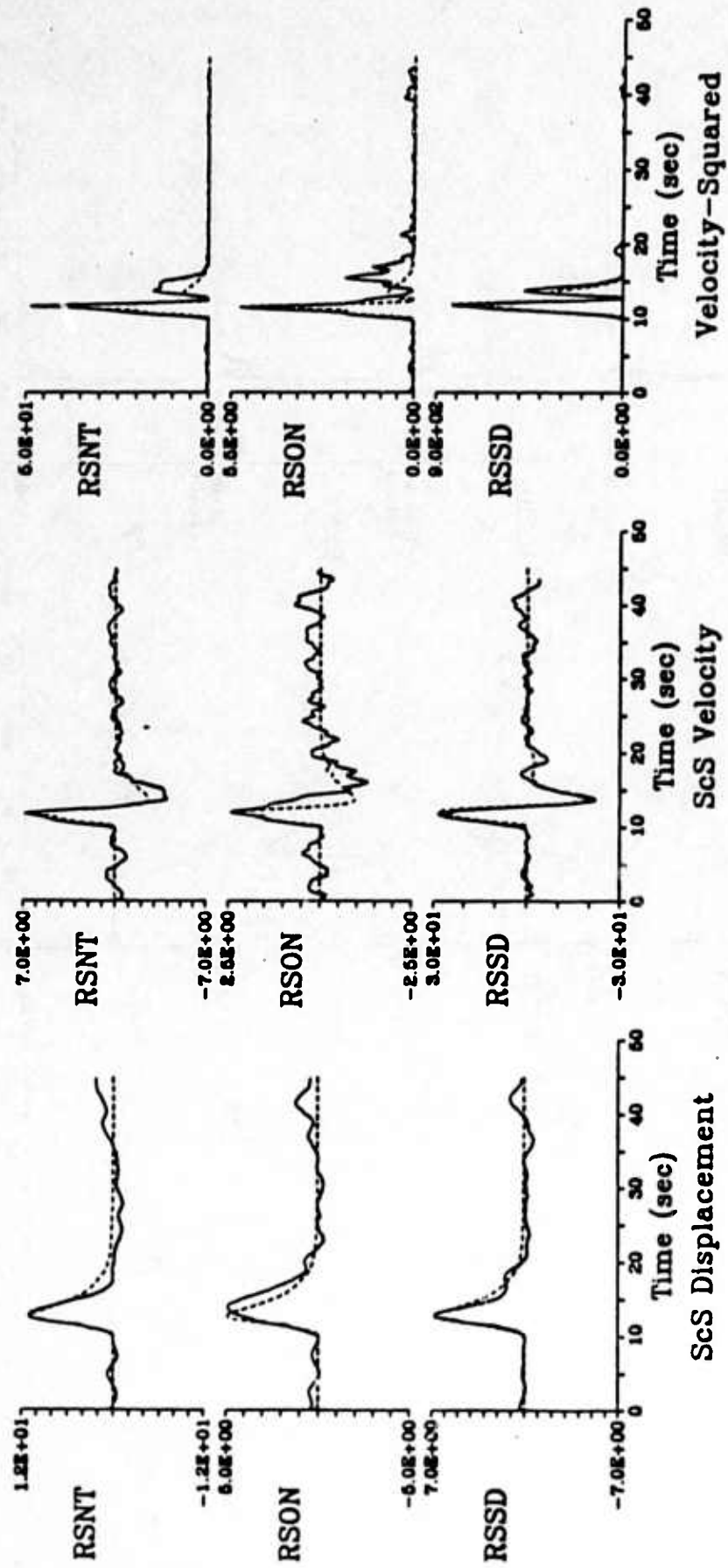


Figure 3.11. Broadband ScS waves of displacement, velocity and velocity-squared (solid lines). Synthetic records are plotted as dashed lines.

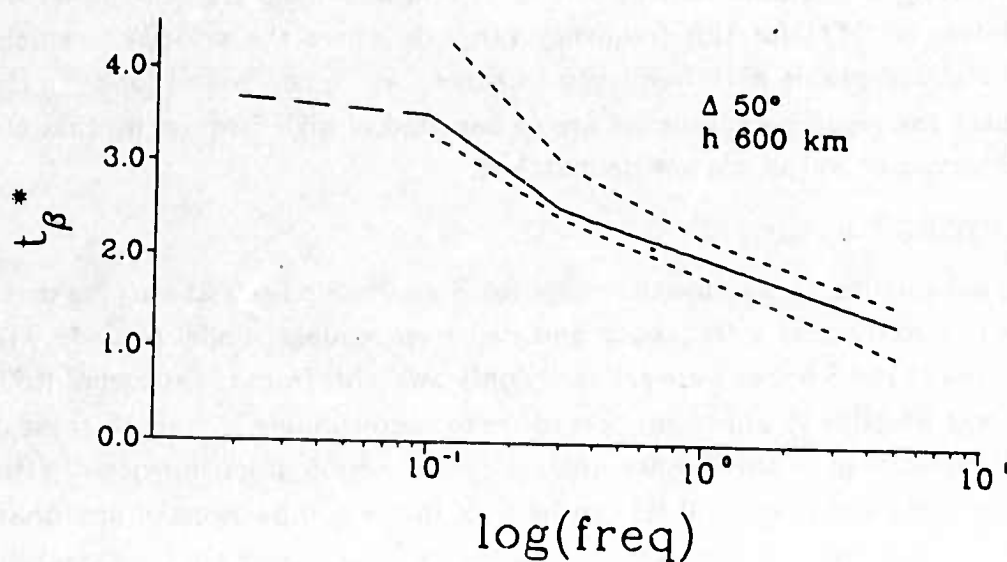


Figure 3.12. The frequency dependence of t_{β}^* for a source at depth 600 km and a distance of 50° (solid line). The long-dashed line between 0.01-0.1 Hz indicates assumed values, as our synthetics are not sensitive to $f \ll 0.1$ Hz. At 0.01 Hz and for a depth of 600 km, we assume the value given by the ABM model which gives t_{β}^* of about 3.7 s. The short-dashed lines bracket the uncertainty of this attenuation operator.

