

Journal of Geophysics Zeitschrift für Geophysik

Volume 57 Number 1 1985

Original investigations

H.-P. Harjes

Global seismic network assessment for teleseismic detection of underground nuclear explosions. I. Model calculations for different amplitude-attenuation curves 1

V.F. Cormier

Some problems with *S*, *SKS* and *ScS* observations and implications for the structure of the base of the mantle and the outer core 14

F. Ziegler, Y.-H. Pao, Y.S. Wang

Transient *SH* waves in dipping layers: the buried line-source problem 23

B. Gelchinsky

The formulae for the calculation of the Fresnel zones or volumes 33

J. Meyer, J.-H. Hufen, M. Siebert

Stromfunktion des erdmagnetischen Hauptfeldes in der Quellschicht an der Kern/Mantel-Grenze 42

T. Schultz-Krutsch, F. Heller

Measurement of magnetic susceptibility anisotropy in Buntsandstein deposits from southern Germany 51

P.Y. Galibert, N. Bonhommet

Deep drilling through the accreting plate boundary of Asal, Southern Afar: palaeomagnetism and magnetic properties of basaltic cores 59

N. Klöcker, H. Lühr, A. Korth, P. Robert

Observation of kinetic Alfvén waves excited at substorm onset 65

Short communication

U.R. Christensen

Velocity-viscosity correlation in convection cells with non-uniform viscosity 72

Indexed in Current Contents

Evaluated for Physics Briefs and INKA-PHYS data base

A-3, T, I

8 Z Nat. 2148

Z 4137

G4

X, 501



Springer International

252 J. Geophys. — ISSN 0340-062X — JGEOD4 57(1) 1-76 (1985) — August 1985

15. Aug. 1985
Niedersächsische Staats-
Universitätsbibliothek
Göttingen

Journal of Geophysics – Zeitschrift für Geophysik

Edited for the Deutsche Geophysikalische Gesellschaft by W. Dieminger, G. Müller, and J. Untiedt

This journal was founded by the Deutsche Geophysikalische Gesellschaft on the initiative of L. Mintrop in 1924 as the *Zeitschrift für Geophysik* and edited by G. Angenheister from Vols. 1–18 (1944). It reappeared in 1954, edited by B. Brockamp from Vols. 19–26 (1960), by W. Dieminger from Vols. 27–33 (1961), by W. Dieminger and J. Untiedt from Vols. 34 (1968), and edited by W. Dieminger, J. Untiedt and G. Müller from Vol. 45 (1979). After Vol. 40 (1970) the title was changed to *Journal of Geophysics – Zeitschrift für Geophysik*.

Published: Vols. 19–39 by Physica-Verlag, Würzburg, from Vol. 40 by Springer-Verlag GmbH & Co. KG Berlin, Heidelberg, New York, Tokyo.

Manuscripts may be addressed to any of the Editors. For addresses see last cover page. Manuscripts should conform with the journal's accepted practice as described in the Instructions to Authors.

The Journal accepts

- Review articles (invited by the editors)
- Original papers
- Short communications
- Letters to the editors
- Book reviews

in the field of Geophysics and Space Physics.

Copyright

Submission of a manuscript implies: that the work described has not been published before (except in the form of an abstract or as part of a published lecture, review, or thesis); that it is not under consideration for publication elsewhere; that its publication has been approved by all coauthors, if any, as well as by the responsible authorities at the institute where the work has been carried out; that, if and when the manuscript is accepted for publication, the authors agree to automatic transfer of the copyright to the society; and that the manuscript will not be published elsewhere in any language without the consent of the copyright holders. All articles published in this journal are protected by copyright, which covers the exclusive rights to reproduce and distribute the article (e.g., as offprints), as well as all translation rights. No material published in this journal may be reproduced photographically or stored on microfilm, in electronic data bases, video disks, etc.,

without first obtaining written permission from the publisher.

The use of general descriptive names, trade names, trademarks, etc., in this publication, even if not specifically identified, does not imply that these names are not protected by the relevant laws and regulations.

While the advice and information in this journal is believed to be true and accurate at the date of its going to press, neither the authors, the editors, nor the publisher can accept any legal responsibility for any errors or omissions that may be made. The publisher makes no warranty, express or implied, with respect to the material contained herein.

Special regulations for photocopies in the USA:

Photocopies may be made for personal or in-house use beyond the limitations stipulated under Section 107 or 108 of U.S. Copyright Law, provided a fee is paid. This fee is US \$0.20 per page, or a minimum of US \$1.00 if an article contains fewer than five pages. All fees should be paid to the Copyright Clearance Center, Inc., 21 Congress Street, Salem, MA 01970, USA, stating the ISSN 0340-062X, the volume, and the first and last page numbers of each article copied. The copyright owner's consent does not include copying for general distribution, promotion, new works, or resale. In these cases, specific written permission must first be obtained from the publisher.

Other regulations: Authors publishing in this journal can, under certain conditions, benefit from library and photocopy fees collected by VG WORT. Authors of German nationality and those resident in the Federal Republic of Germany or Berlin (West), as well as citizens of Austria, Switzerland and member countries of the European Community, may apply to Verlagsgesellschaft WORT, Abteilung Wissenschaft, Goethestraße 49, D-8000 München 2, for detailed information.

Subscription information

Volumes 57 + 58 (3 issues each) will appear in 1985.

Members. Members of the Deutsche Geophysikalische Gesellschaft are entitled to purchase the Journal for their own use at a privilege price of DM 98.00 payable with the Membership dues. Orders should be sent to the Society's office at the following address: Postfach 230, D-3392 Clausthal-Zellerfeld.

North America. Annual subscription rate: Approx. US \$143.00 (single issue price: Approx. US \$28.00), including carriage charges. Subscriptions are entered with prepayment only. Orders should be addressed to: Springer-Verlag New York Inc. Service Center Secaucus 44 Hartz Way Secaucus, NJ 07094, USA Tel. (201) 348-4033, Telex 0023-125994.

All other countries. Annual subscription rate: DM 368.00 plus carriage charges. Airmail delivery on request only. Volume price: DM 184.00, single issue price: DM 73.60, plus carriage charges. Carriage charges for SAL (Surface Airmail Lifted) to Japan, India, Australia and New Zealand are available on request. Orders can either be placed with your book-dealer or sent directly to: Springer-Verlag Heidelberger Platz 3 D-1000 Berlin 33 Tel. (0)30/8207-1, Telex 1-83319.

Changes of address: Allow six weeks for all changes to become effective. All communications should include both old and new addresses (with Postal Codes) and should be accompanied by a mailing label from a recent issue.

Back volumes: Prices are available on request.

Microform: Microform editions are available from: University Microfilms International 300 N. Zeeb Road Ann Arbor, MI 48106, USA

Production

Springer-Verlag
Journal Production Department II
Postfach 105280
D-6900 Heidelberg 1
Federal Republic of Germany
Tel. (0)6221/487-342, Telex 4-61 690.

Responsible for advertisements

Springer-Verlag
E. Lückermann
Kurfürstendamm 237
D-1000 Berlin 15
Tel. (0)30/882 1031, Telex 1-85411.

Printers

Universitätsdruckerei H. Stürtz AG Würzburg
© the Deutsche Geophysikalische Gesellschaft, Clausthal-Zellerfeld, 1985
Springer-Verlag GmbH & Co. KG
D-1000 Berlin 33
Printed in Germany



Springer International

Journal of **Geophysics** Zeitschrift für **Geophysik**

Volume 57 1985

Managing Editors **W. Dieminger, G. Müller, J. Untiedt**

Editorial Board

K. M. Creer Edinburgh, Scotland
W. Dieminger Lindau, F.R.G.
C. Kisslinger Boulder, Colorado
Th. Krey Hannover, F.R.G.
G. Müller Frankfurt, F.R.G.
G. C. Reid Boulder, Colorado
U. Untiedt Münster, F.R.G.
S. Uyeda Tokyo, Japan

Advisory Board

G. Angenheister , München	St. Müller , Zürich
A. A. Ashour , Cairo	F. Neubauer , Köln
H.-J. Burkhardt , Berlin	H. Neugebauer , Clausthal-Zellerfeld
K. Fuchs , Karlsruhe	F. Rummel , Bochum
N. Fukushima , Tokyo	S. Saxov , Aarhus
H. Gebrande , München	U. Schmucker , Göttingen
V. Haak , Berlin	M. Siebert , Göttingen
W. Jacoby , Mainz	H. Soffel , München
T. Kirsten , Heidelberg	H. Stiller , Potsdam
I. P. Kosminskaja , Moskwa	P. Weidelt , Braunschweig
G. Kremser , Katlenburg-Lindau	J. Zschau , Kiel
W. Lowrie , Zürich	



Springer International

Journal of Geophysics – Zeitschrift für Geophysik

This journal was founded by the Deutsche Geophysikalische Gesellschaft on the initiative of L. Mintrop in 1924 as the *Zeitschrift für Geophysik* and edited by G. Angenheister from Vol. 1–18 (1944). It reappeared in 1954 edited by B. Brockamp from Vol. 19–26 (1960), and edited by W. Dieminger from Vol. 27–33 (1961), and from Vol. 34 (1968) edited by W. Dieminger and J. Untiedt. After Vol. 40 (1970) the title was changed to *Journal of Geophysics – Zeitschrift für Geophysik*.

Published: Vols. 19–39 by Physica-Verlag, Würzburg, from Vol. 40 by Springer Berlin, Heidelberg, New York, Tokyo.

Copyright

Submission of a manuscript implies that the work described has not been published before (except in the form of an abstract or as part of a published lecture, review or thesis), that it is not under consideration for publication elsewhere, that its publication has been approved by all the authors and by the responsible authorities – tacitly or explicitly – in the laboratories where the work was carried out and that, if accepted, it will not be published elsewhere in the same form, in either the same or another language, without the consent of the copyright holders. By submitting a manuscript, the authors agree that the copyright for their article is transferred to the publisher if and when the article is accepted for publication. The copyright covers the exclusive rights to reproduce and distribute the article, including reprints, photographic reproductions, microform, electronic data-base, video-disks, or any other reproductions of similar nature, and translations.

Photographic reproduction, microform, electronic data-base, video-disks, or any other reproduction of text, figures, or tables from this journal is prohibited without permission obtained from the publisher.

The use of general descriptive names, trade names, trade marks, etc., in this publication, even if the former are not specifically identified, is not to be interpreted as exempt from the relevant protective laws and regulations and may accordingly be used freely by anyone.

Special Regulations for the USA

The Article Fee Code on the first page of an article in this journal indicates the copyright owner's consent that in the USA copies may be made for personal or internal use, provided the stated fee for copying beyond that permitted by Section 107 or 108 of the United States Copyright Law is paid through the Copyright Clearance Center, Inc., 21 Congress Street, Salem, Mass. 01970, USA.

If a code does not appear, copies of the article may be made without charge, provided permission is obtained from the publisher.

The copyright owner's consent does not extend to copying for general distribution, for promotion, for creating new works, or for resale. Specific written permission must be obtained from the publisher for such copying.

Other Regulations

Authors of this journal can benefit from library and photocopy fees collected by VG WORT if certain conditions are met. Authors of German nationality and those resident in the Federal Republic of Germany or Berlin (West), as well as citizens of Austria, Switzerland and member countries of the European Community, may apply to Verwertungsgesellschaft WORT, Abteilung Wissenschaft, Goethestraße 49, D-8000 München 2, for detailed information.

Springer-Verlag Berlin Heidelberg New York Tokyo

Printed in Germany by Universitätsdruckerei H. Stürtz AG Würzburg

© Deutsche Geophysikalische Gesellschaft, Clausthal-Zellerfeld, 1985

Author Index

- | | | | |
|----------------------|-----------------------|--------------------|-------------------------|
| Aizenberg, A.M. 90 | Galibert, P.Y. 59 | Klöcker, N. 65 | Schultz-Krutisch, T. 51 |
| Akasofu, S.-I. 217 | Geichinsky, B. 33 | Korth, A. 65 | Siebert, M. 42 |
| Alekseev, A.S. 164 | Gowin, J. 137 | Kümpel, H.-J. 184 | Stammler, W. 172 |
| | | Kurita, K. 130 | Stiller, M. 137 |
| | Habermann, R.E. 172 | | |
| Beamish, D. 207 | Harjes, H.-P. 1 | Lee, L.C. 217 | Vasco, D.W. 178 |
| Bonhommet, N. 59 | Hayakawa, M. 130 | Lohr, G. 184 | Volkov, M.A. 125 |
| Bortfeld, R.K. 137 | Heller, F. 51 | Lühr, H. 65 | |
| | Hufen, J.-H. 42 | | |
| | | Maltsev, Yu.P. 125 | Wang, Y.S. 23 |
| Christensen, U.R. 72 | Kan, J.R. 217 | Meyer, J. 42 | Weidelt, P. 191 |
| Cormier, V.F. 14 | Kerner, C. 77 | | Wolf, D. 106 |
| | Kerth, M. 118 | Pao, Y.-H. 23 | Wyss, M. 172 |
| | Kind, R. 172 | | |
| Dresen, L. 77 | Klem-Musatov, K.D. 90 | Robert, P. 65 | Zherniak, G.F. 164 |
| | | | Ziegler, F. 23 |

Subject Index

Africa

Deep drilling trough the accreting plate boundary of Asal, Southern Afar: palaeomagnetism and magnetic properties of basaltic cores (Galibert, P.Y., Bonhommet, N.) 59

Alfvén waves

Observation of kinetic Alfvén waves excited at substorm onset (Klöcker, N., et al.) 65

Aquifer permeability

In-situ permeability from non-dilatational soil deformation caused by ground-water pumping – a case study (Kümpel, H.-J., Lohr, G.) 184

Birkeland currents

Finite parallel conductivity in the open magnetosphere (Volkov, M.A., Maltsev, Yu.P.) 125

On the origin of the cusp field-aligned currents (Lee, L.C., et al.) 217

Book reviews 222

British Isles

The frequency characteristics of anomalous vertical fields observed in the British Isles (Beamish, D.) 207

Coal seam waves

The influence of dirt bands and faults on the propagation of Love seam waves (Kerner, C., Dresen, L.) 77

Convection

Velocity-viscosity correlation in convection cells with non-uniform viscosity (Christensen, U.R.) 72

DEKORP

First results and preliminary interpretation of deep-reflection seismic recordings along profile DEKORP 2-South (Bortfeld, R.K., et al.) 137

Earthquake localization

The resolution of the Graefenberg array for earthquake locations in the eastern Mediterranean (Habermann, R.E., et al.) 172

Earth's core

Some problems with S, SKS and ScS observations and implications for the structure of the base of the mantle and the outer core (Cormier, V.F.) 14

Stromfunktion des erdmagnetischen Hauptfeldes in der Quellschicht an der Kern/Mantel-Grenze (Meyer, J., et al.) 42

Earth's crust

First results and preliminary interpretation of deep-reflection seismic recordings along profile DEKORP 2-South (Bortfeld, R.K., et al.) 137

Earth's mantle

Some problems with S, SKS and ScS observations and implications for the structure of the base of the mantle and the outer core (Cormier, V.F.) 14

Electromagnetic induction

Construction of conductance bounds from magnetotelluric impedances (Weidelt, P.) 191

The frequency characteristics of anomalous vertical fields observed in the British Isles (Beamish, D.) 207

Extremal models

Construction of conductance bounds from magnetotelluric impedances (Weidelt, P.) 191

Fresnel zones

The formulae for the calculation of the Fresnel zones or volumes (Gelchinsky, B.) 33

Geomagnetic main field

Stromfunktion des erdmagnetischen Hauptfeldes in der Quellschicht an der Kern/Mantel-Grenze (Meyer, J., et al.) 42

GEOS-2

Observation of kinetic Alfvén waves excited at substorm onset (Klöcker, N., et al.) 65

Germany

Measurement of magnetic susceptibility anisotropy in Buntsandstein deposits from southern Germany (Schultz-Krutisch, T., Heller, F.) 51

A palaeomagnetic study of Turonian carbonates from the southeastern Münsterland area, NW Germany (Kerth, M.) 118

First results and preliminary interpretation of deep-reflection seismic recording along profile DEKORP 2-South (Bortfeld, R.K., et al.) 137

Graefenberg array

The resolution of the Graefenberg array for earthquake locations in the eastern Mediterranean (Habermann, R.E., et al.) 172

Induction arrows

The frequency characteristics of anomalous vertical fields observed in the British Isles (Beamish, D.) 207

Inverse problems

Inverse dynamic problems and seismic methods for determination of the structure of a medium (Alekseev, A.S., Zherniak, G.F.) 164

Extremal inversion of vertical displacements, Long Valley Caldera, California 1982/1983 (Vasco, D.W.) 178

Construction of conductance bounds from magnetotelluric impedances (Weidelt, P.) 191

Isostasy

The normal modes of a layered, incompressible Maxwell half-space (Wolf, D.) 106

Magnetic susceptibility

Measurement of magnetic susceptibility anisotropy in Buntsandstein deposits from southern Germany (Schultz-Krutsch, T., Heller, F.) 51

Magnetosphere-Ionosphere coupling

Observation of kinetic Alfvén waves excited at substorm onset (Klößner, N., et al.) 65

Finite parallel conductivity in the open magnetosphere (Volkov, M.A., Maltsev, Yu.P.) 125

On the origin of the cusp field-aligned currents (Lee, L.C., et al.) 217

Magnetotellurics

Construction of conductance bounds from magnetotelluric impedances (Weidelt, P.) 191

Mediterranean

The resolution of the Graefenberg array for earthquake locations in the eastern Mediterranean (Habermann, R.E., et al.) 172

Normal modes

The normal modes of a layered, incompressible Maxwell half-space (Wolf, D.) 106

Palaeomagnetism

Deep drilling through the accreting plate boundary of Asal, Southern Afar: palaeo-

magnetism and magnetic properties of basaltic cores (Galibert, P.Y., Bonhommet, N.) 59

A palaeomagnetic study of Turonian carbonates from the southeastern Münsterland area, NW Germany (Kerth, M.) 118

Plasmopause

Evaluation of the effectiveness of theoretical model calculation in determining the plasmopause structure (Kurita, K., Hayakawa, M.) 130

Pump tests

In-situ permeability from non-dilatational soil deformation caused by groundwater pumping – a case study (Kümpel, H.-J., Lohr, G.) 184

Reflection seismics

First results and preliminary interpretation of deep-reflection seismic recordings along profile DEKORP 2-South (Bortfeld, R.K., et al.) 137

Inverse dynamic problems and seismic methods for determination of the structure of a medium (Alekseev, A.S., Zherniak, G.F.) 164

Rock magnetism

Measurement of magnetic susceptibility anisotropy in Buntsandstein deposits from southern Germany (Schultz-Krutsch, T., Heller, F.) 51

Deep drilling through the accreting plate boundary of Asal, Southern Afar: palaeomagnetism and magnetic properties of basaltic cores (Galibert, P.Y., Bonhommet, N.) 59

Seismic detection

Global seismic network assessment for teleseismic detection of underground nuclear explosions. I. Model calculations for different amplitude-attenuation curves (Harjes, H.-P.) 1

Seismic image inversion

Inverse dynamic problems and seismic methods for determination of the structure of a medium (Alekseev, A.S., Zherniak, G.F.) 164

Seismic waves

Some problems with S, SKS and ScS observations and implications for the struc-

ture of the base of the mantle and the outer core (Cormier, V.F.) 14

Transient SH waves in dipping layers: the buried line-source problem (Ziegler, F., et al.) 23

The formulae for the calculation of the Fresnel zones or volumes (Gelchinsky, B.) 33

The influence of dirt bands and faults on the propagation of Love seam waves (Kerner, C., Dresen, L.) 77

Seismic modelling by methods of the theory of edge waves (Klem-Musatov, K.D., Aizenberg, A.M.) 90

Soil deformation

In-situ permeability from non-dilatational soil deformation caused by groundwater pumping – a case study (Kümpel, H.-J., Lohr, G.) 184

Spherical harmonic analysis

Stromfunktion des erdmagnetischen Hauptfeldes in der Quellschicht an der Kern/Mantel-Grenze (Meyer, J., et al.) 42

Synthetic seismograms

Transient SH waves in dipping layers: the buried line-source problem (Ziegler, F., et al.) 23

Seismic modelling by methods of the theory of edge waves (Klem-Musatov, K.D., Aizenberg, A.M.) 90

Underground nuclear explosions

Global seismic network assessment for teleseismic detection of underground nuclear explosions. I. Model calculations for different amplitude-attenuation curves (Harjes, H.-P.) 1

Uplift

Extremal inversion of vertical displacements, Long Valley Caldera, California 1982/1983 (Vasco, D.W.) 178

VLF/ELF emissions

Evaluation of the effectiveness of theoretical model calculation in determining the plasmopause structure (Kurita, K., Hayakawa, M.) 130

Viscosity

Velocity-viscosity correlation in convection cells with non-uniform viscosity (Christensen, U.R.) 72

Global seismic network assessment for teleseismic detection of underground nuclear explosions

I. Model calculations for different amplitude-attenuation curves

H.-P. Harjes

Geophysical Institute, Ruhr-University, D-4630 Bochum, Federal Republic of Germany

Abstract. The detection capability of a global seismic network is examined on the basis of a probability model. Given the location of seismograph stations with known background noise level, a worldwide grid of epicentres and amplitude-distance attenuation curves, the detection capability is expressed by the magnitude corresponding to a fixed probability that a specified minimum number of stations detect an event.

A globally distributed network – composed by an international group of seismologists (network III from CCD/558, 1978) – is selected as a model. These stations are judged to produce the best seismological results currently achievable for teleseismic detection. Multi-wave detection criteria are applied which take variation of attenuation for different wave types into consideration. The extension of amplitude-attenuation curves to include core phases is investigated and effects of regional attenuation are studied. Depending on the detection criterion and attenuation curve, magnitude thresholds of a 50-station network can vary significantly.

Magnitude thresholds for this hypothetical network – requiring a 90% probability of at least four detecting stations – range from $3.4 \leq m_b \leq 3.6$ for Scandinavia and Europe, from $3.7 \leq m_b \leq 3.9$ for North America, Asia and Arctica, from $3.8 \leq m_b \leq 4.0$ for South America, Africa and Antarctica and $3.9 \leq m_b \leq 4.2$ for Australia, New Zealand and the Pacific.