are much stronger. In Figure 3.10, the S wave displacement at RSNT is characterized by a steep rise in displacement followed by a very slowly decaying tail. Frequencies near 0.1 Hz control the rate of falloff for most of the long tail. Using a preferred value of 3.5 s at 0.1 Hz, the tail of the synthetic slightly underfits the tail of the data. The underfit compensates for what, in our estimation, is the level of coda noise. The upper bound at 0.1 Hz is a value which generates a synthetic tail that overfits the observed decay in displacement. The lower bound, about 3.2 s, is a value where the tail of the synthetic displacement is obviously too sharp to be compensated by the addition of noise on the same level as the coda. By emphasizing radiated energy, the v^2 record provides constraints for the higher frequencies. In v^2 , the rise time in displacement becomes a pulse in which is contained almost all the energy of the body wave. The remainder of the energy is concentrated in the second v^2 pulse of similarly short duration but markedly smaller amplitude. The slowly decaying tail of the S wave displacement apparently contributes virtually no energy to the v^2 record. The duration in which the preponderance of energy is concentrated provides strong constraints on frequencies at and above 0.3 Hz. The lower bound of the uncertainties in $t_{\beta}^*(f)$ for this frequency range is where the velocity seismograms begin to exhibit unacceptable high-frequency features. An upper bound for t_{β}^* is chosen at the value where the resulting synthetics are so depleted of high frequencies that the durations of the observed v^2 pulses cannot be matched.

3.7 A MODEL OF $Q_{\beta}(r, f)$

The best fitting $t_{\beta}^*(f)$ measurements for S and ScS phases at varying distance ranges were used to determine a frequency and depth dependent model of $Q_{\beta}(r, f)$, where r is radius. Broadband S waves were primarily only available from the subset of RSTN stations. Thus, it was possible to apply our procedure for determining t_{β}^* only to these stations. As previously described in the P-wave analysis, some regional dependence of attenuation for frequencies near and above 1.0 Hz can be seen in the comparisons of broadband P-waves. At these frequencies, the attenuation of S waves is so strong that regional differences in t_{β}^* would not be discernible. It may be assumed that the S wave paths available from the RSTN stations traversed a region of nearly uniform attenuation. No attempt was made to invert for regional differences among the S wave paths to the RSTN stations.

Within the bounds of the source-attenuation tradeoff, no evidence for bulk attenuation was found for any pair of P and S propagation paths. For the frequencies for which we had the greatest constraints ($f \gtrsim 0.1$ Hz), our t^* operators favor $t_{\beta}^* \sim 4.5t_{\alpha}^*$. Synthetics computed using smaller ratios of $\frac{t_{\beta}^*}{t_{\alpha}^*}$ fit the observed data worse than synthetics using higher ratios. This would be consistent with the assumption that attenuation occurs in pure shear. Along paths for which the ratio $\frac{\alpha}{\beta}$ is constant, the assumption of pure shear in attenuation leads to the relation

$$t_{\beta}^* = \left(\frac{3}{4}\right)\left(\frac{\alpha}{\beta}\right)^3 t_{\alpha}^* \quad (3.7)$$

In the most highly attenuating region of our $Q_\beta(r, f)$ model of the upper mantle, the isotropic PREM model (Dziewonski and Anderson, 1981) has $(\frac{\alpha}{\beta}) = 1.81$ at 1 Hz. Substituting this in (7) gives $t_\beta^* = 4.5t_\alpha^*$.

There are many possible parameterizations of $Q_\beta(r, f)$ that are consistent with the observed $t_\beta^*(f)$ functions. A general requirement of the data is that the $t_\beta^*(f)$ decreases gradually with increasing frequency in the 0.1 to 1 Hz band. The rate of this decrease is slower than that predicted by a simple absorption band cut-off, in which the cut-off frequency is constant with depth, and above which Q increases as the first power of frequency. To fit the observed $t_\beta^*(f)$ curves within the estimated error bounds, either this cut-off frequency must vary with depth, or the power n in a Q frequency dependence $Q = Cf^n$ varies with depth, or both. These are, however, the general and expected features of all models that seek to tie a $Q_\beta(r, f)$ model to a rate-activated process that depends on the temperature and pressure profiles of the Earth's mantle. Such models have been proposed by Lundquist and Cormier (1980) and Anderson and Given (1982), with parameters adjusted to fit seismic data over the frequency band from Chandler wobble to 10 Hz body waves. In these models the frequency and depth dependence of Q_β is related to the (P,T) profiles of the mantle through an assumed dependence of the longest relaxation time τ_1 (or high frequency cut-off $f_c = \frac{1}{2\pi\tau_1}$) of the form

$$\tau_1 = \tau_0 \exp \left[\frac{(E^* + PV^*)}{RT} \right] \quad (3.8)$$

where E^* and V^* are an activation energy and volume change, respectively, associated with the movement of Avogadro's number of defects over their respective potential barriers. The higher time constant τ_2 (or low frequency cut-off) of the relaxation spectrum of the mantle can be constrained to some extent by the gravest modes of oscillation of the Earth and by the period and attenuation of the Chandler wobble. Minster and Anderson (1981) suggest that a five decade band of frequencies is a typical width to expect for the relaxation spectrum of the mantle, and the model of Anderson and Given (1982) fixed the ratio $\frac{\tau_2}{\tau_1}$ to 10^5 .

A model, which is consistent with the measured $t_\beta^*(f)$ curves in the present study, can be found by perturbing the parameters of either the double absorption band model of Lundquist and Cormier (1980) or the parameters of the model by Anderson and Given (1982). Table 3.2 lists one of the possible adjustments in the Anderson and Given model that will produce good fits to the $t_\beta^*(f)$ curves. This model, although giving a relatively complex depth and frequency dependence, is specified by a small number of parameters: three layers; a value of τ_1 in each layer; a fixed $\frac{\tau_2}{\tau_1}$ ratio; a power law for the frequency dependence of Q_β between τ_2 and τ_1 ; and a minimum $Q_\beta = 80$ and $Q_K = \infty$. A search for fits to observed t_β^* curves consisted in simply tabulating S, ScS, and sS travel times through the three layers and testing various assumptions for a power law n in a relation

$Q_\beta = C f^n$. The value of n was allowed to vary in each layer and allowed to take on different values in different frequency bands in the range $|n| \leq 1$. The starting model of $Q_\beta(r, f)$ was taken to be that of Anderson and Given. Since this model was designed to fit a broad band of low frequency data, including free oscillations and Chandler wobble, an attempt was made to keep the perturbations to this model small and confined to the mid and higher frequency band of body waves constrained by our data. Even with the small number of free parameters, however, many possible combinations can be found that will fit the observed $t_\beta^*(f)$ curves. In particular, the power law between τ_2 and τ_1 is rather weakly constrained as well as the frequency of transition between a specified power law and a f^{-1} decay of the attenuation. The frequency dependence of attenuation in the uppermost 400 km of the mantle was constrained by measurements of t_β^* from sS waveforms. It was found that crustal reverberations at the bounce point of sS can strongly affect the long period tail of sS, introducing some trade-off between crustal structure at the surface bounce point and the frequency dependence of attenuation in the band up to 1 Hz in the asthenosphere. In the uppermost 400 km of the mantle, the model given in Table 3.2 was chosen to be consistent with the Q model for the same region given by Anderson and Given and the double absorption band model of Lundquist and Cormier. In all of these models, the relaxation spectrum in the uppermost 200 - 400 km of the mantle spans most of the body-wave band from 0.01 Hz up to several Hz. The frequency band of attenuation in the uppermost mantle may be related to the rapid increase in temperature in the upper 100 km of mantle, which may move τ_1 towards low values, and hence the cut-off frequency to higher values (Figure 3.13).

In the mid to lower mantle, if E^* and V^* are assumed to be approximately constant with depth, then the effect of pressure will dominate that of temperature in equation (3.7). For values of V^* greater than $4 \text{ cm}^3/\text{mole}$, increasing pressure with depth will move τ_1 towards higher values (Figure 3.13). This moves the high frequency cut-off of the relaxation spectrum through the body-wave band towards lower frequencies. At the base of the mantle, the effect of pressure may be sufficient to move the high frequency cutoff of the relaxation spectrum to a frequency value below the body wave band. A low Q zone in the body wave band may still be possible at the base of the mantle if the effect of pressure is overridden by either a thermal boundary layer of the types proposed by Jeanloz and Richter (1978) and Jarvis and Peltier (1982) or by a compositional change that changes E^* and V^* . The S and ScS paths investigated in this study do not show any evidence of a zone of increased attenuation at the base of the mantle beneath the Canadian Shield. Rise times are nearly identical for S and ScS at all stations and distance ranges (Figure 3.14). Rise times are controlled by the phase velocity spectrum at $f \gtrsim 0.3$ Hz. Thus for these paths, at least, there are no measurable differences in the attenuation of 0.3 Hz and higher frequency S waves traversing the lower 1000 km of the mantle versus those bottoming higher in the mantle.