Key words: Seismic detection probability – Magnitude threshold – Amplitude-attenuation curve – Global networks – Underground nuclear explosions

Introduction

“Although seismological capacity for identifying underground nuclear explosions may now be secondary to the political will of parties engaged in Comprehensive Test Ban negotiations it is still important to present the clearest possible evaluation of the role seismology might play should a Comprehensive Test Ban become reality.”

This quotation from a 12-year-old paper (Marshall and Basham, 1972) is still an adequate description of the general purpose of studies on seismic verification of nuclear test ban treaties. In a more specific sense, we

want to assess the detection threshold of a network of modern seismic stations. Detection thresholds will be given in terms of magnitude. Therefore, thresholds described herein apply to both shallow earthquakes and underground explosions without regard to source type. The important questions of source identification and yield estimation are not addressed in this paper.

Previous detection studies include the SIPRI-report (Davies, 1969), an analysis initiated by the United Nations (Basham and Witham, 1970) and a report of a group of seismologists to the Conference on Disarmament in Geneva (CCD/558, 1978). All of these assessments present conceptually similar schemes whereby worldwide existing seismological facilities are applied to a straightforward statistical estimation model. Given a globally distributed network of seismograph stations with known background noise level, a set of epicentre locations and standard amplitude-distance attenuation curves for seismic waves, the detection capability is expressed by the magnitude corresponding to a fixed probability that a specified minimum number of stations detect an event.

A second group of papers (Kelly and Lacoss, 1969; Report US/GSE/7, 1980) describe a different approach to examine the detection capability of a seismic network by including average worldwide seismicity. Using known earthquake recurrence rates, a synthetic list of events is produced as a reasonable approximation to those actually observed in a specific time interval. Keeping station parameters unchanged, this method allows an independent check on the results of studies of the first kind.

Finally, a third procedure starts from real data collected during special experiments (Lacoss et al., 1974) or published in bulletins by international agencies [e.g. International Seismological Centre (ISC) in Newbury, UK] and estimates detection thresholds with the use of Gaussian or maximum-likelihood techniques (Ringdal et al., 1977; Ringdal, 1984).

Estimates of the capabilities of seismic networks differ substantially as a result of these various approaches. In general the detection threshold increases in the order of the described procedures. Some reasons for these differences are obviously due to the difference between operational station performance used in the last approach and the idealized assumptions based on pure noise statistics which are input to the process

mentioned first. Smaller discrepancies simply reflect the difficulty in making this type of estimate and should be kept in mind in judging the accuracy of the results.

The main purposes of this study are:

i) To examine recently published approaches to network detection capability estimation by using multiwave detection criteria. A computer coded version of this procedure (Ciervo et al., 1983) was made available at the Center for Seismic Studies (CSS) in Arlington, VA. This code, called "Seismic Network Assessment Program for Detection (SNAPD)", not only models the propagation of P waves which were employed in previous programs (Wirth, 1977) but takes all relevant seismic phases into account and calculates wave attenuation and travel time as a function of regional media characteristics and event type.

ii) To study the influence of geophysical input parameters on the outcome. These parameters include, especially, the amplitude-attenuation curves at teleseismic distances and also the extension to core phases. Variations of attenuation in tectonic and stable areas at regional distances are of importance for detailed epicentre-station configurations. Special attention will be given to amplitude-attenuation curves derived from seismograms of underground nuclear explosions; otherwise earthquake data are included using shallow events.

Analysis of network detection probabilities

The statistical model and basic computational procedures are described in this section.

The model includes various parameters to be known at the beginning. The most important are

- Seismic station locations and their noise statistics (mean and variance)

- Amplitude-distance relations for several phases (besides P waves the prominent regional phases, Pg and Lg , are used)

- Signal variance

- Signal-to-noise ratio required for detection.

The procedure is then to do the following:

- Select a source location and compute detection probabilities for each station as a function of event magnitude.

- Find the lowest magnitude for which there is a defined probability to meet the specific detection criterion. (The most frequent criterion for global detection studies asks for a 90% probability of detecting P waves by at least four stations.)

First a single station is considered and the probability that it detects a certain wave is derived. Then the multiwave detection probability and the network capability are defined. We closely follow the notation by Ciervo et al. (1983). Further details can be found in Wirth (1977), Elvers (1980) and Evernden (1969a, b, 1975).

p_{ijk} denotes the probability that wave k propagated from epicentre j will be observed at station i . It is given by

$$p_{ijk} = R_i \Phi \left[\frac{\log A_{ijk}^{(\alpha)} - (\mu_{ik} + \log r_{ik})}{\sqrt{\sigma_{n_{ik}}^2 + \sigma_{s_{ik}}^2 + \sigma_{b_k}^2}} \right], \quad (1)$$

where

$$\Phi(x) \equiv \int_{-\infty}^x e^{-\frac{y^2}{2}} \frac{dy}{\sqrt{2\pi}} \quad (2)$$

is the normal cumulative probability function.

In Eq. (1), signal and noise are assumed to be log-normally distributed (Freedman, 1967) and the log of the noise amplitude has expectation μ_n and variance σ_n^2 , and the variance of the log of the signal is σ_s^2 .

$\sigma_{b_k}^2$ defines the additional variance of the log signal amplitude for wave k , given an m_b value; hence $\sigma_{b_k} \equiv 0$ if k denotes the P wave. A station i is supposed to detect wave k provided that the ratio of signal to noise is at least r_{ik} .

Given an event at epicentre j of magnitude m_b and distance Δ_{ij} from station i , the amplitude of wave k at that station is calculated as

$$\log A_{ijk}^{(\alpha)} = m_{k_j} + b_k^{(\alpha)}(\Delta_{ij}) + c_k^{(\alpha)} \log(\Delta_{ij}) + \varepsilon_{ijk}, \quad (3)$$

for both stable ($\alpha = S$) and tectonic ($\alpha = T$) media.

For waves other than conventionally used P -phase, m_b has to be converted into an adequate magnitude m_k given the respective regression formula

$$m_k = [KE]_k + [KM]_k m_b. \quad (4)$$

The b_k and c_k in Eq. (3) are attenuation table entries and ε_{ijk} is the epicentre-station calibration term for wave k .

If wave k does not require regional attenuation or if $\Delta > 25^\circ$ then $\log A_{ijk}$ is computed directly from Eq. (3) using a stable medium attenuation table. Otherwise, for regional distances ($\Delta \leq 25^\circ$)

$$\log A_{ijk} = (1 - w_{ij}) \log A_{ijk}^{(S)} + w_{ij} \log A_{ijk}^{(T)}, \quad (5)$$

where w_{ij} is the regional path weight; i.e. the ratio of the length of the wave path in tectonic media to the total great-circle path length Δ_{ij} . Especially if the epicentral path is assumed to have passed a region that severely attenuates Lg waves, then $\log A_{ij(Lg)} = -\infty$.

If the attenuation table entries are δ_k , b_k , c_k , then, for $\delta_{k-1} < \Delta_{ij} < \delta_k$,

$$b_{\Delta_{ij}} \equiv b_k \quad \text{and} \quad c_{\Delta_{ij}} \equiv c_k \quad (6)$$

except that if $c_k = 0$, linear interpolation is used for b :

$$b_{\Delta_{ij}} = b_k - (\delta_k - \Delta_{ij}) [b_k - b_{k-1}] / (\delta_k - \delta_{k-1}). \quad (7)$$

The station's probability of detection p_{ijk} , given by Eq. (1), is influenced by the reliability R_i of its operation. p_{ijk} includes, therefore, a factor R_i , $0 < R_i \leq 1$. R_i is of course dependent on a number of local circumstances which are not well known. Usually we set $R_i = 1$, assuming perfect operation.

Given the probability for a single station to detect an individual wave, we have to develop a procedure for multiwave network detection criteria that uses combinations of dependent wave arrivals at individual stations. An essential feature in the development of the model is the assumption that a minimum of four phases (not more than two of which are recorded at the same station) are required for detection assessment. Re-

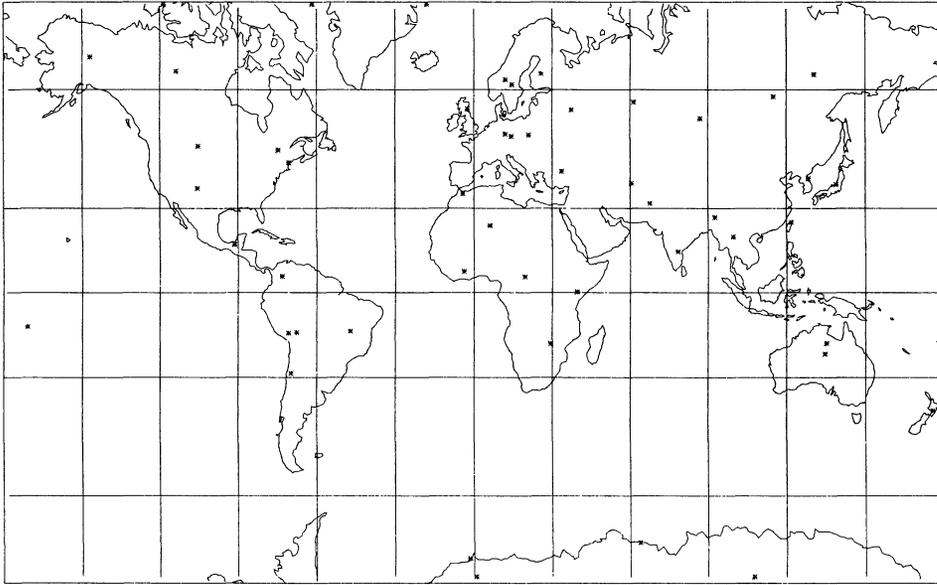


Fig. 1. Global seismic network (50 stations from CCD/558, 1978)

laxations of this requirement are possible and will be used when regional detection probabilities are estimated.

As an example, we illustrate the detection criterion

$$(Pg \cap Lg)/2 \cup P/4, \quad (8)$$

which literally means that a network detection consists of at least a two-station detection of Pg and Lg regional waves or at least a four-station detection of P waves or both. Expression (8) consists of two subcriteria

$$D^1 = (Pg \cap Lg)/2, \quad D^2 = P/4$$

which specify different wave combinations.

Detection criteria such as Eq. (8) have to be decomposed and reduced to a set of canonical probabilities. Therefore, the logical expression (8) is transformed into an algebraic expression involving the marginal probabilities of independent subcriteria and the joint probabilities of dependent pairs of subcriteria. This transformation eliminates all logical "or" (\cup) among subcriteria by use of the elementary rule

$$\text{prob}[D^1/n \cup D^2/m] = \text{prob}[D^1/n] + \text{prob}[D^2/m] - \text{prob}[D^1/n \cap D^2/m], \quad (9)$$

where, if D^1 and D^2 are independent wave criteria

$$\text{prob}[D^1/n \cap D^2/m] = \text{prob}[D^1/n] \text{prob}[D^2/m]. \quad (10)$$

In the case of Eq. (10), independent calculations for each individual wave define the probability $\text{prob}[D^1_{(n':N)}]$ that exactly n' out of N stations detect wave combination D^1 . If, however, the detection subcriteria D^1 and D^2 have waves in common, like

$$(Lg/2 \cap P/1) \cup P/4,$$

then computation of the joint probability, Eq. (10), does not split into independent probabilities.

Clearly the probability $\text{prob}[D^1/n]$ of at least n detecting stations follows as:

$$\text{prob}[D^1/n] = \sum_{n'=n}^N \text{prob}[D^1_{(n':N)}]. \quad (11)$$

Thus we need only compute the probability of exactly n' detecting stations.

In a concluding step, a binary search is used to find the magnitude value that results in p_t , the threshold probability for network detection.

Let m_i be a sequence of test magnitudes $m_{\min} \leq m_i \leq m_{\max}$ such that each m_i results in a network detection probability p_i . Initially $m_1 = m_{\min}$. If $p_1 \geq p_t$, the search is terminated; otherwise $m_2 = m_{\max}$ and if $p_2 \leq p_t$, the search is also terminated.

Assuming $p_1 < p_t < p_2$,

$$m_i = (m_{i-1} + m')/2, \quad i = 1, 2, 3, \dots$$

where, if p_{i-1} is less than p_t , m' is the last test magnitude for which the corresponding network probability is greater than p_t , and vice versa. The search is continued until $i = i^*$ is reached such that

$$|p_{i^*} - p_t| \leq \varepsilon,$$

corresponding to a threshold magnitude m_{i^*} .

Detection capability of a global network

In selecting a network for detection of seismic events based upon existing seismograph stations, it is desirable to

- arrive at a relatively uniform geographical distribution of stations
- select stations with modern instrumentation and optimum detection capabilities.

With these criteria in mind, a selection of relatively few stations is considerably more effective than using all (about 1,000) stations that routinely report to one of

Table 1. Stations used in network detection capability computations with noise statistics (mean and standard deviation of logarithms in nm) from CCD/558 (1978)

Station code	LAT + = N	Long + = E	SP-noise	
			Mean	STD
AFI	-13.91	-171.78	1.60	0.50
ALE	82.48	-62.40	0.70	0.40
ANMO	34.93	-106.45	0.30	0.20
ANTO	39.92	32.82	0.30	0.20
ARE	-16.46	-71.49	0.85	0.45
ASP	-23.68	133.90	0.48	0.30
BDF	-15.39	-47.54	0.48	0.30
BNG	4.37	18.57	0.00	0.15
BOCO	4.62	-74.07	0.30	0.20
BOD	57.85	114.18	0.70	0.40
BUL	-20.14	28.61	0.60	0.35
CMT	18.79	-98.98	0.30	0.20
COL	64.90	-147.78	0.70	0.40
COM	16.15	-92.07	1.40	0.50
DAG	76.77	-18.77	1.08	0.50
EKA	55.33	-3.16	0.90	0.45
ELT	53.25	86.27	0.70	0.40
GAC	45.70	-75.48	0.30	0.20
GBA	13.60	77.40	1.18	0.50
GRF	49.69	11.21	0.30	0.20
HFS	60.13	13.70	0.00	0.15
IFR	33.31	-5.07	0.90	0.45
JYS	62.17	24.87	0.30	0.20
KHC	49.13	13.58	0.48	0.30
KIC	6.36	-4.74	0.48	0.30
KSR	38.00	128.00	0.48	0.30
LAO	46.68	-106.22	-0.40	0.15
MAIO	36.31	59.59	0.30	0.20
MAT	36.54	138.21	1.00	0.50
MAW	-67.60	62.88	1.00	0.50
MBC	76.24	-119.36	0.78	0.45
NAO	61.04	11.22	-0.10	0.14
NIE	49.41	20.31	0.70	0.40
NIK	-1.27	36.80	0.30	0.20
OBN	55.17	36.60	0.78	0.45
PNS	-16.27	-68.47	0.48	0.30
QUE	30.18	66.95	1.00	0.50
SBA	-77.85	166.76	1.48	0.50
SHL	25.57	91.88	0.30	0.20
SNA	-70.32	-2.33	1.20	0.50
SPA	-89.90	0.10	0.90	0.45
SVE	56.80	60.63	0.78	0.45
TAM	22.79	5.52	0.60	0.35
TATO	25.03	121.52	0.60	0.35
TLL	-30.10	-70.48	1.08	0.50
WEL	-41.29	174.78	1.48	0.50
WES	42.38	-71.32	1.18	0.50
WRA	-19.95	134.35	0.30	0.20
YAK	62.02	129.72	0.70	0.40
YKA	62.49	-114.60	0.48	0.30

the international data centres. In a previous report (CCD/558, 1978) a network of 50 stations was composed which was judged to produce the best seismological results currently achievable for teleseismic detection of seismic events.

These stations, whose geographical distribution is shown in Fig. 1 and whose coordinates and further parameters are given in Table 1, are used as a reference

network in our study. The global station distribution is not as uniform as desirable, having 36 stations in the northern hemisphere compared to 14 stations in the southern hemisphere or 31 stations in the eastern hemisphere compared to 19 in the western hemisphere, but it reflects to some extent the distribution of land masses on earth.

Mean value and standard deviation of logarithmic station noise which are used to calculate station detection probabilities from Eq. (1) are also included in Table 1 (columns 4 and 5). Due to the quoted report (CCD/558, 1978), average noise levels were partly derived from published noise power spectra, partly estimated from magnification curves of seismographs. Because of lack of adequate measurements, an adhoc procedure was used to estimate the variance of the station noise: those stations with higher noise level were also assigned greater variance.

As can be seen from Eq. (1), it is the combined effect of variance in noise amplitude and signal amplitude which influences the station detection threshold: it decreases as the denominator in Eq. (1) is increased, provided that the detection probability is less than 0.5, and vice versa. For the network, the detection threshold generally decreases when signal or noise variance is increased. For the 50-station network – introduced in Fig. 1 and Table 1 – a constant value for the standard deviation of signal amplitude (0.2 in logarithmic units) was used. Test runs showed that doubling this parameter to 0.4 result in a very small difference of the network detection threshold (not exceeding 0.1 magnitude unit). Consequently, network capability is not very sensitive to this parameter.

There are two other input parameters to Eq. (1) which are to be assumed in an adhoc manner: the reliability factor R describing grossly the station operation (up-time), is set in our calculations to 1.0. Earlier studies (Ringdal et al., 1977) have estimated this parameter to range from 0.8 to 1.0 for most of the stations selected for our network.

A minimum signal-to-noise ratio has to be chosen to detect seismic signals emerging from background noise. Throughout this study a s/n ratio [r_{ik} in Eq. (1)] of 1.5 was chosen. The capability results are easily transformed to correspond to other s/n ratios, since this parameter occurs as a difference ($m - \log r_{ik}$) in Eq. (1). Thus a simple relationship exists between the chosen value of r_{ik} and the corresponding magnitude level. If r_{ik} is increased from 1.5 to 3.0, for example, the threshold will be increased by $(\log 3 - \log 1.5) = 0.3$ magnitude units.

Amplitude-attenuation curves at teleseismic distances: $25^\circ < \Delta < 100^\circ$

Several investigations of amplitude attenuation with distance have been made since the pioneering work of Gutenberg and Richter (1956). Following their paper, we will summarize amplitude-distance relations from Eq. (3) in the form

$$B(\Delta) = b(\Delta) + c \cdot \log(\Delta).$$

So we can interpret B values in terms of magnitude units.

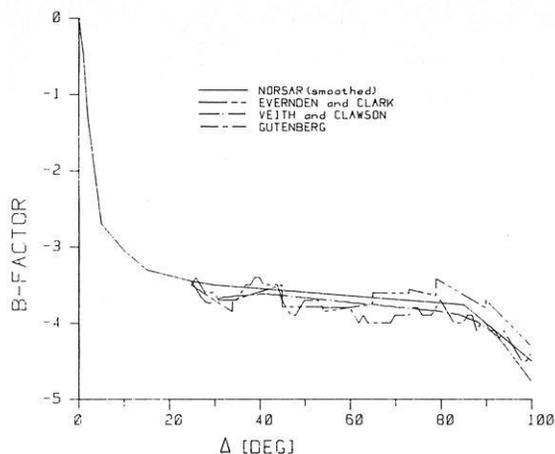


Fig. 2. Teleseismic attenuation curves for *P*-wave amplitudes

In Fig. 2, the results of Evernden and Clark (1970), Veith and Clawson (1972) and NORSAR (Ringdal and Fyen, 1979) are compared to the Gutenberg and Richter curve. Whereas Evernden and Clark and Veith and Clawson use mostly explosions and a station network of mainly LRSM-stations in the US, the Ringdal-Fyen curve is derived from ISC bulletin data (1971–1976) of 136 globally distributed WSSN stations.

All curves have been arbitrarily connected at regional distances to focus upon the differences in the teleseismic window. The principal difference between Veith-Clawson and NORSAR curves on the one hand and Gutenberg-Richter and Evernden-Clark curves on the other hand appears in the fact that the former indicate smooth amplitude variations from mantle discontinuities, whereas the latter indicate several step-wise changes in amplitude as a function of distance.

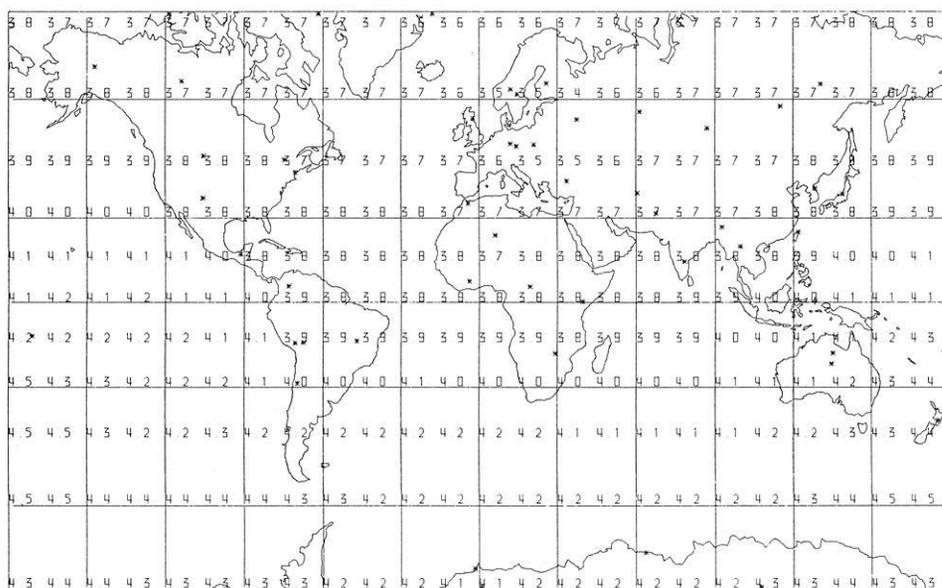


Fig. 3. 50-station network detection capability corresponding to a 90% probability of (at least) 4 detecting stations. NORSAR attenuation curve is used

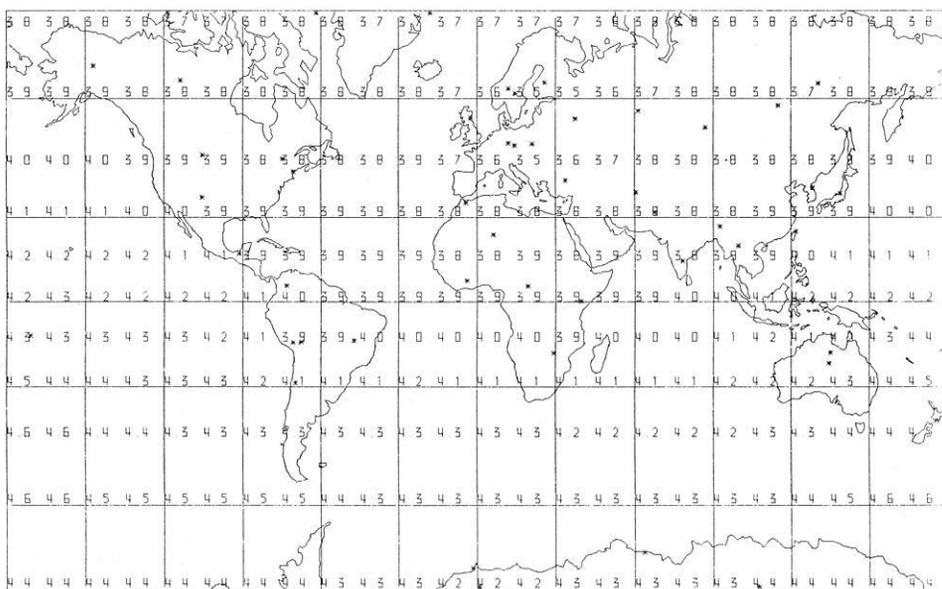


Fig. 4. 50-station network detection capability using Veith and Clawson attenuation curve. Other parameters are the same as in Fig. 3

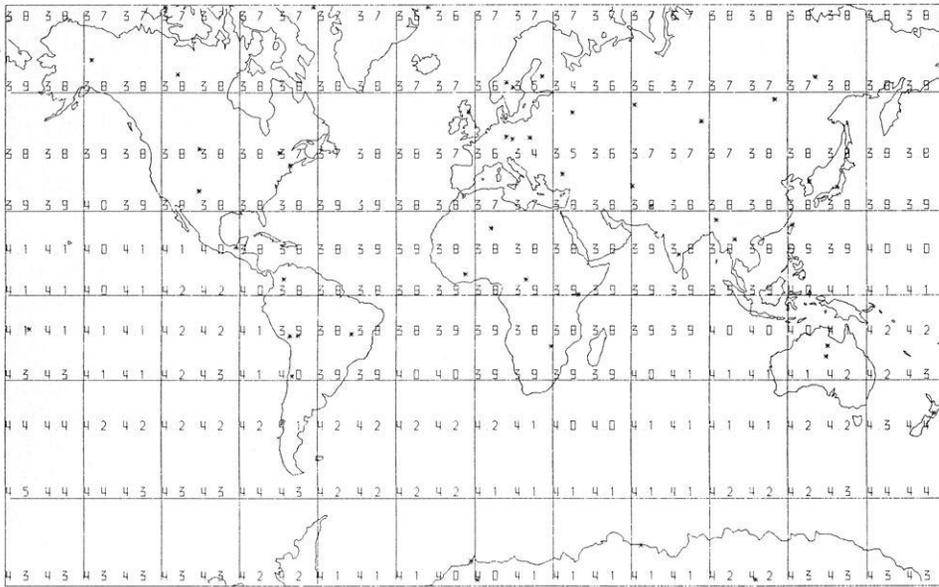


Fig. 5. 50-station network detection capability using Evernden and Clark attenuation curve. Other parameters are the same as in Fig. 3

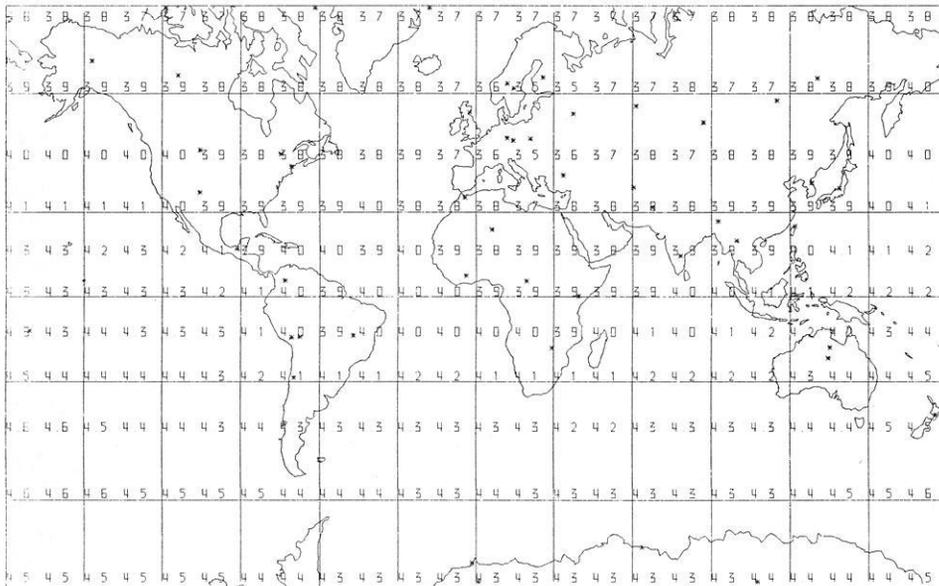


Fig. 6. 50-station network detection capability using Gutenberg and Richter attenuation curve. Other parameters are the same as in Fig. 3

The general difference between Veith-Clawson and Norsar attenuation curves is nearly a constant 0.1 magnitude unit over the whole teleseismic range. This may result from the fact that the Veith-Clawson curve was corrected for surface focus events; and the NOR-SAR curve, in comparison, comprises measurements from shallow earthquakes. Additionally, we have smoothed the minimum at 30° distance for NOR-SAR data. The obvious distinction between the curves of Gutenberg-Richter and Evernden-Clark appears in the far teleseismic portion where a difference of 0.3 magnitude units can be found. This is a consequence of the way Evernden and Clark have chosen to normalize their data.

Some of these differences can be seen in the network detection capability which is shown in Figs. 3–6. Using a 15° epicentre grid, the 90% probability of at least four detecting stations was calculated. For the NOR-SAR attenuation function (Fig. 3), which will be

used as a reference curve in this study, the magnitude threshold is estimated to be from m_b 3.4–3.7 in Europe and Scandinavia, m_b 3.7–3.8 in North America, Asia and Arctica, m_b 3.7–4.0 in Africa and most parts of South America, whereas we get values up to $m_b = 4.5$ in the Pacific region. The slight difference between the eastern and western hemisphere (about 0.2 magnitude units) as well as the large difference of more than one magnitude unit between the northern and southern hemisphere mainly result from the station distribution of the network. The high station noise at the sites in the Pacific (New Zealand and Samoa) gives an additional contribution to the low detection capability in the southern hemisphere.

As expected from the preceding discussion of attenuation curves (Fig. 2), detection thresholds for the network increase globally by about 0.1 magnitude units using the Veith-Clawson data. The Evernden-Clark curve has a remarkable effect in lowering the detection

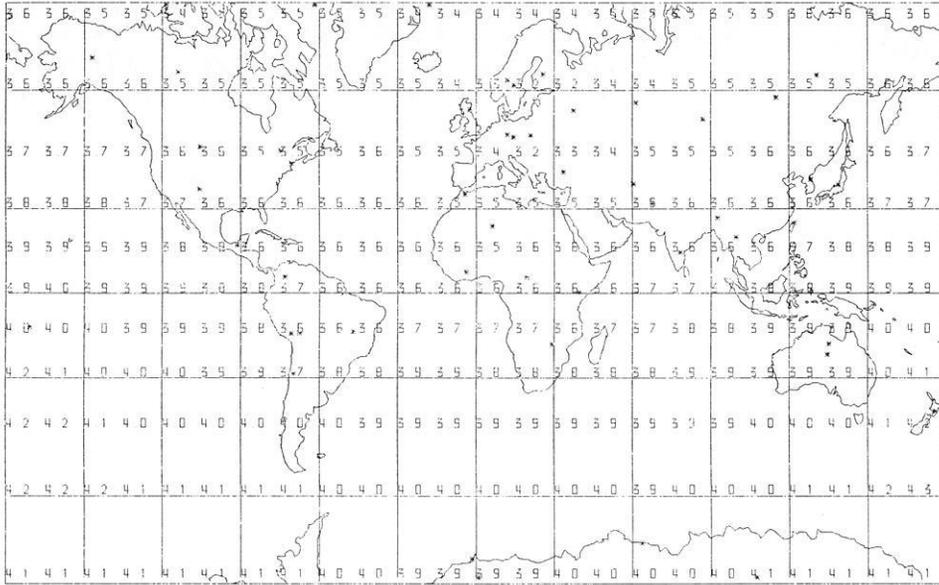


Fig. 7. 50-station network detection capability for 30% probability of 4 detecting stations. NORSAR attenuation curve is used

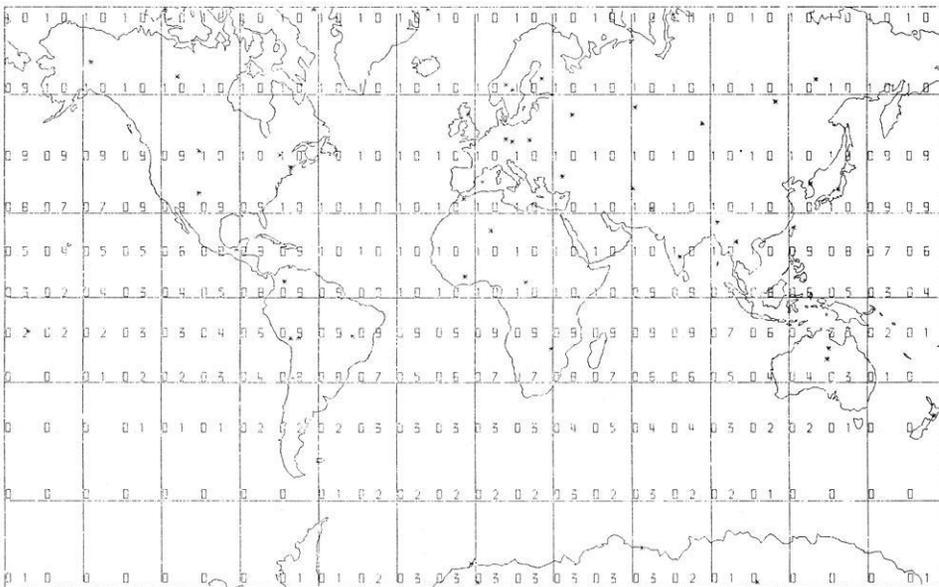


Fig. 8. 50-station network detection probability at fixed magnitude $m_b = 4$. NORSAR attenuation curve is used

threshold in the southern hemisphere by about 0.3 magnitude units (Fig. 5), again not unexpected from the shape of the attenuation curve which shows pronounced lower B values at far teleseismic distances than any other attenuation curve in Fig. 2. Finally, the Gutenberg-Richter curve yields detection thresholds (Fig. 6) very similar to the Veith-Clawson curve (Fig. 4).

We can conclude from these calculations that step-wise changes or fluctuations in the attenuation curve do not significantly affect the global detection capability of a 50-station network. Of course these discontinuities have remarkable focussing-defocussing effects for specific epicentre-station configurations, but these are smoothed and can not be resolved by global grids of 15° size. These peculiarities are better implemented by use of epicentre-station calibration factors expressed by ε_{ijk} in Eq. (3).

For all computations we kept the probability level at 90%. In changing this parameter one can significantly

influence detection thresholds (of course the seismological capability is not changed at all). At the 30% probability level (Fig. 7) we get 0.2–0.4 magnitude units lower thresholds compared to the commonly accepted 90% probability level.

Another way of demonstrating this difference is to calculate the network detection probability for a fixed magnitude value. Figure 8 shows the global distribution of probabilities to detect a magnitude 4 event. Besides Antarctica, New Zealand, the Pacific islands and the tip of South America, the chance of detecting events on land down to this size at (at least) four stations is higher than 80%.

Extension of amplitude-attenuation curves beyond 100° distance

The most prominent result of the last section is the clear difference in detection capability between the north-

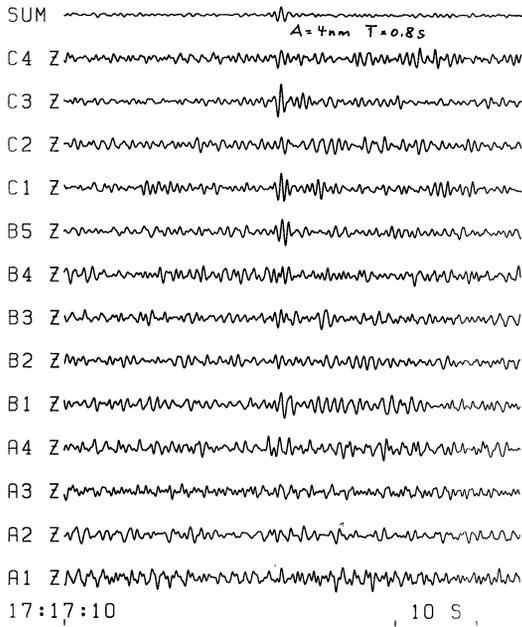


Fig. 9. French nuclear test recorded with Graefenberg array seismometers. Distance to test site is 143.7° . Amplitude on summation trace is 4 nm at 0.8 s

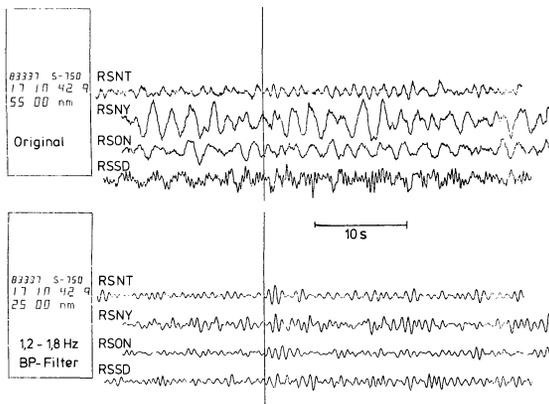


Fig. 10. RSTN-recordings of French nuclear test from previous figure. Upper traces are original data, lower traces are bandpass-filtered

ern and southern hemisphere regardless of what type of attenuation function was applied in the teleseismic window. To reduce the detection threshold in the southern hemisphere, a substantial increase in the number of stations in that region has often been recommended. Although this is a solution in principle, there are several practical problems with its realization. The

major part of the southern hemisphere is covered by deep ocean areas where installation and maintenance of seismographs is still difficult and expensive. Islands are known for a high microseismic noise level and poor detection capability.

An attractive alternative for improving the detection capability in the southern hemisphere is provided by the use of signals which have travelled through the earth's core and are routinely detected at stations beyond 100° distance. In a specific window, i.e. $142^\circ < \Delta < 152^\circ$, these refracted core phases offer even better detection possibilities than earlier described direct *P* waves. Numerous studies have shown that observation of high amplitudes for various *PKP* branches can be a powerful tool in lowering detection thresholds (Blandford and Sweetser, 1973; Quamar, 1973). This can easily be demonstrated with a seismogram of a French nuclear underground explosion exploded at Mururoa atoll (21°S , 140°W) and recorded with the Graefenberg array in Germany ($\Delta = 143.7^\circ$). Figure 9 shows a recording at all 13 vertical elements of the array (for a more detailed description, see Harjes and Seidl, 1978) and the beam-trace on top from which a displacement amplitude of 4 nm at a period of 0.8 s was measured.

In comparison, Fig. 10 shows recordings of the same event by the RSTN-stations in North America. These stations (Engdahl et al., 1982) are new borehole installations and include seismometers (Geotech S-750) with high sensitivity in the short-period band. Traces are aligned to the theoretical arrival time of the *P* wave which is marked by the cursor line. The stations have a distance of 73° – 88° from the event. Neither the original (upper part) nor the narrow-band filtered (lower part) traces meet the detection requirement set in our calculations. Correspondingly, this event was not reported by international data centres, which restrict their event-defining association process to *P* arrivals within 100° distance.

Table 2 shows the summary of station reportings for this event available from GTS/WMO-channels at the CSS. The association program implemented at the CSS found the questionable event by using *PKP* observations from three stations in Europe (including GRF). With the detection criterion we applied in the preceding paragraph (at least four *P* detections), this event would have been missed. Consequently, we amended amplitude-distance curves beyond 100° as shown in Fig. 11. There are two curves from different sources (Blandford and Sweetser, 1973; Ringdal, personal communication, 1984) which show a great similarity, although Blandford and Sweetser's curve is based on a

Table 2. Association result for French explosion using all WMO-messages at CSS

12/3/1983	OT = 16:58:3, 5	LAT = -20.95	LON = -139.87	H = 9 km	Tuamotu archipel	
Station	Arrival time	Phase	δt [sec]	Δ [Deg]	AMP [nm]	<i>T</i> [sec]
ALQ	17:8:38.0	P	-1.4	64.0	3.4	0.9
CTA	17:9:8.0	P	-0.5	68.6		
YKA	17:10:44.5	P	1.4	85.5	3.8	0.8
GRF	17:17:36.1	PKPAB	-0.3	143.3	4.0	0.8
PRU	17:17:39.0	PKPDF	0.3	144.7		
KHC	17:17:39.5	PKPDF	0.5	144.8		

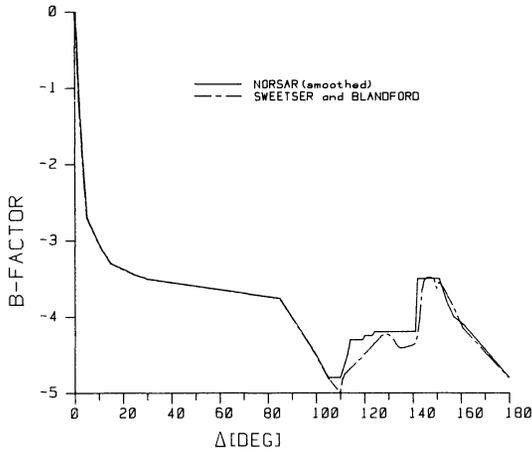


Fig. 11. Extension of amplitude-attenuation curve beyond 100° using *PKP* phases

more general data base (ISC) than Ringdal's data which are in use at the NORSAR array.

Detection threshold estimates for the 50-station network calculated with these attenuation curves differ less than 0.1 magnitude units. Israelson (1984) recommended the requirement of at least one *P* detection in addition to *PKP* arrivals to define an event to avoid large location errors because the stations observing *PKP* might be clustered in a narrow geographical area. Figure 12 gives the network detection capability for detecting at least four *P* or *PKP* arrivals, one of which has to be a direct *P* wave i.e. observed at a station within less than 100° distance of the event. This result can directly be compared with Fig. 3 because the attenuation functions are identical up to 100° distance.

The inclusion of *PKP* phases yields a large decrease of the detection threshold for the southern hemisphere (0.3–0.5 magnitude units) and divides the difference in detection capability between northern and southern hemispheres by half compared to the use of only *P* arrivals within 100° distance.

Variation of amplitude-attenuation curves at regional distances

In discussing the influence of varying teleseismic attenuation functions on the network detection capability, we always used the same curve for regional distances ($\Delta < 25^\circ$).

Although the teleseismic part (including *PKP*) will be most important for a global network, regional differences will have some effect in areas (Europe and Scandinavia) where the 50-station network is highly represented. These areas give us the opportunity to study the effect of different regional attenuation functions which, in a gross sense, represent "stable" (i.e. high *Q*) and "tectonic" (i.e. low *Q*) provinces.

The distinction was initially introduced for the North American continent, taking into account the different crustal structure in western and eastern US (Evernden, 1967). In this section we discuss only the influence of variations of P_n wave which is seen as the first arrival on regional seismograms recorded at distances greater than 1° . To emphasize the difference, rather extreme representatives of published attenuation curves have been used, namely Evernden's "8.5" curve (Evernden, 1967) derived from data in the eastern US, $b = -0.83$, $c = -2$ in Eq. (3), compared to a P_n -attenuation curve derived from data in southwestern US (Der et al., 1982) which yield $b = 0.134$ and $c = -3.803$. Thus the main difference between these two types of regional attenuation curves is that the "stable" P_n drops off as the square of the distance, while the "tectonic" curve decreases more rapidly, almost with the fourth power of the distance.