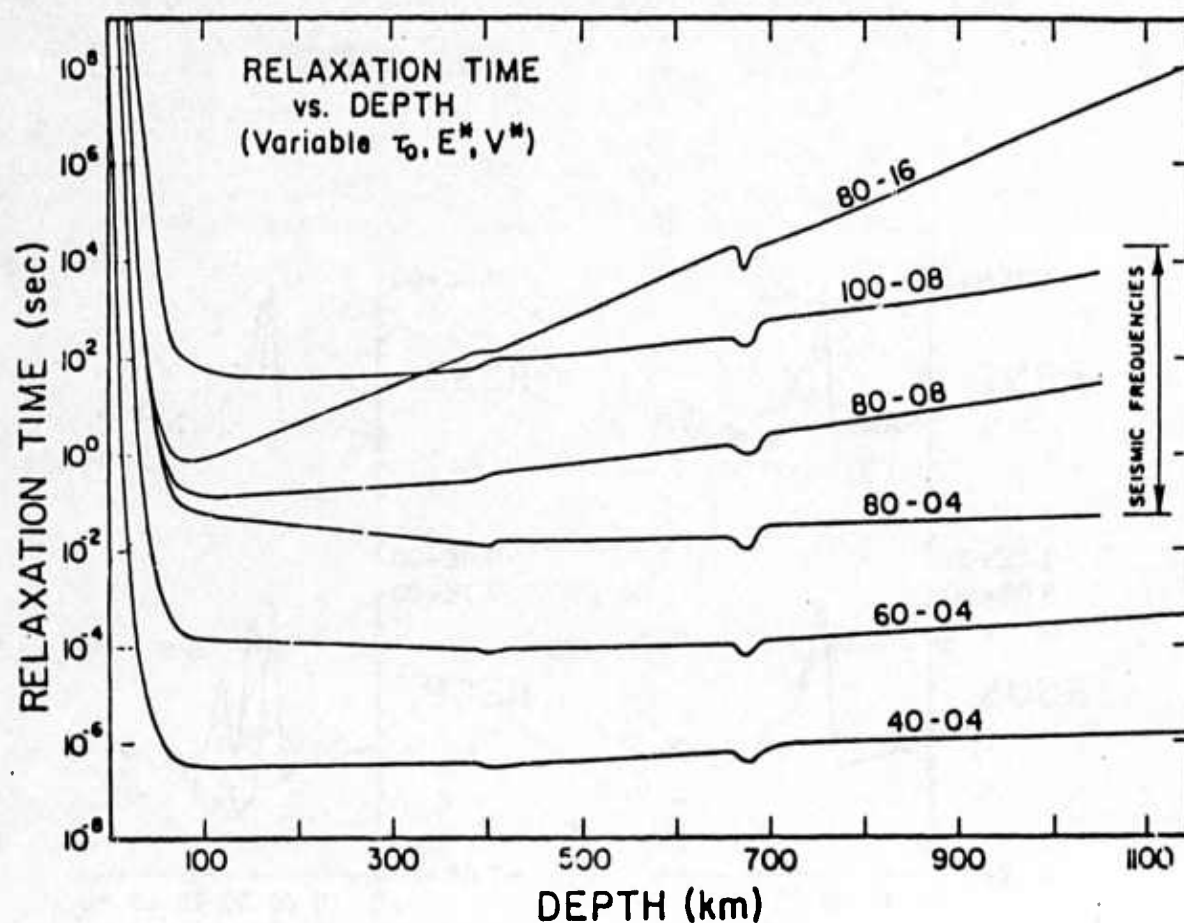


Figure 3.13. Possible τ_1 behavior as a function of depth for (E^*, V^*) pairs. E^* and V^* are assumed constant with depth. $\tau_1 = \tau_0 \exp \left[\frac{(E^* - PV^*)}{RT} \right]$, with the T profile taken from Stacey (1977) and the P profile taken from Anderson and Hart (1977).

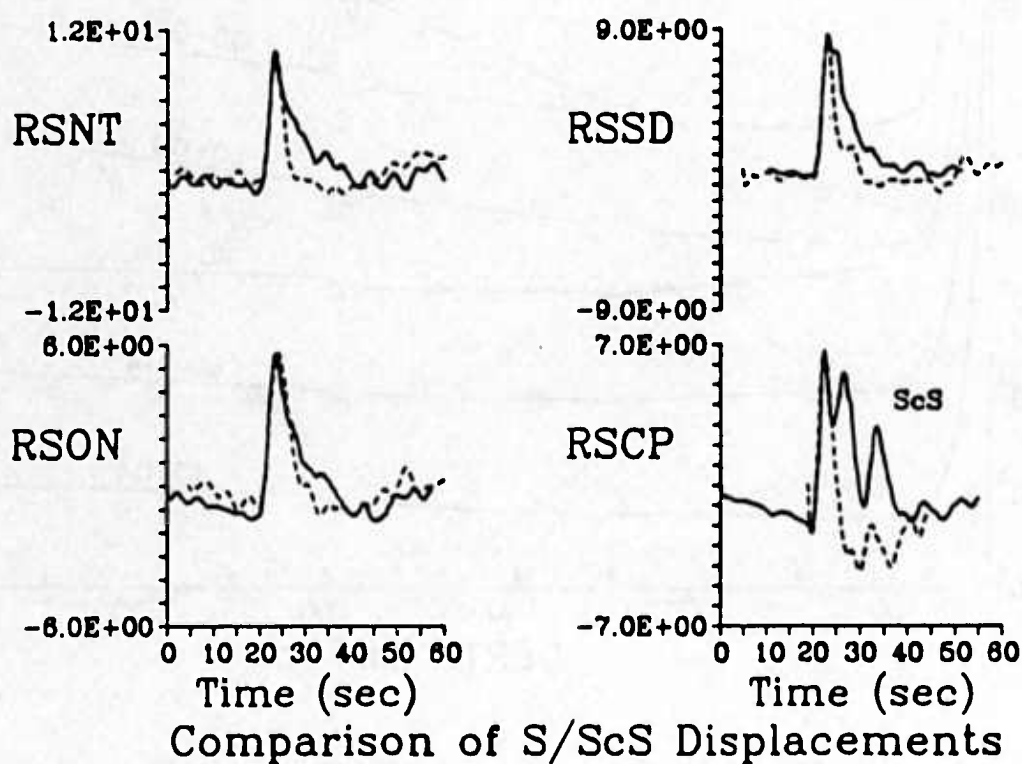


Figure 3.14. A comparison of broadband S and ScS displacement from the deep earthquake of 20 April 1984. At each station the ScS displacement (dashed line) is time shifted and normalized to the same amplitude as the S displacement (solid line). The S waves generally have a much longer tail than the ScS waves. The difference decreases with distance from the hypocenter. The distances are 50° for RSNT; 66° for RSSD; 67° for RSON; and 82° for RSCP.

At frequencies less than 0.3 Hz, however, evidence was found that may support significant differences in the attenuation of S versus ScS waves. A conspicuous difference is observed in the decay of the long period tail of the displacement pulses between the S and ScS phases (Fig. 3.14). The direct S seems to have a much longer tail relative to the ScS phases. This difference disappears as range increases toward 82° . The S waves at GRF for this event were of poor quality because GRF was on an SH nodal line for both S and ScS. Fortunately, S waves were available for another earthquake that occurred in nearly the same region and that had nearly the same type of source function. The relative behavior of the S versus ScS waveforms from this earthquake is also observed at the GR π array (Figure 3.15).

The mechanisms that may be responsible for the differences in the long period tail of S versus ScS include (1) crustal reverberations and reverberations of ScS in layered structure at the base of the mantle, (2) anisotropic scattering, (3) multipathing, and (4) the depth and frequency dependence of intrinsic attenuation. Of these, intrinsic attenuation appears to be the most likely explanation. The feasibility of each mechanism is summarized as follows:

Crust and D'' Reverberations. As the angle of incidence on the receiver crust is nearly identical for S and ScS in the distance range investigated, the difference in the long period tail cannot be simply explained by the effects of the crustal response. Other possible mechanisms for generating arrivals that may interfere with S and ScS may be multipathing in laterally heterogeneous structure at the base of the mantle (Cormier, 1985) or postcritical reflection from a velocity discontinuity near the core-mantle boundary (Lay and Helmberger, 1983). For the distances of RSNT (50°), RSSD (66°), and RSON (66°), the S wave bottoms far above any such structure. Such structures could only affect the ScS wave. Our calculations show that the reverberations associated with ScS in a layered D'' region would not be sufficient to interfere destructively with any long-period tail of ScS between 50° and 66° . The pulse between S and ScS at RSCP (82°) could possibly be explained as a postcritical reflection or a multipath arrival from some structure near the core-mantle boundary. Because these possibilities themselves require further confirmation we chose not to model the intermediate pulse at RSCP for the time being.

Anisotropic Scattering. Anisotropic scattering may operate over the mid- and deep mantle depths where the ray path of S most strongly differs from that of ScS. In this mechanism, the spatial distribution of scattering heterogeneities at these depths must be such as to more strongly scatter the low frequency S waves incident at the S wave angles than those incident at the ScS angles. Chowdury and Frasier (1978) proposed anisotropic scattering as the mechanism that accounted for amplitude variations in short period P and PcP, which could not be explained by intrinsic attenuation. More recently, studies by Bache et al. (1985) and Taylor et al. (1985) suggest that scattering becomes increasingly important for

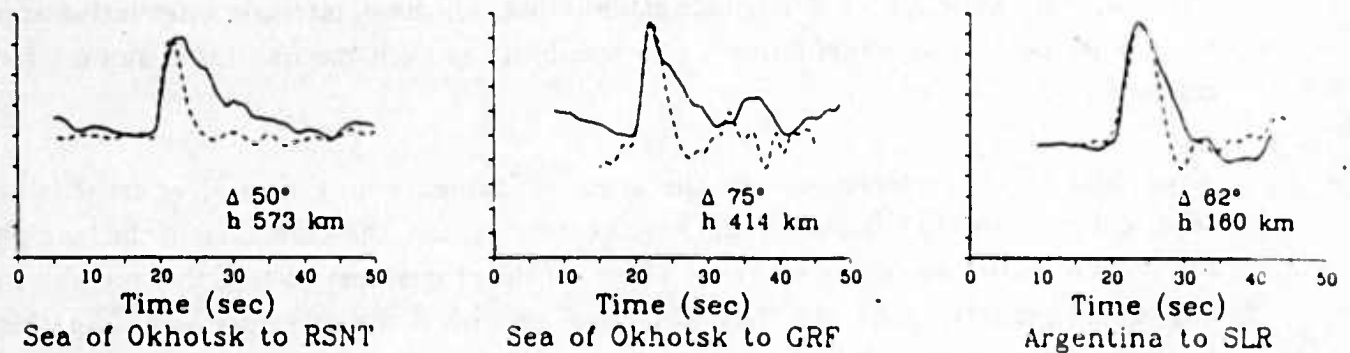


Figure 3.15. A comparison of broadband S and ScS displacements from other earthquakes with depths ranging from 160 to 573 km. (Left) The ScS displacement (dashed line) from station RSNT has been time shifted and normalized to the amplitude of S (solid line). The station is 50° from an event in the Sea of Okhotsk (February 1, 1984, $h = 573 \text{ km}$). (Middle) S and ScS at station A1 of the GRF array ($\Delta 75^\circ$) from an event in the Sea of Okhotsk (April 23, 1984, $h = 414 \text{ km}$). The ray path to GR is nearly normal to the orientation of the Kuril-Kamachatka slab. (Right) S and ScS displacements at SLR ($\Delta 82^\circ$) from an event in Catamarca Province, Argentina (June 10, 1985, $h = 160 \text{ km}$). The ray path is nearly normal to the orientation of the South American slab.

pulse attenuation as frequency increases above 1 Hz. This can account for pulse dispersion in the reverse sense of that expected of intrinsic attenuation; and for apparent P attenuation becoming more nearly equal to apparent S attenuation. Our analysis of S data, however, is restricted to frequencies less than about 2 Hz, where pulse and rise-time coherence suggest that the effects of scattering do not dominate the effects of intrinsic attenuation. In this frequency band, the P and S waveforms analyzed did not evidence any significant amount of reverse dispersion or strong departures from $t_{\beta}^* = 4.5t_{\alpha}^*$.