Another peculiarity at regional distances is the relative maximum in the amplitude-distance curve as a result of the 20° discontinuity which is observed worldwide with differing prominence (Gutenberg and Richter, 1956; Veith and Clawson, 1972). To emphasize also the effect of the 20° discontinuity the Veith-Clawson curve has been appended to Evernden's curve in the distance range $17^\circ < \Delta < 25^\circ$. On the opposite side,

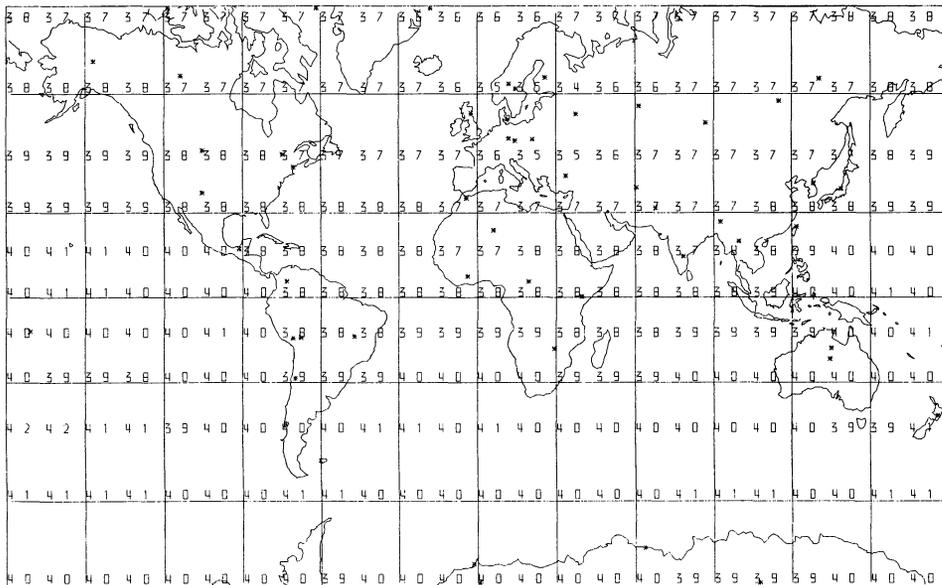


Fig. 12. 50-station network capability for (at least) 4 (*P* or *PKP*) detections. At least 1 station has to detect *P*. Other parameters are the same as in Fig. 3

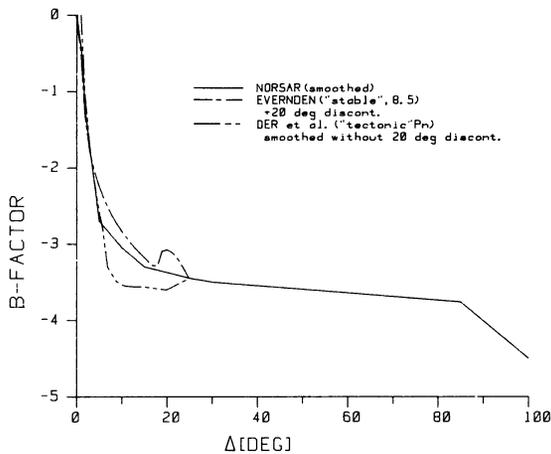


Fig. 13. Regional variation of attenuation curves to model "stable" and "tectonic" crustal wave paths

the "tectonic" attenuation curve has been smoothly connected to teleseismic distances without consideration of a 20° discontinuity.

Our final regional extensions of attenuation functions are plotted in Fig. 13. It is again noted that these curves are artificial compositions to show the most pronounced effect on network detection capability. The result can be seen in Figs. 14 and 15 which are to be compared to Fig. 12. Using specified regional attenuation curves generally increases the influence of station distribution. Global differences of detection thresholds are pronounced in Fig. 14, which shows the influence of a "stable" (Evernden "8.5" + 20° disc) regional attenuation curve. At close-to-station distances, detection thresholds are decreased by up to 0.2 magnitude units as compared to Fig. 12. On the contrary, detection thresholds are increased for most parts of Eurasia and America by up to 0.2 magnitude units for a "tectonic" attenuation curve (Fig. 15). Worldwide, this attenuation curve smooths the difference in detection capability between northern and southern hemispheres; certainly a result of our station distribution.

In summary, regional variations of *P*-wave amplitude-attenuation curves can change the detection threshold of a global 50-station network by as much as ± 0.2 magnitude units.

Regional detection probabilities of a global network

Although this paper mainly deals with teleseismic detection capabilities, it is interesting to investigate to what extent the magnitude threshold of a global network will be influenced by including phases other than P_n phases. It should be mentioned at the beginning that this section is intended as an amendment to teleseismic capabilities. To study the full potential of regional waves for detection purposes a regional station network has to be introduced. This is beyond the scope of this report and for details on this matter we refer to excellent recent review articles (Pomeroy et al., 1982; Blandford, 1981).

It is well known that the largest amplitudes on a short-period regional seismogram occur within the *Lg* wavetrain which may be interpreted as a superposition of a large number of higher mode Love and Rayleigh waves – *Lg* amplitudes can be 10 times larger than the maximum amplitude of the first arrival (*P* wave) at the same distance and in the same period band around 1 s. The actual amplitude can drastically vary due to local geology. An extensive literature exists on amplitude-distance attenuation curves for *Lg* for various regions of the world (for reference, see the above-mentioned review articles).

Again we extract two extreme representatives of published curves to examine the effect on the detection capability. As representing "stable" regions with low attenuation, we choose an attenuation curve published by Nuttli (1973). Nuttli derived an amplitude decay with distance proportional to $\Delta^{-5/3}$ ($0.5^\circ < \Delta < 40^\circ$) corresponding to the shape of the well-known "Prague"-formula (Vanek et al., 1962) adopted by IASPEI to be used for teleseismic Rayleigh wave observations. However, Nuttli derived his curve from observations of 1 s

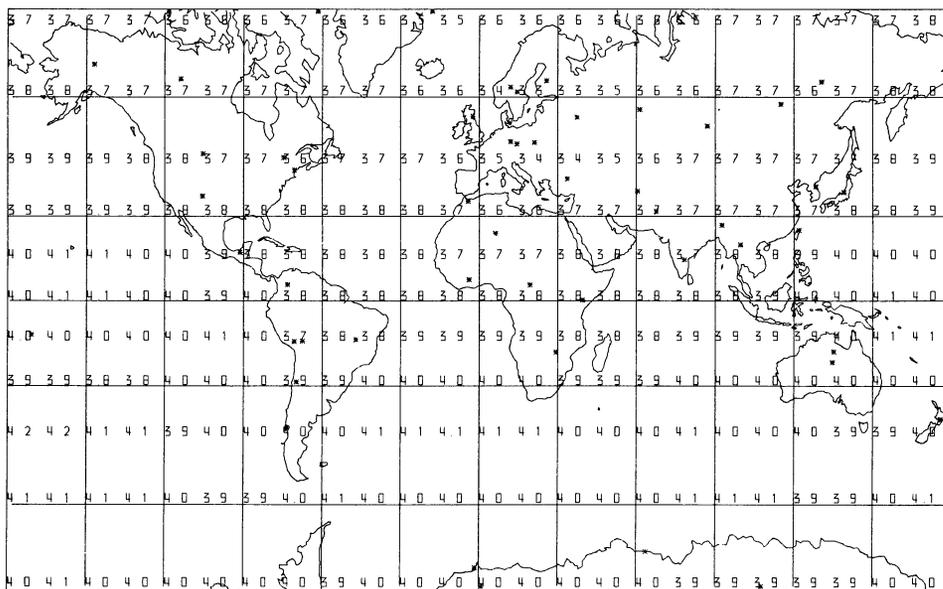


Fig. 14. 50-station network detection capability using "stable" attenuation curve from Fig. 14. Other parameters are the same as in Fig. 3

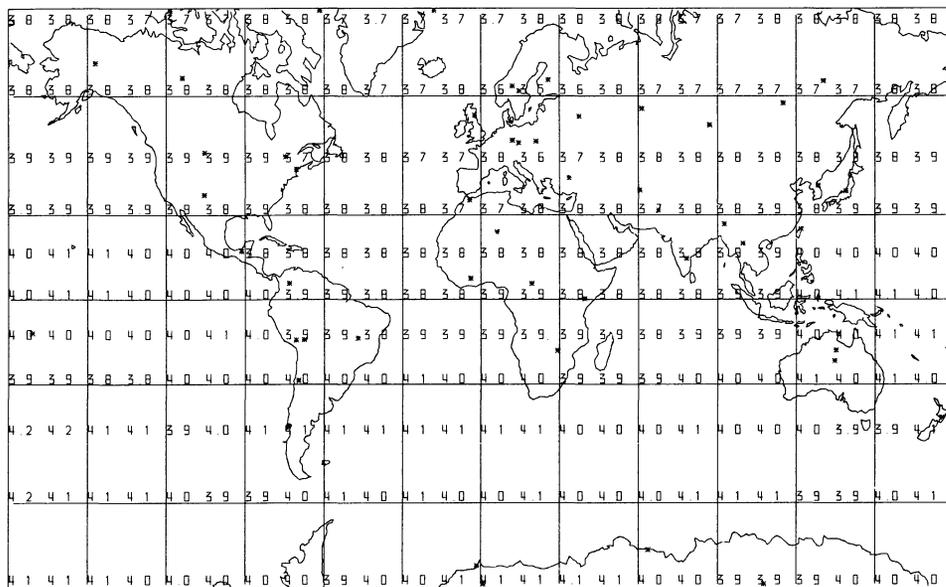


Fig. 15. 50-station network detection capability using “tectonic” curve from Fig. 14. Other parameters are the same as in Fig. 3

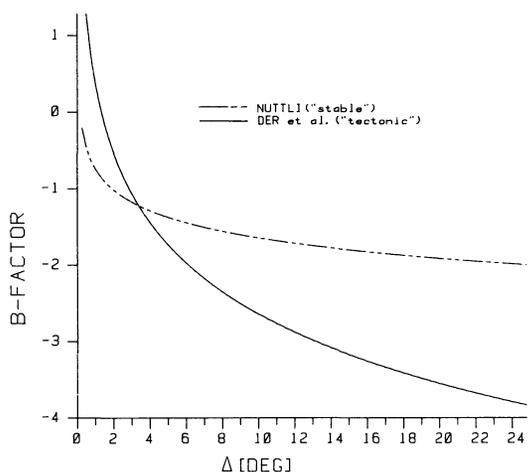


Fig. 16. Amplitude attenuation curves for L_g waves in “stable” and “tectonic” areas

L_g waves in the eastern US, whereas the “Prague”-formula is valid for Rayleigh waves around 20 s.

For crustal structures in “tectonic” provinces like the western US, an amplitude-distance decay for L_g proportional to Δ^{-3} was observed (Der et al., 1982). The selected attenuation curves are plotted in Fig. 16. In estimating detection capabilities by use of L_g waves it has to be mentioned that its amplitude not only varies regionally but it can also be totally suppressed. Representing a wave guide phenomenon, L_g propagation is seriously affected by variations in crustal thickness (mountains, ocean-continent and/or tectonic boundaries). To take these inefficient wavepaths into account a much finer grid than the $15^\circ \times 15^\circ$ grid used in our calculations has to be applied.

We want to examine whether inclusion of L_g waves has an effect on a global network at all. If there appear regional improvements of the detection capability, these have to be verified by considering the corresponding regional crustal structure. Because we want to insist on teleseismic detections we use L_g detections only if at

least one station of the network observed a teleseismic P wave ($\Delta > 25^\circ$). So we required at least two L_g detections and one teleseismic P detection or four P detections. As L_g is observed on all three components of ground motion, an azimuth estimate can be calculated and two stations are sufficient to roughly associate the event origin. An event is declared if at least one P observation at teleseismic distances confirms this association.

Figures 17 and 18 show the detection capability using this criterion for Nuttli’s and Der et al.’s attenuation function, respectively. In comparison to Figs. 14 and 15, which show the corresponding P -wave detection results, thresholds are lowered by 0.1–0.3 magnitude units. Again it should be emphasized that these improvements are irrelevant if they occur in oceanic areas because L_g waves disappear after crossing approximately 100 km of oceanic structure. Restricting the evaluation to continental areas only, it can be seen from Figs. 17 and 18 that the detection threshold is mostly influenced in the southern hemisphere where occasionally two stations are located at regional distances. In Europe and Scandinavia we already reach a high capability for four P detections which is not significantly improved by the additional (two L_g and one teleseismic) detection probability.

Finally, it might be mentioned that the detection difference caused by the difference of the two attenuation curves (Fig. 16) is only marginal for the selected detection criteria and station spacing of our global network.

Conclusions

Estimation techniques to examine seismic network detection capabilities, including multiwave criteria, are well established. For the hypothetical 50-station network studied in this report, best estimates of magnitude thresholds requiring a 90% probability of at least four detecting stations range from

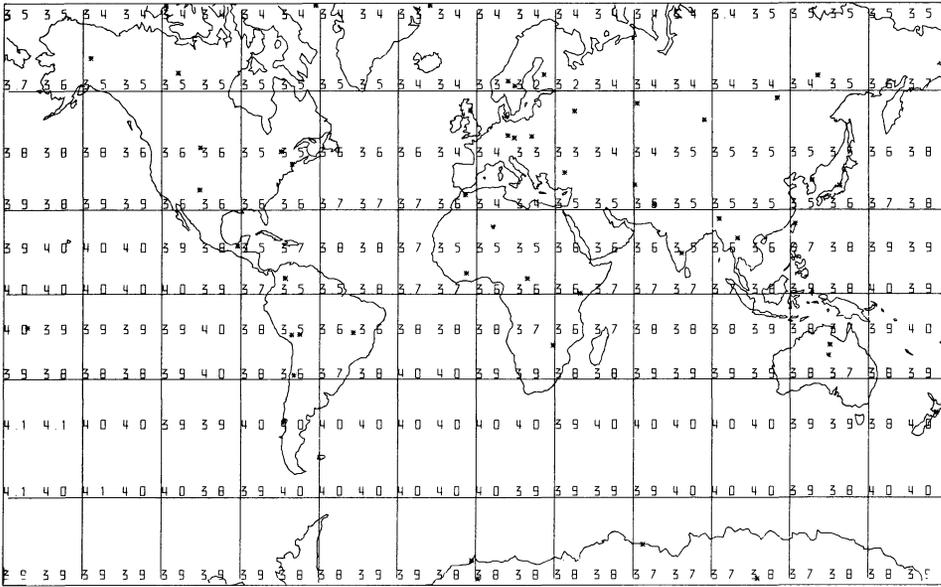


Fig. 17. 50-station network capability to detect (at least) L_g waves at 2 stations and P waves at 1 station or P waves at 4 stations. "Stable" attenuation curve for L_g waves is used

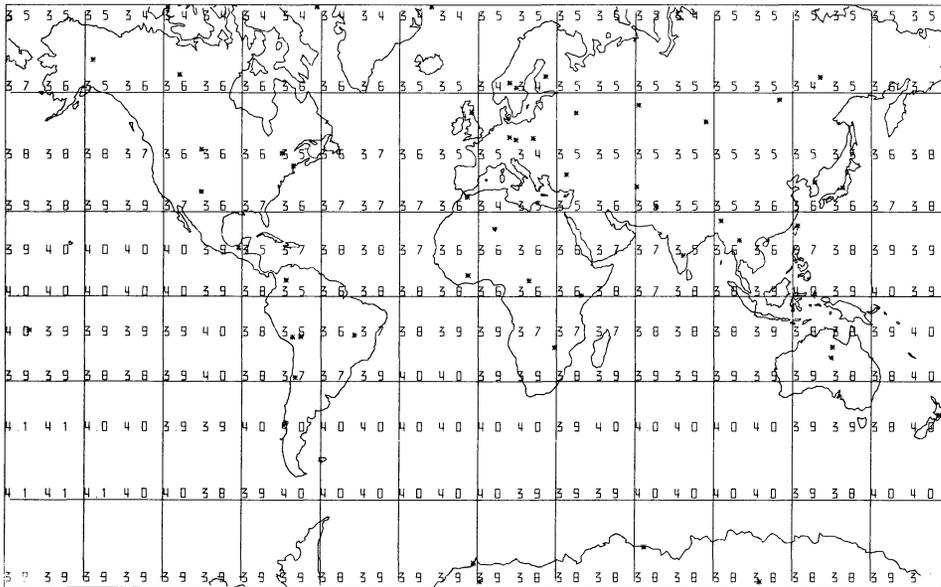


Fig. 18. 50-station network capability to detect (at least) L_g waves at 2 stations and P waves at 1 station or P waves at 4 stations. "Tectonic" attenuation curve for L_g waves is used

$3.4 \leq m_b \leq 3.6$ for Scandinavia and Europe,
 $3.7 \leq m_b \leq 3.9$ for North America, Asia and Arctica,
 $3.8 \leq m_b \leq 4.0$ for South America, Africa and Antarctica,
 $3.9 \leq m_b \leq 4.2$ for Australia, New Zealand and the Pacific.

Of course it has to be kept in mind that the results are based upon the particular input values chosen. In this report we have mainly studied the influence of different amplitude-attenuation curves and various wave types on the detection threshold of a global network. The essential conclusion from our calculations is a strong recommendation to extend the P -wave attenuation curves to PKP phases. This yields a large decrease of the detection threshold for the southern hemisphere and will reduce the difference in present detection capabilities between northern and southern hemispheres.

Emphasis is stressed upon teleseismic capabilities, regional effects are limited by the station spacing with-

in our network. Varying some input parameters, e.g. signal-to-noise ratio or station reliability, the estimates of detection threshold can easily be re-evaluated; the influence of others, e.g. signal variance or detection probability, can generally be estimated without reference to the specific configuration of the network.

Certainly, estimates of detection thresholds strongly depend on station locations and their noise statistics. As a starting model we have restricted our calculations to the hypothetical network composed by the Geneva group (CCD/558, 1978). In a subsequent paper we will compare these results with actual station reportings from a 2month data exchange experiment which included noise measurements. For the first time these data will allow the derivation of noise statistics on a global scale for simultaneously reported seismic events.

Acknowledgements. This work was carried out during a visit to the Center for Seismic Studies (CSS) in Rosslyn, Virginia.

The author is especially indebted to R. Alewine and A. Kerr from the U.S. Defense Advanced Research Projects Agency for their kind invitation. The support of C.F. Romney and the entire staff at the CSS is gratefully acknowledged. Many of the ideas in this paper resulted directly or indirectly from discussions with F. Ringdal to whom the author is deeply indebted. He further wishes to thank R. Suey, Pacific-Sierra Research, for his support in making SNAP/D run at the CSS, and W.J. Hannon, Lawrence Livermore Laboratory, for a critical review of the manuscript.

References

- Basham, P.W., Whitham, K.: Seismological detection and identification of underground nuclear explosions. Report, EPB, Department of Energy, Mines and Resources, Ottawa, Canada, 1970
- Blandford, R.R.: Seismic discrimination problems at regional distances. NATO, Advanced Study Institutes, Series C, 74. Dordrecht, Holland: D. Reidel Publ. Comp., 1981
- Blandford, R.R., Sweetser, E.I.: Seismic distance-amplitude relations for short period *P*, *P*_{diff}, *PP* and compressional core phases for delta >90 deg. Report, Teledyne-Geotech, SDAC-TR-73-9, Alexandria, VA, 1973
- CCD/558: Report to the Conference of the Committee on Disarmament of the AdHoc Group of Scientific Experts to Consider International Cooperative Measures to Detect and to Identify Seismic Events. Geneva, Switzerland, 1978
- Ciervo, A.P., Sanemitsu, S.K., Snead, D.E., Suey, R.W.: Seismic network assessment program for detection. Report 1027A, Pacific-Sierra Research, Los Angeles, CA, 1983
- Davies, D. (Rapporteur): Seismic methods for monitoring underground explosions. SIPRI, Stockholm, Sweden, 1969
- Der, Z.A., O'Donnell, A., McElfresh, T.W., Julita, R., Burnett, J.A., Marshall, M., Silk, M., Gordon, E.: A study of seismic wave propagation at regional distances in five areas of the world. Report, Teledyne-Geotech, VSC-TR-82-14, Alexandria, VA, 1982
- Elvers, E.: The capability of a network of seismological stations to detect events and to obtain identification parameters. Report, FOA, C 20231-T1, Stockholm, Sweden, 1980
- Engdahl, E.R., Peterson, J., Orsini, N.A.: Global digital networks - Current status and future directions. Bull. Seismol. Soc. Am. **72B**, S243-S260, 1982
- Evernden, J.F.: Magnitude determination at regional and near-regional distances. Bull. Seismol. Soc. Am. **57**, 591-639, 1967
- Evernden, J.F.: Precision of epicentres located by small numbers of worldwide stations. Bull. Seismol. Soc. Am. **59**, 1365-1398, 1969a
- Evernden, J.F.: Identification of earthquakes and explosions by use of teleseismic data. J. Geophys. Res. **74**, 3828-3856, 1969b
- Evernden, J.F.: Further studies on seismic discrimination. Bull. Seismol. Soc. Am. **65**, 359-392, 1975
- Evernden, J.F., Clark, D.M.: Study of teleseismic *P*. II. Amplitude data. Phys. Earth Planet. Int. **4**, 24-31, 1970
- Freedman, H.: Estimating earthquake magnitude. Bull. Seismol. Soc. Am. **57**, 747-760, 1967
- Gutenberg, B., Richter, C.F.: Magnitude and energy of earthquakes. Ann. Geofis. **9**, 1-15, 1956
- Harjes, H.-P., Seidl, D.: Digital recording and analysis of broad-band seismic data at the Graefenberg (GRF)-array. J. Geophys. **44**, 511-521, 1978
- Israelson, H.: Experiments with a computer program for automatic association of seismic events. Unpublished Manuscript, 1984
- Kelly, E.J., Lacos, R.T.: Estimation of seismicity and network detection capability. Technical Note 1969-41, Lincoln Lab, MIT, Mass., 1969
- Lacos, R.T., Needham, R.E., Julian, B.R.: International seismic month event list. Technical Note 1974-14, Lincoln Lab., MIT, Mass., 1974
- Marshall, P.D., Basham, P.W.: Discrimination between earthquake and underground explosions employing an improved Ms scale. Geophys. J. R. Astron. Soc. **28**, 431-458, 1972
- Nuttli, O.W.: Seismic wave attenuation and magnitude relations for eastern North America. J. Geophys. Res. **78**, 876-885, 1973
- Pomeroy, P.W., Best, W.J., McEvelly, T.V.: Test ban treaty verification with regional data - A review. Bull. Seismol. Soc. Am. **72B**, S89-130, 1982
- Quamar, A.: Revised velocities in the Earth's core. Bull. Seismol. Soc. Am. **63**, 1073-1105, 1973
- Ringdal, F.: Study of magnitudes, seismicity and earthquake detectibility using a global network. NORSAR Semiann. Techn. Summ., Kjeller, Norway, 1984
- Ringdal, F., Fyen, J.: Analysis of global *P*-wave attenuation characteristics using ISC data files. NORSAR Semiann. Techn. Summ., Kjeller, Norway, 1979
- Ringdal, F., Huseby, E.S., Fyen, J.: Earthquake detectibility estimates for 478 globally distributed seismograph stations. Phys. Earth Planet. Int. **15**, 24-32, 1977
- US/GSE/7: Investigation of the properties of CD-network III using synthetic data. Report of the United States Delegation, Geneva, Switzerland, 1980
- Vanek, J.A., Zatopek, V., Karnik, N.V., Kondorskaya, Y.V., Rizinichenko, E.F., Savarenski, E.L., Solov'ev Shebalin, N.V.: Standardization of magnitude scales. Bull. (Izvest.) Acad. Sci. USSR, Geophys. Ser. **3**, 108, 1962
- Veith, K.F., Clawson, G.E.: Magnitude from short-period *P*-wave data. Bull. Seismol. Soc. Am. **62**, 435-452, 1972
- Wirth, M.H.: Estimation of network detection and location capability. Report, Teledyne-Geotech, Alexandria, VA, 1977

Received October 18, 1984; Revised Version February 21, 1985
Accepted February 25, 1985

Some problems with *S*, *SKS* and *ScS* observations and implications for the structure of the base of the mantle and the outer core

V.F. Cormier

Department of Earth, Atmospheric and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, MA 02139, U.S.A.