Multipathing. W. Chan and P. Silver (personal communication) have suggested that the long-period tail of S may be due to multipathing in the vicinity of the descending Kuril-Kamchatka slab. This is likely to be a frequency dependent effect, in which longer period waves radiated from the source region average over the locally heterogeneous structure. Thus, it is difficult to evaluate this hypothesis without performing a synthesis of body waves in a three dimensional structure by a method capable of including the effects of diffraction and spatial averaging over a wavelength. An observation that must be satisfied is that the long-period tail of S does not appear to be affected very much by the take-off azimuth with respect to the strike of the slab. The long period tail is observed at the GRF array (Figure 3.15), for a path that takes off at an azimuth nearly perpendicular to the strike of the slab, and at the RSTN stations (Figure 3.14), for paths that take off at azimuths sub-parallel to the slab. The long-period tail of S was also observed at GRF and RSTN stations for events that differed in depth by as much as 150 km and in relative location by 4° . The azimuthal and depth independence of these observations may prove difficult to satisfy with any model of slab multipathing.

Intrinsic Attenuation. The model of $Q_{\beta}(r, f)$ given in Table 3.2 was designed to match the observed similarities between S and ScS as well as their differences in the long-period tail. The theoretical t_{β}^* 's for 66° and a depth of 600 km, for example, are shown in Figure 3.16. Note that t_{β}^* of S differs from the t_{β}^* of ScS in the low frequency band but is nearly identical in the band above 0.3 Hz. The comparison of S and ScS displacements (Figure 3.14) shows the design criteria. The rise times of S and ScS are nearly indistinguishable. Our modeling shows that the high frequency band ($f \gtrsim 0.3$ Hz) controls the rise times, while frequencies below about 0.3 Hz control the decay of the tail. In order to match our best-fitting t_{β}^* curves, our model of $Q_{\beta}(r, f)$ eliminates most attenuation in the body wave band of seismic frequencies in the lower 1000 km of the mantle. It concentrates attenuation at mid-mantle depths (400 to 2000 km), primarily in the low frequency end of the body wave frequencies. A schematic of the earth using our model of $Q_{\beta}(r, f)$ illustrates how our body wave data are satisfied (Figure 3.17). ScS waves spend less time at mid-mantle depths compared to S waves, and hence suffer less attenuation at frequencies less than 0.3 Hz. Both S and ScS waves, however, spend nearly equal times in the upper mantle, which primarily controls the attenuation of frequencies greater than 0.3 Hz, and hence the rise times of S and ScS are nearly identical. The $Q_{\beta}(r, f)$ model of Table 3.2, which predicts

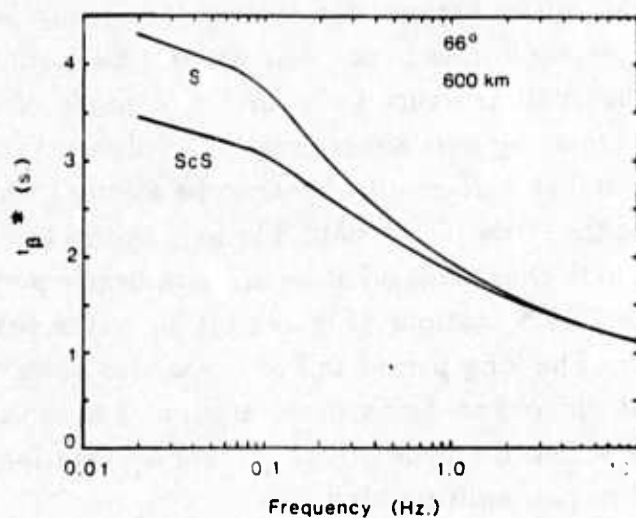


Figure 3.16. The frequency dependence of $t_{\beta}^*(f)$ for S and ScS at 66° for a deep focus earthquake predicted from a $Q_{\beta}(\tau, f)$ model of the type shown in Figure 17.

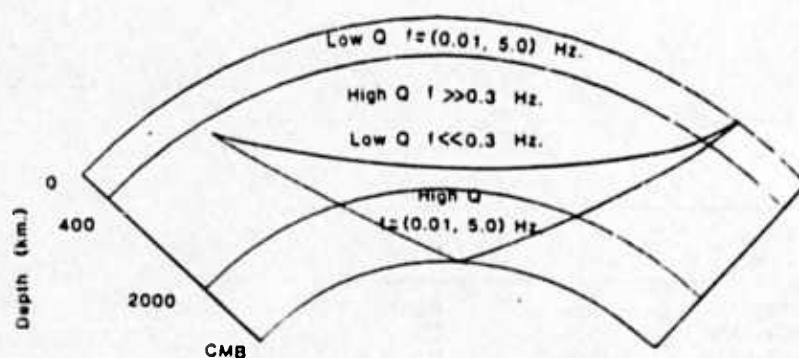


Figure 3.17. A schematic cross-section of the earth showing how the ray paths of S and ScS sample a $Q_p(\tau, f)$ model.

Radius, km	r_1 , s	Q_β (at 0.01 Hz)	$\min Q_\beta$
Moho-5971	< 0.16	100	100
5971-4371	1.60	250	114
4371-CMB	> 16.00	> 100	unknown

Table 3.2 A Model of $Q_\beta(r, f)$

this effect, is consistent with a thermally activated model of attenuation with V^* greater than $4\text{cm}^3/\text{mole}$, in which the effect of pressure overrides that of temperature as depth increases, and thereby moves f_c through the body wave band towards lower frequency.

3.8 CONCLUSIONS

The analysis in the time domain of broadband P and S waves from deep focus earthquakes can significantly constrain the frequency dependence of the intrinsic Q of the earth. For the paths investigated in this study, the effects of intrinsic attenuation seem to dominate any effects of scattering attenuation on pulse shapes from low frequencies up to at least 1.0 Hz. The gradual decrease in intrinsic anelasticity throughout the body wave band very closely follows the shape of $t^*_\alpha(f)$ curves given in Der et al. (1985), which they obtained from a summary of many spectral and time domain measurements of apparent attenuation. The gradual decrease of t^* as a function of frequency cannot be matched by any simple parameterization of the frequency dependence of Q, such as a single cutoff frequency in a relaxation spectrum which is constant with depth. Although S and ScS displacement pulses had nearly equal rise times, the pulse width of S was substantially broader than the pulse width of ScS. Such a feature can be matched by a model of $Q_\beta(r, f)$ in which the mid-mantle between depths of 400 and 2000 km attenuates primarily frequencies less than about 0.3 Hz, although frequencies greater than 0.3 Hz are also attenuated. S and ScS waveforms from deep earthquakes beneath the Sea of Okhotsk observed at both RSTN stations in North America and at the Graefenberg array in Germany do not indicate the existence of a low Q zone at the base of the mantle along these paths. All waveforms can be explained by a model that has little attenuation throughout the body wave band at depths exceeding 2000 km. Revisions in the high frequency cutoff parameter in the models of $Q_\beta(r, f)$ proposed by Lundquist and Cormier (1980) and Anderson and Given (1980) are all that are required to explain the pulse shapes of S, sS and ScS phases satisfactorily. This suggests that the basic form of $Q_\beta(r, f)$ is consistent with the hypothesis of these papers, i.e., that Q_β is due to thermally activated mechanisms of intrinsic anelasticity, with depth dependence controlled by the temperature and pressure profiles of the mantle. Interpretation of the best-fitting $Q_\beta(r, f)$ models as a rate activated process suggests that the effect of temperature dominates the effect of pressure in the upper mantle, but that pressure dominates temperature in the lower mantle.

Broadband pulse modeling combined with realistic source constraints offers a new means of determining the frequency dependence of attenuation in the earth. Broadband pulses contain information on the dispersive property of intrinsic attenuation that can be used to retrieve the frequency dependence of t^* . The results reported here depend on a source analysis of P and S waves and on the existence of broadband digital recordings of horizontal channel data. This configuration of instruments has, until quite recently, been rare. The GDSN day-tape has not contained short-period or intermediate period data until RSTN data were incorporated into the day-tape. In the future, greater global coverage

by broadband digitally recording stations can permit the mapping of the frequency and regional dependence of attenuation observed from body waves.

3.9 REFERENCES

- Anderson, D.L., and J.W. Given, Absorption band Q model for the Earth, *J. Geophys. Res.*, **87**, 3893-3904, 1982.
- Anderson, D. L., and R. S. Hart, The effect of attenuation on gross Earth models, *J. Geophys. Res.*, **82**, 1647-1654, 1977.
- Anderson, D.L., and R. S. Hart, Q of the Earth, *J. Geophys. Res.*, **83**, 5869-5882, 1978.
- Bache, T.C., P.W. Marshall, and L.B. Bache, Q for teleseismic P waves from central Asia, *J. Geophys. Res.*, **90**, 3575-3587, 1985.
- Boatwright, J., A spectral theory for circular seismic source; simple estimates of source dimension, dynamic stress drop and radiated seismic energy, *Bull. Seismol. Soc. Am.*, **70**, 1-27, 1980.
- Boatwright, J., The effect of rupture complexity on estimates of source size, *J. Geophys. Res.*, **89**, 1132-1146, 1984.
- Boore, D. M. and W. B. Joyner, The influence of rupture incoherence on seismic directivity, *Bull. Seismol. Soc.*, **68**, 283-300, 1978.
- Chapman, C. H., A new method of computing synthetic seismograms, *Geophys. J. R. astr. Soc.*, **54**, 481-518, 1978.
- Chowdhury, D.K., and C.W. Frasier Evidence for anisotropic scattering of short period P waves in the upper mantle, *Bull. Seismol. Soc. Am.*, **68**, 609-618, 1978.
- Choy, G. L., Experiments with SRO and GRF-array data. *Open-File Report U.S. Geol. Surv.*, **82-216**.
- Choy, G. L. and J. Boatwright. The rupture characteristics of two deep earthquakes inferred from broadband GDSN data, *Bull. Seismol. Soc. Am.*, **71**, 691-711.
- Choy, G. L. J. Boatwright. J. W. Dewey and S. A. Sipkin. A teleseismic analysis of the New Brunswick earthquake of January 9, 1982. *J. Geophys. Res.*, **88**, 2199-2212, 1983.
- Cormier, V. F., The effect of attenuation on seismic body waves, *Bull. Seismol. Soc. Am.*, **72**, s169-s200, 1982.
- Cormier, V. F., Some problems with S, SKS, and ScS observations and implications for the structure of the base of the mantle and outer core. *J. Geophys.*, **57**, 14-22, 1985.
- Cormier, V. F. and P.G. Richards. Comments on damping of core waves by Anthony Qamar and Alfredo Eisenberg. *J. Geophys. Res.*, **81**, 3066-3068, 1976.
- Der, Z. A. and T. W. McElfresh. The relationship between anelastic attenuation and regional amplitude anomalies of short-period P waves in North America. *Bull. Seismol. Soc. Am.*, **67**, 1303-1317, 1977.