Abstract. Complicated radially symmetric models of the seismic velocity structure at the base of the mantle (Bullen's D'' region) and the uppermost outer core have been inferred from analyses of the waveforms and relative amplitudes of *S*, *SKS* and *ScS* phases. Using radially symmetric structure, it has been difficult to construct physically realizable models of the rheology of D'' that simultaneously satisfy *P* and *S* amplitudes and slownesses in the core shadow. These data are reviewed in the light of an increasing body of evidence that the structure of D'' is characterized by heterogeneities having a broad spectrum of scale lengths.

Depending on the region and range interval of D'' sampled, *S* waveforms can be found that support either a radially simple or complex model of D'' . The complex models have one or more first-order discontinuities in velocity. The particle motion measured by three-component recordings of some *S*+*ScS* waveforms is consistent with a discontinuous increase in *S* velocity 250–300 km above the core-mantle boundary. The observed particle motion in these examples cannot readily or alternatively be explained by either general anisotropy or by strong lateral velocity gradients in D'' . Sufficient variability in *S* waveforms and travel times exists, however, that any radially symmetric model having a strong degree of complexity should be accepted with caution until all of the competing effects of lateral heterogeneity and possible anisotropy in D'' are fully investigated. The distribution and scale lengths of heterogeneities in D'' may account for regional differences in the properties of D'' inferred from waveform data, including features that mimic intrinsic attenuation and anisotropy

Key words: Lower mantle – *S* waves – Earth structure

Introduction

Almost every conceivable variation of *P* and *S* velocities has been proposed for the lowermost 200 km of the mantle [Bullen's (1950) D'' region]. The data fit by the various models have included the amplitudes, waveforms, travel times and apparent slownesses of core-diffracted *P* and *S* waves (e.g. Doornbos and Mondt, 1979; Ruff and Helmberger, 1982) and *S*, *SKS* and *ScS* waves (e.g. Mitchell and Helmberger, 1973; Lay and Helmberger, 1983a, b). The only common feature among the many different studies is a recognition that the behaviour of body waves sampling

the base of the mantle cannot be explained by simple extrapolation of the *P* and *S* velocity profile in the region immediately above D'' .

This extrapolation is performed using the theory of finite strain and the assumptions of homogeneity of composition and phase and an adiabatic gradient in temperature. The departure of a velocity profile in D'' from this extrapolation can be explained by relaxing any single one or combination of these assumptions. Independent geophysical evidence suggests that the assumption to be relaxed is adiabaticity. Estimates of the geotherm of the lower mantle and the heat flux from the core point to the existence of a thermal boundary layer at the base of the mantle (Jeanloz and Richter, 1979). Hence, many interpretations of the velocity profiles of D'' have attempted to determine a plausible thermal structure to account for specific models. Velocity models having either a simple reduction or reversal in gradient with depth can usually be shown to be consistent with compositional homogeneity and the effects of a single thermal boundary layer at the core-mantle boundary (Cleary, 1974; Jones, 1977). More complicated models having first-order discontinuities or zones of rapid velocity increase or decrease have been taken to indicate either compositional changes (Anderson and Hanks, 1972; Ruff and Anderson, 1980) or the existence of multiple thermal boundary layers (Ruff and Helmberger, 1982).

Many of these D'' models and their interpretations have ignored a growing body of evidence that D'' is characterized by a heterogeneous velocity structure. This evidence includes studies of the short-period precursors to the *PKP-DF* branch (Haddon and Cleary, 1974; Husebye et al., 1976; Haddon, 1982) and studies that have inverted large catalogues of *P* travel times for properties of the lowermost mantle (Comer and Clayton, 1984; Dziewonski, 1984). A significant result of both types of studies is that D'' seems to possess 2%–3% variations in velocity over a broad spectrum of characteristic scale lengths, from a few tens of kilometres to over 1000 km. This paper will review both the radially symmetric and laterally heterogeneous models that have been proposed for D'' and consider how the distribution of heterogeneities in D'' might affect observations of long-period body waves, specifically *S* waves. The example data and synthetic calculations will concentrate on the results of Lay and Helmberger (1983a, b). The intent of the paper, however, is to emphasize that the search for a D'' model must include all of the possible effects of velocity heterogeneity on body waves sampling D'' .

P, *S* and *K* velocity profiles

Simple radially symmetric

Figure 1 summarizes two *P* and *S* velocity profiles proposed for the base of the mantle. The models are chosen to be representative of simple forms. The PREM profiles (Dziewonski and Anderson, 1981) were constructed to satisfy a large data set of normal mode eigenfrequencies and *P* and *S* travel times. The feature that defines *D''* in PREM is a second-order discontinuity in the *P* and *S* velocity profiles 150 km above the core-mantle boundary. Dziewonski and Anderson (1981) note that this feature is primarily dictated by a sudden change in slope of $\frac{dT}{d\Delta}$ of *P* waves at 90° .

The scatter in *S* travel times precluded an inversion for such fine scale features and the gradient change shown in Fig. 1 followed from the assumption that second-order discontinuities in the *S*-velocity profile exist at the same depths as in the *P*-velocity model.

Doornbos and Mondt (1979) prefer negative rather than reduced positive velocity gradients in *D''*. They investigated perturbations of PEM (Dziewonski et al., 1975) needed to obtain agreement the spectral decay and apparent $\frac{dT}{d\Delta}$ s of

core-diffracted *P* and *S* waves. Model PEM-L01 from this study has reversals of both *P*- and *S*-velocity profiles starting 75 km above the core-mantle boundary. The gradients with depth for *P* and *S* velocity are both equal to -0.0019 s^{-1} . This gradient, which is positive with respect to radius *r*, is subcritical for *P* waves and close but less than the critical gradient for *S* waves ($\frac{d\beta}{dr} = 0.002 \text{ s}^{-1}$) in *D''*. Thus, turning rays can exist for both *P* and *S* waves in *D''*. Stacey and Loper (1983) have noted that a model having negative velocity gradients in *D''* such as these would require a compositional variation to be consistent with a plausible thermal structure. They also criticized the equivalence of *P* and *S* gradients as being inconsistent with second-order elasticity theory, which requires an increase

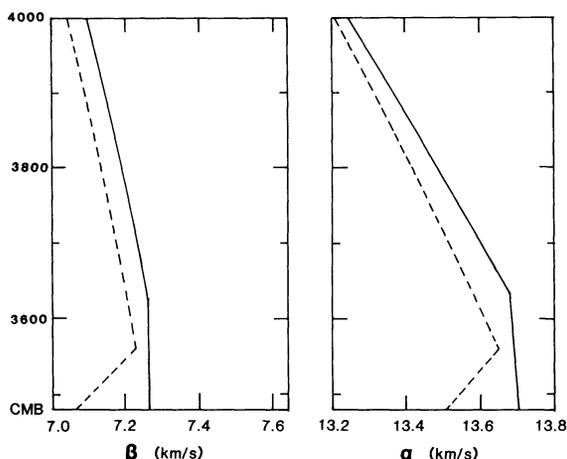


Fig. 1. The simplest models of *D''* have reduced or reversed velocity gradients in the lowermost 50–200 km of the mantle. *Solid lines* are the isotropic PREM velocities at 1 Hz (Dziewonski and Anderson, 1981). *Dashed lines* are model PEM-L01 (Doornbos and Mondt, 1979). The lower velocities of PEM relative to PREM above *D''* are due to the velocity dispersion of intrinsic anelasticity. PEM is referenced to 0.005 Hz rather than to 1 Hz

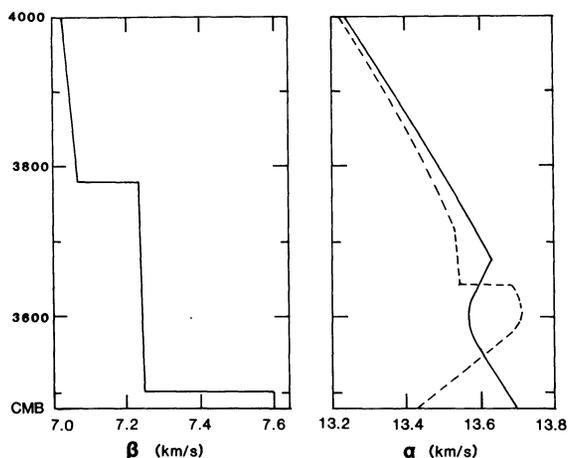


Fig. 2. Radially complex models have also been proposed for *D''*. The *S* model is by Lay and Helmberger (1983a). The thin (20 km thick) high *S* velocity layer at the base of the mantle was originally proposed by Mitchell and Helmberger (1973). The *P*-velocity profiles shown are (*solid line*) by Ruff and Helmberger (1982) and (*dashed line*) by Wright and Lyons (1981)

in Poisson's ratio with depth. This does not, however, allow for the effects of temperature.

Slightly positive *P* and *S* velocity gradients in *D''* are favoured by Mula and Müller (1980) and Mula (1981). Their analysis of diffracted *P*- and *S*-wave amplitudes and slownesses could not exclude a small negative gradient in *P* velocity of the size and width favoured by Doornbos and Mondt. They concluded, however, that the most reliable *S* data best agreed with a small positive rather than negative gradient in *S* velocity.

Complex radially symmetric

Another class of *D''* models have velocity discontinuities or narrow zones in which *P* or *S* velocity rapidly increases or decreases (Fig. 2). Ruff and Helmberger (1982) fit small scale features of the amplitude curve of short-period *P*+*PcP* and diffracted *P* with a model in which *P* velocity rapidly decreases in a narrow zone 160 km above the core-mantle boundary and then smoothly increases. Although this study used a normalization and averaging procedure to reduce the amplitude fluctuations due to near receiver structure, some fluctuations may remain due to variations in source locations and near-source structure.

Wright and Lyons (1981) and Wright et al. (1985), using a treatment of array data to resolve interfering multiples in short-period *P* waveforms, favour a *P* structure having a rapid or discontinuous increase in *P* velocity 160 km above the core-mantle boundary. This is followed by a smooth decrease in *P* velocity toward the core-mantle boundary. A feature similar to this is found in Lay and Helmberger's (1983a, b) SLHO models of *S* velocity. A discontinuity is introduced 250–300 km above the core-mantle boundary to produce a triplication between *SH* and *ScSH* in the distance range 70° – 85° . Without introducing lateral heterogeneity, however, this model cannot fit the *S* waveforms of specific data sets in the longer distance range beyond 100° (Schlittenhardt et al., 1985). A similarly shaped *P* profile cannot fit *P* waveforms unless the jump in *P* velocity is smaller by one-half or less in per cent than

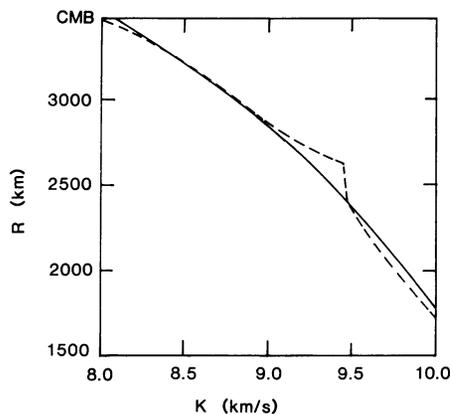


Fig. 3. Kind and Müller proposed a region of rapid velocity increase in the outer core to fit *SKS/SKKS* amplitude ratios (dashed line). The PREM velocity profile in the outer core is shown by the solid line

that in *S* velocity (Schlittenhardt et al., 1985; Lay and Young, submitted for publication). This *P*-velocity jump is closer in size to the one originally proposed by Wright and Lyons (1981) than the one most recently proposed by Wright et al. (1985).

Figure 3 compares the *K*-velocity profile of PREM with one determined by Kind and Müller (1977). Discussion of this model is included here because it is another example of the distinctive behaviour of *S* waves interacting with the *D''* region. Kind and Müller found that standard Earth models having a smooth profile of *K* velocity in the outer core predict *SKS/SKKS* ratios that are too large compared with those observed in the 100°–110° range. Such models also predict differential travel times $T_{SKS} - T_{SKKS}$ that are too small at ranges greater than 105°. In order to fit travel times and amplitudes, Kind and Müller introduce the structure shown in Fig. 3, which generates a short triplication of the *SKS* travel-time curve and destructive interference of phases in the *SKS* waveform. Schweitzer and Müller (personal communication) have now tested whether the inner core phases, which were neglected in the earlier study, can account for the observed amplitudes and travel times. They have concluded that inner core phases do not contaminate the *SKS* and *SKKS* waveforms.

Radially asymmetric

Lateral variations in the seismic velocities of the lower mantle have been found to be significant and resolvable in inversions of large catalogues of *P* travel times reported by the ISC (Dziewonski, 1984; Comer and Clayton, 1984). These inversions are aimed at retrieving radially asymmetric perturbations to a best-fitting spherically symmetric Earth structure. The strongest heterogeneities obtained in these studies are concentrated near 5700 km and the lowermost mantle. The velocity perturbations are as large as 0.1–0.2 km/s in *D''*, which may represent a smoothing-over of a more intense variation concentrated in either a narrower depth range or over smaller lateral scale lengths. Disregarding this possibility, these variations are nearly as large as those proposed in some of the complex, radially symmetric models of *D''*.

Anomalies in *S* travel time have generally been found to correlate with those in *P* travel time (Hales and Roberts,

1970) and this appears to also be true in the lower mantle. Schweitzer (1984) has compared the travel times of *SKS*, *SKKS* and *P4KP* predicted by the core model of Kind and Müller with observed travel times sorted into different regions. He found that a radially symmetric model cannot simultaneously satisfy observations in different regions and concluded that lateral heterogeneity must exist in either the outer core and/or the lower mantle. If the heterogeneity is placed in the lower mantle, the sense of the perturbations in *S* velocity correlate with the perturbations in *P* velocity found in the Dziewonski (1984) study. Schweitzer also reported regional variations in the *SKS/SKKS* ratio. Together with the variations in $T_{SKS} - T_{SKKS}$ times, these amplitude ratios are consistent with the sense of those expected for focussing and defocussing by lateral velocity variations; i.e. slow travel times correlate with high amplitude and fast travel times correlate with low amplitude.

Comparison of *S* data and *S* synthetics

Features of *SH* and *SV* waveforms predicted by PREM

Synthetic *S* waves have been computed using the isotropic PREM velocities and the techniques described in Choy (1977) and Cormier and Richards (1977). Profiles of the radial, transverse and an intermediate component of horizontal ground motion (Figs. 4 and 5) were calculated for the depth and focal mechanism of the 1970 September 5, Sea of Okhotsk event included in the studies of Mitchell and Helmberger (1973) and Lay and Helmberger (1983a, b). All components were calculated for a profile of stations along an azimuth at the epicentre representative of North American stations. The synthesis included the infinite set of phases *SKS*+*SKKS*+..., which form an interference head wave along the underside of the core-mantle boundary.

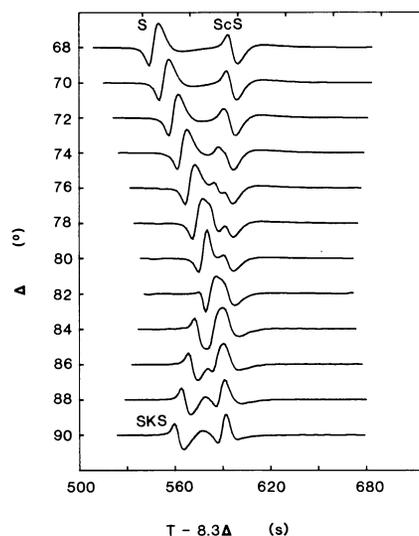


Fig. 4. Synthetic *SV* waveforms calculated in the isotropic 1 Hz PREM model for a point double couple source at 540 km depth and a LP-WWSSN instrument response. The focal mechanism was taken to be that determined by Strelitz (1975) for the September 5, 1970 Sea of Okhotsk earthquake. The takeoff azimuth was chosen to be 40°, which is representative of North American stations from the Sea of Okhotsk

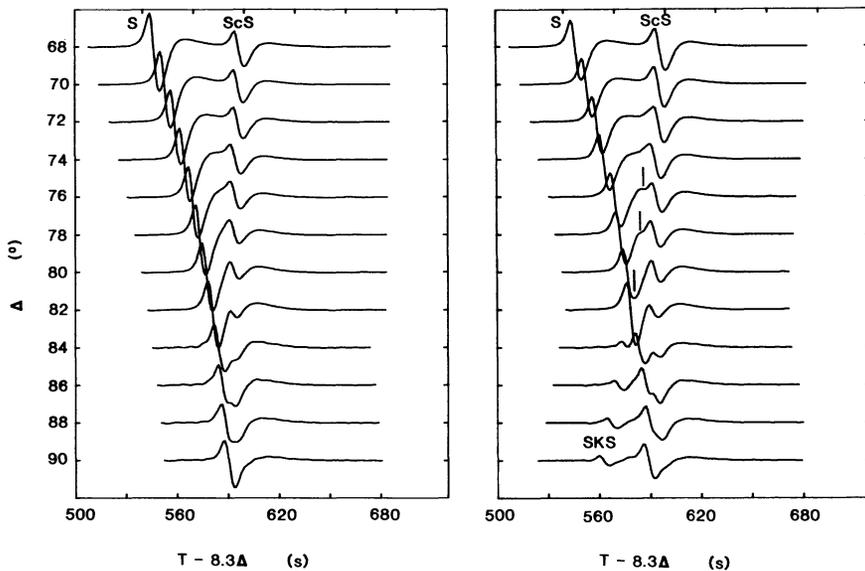


Fig. 5. Synthetic *SH* waveforms (*left*) and synthetic *SH* waveforms contaminated by *SV* energy (*right*) by rotating into an azimuth -20° from the true azimuth of arrival. The source is the September 5, 1970 Sea of Okhotsk earthquake described in Fig. 4. Note the perturbation in the waveforms at the *right* due to *SKS* starting at about 76°

Note that the *SKS* phase arrives between the *S* and *ScS* phases in the distance range 66° – 80° and crosses over the *S* phase at 82° . Also note (Fig. 5) that for this focal mechanism at least, the *SKS* phase will be visible on a horizontal component even within 10° – 15° of the true transverse component. It will be difficult to diagnose the perturbation marked in Fig. 5 as *SKS* on the basis of its measured slowness. This is because its small amplitude will make it hard to accurately correlate a peak or trough using a sparse long-period array. These features motivated a series of experiments, which are described in the following subsection, to determine whether the phase observed between transverse *S* and *ScS* by Lay and Helmberger (1983a, b) could be the *SKS* phase rather than a branch of a triplication due to a discontinuity near to core-mantle boundary.

SKS contamination of the transverse component

Laterally heterogeneous structure along the ray path can modify both the orientation of the *S* polarization vector and the azimuth of the ray arrival (Cormier, 1984). When this occurs, two mechanisms exist by which *SKS* energy can be observed on the transverse component. First, deviations in ray path can cause the rotation of horizontal components using the great circle azimuth to contaminate the theoretical transverse component with *SKS* energy. Second, even for an *S* wave that has left the source as a pure *SH* wave, deviations in the *S* polarization vector can introduce an effective *SV* component of particle motion at the point of incidence of the *S* wave on the core-mantle boundary. This mechanism can produce *SKS* energy at azimuths that are nodal for *SV* radiation. Whether or not either mechanism operates, the *SKS* phase can be minimized by resolving the NS and EW components into radial and transverse components using an apparent rather than a great circle azimuth. The apparent azimuth can be determined either from an array measurement of vector slowness or from the orientation of the ellipsoid of particle motion.

As an experiment to see how the rotation of horizontal components can affect on *SH* or *SV* waveform, several events studied by Lay and Helmberger were digitized and rotated using a variable azimuth. Figure 6 shows the results

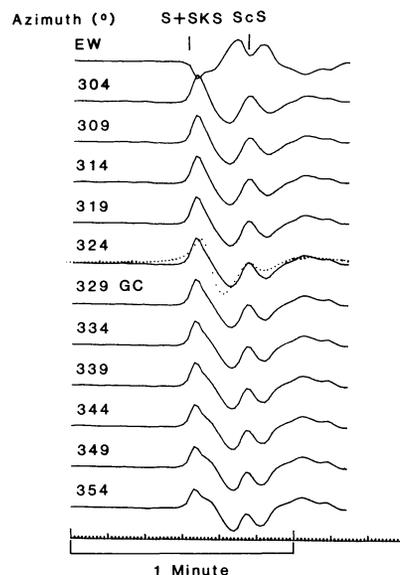


Fig. 6. *S* waveforms at station ATL recorded from the September 5, 1970 Sea of Okhotsk earthquake. ATL is 81.4° from this event. *S* has been rotated into *SH* assuming a back azimuth that varies about the true great circle back azimuth of 329° . A synthetic *SH* (dotted) calculated in PREM fits most of the traces equally well until a back azimuth of 334° and greater, where an inflection due to the arrival of *SKS* begins to be amplified. At 81.4° ATL is close to the cross over distance of *S* and *SKS*. The time difference between *S* and *SKS* agrees well with the predictions of PREM

of one such experiment. The *S* waveforms recorded by station ATL from the September 5, 1970 event were rotated using a sequence of different back azimuths. Note that beginning at the great circle azimuth of 329° and increasing towards rotations using larger back azimuths, a double inflection in the first downswing of the pulse is seen. This is generated by the interference of the *SKS* phase with the *S* phase and may be mistakenly identified as a branch of a new triplication. In this example, the waveforms and relative arrival times of all phases agree with those predicted by PREM, including the *S*–*SKS* time difference.