- Der, Z. A., T. W. McElfresh and A. O'Donnell, An investigation of the regional variations and frequency dependence of anelastic attenuation in the mantle under the United States in the 0.5-4.0 Hz band, *Geophys. J. R. Astron. Soc.*, **69**, 67-100, 1982.
- Der, Z. A., and A. C. Lees, Methodologies for estimating $t^*(f)$ from short-period body waves and regional variations in $t^*(f)$ in the United States, *Geophys. J. R. Astron. Soc.*, **82**, 125-140, 1985.
- Der, Z. A., A. C. Lees, and V. F. Cormier, Frequency dependence of Q in the mantle underlying the shield areas of Eurasia, Part III. The Q model, *Geophys. J. R. Astron. Soc.*, submitted, 1985.
- Dziewonski, A. M. and D. L. Anderson, Preliminary reference earth model, *Phys. Earth and Planet. Int.*, **25**, 297-356, 1981.
- Engdahl, E. R., J. Peterson and N. A. Orsini, Global digital networks—Current status and future directions, *Bull. Seismol. Soc. Am.*, **72**, s243-s259, 1982.
- Ganse, R. and C. R. Hutt (1982). Directory of World Seismograph Stations, *World Data Center A for Solid Earth Geophysics, Report SE-32*, Boulder, Colorado.
- Gurtin, M. E., and L. Herrera, On dissipation inequalities and linear viscoelasticity, *Quart. Appl. Math.*, **23**, 235-245, 1965.
- Hanks, T. C., The corner frequency shift, earthquake source models, and Q, *Bull. Seismol. Soc.*, **71**, 597-612.
- Harjes, H.-P., and D. Seidl (1978). Digital recording and analysis of broadband seismic data at the Graefenberg (GRF)-array, *J. Geophys. Res.*, **44**, 511-523, 1978.
- Harvey, D. and G. L. Choy, Broad-band deconvolution of GDSN data, *Geophys. J. R. astr. Soc.*, **69**, 659-668, 1982.
- Jarvis, G.T., and W.D. Peltier, Mantle convection as a boundary layer phenomenon, *Geophys. J. R. Astron. Soc.*, **68**, 389-427, 1982.
- Jeanloz, R., and F.M. Richter, Convection, composition, and the thermal state of the lower mantle, *J. Geophys. Res.*, **84**, 5497-5504, 1979.
- Kennett, B. L. N., The effects of attenuation on seismograms, *Bull. Seismol. Soc. Am.*, **65**, 1643-1652, 1975.
- Lay, T., and D.V. Helmberger, A lower mantle S wave triplication and the shear velocity structure of D'' , *Geophys. J. R. Astron. Soc.*, **75**, 799-838, 1983.
- Lundquist, G.M., and V.F. Cormier, Constraints on the absorption band model of Q, *J. Geophys. Res.*, **85**, 5244-5256, 1980.
- Minster, J.B., and D.L. Anderson, A model of dislocation-controlled rheology for the mantle, *Philos. Trans. R. Soc. London*, **299**, 319-356, 1981.

- Minster, J.B., Transient and impulse responses of a one-dimensional linearly attenuating medium, I, Analytical results, *Geophys. J. R. Astron. Soc.*, 52, 479-501, 1978.
- Owens, T.J., S. R. Taylor, and G. Zandt, Crustal structure beneath RSTN stations inferred from teleseismic P waveforms: Preliminary results at RSCP, RSSD and RSNY, *Lawrence Livermore Laboratory, UCID-19859*, 1983.
- Robinson, E. A., *Statistical Communication and Detection*, C. Griffin and Co., London, 1967.
- Sato, T. and T. Hirasawa, Body wave spectra from propagating shear cracks, *Bull. Seismol. Soc. Am.*, 64, 1621-1627, 1973.
- Sipkin, S. A. and T. H. Jordan, Frequency dependence of Q_{ScS} , *Bull. Seismol. Soc. Am.*, 69, 1055-1079, 1979.
- Stacey, F.D., *The Physics of the Earth*, John Wiley, New York, 1977.
- Taylor, S.R., B.P. Bonner, and G. Zandt, Attenuation and scattering of broadband P and S waves across North America, *J. Geophys. Res.*, submitted, 1985.
- Ulug, A. and H. Berckhemer, Frequency dependence of Q for seismic body waves in the earth's mantle, *J. Geophys.*, 56, 9-19, 1984.