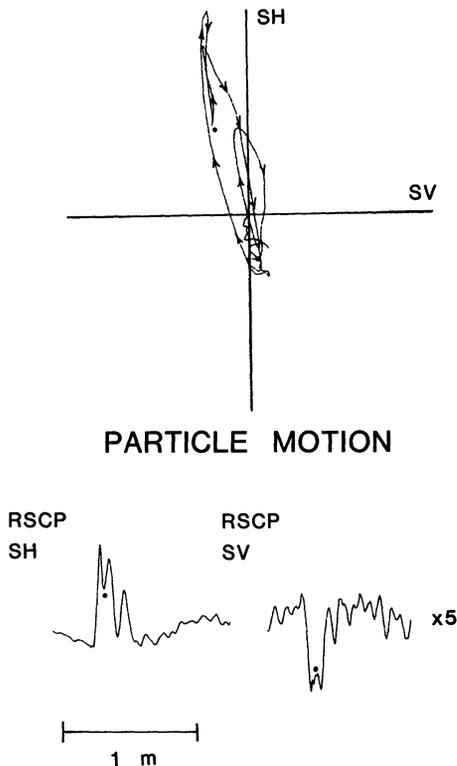


Fig. 7. Particle motion and S waveforms observed at station RSCP from a Sea of Okhotsk deep focus earthquake (April 20, 1984; 06 31 10.6 GMT; 50.12N, 148.74E; 582 km deep). The feature marked with the *dot* is not predicted by models of D' such as PREM, which has a smooth profile in S velocity in the lower mantle. The particle motion remains nearly pure SH across a time window that includes the anomalous feature in the waveforms

Similar, more extensive tests (Lay and Young, submitted for publication), however, are not always successful in explaining the $S+ScS$ waveform observed on the transverse component. In many cases, good matches between observed and synthetic waveforms can be obtained using the PREM model only if so much SKS contamination is introduced as to be grossly inconsistent with both the focal mechanisms of the events and any likely lateral heterogeneity.

More precise tests of SKS contamination are possible with digitally recorded waveforms. The RSTN digital network, which is one of the few digital networks to have horizontal components in the broad and short-period bands, offers an opportunity to examine S particle motion. Broadband displacements from a deep focus earthquake beneath the Sea of Okhotsk, which occurred on April 20, 1984, were constructed from mid- and short-period recordings using the procedure described in Harvey and Choy (1982). From the analysis of short-period and broadband P waves, the source time function of this event is estimated to have a duration of approximately 2 s. Thus, at the time scale shown in Figs. 7–9, any complexity seen in the $S+ScS$ waveforms is due to the effects of earth structure rather than the source. The particle motion observed at station RSCP (82°) is shown in Fig. 7. The anomalous feature predicted by the D' triplication of Lay and Helmberger is marked by a dot in the SH and SV waveforms and on the particle motion plot. The focal mechanism of this event and the azimuth of RSCP is such that SV radiation is nodal, with the SV waveform being about a factor of five smaller

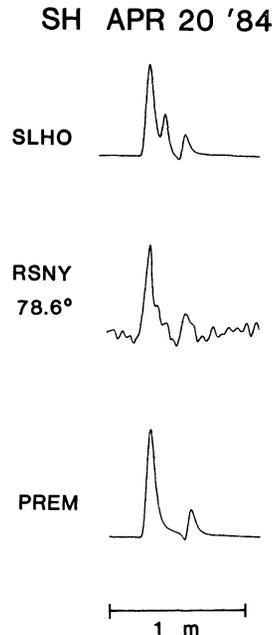


Fig. 8. A comparison of an $S+ScS$ waveform observed on the transverse component of station RSNY from the deep focus earthquake described in Fig. 7 with synthetic seismograms calculated in the PREM and SLHO models

than the SH waveform. The particle motion remains nearly pure SH in the vicinity of the anomalous feature between the S and ScS waves. It does not suggest any SKS contamination, which would have signified itself in the plot of particle motion by an abrupt change to a more SV -like particle motion at the arrival of the anomalous feature.

In summary, some $S+ScS$ waveforms can be found in the 75° – 82° range that are consistent with PREM, allowing for small amounts of SKS contamination. In many other cases, however, the waveforms are more consistent with the prediction of the SLHO model, having particle motions that cannot be explained by SKS contamination.

Comparison of data with PREM and SLHO synthetics

Using a source time function and focal mechanism of the April 20, 1984 event (Choy, personal communication), S waveforms were synthesized in the PREM and SLHO models and compared with those observed at the RSTN stations. All of the synthetics include an attenuation operator with $t^*_\beta=2$, constant across the frequency band. In a model of the SLHO type, the interference of multiple branches of the travel-time curve can cause a broadening in the observed pulse. The relatively broad base of the direct S pulse at RSNY (78.6°) in Fig. 8 contains a suggestion of the intermediate phase predicted by the SLHO model. The size and sharpness of the velocity jump above the core-mantle boundary can be easily adjusted in an SLHO type model to obtain a better match to the RSNY waveform than that shown in Fig. 8. A good fit to the RSNY waveform can also be obtained with the PREM model by modifying the attenuation operator. In order to match the steep rise time of the waveform and to be consistent with the source duration inferred from P waveforms in the short-period band, the required attenuation operator must be strongly frequency dependent. At long periods, however, such an oper-

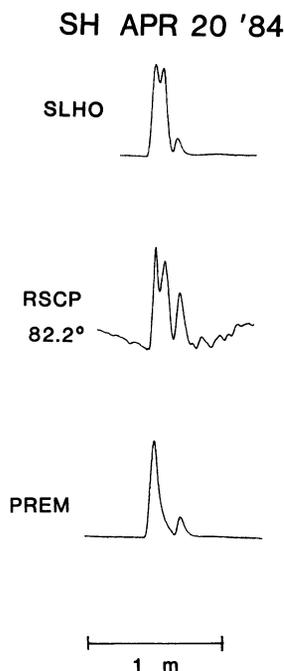


Fig. 9. A comparison of an $S+ScS$ waveform observed on the transverse component of station RSCP from the deep focus earthquake described in Fig. 7 with synthetic seismograms calculated in the PREM and SLHO models

ator predicts an S -wave attenuation that is significantly larger ($t^*_\beta > 6$) than the S attenuation typically measured from a shallow focus earthquake (e.g. Brudick, 1978). A body wave from a deep-focus earthquake, such as the one studied here, is instead typically observed to be attenuated only by one half as much as one from a shallow-focus earthquake (e.g. Der et al., 1982). In this sense, an SLHO type model provides a better fit to the broad base of the waveform at RSNY than the PREM model coupled with an unreasonable attenuation model.

At station RSCP (82°), an SLHO type model provides unequivocally a better fit to the waveform than does PREM (Fig. 9). Again some slight adjustments in the size and depth of the D'' discontinuity are all that is necessary to obtain a good match.

A troubling feature remains in the waveform comparisons shown in Fig. 8: the observed ScS waveform at RSNY also appears to be broadened relative to the prediction of either the SLHO or PREM models. The broadening of the ScS phase might be dismissed as a consequence of its low signal-to-noise ratio. Similar broadening, however, is observed not only in the direct S and ScS at this station, but also at other RSTN stations at closer ranges. In these ranges, the pulse broadening cannot be easily explained by either the interference of triplicated branches or by an attenuation operator. Two possibilities that remain to be investigated are: (1) multi-pathing due to lateral rather than vertical heterogeneity and (2) shear-wave splitting due to anisotropy. The Discussion section outlines these possibilities in greater detail.

Waveform stability across an array

In other tests of the SLHO model, the stability and coherence of $S+ScS$ waveforms were observed across the Grae-

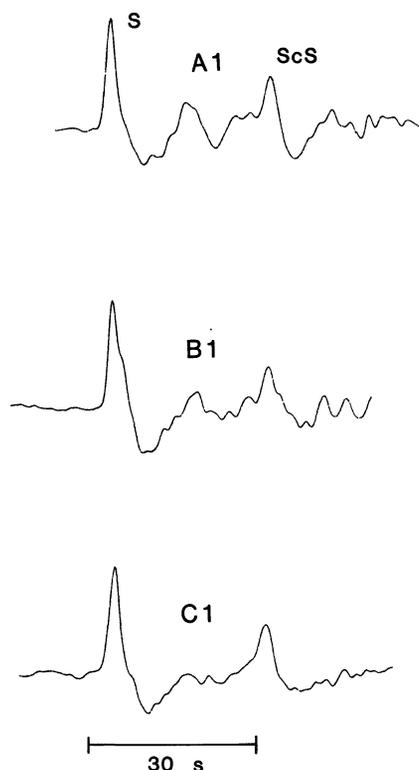


Fig. 10. Transverse $S+ScS$ waveforms observed at stations A1, B1 and C1 of the Graefenberg array from a deep focus earthquake beneath the Sea of Okhotsk (April 23, 1984; 21 40 35.5; 47.45N, 146.69E; 414 km deep). The data have been processed to simulate the response of a Kirnos broadband seismograph. The pulse between S and ScS is largest at A1, smaller at B1 and unidentifiable at C1

fenberg array from another deep focus event beneath the Sea of Okhotsk. The centre elements, A1, B1 and C1, of three sub-arrays are approximately 50 km apart and the epicentral distance to any of the stations differs by no more than several tenths of a degree from 75°. Note that among the variations in the waveforms (Fig. 10) the phase between S and ScS is quite large on A1, smaller on B1 and practically absent on C1. The waveforms at B1 and C1 are generally more complex with a higher coda level than A1, suggesting that the crust may introduce more scattering and multi-pathing along the paths to A1 and B1 than the path to C1. In this example, the identification of the phase between S and ScS as a branch from a lower mantle triplication would be somewhat suspect unless one could demonstrate that the crustal effects are either unimportant along the paths to A1 and B1 or that they operate in such a way as to selectively defocus the intermediate phase along the path to C1.

Discussion

The scale lengths of heterogeneities and S polarization

The preceding examples show that SKS generally does not substantially contaminate the transverse component of long-period records. In the short-period band, however, Murtha (1985) has found many examples of SKS on the transverse component. The variation of SKS contamination with frequency suggests a causal relation with the scale lengths and

intensities of velocity heterogeneities of the earth. The possible scale lengths and intensities responsible for any *SKS* contamination are limited. This is because some distributions of heterogeneity are too intense to be consistent with the agreement in the long-period band of *S*-wave polarizations with focal mechanism solutions. Ray tracing experiments with *S* waves indicate that in order to strongly affect *S* polarization, heterogeneity in D'' must be significantly stronger than the 1% velocity fluctuations determined from inversions of large catalogues of *P* travel times. The ray tracing equations for *S*-wave trajectory and polarization have been integrated (Cormier, 1984) through both Dziewonski's heterogeneous model and models of descending slabs. These tests indicate that a large degree (greater than 10% velocity fluctuations) of coherent heterogeneity in D'' having long scale lengths (1000 km and greater) are required to produce observations of *SKS* on the transverse component. The only distribution consistent with any significant observations of *SKS* on the transverse component is a concentration of the heterogeneity in D'' . Such a distribution would not affect *S* waveforms and polarizations measured in the long-period band at distances less than the *SKS* crossover (82°).

Ray tracing of the directly transmitted *S* wave, however, neglects the effects of multiple scattering and mode conversions by smaller scale heterogeneities. In the crust and lithosphere (Aki, 1980) and in the D'' region (Haddon and Cleary, 1974), the characteristic scale lengths of these smaller scale heterogeneities are estimated to be of the order of 10 km and the associated velocity fluctuations are estimated to be 10% or less. In order to explain the degree of *SKS* contamination observed on short-period records (Murtha, 1985) one must appeal either to the effects of scattering, in which multiple *SH* to *P* and *SV* mode conversions occur at sharp impedance contrasts, or to the effects of anisotropy.

General anisotropy from D'' heterogeneities

Any consideration of the distribution of heterogeneities in D'' should also include the constraints provided by analyses of *PKP-DF* precursors. The behaviour of the complex, short-period bundle of energy, which arrives before the *PKIKP* (*PKP-DF*) phase in the distance range (125° – 143°) has been explained by seismic scattering of the *PKP-AB* and *-BC* phases by structural heterogeneity in D'' or by bumps on the core-mantle boundary (e.g. Haddon and Cleary, 1974; Husebye et al., 1976; Doornbos, 1978). This theory accounts for the detailed behaviour of the travel times, energy and frequency content and the time-distribution of apparent azimuth and slowness of the precursor wavetrain. The studies of the *PKP-DF* precursors have established the statistical properties of heterogeneities in D'' or bumps on the core-mantle boundary required to generate the amount of precursor energy observed. For heterogeneity in D'' , these are a 2%–3% velocity fluctuation having scale lengths of several tens of kilometres. For bumps, these are heights of several hundreds of metres and having similar scale lengths. Significantly, an analysis and review by Haddon (1982) has demonstrated that more than one concentration of scale length of heterogeneity is needed to explain some of the precursor properties. Haddon finds that one type of precursor has an instantaneous arrival direction that systematically migrates with time along the precursor

wavetrain. He suggests that this implies distributions of heterogeneities that are strongly directional in character, distributed along bands or lines having scale lengths up to 500 km.

Given the existence of this type of distribution of heterogeneities, it may be important to consider whether it may exhibit long-wavelength anisotropy. This is the anisotropy when a body wave averages over the elastic properties of a heterogeneous medium (Backus, 1962). Some examples discussed by Crampin (1984) include "periodic thin layering" and "checkerboard anisotropy". The directionality in the distribution of heterogeneities in D'' inferred from short-period *PKP-DF* precursors may effectively appear to be an anisotropic layer to long-period *S* waves.

This anisotropy may not only account for the apparent difference in arrival times on radial and transverse components observed for some *ScS* waves sampling D'' (Mitchell and Helmberger, 1973; Lay and Helmberger, 1983a; Fukao, 1984), but it may also account for apparent polarization and particle motion anomalies and fluctuations in amplitude and pulse width. The mechanism that produces fluctuations in amplitude and pulse width is described by Crampin (1981). Upon entering an anisotropic region, an *S* wave is split into two pulses. This splitting or birefringence is described by two fixed, orthogonal polarizations. These polarizations are determined from an eigenvector problem, whose solution depends on the elastic constants of the anisotropic region. For a region possessing general anisotropy, the fixed polarizations that travel at different velocities are not, in general, parallel to either the *SH* or *SV* components of particle motion. Thus both *SH* and *SV* waves will split in the anisotropic region and will be, thereafter, preserved along the remainder of the ray path. For weak anisotropy, in which the split pulses cannot be resolved by the instrument response, the split pulses either destructively interfere or can appear as a broader single pulse of diminished amplitude. The broadening may be mistakenly interpreted as the effect of a zone of increased intrinsic attenuation. Lateral variations in shear-wave splitting may then be an alternative mechanism that explains the low *ScSH/SH* ratios of Mitchell and Helmberger (1973), which they model with the effects of a thin-high-velocity layer at the core-mantle boundary, and the variations in the *SKS/SKKS* ratios of Kind and Müller (1977) and Schweitzer (1984).

Bolt and Niazi (1984) found a suggestion of azimuthal anisotropy in the slowness of diffracted *S* waves. Such observations, together with the differing arrival times of radial and transverse *ScS* reported in several studies, indicate that the possibility of general anisotropy in the D'' should not be dismissed until further study. The existence of general anisotropy can be diagnosed from the behaviour of the particle motion of *S* waves. In the example investigated in this paper (Fig. 10), the particle motion appeared to be nearly linear, having little evidence of either the elliptic or cruciform pattern characteristic of general anisotropy. Before any firm conclusions can be reached, however, more particle motions should be examined and compared with those expected for a radially homogeneous model of D'' .

Another important model of D'' to investigate is one having transverse isotropy, in which *SH* and *SV* velocities are separate functions of ray parameter. This is a less radical generalization than a generally anisotropic model. It may be used to test the idea of Stacey and Loper (1983) of

return convective flow being dominantly concentrated in a thin (order of 10 km) radial shell in the lowermost mantle.

Conclusions

Inversions of the travel times of P waves sampling D'' and studies of the detailed behaviour of short period precursors to $PKP-DF$ indicate that the heterogeneous velocity structure of D'' is characterized by a broad spectrum of scale lengths, from several tens of kilometres to 1000 km. Precursor studies suggest that the distribution of the longer scale lengths are highly directional. One possibility, suggested by Haddon (1982), is that the heterogeneities may be either bands of corrugations on the core-mantle boundary or convective rolls having a short 10–20 km scale length and a longer 500–1000 km scale length concentrated in particular directions. The directionality of the longer scale lengths may reflect patterns of horizontal flow concentrated at the base of a lower mantle convection cell. Whatever the exact cause of this directionality, its existence suggests that it may be worthwhile to consider the possibility of long-wavelength anisotropy for S waves traversing D'' .

Smaller scale heterogeneity (10 km or smaller), distributed throughout the earth, may account for the anomalies observed in S and SKS polarization on short-period instruments. These anomalies are most likely generated by multiple scattering and mode conversions rather than by changes in ray path or polarization vector of the directly transmitted waves. Longer scale heterogeneity (1000 km or greater), with velocity perturbations of the order of several per cent, cannot produce large anomalies in S polarization. Such anomalies may be able to focus or defocus S and SKS phases traversing the deep mantle, but the per cent velocity perturbations required to produce the observed amplitude fluctuations in the long-period band are as yet unknown.

Without at least perturbing a radially symmetric model with lateral heterogeneity at the longer scale lengths of several 1000 km, a model of D'' has yet to be found that can satisfy all of the body wave data. Lay and Helmberger (1983a, b), for example, have found that data require some heterogeneity in the depth of their D'' discontinuity. Based on S waveform data at longer distance ranges and the decay constant of diffracted P waves, Schlittenhardt et al. (1985) have shown that if Lay and Helmberger's velocity discontinuity exists, it is unlikely to be a global feature. In order to be consistent with a global data set of waveforms, travel times and slownesses, the properties of this discontinuity can be summarized as follows: (a) if such a discontinuity exists, it varies laterally in depth as well as in the size of its jump; (b) the scale lengths of these variations may be as short as 500–1000 km; (c) the per cent jump in P velocity must be one-half or less than in S velocity; (d) the relative size of the P versus S velocity jump suggests a thermal effect, but its sharpness suggests a compositional change.

Although differences in radially symmetric models may be reconciled by appealing to heterogeneity at scale lengths of several 1000 km, such an appeal may neglect potentially large effects of smaller scale heterogeneities. The effects of these smaller scale heterogeneities on long-period body waves have not yet been calculated, nor have data yet been analysed in ways that would highlight their potential effects on long-period S waves. A key test would be to examine the three-component particle motion of S waves traversing D'' on broad-band, three-component, digitally recorded

seismic stations. These data can then be compared with synthetic particle motion calculated in both spherically symmetric models and heterogeneous/anisotropic models. Another important test of the effects of velocity heterogeneity at the smaller scale lengths would be to see how well a model of intrinsic attenuation together with strong constraints on the source-time function can account for the pulse broadening and rise times of S , SKS and ScS phases traversing the deep mantle. If such a model cannot successfully match the observed waveforms, then it may be possible that the heterogeneities of the mantle are introducing a stochastic dispersion of the type that has been found in numerical modelling of wave propagation in randomly heterogeneous media (e.g. Richards and Menke, 1983; Frankel and Clayton, 1984). The pulse broadening and waveform complexity due to this effect may mimic the type due to closely spaced triplications in a travel-time curve.

Finally, the stability and coherence of the waveform features used to infer any complex structure of D'' should be observed across densely spaced arrays. These experiments can be used to check whether a waveform feature is truly due to deep structure rather than local structure beneath a receiver.

Acknowledgements. Critical reviews by Thorne Lay, Brian Mitchell, Gerhard Müller, Ray Haddon, Paul Richards and J. Schweitzer were valuable in developing the ideas of this study. I also thank P. Murtha and J. Schweitzer for copies of their theses and J. Schlittenhardt and T. Lay for copies of their papers in preparation and in press. RSTN and Graefenberg data were obtained with the assistance of George Choy and Rainer Kind. J. Bullitt assisted with data processing and preparation of figures. This research was supported by the National Science Foundation under grant EAR-82-18737.

References

- Aki, K.: Scattering and attenuation of shear waves in the lithosphere. *J. Geophys. Res.* **85**, 6496–6504, 1980
- Anderson, D.L., Hanks, T.: Formation of the Earth's core. *Nature* **237**, 387–388, 1972
- Backus, G.E.: Long wave elastic anisotropy produced by horizontal layers. *J. Geophys. Res.* **67**, 4427–4440, 1962
- Bolt, B.A., Niazi, M.: S velocities in D'' from diffracted SH-waves at the core boundary. *Geophys. J. R. Astron. Soc.* **79**, 1984
- Bullen, K.E.: An Earth model based on a compressibility-pressure hypothesis. *Mon. Not. R. Astron. Soc. Geophys. Suppl.* **6**, 50–59, 1950
- Burdick, L.J.: t^* for S waves with a continental ray path. *Bull. Seismol. Soc. Am.* **68**, 1013–1030, 1978
- Choy, G.L.: Theoretical seismograms of core phases calculated by a frequency-dependent full wave theory, and their interpretation. *Geophys. J. R. Astron. Soc.* **51**, 275–312, 1977
- Cleary, J.R.: The D'' region. *Phys. Earth Planet. Int.* **9**, 13–27, 1974
- Comer, R.P., Clayton, R.W.: Tomographic reconstruction of lateral velocity heterogeneity in the earth's mantle. *EOS Trans. Am. Geophys. Un.* **65**, 236, 1984
- Cormier, V.F.: The polarization of S waves in a heterogeneous isotropic whole Earth model. *J. Geophys.* **56**, 20–23, 1984
- Cormier, V.F., Richards, P.G.: Full wave theory applied to a discontinuous velocity increase: the inner core boundary. *J. Geophys.* **43**, 3–31, 1977
- Crampin, S.: A review of wave motion in anisotropic and cracked elastic-media. *Wave Motion* **3**, 343–391, 1981
- Crampin, S.: Seismic anisotropy – the state of the art: II. *Geophys. J. R. Astron. Soc.* **76**, 1–16, 1984
- Der, Z.A., McElfresh, T.W., O'Donnell, A.: An investigation of

- the regional variations and frequency dependence of anelastic attenuation in the United States in the 0.5–4 Hz. band. *Geophys. J. R. Astron. Soc.* **69**, 67–100, 1982
- Doornbos, D.J.: On seismic-wave scattering by a rough core-mantle boundary. *Geophys. J. R. Astron. Soc.* **53**, 643–662, 1978
- Doornbos, D.J., Mondt, J.D.: P and S waves diffracted around the core and the velocity structure at the base of the mantle. *Geophys. J. R. Astron. Soc.* **57**, 381–395, 1979
- Dziewonski, A.M.: Mapping the lower mantle: determination of lateral heterogeneity up to degree 6. *J. Geophys. Res.* **89**, 5929–5952, 1984
- Dziewonski, A.M., Anderson, D.L.: Preliminary reference Earth model. *Phys. Earth Planet. Int.* **25**, 297–356, 1981
- Dziewonski, A.M., Hales, A.L., Lapwood, E.R.: Parametrically simple Earth models consistent with geophysical data. *Phys. Earth Planet. Int.* **10**, 12–48, 1975
- Frankel, A., Clayton, R.W.: A finite difference simulation of wave propagation in two-dimensional media. *Bull. Seismol. Soc. Am.* **74**, 2167–2186, 1984
- Fukao, Y.: Evidence from core-reflected shear waves for anisotropy in Earth's mantle. *Nature* **309**, 695–698, 1984
- Haddon, R.A.W.: Evidence for inhomogeneities near the core-mantle boundary. *Phil. Trans. Roy. Soc. Lond.* **A-306**, 61–70, 1982
- Haddon, R.A.W., Cleary, J.R.: Evidence for scattering of seismic PKP waves near the mantle-core boundary. *Phys. Earth Planet. Int.* **8**, 211–234, 1974
- Hales, A.L., Roberts, J.: The travel times of S and SKS. *Bull. Seismol. Soc. Am.* **60**, 461–489, 1970
- Harvey, D., Choy, G.L.: Broadband deconvolution of GDSN data. *Geophys. J. R. Astron. Soc.* **69**, 659–668, 1982
- Husebye, E.S. King, D.W., Haddon, R.A.W.: Precursors to PKIKP and seismic wave scattering near the mantle-core boundary. *J. Geophys. Res.* **81**, 1870–1882, 1976
- Jeanloz, R., Richter, F.M.: Convection, composition, and the thermal state of the lower mantle. *J. Geophys. Res.* **84**, 5497–5504, 1979
- Jones, G.M.: Thermal interaction of the core and the mantle and long-term behavior of the geomagnetic field. *J. Geophys. Res.* **82**, 1703–1709, 1977
- Kind, R., Müller, G.: The structure of the outer core from SKS amplitudes and travel times. *Bull. Seismol. Soc. Am.* **67**, 1541–1554, 1977
- Lay, T., Helmberger, D.V.: A lower mantle S-wave triplication and the shear velocity structure of D''. *Geophys. J. R. Astron. Soc.* **75**, 799–838, 1983a
- Lay, T., Helmberger, D.V.: The shear-wave velocity gradient at the base of the mantle. *J. Geophys. Res.* **88**, 8160–8170, 1983b
- Mitchell, B.J., Helmberger, D.V.: Shear velocities at the base of the mantle from observations of S and ScS. *J. Geophys. Res.* **78**, 6009–6020, 1973
- Mula, A.H.: Amplitudes of diffracted long-period P and S waves and the velocities and Q structure at the base of the mantle. *J. Geophys. Res.* **86**, 4999–5011, 1981
- Mula, A.H., Müller, G.: Ray parameters of diffracted long period P and S waves and the velocities at the base of the mantle. *Pure Appl. Geophys.* **118**, 1270–1290, 1980
- Murtha, P.E.: Seismic velocities in the upper part of the Earth's core. Ph.D. thesis, University of California, Berkeley, 1985
- Richards, P.G., Menke, W.: The apparent attenuation of a scattering medium. *Bull. Seismol. Soc. Am.* **61**, 1675–1692, 1983
- Ruff, L.J., Anderson, D.L.: Core formation, evolution and convection: a geophysical model. *Phys. Earth Planet. Int.* **21**, 181–201, 1980
- Ruff, L.J., Helmberger, D.V.: The structure of the lowermost mantle determined by short-period P-Wave amplitudes. *Geophys. J. R. Astron. Soc.* **68**, 95–120, 1982
- Schlittenhardt, J., Schweitzer, J., Müller, G.: Evidence against a discontinuity at the top of D''. *Geophys. J. R. Astron. Soc.* 1985 (in press)
- Schweitzer, J.: Laufzeiten und Amplituden der Phasen SKS und SKKS und die Struktur des äußeren Erdkerns, Diplomarbeit, Institut für Meteorologie und Geophysik der Johann Wolfgang Goethe Universität, Frankfurt a.M., 1984
- Stacey, F.D., Loper, D.E.: The thermal boundary layer interpretation of D'' and its role as a plume source. *Phys. Earth Planet. Int.* **33**, 44–55, 1983
- Strelitz, R.: The September 5, 1970 Sea of Okhotsk earthquake: a multiple event with evidence of triggering. *Geophys. Res. Lett.* **2**, 124–127, 1975
- Wright, C., Lyons, J.A.: Further evidence for radial velocity anomalies in the lower mantle. *Pageoph.* **119**, 137–162, 1984
- Wright, C., Muirhead, K.J., Dixon, A.E.: The P wave velocity structure near the base of the mantle. *J. Geophys. Res.* **90**, 623–634, 1985

Received November 8, 1984; Revised version February 18, 1985
Accepted February 25, 1985

Transient *SH* waves in dipping layers: the buried line-source problem

F. Ziegler¹, Y.-H. Pao² and Y.S. Wang^{2*}

¹ Institut für Allgemeine Mechanik, Technische Universität Wien, Vienna, Austria, A-1040

² Department of Theoretical and Applied Mechanics, Cornell University, Ithaca, New York 14853, USA

Abstract. The theory of generalized rays is applied to analysing transient waves in a layered half-space with non-parallel interfaces. The propagation, transmission, reflection and refraction of *SH* waves generated by a line source which is buried in the underlying half-space (bedrock) of a three-layer model is considered, each of the two overlying layers having a different dip angle.

Generalized ray integrals for multiply refracted rays are formulated by using three rotated coordinate systems, one for each interface. Through a series of transformations of the local slowness and the application of Snell's law, all ray integrals are expressible in a common slowness variable. The arrival time of each ray undergoing multiple transmissions and reflections before reaching an observation point in the top layer is then determined from the stationary value of the phase function with common slowness of the ray integral. Early arrivals of head waves which follow a ray path refracted at a fast bottom are calculated from proper branch points of the Cagniard mapping. Inverse Laplace transformation of these ray integrals is then completed by Cagniard's method.

Key words: Cylindrical waves – Ray integrals – Ray sorting – Arrival times – Head waves – Divergence effect – Synthetic seismograms

Introduction

Recently the theory of generalized ray integrals, originally developed for layered media with parallel surfaces (see e.g. Müller, 1968a, b, c; Wiggins and Helmberger, 1974; Pao and Gajewski, 1977; Kennett, 1980; and, for a comparison with other theories, Aki and Richards, 1980), was extended to account for a single dipping layer by Ishii and Ellis (1970a, b) and Hong and Helmberger (1977). They showed that the exact solution of *SH* waves in a wedge (Hudson, 1963) can be expanded into a series of integrals. One of them represents the radiation from the apex of the wedge and the others can be identified as generalized ray integrals, representing cylindrical waves which are multiply reflected between two plane nonparallel surfaces. These

ray integrals are then evaluated by applying the Cagniard-de Hoop method and by using a first motion approximation.

Pao and Ziegler (1982) have shown that by expressing the *SH*-source ray integral in two systems of rotated coordinates, one for each of the two nonparallel surfaces, one can construct successively the ray integrals for multi-reflected waves within the wedge. A comparison with the method of images has been included as far as direct rays are concerned. Their approach was recently generalized to a multilayered medium, Ziegler and Pao (1985), when the line source and the receiver are located in the surface layer.

A report on related site effects has been given by Porceski (1969) and model seismic experiments are described in Drimmel et al. (1973) and in the dissertation by Wiedmann (1983).

We consider a three-layered medium, each of the overlying layers having a different dip angle and non-common apexes. A line source which generates a transient *SH* wave is placed beneath the two top layers. The observational station can be anywhere in the medium, but our interest is confined mainly to that in the surface layer or on the top surface.

Waves emitted by the line source which are incident at the lowest interface (3), see Fig. 1, are represented by $\phi_{\text{inc}}(\mathbf{x}, t)$, the transmitted waves in layer 2 will be denoted $\psi(\mathbf{x}, t)$ and those further transmitted to the top layer 1 are denoted $\phi(\mathbf{x}, t)$. For *SH* waves in solids, ϕ or ψ is the displacement component in the direction of the line source. In a fluid medium, ϕ or ψ is the wave potential, the gradient of which is the velocity. The Laplace-transformed wave potentials are denoted by $\bar{\phi}(\mathbf{x}, s)$ and $\bar{\psi}(\mathbf{x}, s)$. They are solutions to the homogeneous reduced wave equations, except $\bar{\phi}_{\text{inc}}(\mathbf{x}, s)$ which satisfies the inhomogeneous Helmholtz equation

$$\nabla^2 \bar{\phi}_{\text{inc}} - b^2 s^2 \bar{\phi}_{\text{inc}} = -b^2 \bar{f}(s) \delta(x) \delta(z - z_0),$$

where $\nabla^2 = \partial^2/\partial x^2 + \partial^2/\partial z^2$, and $b = c_3^{-1}$ is the slowness (reciprocal of wave speed) of the underlying half-space and $\bar{f}(s)$ the Laplace transform of the time function $f(t)$ of the line source which is located at $x=0, z=z_0$.

In the following section we construct the generalized ray integrals of the waves transmitted through interface (3) into the intermediate layer 2 from the Weyl-Sommerfeld integral representation of the incident

* On leave from Hunan University, China

wave $\bar{\phi}_{\text{inc}}(\mathbf{x}, s)$. These waves are multiply reflected within layer 2 as well as refracted back to the underlying half-space and partially transmitted to the surface layer 1. Multiple reflections of waves transmitted to the top layer are considered in a subsequent section, followed by a discussion of higher order wave trains with more than two transmissions and propagating towards the receiver station in the surface layer.

Application of Cagniard's method (1962) to the inverse Laplace transformation of these ray integrals is then briefly discussed. Special attention is given to the calculation of the stationary point of the phase function and the branch points of the common slowness variable for all rays. The former yields the arrival time of direct rays and the latter is connected with head waves of refracted rays. Receiver stations are assumed at down-dip and up-dip locations.

Synthetic seismograms at those observational points located at the free surface are presented for the simplified case of one dipping top layer and a triangular source time function of a single pulse, thus the apparent source ray is transmitted only once through the interface before reaching the receiver. A comparative study is made to the signals arriving at the surface of a half-space with a parallel top layer of constant thickness equal to that of the dipping layer at source location. The final section contains the conclusions.

Source ray and transmission into the intermediate layer 2

The configuration of the three-layer half-space is shown in Fig. 1, together with three coordinate systems rotated about the common epicentral origin 0. Interface (2) is inclined against the free surface (1) by angle α and interface (3) is inclined against (2) by angle β and, therefore, by $\gamma = (\alpha + \beta)$ against the top surface.

Material of the surface layer has shear modulus μ , slowness $a = c^{-1}$ and vertical thickness h measured along the z -axis through source point S . Material of layer 2 or 3 has shear modulus μ_j ($j=2, 3$) and slowness $a_2 = c_2^{-1}$, $b = a_3 = c_3^{-1}$. The vertical thickness along the z -axis of the intermediate layer 2 is h_2 .

The axes parallel to the interfaces (1), (2) and (3) are denoted x , x' and x'' , respectively. A line source S is located in the underlying half-space 3 at depth $z_0 > H$, measured from the free surface (1), and $H = h + h_2$. The final position of the observation point (x, z) will be in the top layer and a ray path from the source to the receiver passes through all three layers.

The equation of each plane surface is:

$$\begin{aligned} \text{Free surface (1):} \quad & z = z_1 = 0 \\ \text{Interface (2)} \quad & z' = z'_2 = h \cos \alpha = x_0 \sin \alpha \\ \text{Interface (3):} \quad & z'' = z''_3 = H \cos \gamma, \quad H = h + h_2, \quad \gamma = \alpha + \beta. \end{aligned} \quad (1)$$

We shall use the tip-distance measured along interface (2)

$$d' = (h/\sin \alpha) - h_2 \cos \gamma / \sin \beta, \quad (2)$$

see Fig. 1, and

$$x''_0 = -h \sin \gamma + h_2 \cos \gamma \cos \beta / \sin \beta, \quad (3)$$

to be convenient geometric parameters.

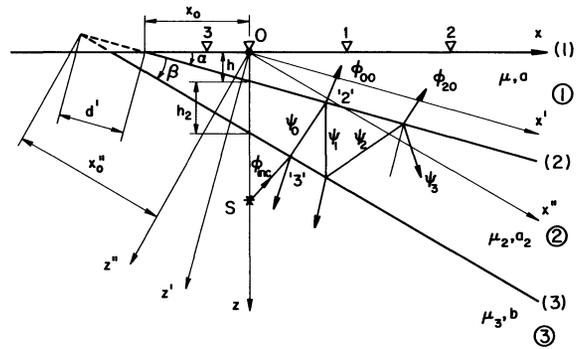


Fig. 1. Geometry of a half-space with two dipping layers. Source S located in the bedrock 3 at $x=0$ and $z=z_0 > H$, $H = h + h_2$. Also shown are: the source ray ϕ_{inc} , a direct ray path ψ_j , $j=3$, in the intermediate layer 2 and two apparent source rays in the surface layer, ϕ_{00} and ϕ_{20} . $\nabla \dots$ Observation stations on top surface, 0 ... epicentre, 1 and 2 down-dip locations, 3 up-dip location

The angles α and β are positive when rotated clockwise, negative when rotated counter-clockwise. At $z = z_1 = 0$, the shear stress (or pressure for a fluid medium) vanishes. At both interfaces, the stress and displacement (or velocity) are continuous and Hooke's law applies to all three layers.

The source ray

The Weyl-Sommerfeld integral representation of the Laplace-transformed displacement for the waves radiated by the line source into infinite homogeneous and isotropic space is (Aki and Richards, 1980, Chapter 6)

$$\bar{\phi}_{\text{inc}}(s) = \bar{F}(s) \int_{-\infty}^{\infty} S(\xi'') e^{sg(\xi'')} d\xi'' \quad (4)$$

where

$$\begin{aligned} \bar{F}(s) &= b^2 \bar{f}(s) / 4\pi, \\ S(\xi'') &= 1/\chi, \quad \chi = (b^2 + \xi''^2)^{\frac{1}{2}} \end{aligned} \quad (5)$$

and the phase function becomes, noting the source location in doubly primed coordinates,

$$g(\xi'') = i\xi''(x'' - z_0 \sin \gamma) - \chi|z'' - z_0 \cos \gamma|. \quad (6)$$

We call $\bar{\phi}_{\text{inc}}$ the source ray and $\bar{F}(s)S(\xi'')$ the source function. The variable of integration, ξ'' , is the apparent slowness of waves in the x'' -direction, χ that along the z'' -axis. The phase function shows the projections of the source ray to a receiver station at the point (x'', z'') in bedrock in the x'' - and z'' -directions.

The source ray, when incident at interface (3), is subject to partial reflection and transmission into layer 2. The transmitted ray which may be called an apparent source ray in the intermediate layer 2 is given by the generalized ray integral

$$\bar{\psi}_0(s) = \bar{F}(s) \int_{-\infty}^{\infty} S_0(\xi'') e^{sh_0(\xi'')} d\xi'', \quad (7)$$

where we changed the source function $S(\xi'')$ of the incident source ray to the apparent source function of

the transmitted ray

$$S_0(\xi'') = S(\xi'') T^{(3)}(\xi''), \quad (8)$$

by multiplying with the transmission coefficient of plane SH waves,

$$T^{(3)} = 2\mu_3\chi/(\mu_3\chi + \mu_2\zeta_0''), \quad \zeta_0'' = (a^2 + \xi''^2)^{1/2}. \quad (9)$$

Also, the phase function is changed according to the phase contribution of the incident source ray, $g(\xi'', x_3'', z_3'')$, to a point of refraction at interface (3), say (x_3'', z_3'') , and the phase of the emanating ray to an observation point (x'', z'') within layer 2, $z'' \leq z_3''$, Fig. 1,

$$h_0(\xi'') = g(\xi'', x_3'', z_3'') + i\xi''(x'' - x_3'') + \zeta_0''(z'' - z_3''). \quad (10)$$

Since Snell's law applies to transmission, the apparent slowness ξ'' is common and the unknown coordinate x_3'' cancels. Hence,

$$h_0(\xi'') = -i\xi''z_0 \sin \gamma - \chi(z_0 - H) \cos \gamma - \zeta_0''z_3'' + i\xi''x'' + \zeta_0''z''. \quad (11)$$

The apparent source ray $\bar{\psi}_0$ is incident at interface (2) and subject to reflection and to transmission into the top layer. The refraction should be formulated in primed coordinates. Therefore, we transform the phase function $h_0(\xi'')$ to primed coordinates where, in matrix notation,

$$\mathbf{x}'' = \mathbf{D}_2 \mathbf{x}', \quad \mathbf{D}_2 = \begin{pmatrix} \cos \beta & \sin \beta \\ -\sin \beta & \cos \beta \end{pmatrix}, \quad (12)$$

$$\mathbf{x}'^T = (x', z'), \quad \mathbf{x}''^T = (x'', z'')$$

is to be substituted and require the phase of the ray to be invariant under coordinate rotation through the angle β :

$$h(\zeta_0') = h(\xi'') = -i\xi''z_0 \sin \gamma - \chi(z_0 - H) \cos \gamma - \zeta_0''z_3'' + i\xi''x'' + \zeta_0''z'', \quad (13)$$

where the pair of apparent slownesses of the ray in primed coordinates is derived from the doubly primed pair of apparent slownesses in the doubly primed coordinates by the transformation

$$\xi_0'^T = \xi_0''^T \mathbf{D}_2, \quad \xi_0''^T = (i\xi'', \zeta_0''), \quad \xi_0'^T = (i\xi_0', \zeta_0'), \quad (14)$$

where ζ_0'' is the irreducible radical of Eq. (9).

Hence, the transformed phase function $h(\xi'')$ still contains the same number of two radicals as $h(\xi'')$, namely χ and ζ_0'' , which are apparent slownesses of the transmitted ray path in the z'' -direction in the two media with common interface (3).

We consider multiple reflections within layer 2 first.

Reflections within the intermediate layer 2

The apparent source ray $\bar{\psi}_0$ is subject to partial reflection at interface (2) and the reflected cylindrical wave is given by the generalized ray integral

$$\bar{\psi}_1(s) = \bar{F}(s) \int_{-\infty}^{\infty} S_0(\xi'') R^{(2)}(\xi_1') e^{sh_1(\xi'')} d\xi'' \quad (15)$$

where, $l=1$ above,

$$R^{(2)}(\xi_1') = [-\mu\eta_l' + \mu_2\zeta_l'] / [\mu\eta_l' + \mu_2\zeta_l'], \quad (16)$$

is the reflection coefficient of plane waves at interface (2). Expressed in the local apparent slowness in the primed coordinates:

$$\eta_l' = (a^2 + \xi_l'^2)^{1/2} \quad (17)$$

is an irreducible radical and $\zeta_l' = \sqrt{(a^2 + \xi_l'^2)}$ may be reduced to the pair of slownesses ξ'' and ζ_0'' of the apparent source ray $\bar{\psi}_0$, cf. Eq. (21).

The phase function is changed according to the superposition, $z' > z_2'$,

$$h_1(\xi'') = h_1(\xi_1') = h(\xi', x_2', z_2') + i\xi_1'(x' - x_2') - \zeta_1'(z' - z_2'), \quad (18)$$

where the unknown coordinate x_2' of the point of reflection at interface (2) cancels, since Snell's law holds, see Fig. 1,

$$\xi_1' = \xi'. \quad (19)$$

Hence, the phase function of the once-transmitted and once-reflected ray becomes

$$h_1(\xi'') = -i\xi''z_0 \sin \gamma - \chi(z_0 - H) \cos \gamma - \zeta_0''z_3'' + z_2'(\zeta_0' + \zeta_1') + i\xi_1'x' - \zeta_1'z', \quad (20)$$

where, Eq. (14),

$$\xi_1'^T = \xi_0'^T = \xi_0''^T \mathbf{D}_2, \quad (21)$$

and (x', z') is an observational point in layer 2.

Requiring invariance of phase of this reflected ray under coordinate rotation through the angle β ,

$$i\xi_1'x' - \zeta_1'z' = i\xi_1''x'' - \zeta_1''z'', \quad (22)$$

the constant part remains unaffected, the phase function can be expressed in doubly primed coordinates, thereby preparing it for the next reflection to take place at interface (3):

$$h_1(\xi'') = -i\xi''z_0 \sin \gamma - \chi(z_0 - H) \cos \gamma - \zeta_0''z_3'' + z_2'(\zeta_0' + \zeta_1') + i\xi_1''x'' - \zeta_1''z'', \quad (23)$$

where

$$\xi_2''^T = \xi_1''^T = \xi_1'^T \mathbf{D}_2 = \xi_0''^T \mathbf{D}_2^2. \quad (24)$$

The transformation shows both the diverging effect on a ray pointing down-dip, $\xi'' > 0$, and the steepening effect during reflection at interface (2) for a ray pointing up-dip, $\xi'' < 0$.

Partial reflection at interface (3) renders the ray $\bar{\psi}_2$ in layer 2 and the generalized ray integral becomes

$$\bar{\psi}_2(s) = \bar{F}(s) \int_{-\infty}^{\infty} S_0(\xi'') R^{(2)}(\xi_1') R_{(3)}(\xi_2'') e^{sh_2(\xi'')} d\xi''. \quad (25)$$

The reflection coefficient at interface (3) expressed in local slowness ξ_j'' , $j=2$ above, is given by

$$R_{(3)}(\xi_j'') = (\mu_2\zeta_j'' - \mu_3\chi_j) / (\mu_2\zeta_j'' + \mu_3\chi_j), \quad (26)$$

where

$$\chi_j = (b^2 + \xi_j''^2)^{1/2} \quad (27)$$

denotes an irreducible radical and $\zeta_j'' = \sqrt{(a_2^2 + \xi_j''^2)}$ can be reduced together with ξ_j'' , cf. Eq. (24) for $j=1$ and see Eq. (35).

The phase function

$$h_2(\xi'') = h_1(\xi_1'', x_3'', z_3'') + i\zeta_2''(x'' - x_3'') + \zeta_2''(z'' - z_3'') \quad (28)$$

becomes, under Snell's condition, $\xi_2'' = \xi_1''$, and after the cancellation of the unknown coordinate x_3'' of the point of reflection at interface (3), see Fig. 2,

$$h_2(\xi'') = -i\xi'' z_0 \sin \gamma - \chi(z_0 - H) \cos \gamma - z_3''(\zeta_0 + \zeta_1 + \zeta_2)'' + z_2''(\zeta_0 + \zeta_1) + i\zeta_2'' x'' + \zeta_2'' z''. \quad (29)$$

A transformation to primed coordinates leaves the constant part unchanged and renders

$$i\zeta_2'' x'' + \zeta_2'' z'' = i\zeta_2' x' + \zeta_2' z', \quad (30)$$

where the pair of local apparent slowness in primed coordinates of the last ray segment is transformed according to

$$\xi_2'^T = \xi_2''^T \mathbf{D}_2 = \xi_0'^T \mathbf{D}_2^3. \quad (31)$$

Hence, the cylindrical wave with a ray path showing one transmission and j reflections within layer 2 is given by inference through the generalized ray integral

$$\bar{\psi}_j(s) = \bar{F}(s) \int_{-\infty}^{\infty} S_0(\xi'') \Pi_j e^{sh_j(\xi'')} d\xi''. \quad (32)$$

The product of reflection coefficients is expressed in local apparent slowness in the x' - and x'' -axis, respectively, and is given for an even-numbered ray which is subject to partial transmission into the top layer by

$$\Pi_j = R^{(2)}(\xi_1') R_{(3)}(\xi_2'') R^{(2)}(\xi_3') \dots R_{(3)}(\xi_j''). \quad (33)$$

The phase function expressed in terms of local slowness is

$$h_j(\xi'') = -i\xi'' z_0 \sin \gamma - \chi(z_0 - H) \cos \gamma - z_3'' \sum_{l=0}^j \zeta_l'' + z_2'' \sum_{l=0}^{j-1} \zeta_l' + i\zeta_j' x' + \zeta_j' z', \quad j(\text{zero, even}), \quad (34)$$

and the proper pairs of apparent slowness are related to the pair of slownesses of the apparent source ray $\bar{\psi}_0$ through

$$\xi_{l+1}'^T = \xi_l'^T = \xi_0'^T \mathbf{D}_2^{l+1}, \\ \xi_l''^T = \xi_{l-1}''^T = \xi_0''^T \mathbf{D}_2^l, \quad l(\text{even}). \quad (35)$$

Snell's law is indicated.

The first train of waves in the surface layer 1

The even-numbered rays $\bar{\psi}_j$ are incident at interface (2) and are partially transmitted into the top layer, thereby forming apparent source rays $\bar{\phi}_{j0}$. The first subscript indicates the even number of reflections with-

in layer 2 and the subscript zero identifies the ray to be just transmitted into the surface layer without further reverberations. See Fig. 1 for those ray paths with two transmission and $j=2$ reflections. The ray $\bar{\phi}_{00}$, therefore, denotes the direct ray path from the source to a receiver in layer 1 without any reflection. The generalized ray integral is given by

$$\bar{\phi}_{j0}(s) = \bar{F}(s) \int_{-\infty}^{\infty} S_{0j}(\xi'') e^{sg_{j0}(\xi'')} d\xi'', \quad j=0, 2, 4, \dots \quad (36)$$

The apparent source function becomes

$$S_{0j}(\xi'') = S_0(\xi'') \Pi_j T^{(2)}(\xi_j'').$$

when considering the transmission coefficient of plane waves through interface (2):

$$T^{(2)}(\xi_j'') = 2\mu_2 \zeta_j' / [\mu_2 \zeta_j' + \mu \eta_{j0}'], \quad (37)$$

where

$$\eta_{j0}' = (a^2 + \xi_j'^2)^{1/2} \quad (38)$$

denotes an irreducible radical.

Taking into account that a receiver in the top layer requires $z' < z_2''$, the phase function is expressed in primed receiver coordinates by

$$g'_{j0}(\xi'') = h_j(\xi_j', x_2', z_2') + i\xi_j'(x' - x_2') + \eta_{j0}'(z' - z_2') \quad (39)$$

and, since the unknown coordinate of transmission x_2'' cancels, we have in (x, z) -coordinates of the observation point in the top layer,

$$g'_{j0}(\xi'') = -i\xi'' z_0 \sin \gamma - \chi(z_0 - H) \cos \gamma - z_3'' \sum_{l=0}^j \zeta_l'' + z_2'' \left(\sum_{l=0}^j \zeta_l' - \eta_{j0}' \right) + i\zeta_{j0}' x + \eta_{j0}' z, \quad (40)$$

where the proper pair of slownesses of the apparent source ray in unprimed coordinates is changed according to invariance of phase under coordinate rotation through the angle α

$$i\zeta_j' x' + \eta_{j0}' z' = i\zeta_j x + \eta_{j0} z. \quad (41)$$

Hence, in analogy to Eqs. (14) and (12)

$$\xi_{j0}'^T = \xi_{j0}''^T \mathbf{D}, \quad \mathbf{D} = \begin{pmatrix} \cos \alpha & \sin \alpha \\ -\sin \alpha & \cos \alpha \end{pmatrix}, \quad (42) \\ \mathbf{x}' = \mathbf{D} \mathbf{x}, \quad \xi_{j0}'^T = (i\zeta_j', \eta_{j0}'), \quad \xi_{j0}''^T = (i\zeta_j, \eta_{j0}).$$

In the following subsection we study the generalized ray integral of a ray $\bar{\phi}_{jq}$ undergoing q successive reflections within the top layer.

The ray $\bar{\phi}_{jq}$ of the first wave train

According to the changes of the generalized ray integrals of the apparent source ray $\bar{\psi}_0$ to the ray with $(j+1)$ reverberations $\bar{\psi}_j$, we substitute the product of reflection coefficients

$$\Pi_{jq} = R^{(1)}(\xi_{j1}') R_{(2)}(\xi_{j2}'') R^{(1)}(\xi_{j3}') \\ \dots \begin{cases} R^{(1)}(\xi_{jq}'), & q(\text{odd}) \\ R_{(2)}(\xi_{jq}''), & q(\text{even}) \end{cases} \quad (43)$$

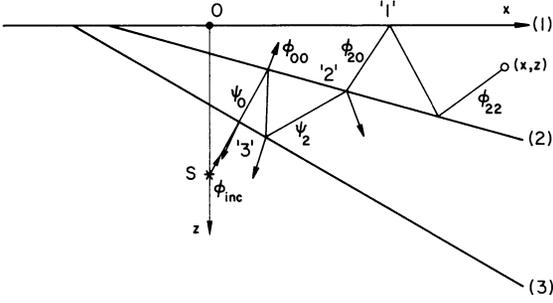


Fig. 2. Direct ray path from source S to observation station (x, z) , ϕ_{jq} of first train of waves in surface layer 1. The ray shown is $j=2, q=2$

where $R_{(2)} = -R^{(2)}$, and change the phase of the apparent source ray $\bar{\phi}_{j0}$ to

$$g_{jq}(\xi'') = -i\xi''z_0 \sin \gamma - \chi(z_0 - H) \cos \gamma - z_3' \sum_{l=0}^j \zeta_l'' + z_2' \left[\sum_{l=0}^j \zeta_l' - \sum_{l=0}^{q^*} \eta_{jl}' \right] + i\xi_{jq}'' x + (-1)^q z \eta_{jq}, \quad (44)$$

where $q^* = q$ (even) and $q^* = q-1$ when q is odd and, hence, find the generalized ray integral

$$\bar{\phi}_{jq}(s) = \bar{F}(s) \int_{-\infty}^{\infty} S_{0j}(\xi'') \Pi_{jq} e^{s g_{jq}(\xi'')} d\xi'', \quad j=0, 2, 4, \dots \quad (45)$$

We note the transformations of proper slowness pairs within layer 1, see Fig. 2,

$$\xi_{jl}'^T = \xi_{j(l-1)}'^T = \xi_{j0}'^T \mathbf{D}^l, \quad l \text{ (even)} \quad (46)$$

$$\xi_{j(q+1)}'^T = \xi_{jq}'^T = \xi_{j0}'^T \mathbf{D}^{q+1}, \quad q \text{ (even)}. \quad (47)$$

Elimination of the local slowness in Eq. (44) and noting the finite sums

$$2 \sum_{l=1}^{j/2} \sin(2l-1)\beta = (1 - \cos j\beta) / \sin \beta$$

$$2 \sum_{l=1}^{j/2} \cos(2l-1)\beta = \sin j\beta / \sin \beta, \quad (48)$$

renders the final form of the phase function in a matrix multiplicative form

$$g_{jq}(x, z, \xi'') = -\xi''^T \mathbf{x}_S'' - \xi_0''^T \mathbf{D}_2^{j+1} \mathbf{T} d' + \xi_{j0}'^T \mathbf{D}^{q+1} \mathbf{x}, \quad (49)$$

where slowness vectors of apparent source rays are related to ξ'' , the integration variable,

$$\xi''^T = (i\xi'', \chi), \quad \xi_0''^T = (i\xi'', \zeta_0''), \quad \xi_{j0}'^T = (i\xi_j', \eta_{j0}'), \quad (50)$$

and

$$\chi = (b^2 + \xi''^2)^{1/2}, \quad \zeta_0'' = (a_2^2 + \xi''^2)^{1/2}, \quad \eta_{j0}' = (a^2 + \xi_j'^2)^{1/2},$$

$$i\xi_j' = \xi_0''^T \mathbf{D}_2^{j+1} \mathbf{T}, \quad \mathbf{T}^T = (1, 0). \quad (51)$$

The point vectors are conveniently defined by

$$\mathbf{x}_S''^T = [(x_0'' + z_0 \sin \gamma), (z_0 - H) \cos \gamma],$$

$$x_0'' = x_0 \cos \gamma - d' \cos \beta, \quad (52)$$

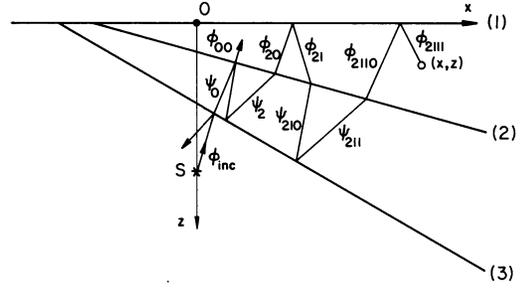


Fig. 3. Direct ray path from source S to observation station (x, z) , ϕ_{jqpn} of second train of waves in surface layer 1. The ray shown is $j=2, q=1, p=1, n=1$

d' of Eq. (2) and

$$\mathbf{x}^T = [(x + x_0), -\varepsilon z]. \quad (53)$$

The exponent q^* equals q (even) with direction factor $\varepsilon = -1$ and equals $(q-1)$ when q is odd, $\varepsilon = +1$.

The second train of waves in the top layer 1

In Fig. 3 we show the ray path of a ray $\bar{\phi}_{jqpn}$ with a last segment in layer 1 which is a result of partial transmission of the ray $\bar{\phi}_{jq}$, j (zero, even), q (odd) through interface (2) and, after $(p+1)$ reverberations within the intermediate layer 2, is further transmitted upwards through interface (2), p (odd). The last index n gives the number of successive reflections considered in the top layer before reaching the observation station (x, z) . Considering $\bar{\psi}_{jq0}$ as an apparent source ray in layer 2, pointing downward, see Fig. 3, and $\bar{\phi}_{jqp0}$ as the apparent source ray of the second train of waves in the surface layer 1, we find the generalized ray integral

$$\bar{\phi}_{jqpn}(s) = \bar{F}(s) \int_{-\infty}^{\infty} S_{0jqpn}(\xi'') \Pi_{jqpn} e^{s g_{jqpn}(\xi'')} d\xi'',$$

$$j=0, 2, 4, \dots, \quad q=1, 3, 5, \dots,$$

$$p=1, 3, 5, \dots, \quad (54)$$

where

$$S_{0jqpn}(\xi'') = S_0(\xi'') \Pi_j T^{(2)}(\xi_j') \Pi_{jq} T_{(2)}(\xi_{jq}') \Pi_{jqp} T^{(2)}(\xi_{jqp}') \quad (55)$$

denotes the apparent source function of the second train of waves in layer 1, and the transmission coefficient from the top layer to the intermediate layer 2 becomes

$$T_{(2)}(\xi_{jq}') = 2\mu\eta_{jq}' / [\mu\eta_{jq}' + \mu_2\zeta_{jq0}']. \quad (56)$$

$$\zeta_{jq0}' = (a_2^2 + \xi_{jq}'^2)^{1/2}, \quad (57)$$

is an irreducible radical. Since the $(j+1)$ reverberations in layer 2 are subsequently followed by further p reflections within the same intermediate layer 2, the product of reflection coefficients is

$$\Pi_{jqp} = R_{(3)}(\xi_{jq1}'') R^{(2)}(\xi_{jq2}'') \dots R_{(3)}(\xi_{jqp}''), \quad p \text{ (odd)} \quad (58)$$

Similarly, we consider n subsequent reflections in the top layer, hence,

$$\begin{aligned} \Pi_{jqpn} &= R^{(1)}(\xi_{jqp1})R_{(2)}(\xi'_{jqp2}) \\ &\cdots \begin{cases} R^{(1)}(\xi_{jqpn}), & n \text{ (odd)} \\ R_{(2)}(\xi'_{jqpn}), & n \text{ (even)}. \end{cases} \end{aligned} \quad (59)$$

Analogous to Eq. (49), the phase function has the final form

$$\begin{aligned} g_{jqpn}(\xi'', x, z) &= -\xi''^T \mathbf{x}_S'' - [\xi''_0^T \mathbf{D}_2^{j+1} - \xi''_{j0}^T \mathbf{D}^{q+1} \\ &\quad + \xi''_{j0}^T \mathbf{D}_2^{p+1}] \mathbf{T} d' + \xi''_{jqp0}^T \mathbf{D}^{n^*+1} \mathbf{x}, \end{aligned} \quad (60)$$

where n^* equals n (even) and $n^* = n - 1$ when n is odd.

It contains five irreducible radicals

$$\begin{aligned} \chi &= (b^2 + \xi''^2)^{\frac{1}{2}}, & \zeta''_0 &= (a_2^2 + \xi''^2)^{\frac{1}{2}}, \\ \eta'_{j0} &= (a^2 + \xi_j'^2)^{\frac{1}{2}}, & \zeta'_{jq0} &= (a_2^2 + \xi_j'^2)^{\frac{1}{2}}, \\ \eta'_{jqp0} &= (a^2 + \xi_{jqp}^2)^{\frac{1}{2}}, \end{aligned} \quad (61)$$

and the local slownesses are expressed in terms of the integration variable ξ'' by successive transformations:

$$\xi''_{jqp}^T = (i \xi'_{jqp}, \zeta'_{jqp}) = \xi''_{jq0}^T \mathbf{D}_2^{p+1}, \quad p \text{ (odd)} \quad (62)$$

$$\begin{aligned} \zeta'_{jq0} &= (a^2 + \xi_j'^2)^{1/2}, \\ \xi''_{jq}^T &= (i \xi'_{jq}, \eta'_{jq}) = \xi''_{j0}^T \mathbf{D}^{q+1}, \quad q \text{ (odd)} \end{aligned} \quad (63)$$

$$\begin{aligned} \eta'_{j0} &= (a^2 + \xi_j'^2)^{1/2} \\ \xi''_j^T &= (i \xi'_j, \zeta'_j) = \xi''_0^T \mathbf{D}_2^{j+1}, \quad j \text{ (even)} \\ \zeta''_2 &= (a_2^2 + \xi''^2)^{1/2}, \end{aligned} \quad (64)$$

Furthermore, we have the pairs of transformations, including Snell's law: in layer 1

$$\xi''_{jqpn}^T = \xi''_{jqp(n-1)}^T = \xi''_{jqp0}^T \mathbf{D}^n, \quad n \text{ (odd)} \quad (65)$$

$$\xi''_{jqp(l+1)}^T = \xi''_{jqpl}^T = \xi''_{jqp0}^T \mathbf{D}^{l+1}, \quad l \text{ (odd)} \quad (66)$$

where apparent source ray slowness is expressed in terms of ξ'' by

$$\begin{aligned} \xi''_{jqp0}^T &= (i \xi'_{jqp}, \eta'_{jqp0}) = (a^2 + \xi_{jqp}^2)^{1/2}, \\ i \xi'_{jqp} &= \xi''_{jq0}^T \mathbf{D}_2^{p+1} \mathbf{T} \\ &= \{[(i \xi''_j, \zeta''_0) \mathbf{D}_2^{j+1} \mathbf{T}, \eta'_{j0}] \mathbf{D}^{q+1} \mathbf{T}, \zeta'_{jq0}\} \mathbf{D}_2^{p+1} \mathbf{T}. \end{aligned} \quad (67)$$

Wave trains in the surface layer of order higher than two are constructed in much the same manner in which Eq. (45) was changed to Eq. (54) and are not recorded here.

In the following section we consider the inverse Laplace transform of the generalized ray integrals and discuss the arrival times of individual direct and refracted rays.

Inverse Laplace transform of generalized ray integrals

The generalized ray integrals of wave trains of first and second order observed in layer 1 are given by Eqs. (45) and (54) in terms of the Laplace-transformed displacements:

$$\bar{\phi}_{j\dots}(s) = \bar{F}(s) \int_{-\infty}^{\infty} S_{0j\dots}(\xi'') \Pi_{j\dots} e^{sg_{j\dots}(\xi'')} d\xi'', \quad (68)$$

The integration can be broken into two parts and from the complex conjugate properties of the reflection and

transmission coefficients we can show that

$$\bar{\phi}_{j\dots}(s) = 2\bar{F}(s) \operatorname{Re} \int_0^{\infty} S_{0j\dots}(\xi'') \Pi_{j\dots} e^{sg_{j\dots}(\xi'')} d\xi'', \quad (69)$$

where Re denotes "the real part of".

In the method of Cagniard (1962) the phase function is changed to t by the mapping

$$t = -g_{j\dots}(\xi'') \quad (70)$$

and the integration over the real variable ξ'' is extended to the complex ξ'' -plane. Assume that t is real and the inverse mapping exists,

$$\xi'' = \hat{\xi}''(t) = g_{j\dots}^{-1}(t). \quad (71)$$

The ray integral is transformed to

$$\bar{\phi}_{j\dots}(s) = 2\bar{F}(s) \operatorname{Re} \int_{t_A}^{\infty} \left\{ S_{0j\dots}[\hat{\xi}''(t)] \Pi_{j\dots} \frac{d\hat{\xi}''}{dt} \right\} e^{-st} dt, \quad (72)$$

where

$$t_A = [-g_{j\dots}(\xi'')]_{\xi''=0}. \quad (73)$$

When $\bar{f}(s) = 1$ in Eq. (5), the inverse Laplace transform of $\bar{\phi}(s)$ is

$$\phi_{j\dots}^{\delta}(t) = (b^2/2\pi) H(t - t_A) \operatorname{Re} \left\{ S_{0j\dots}[\hat{\xi}''(t)] \Pi_{j\dots} \frac{d\hat{\xi}''}{dt} \right\}. \quad (74)$$

$H(t - t_A)$ is the Heaviside step function, vanishing when $t < t_A$. The inverse Laplace transform is thus accomplished by "inspection".

For a general time function $f(t)$, the result is obtained by convolution and the integration may be changed to one over ξ'' :

$$\begin{aligned} \phi_{j\dots}(t) &= (b^2/2\pi) H(t - t_A) \int_0^t f(t - \tau) \\ &\quad \cdot [\operatorname{Re} \int_{\Gamma} S_{0j\dots}(\xi'') \Pi_{j\dots} d\xi''] d\tau. \end{aligned} \quad (75)$$

The integration contour Γ in the ξ'' -plane, the portion OM is along the imaginary ξ'' -axis, is shown in Fig. 4. The point M is at $\xi'' = \xi''_M = \pm iB_M$, where B_M is real and ξ''_M is the root of the equation

$$dt/d\xi'' = 0. \quad (76)$$

The point ξ''_M is a saddle point on the ξ'' -plane and a stationary point for the phase functions $g_{j\dots}$. The upper limit of integration, $\hat{\xi}''(\tau)$, is calculated numerically from Eq. (71) when $t = \tau$. The locus of $\hat{\xi}''(\tau)$ is the contour Γ . Note that the saddle point M can be either below the lowest branch point of the integrand in Eq. (75) as shown in Fig. 4a or above the lowest branch point (Fig. 4b).

Once the wave along each ray path is evaluated, the total response is the summation of rays that have arrived at an observation point (x, z) ,

$$\phi(t) = \Sigma \Sigma \phi_{jq}(t) + \Sigma \Sigma \Sigma \Sigma \phi_{jqpn} + \dots \quad (77)$$

The first term, $j = q = 0$, is the ray from the source to the receiver, traversing the layers, and the double sum represents the first train of waves. The second group

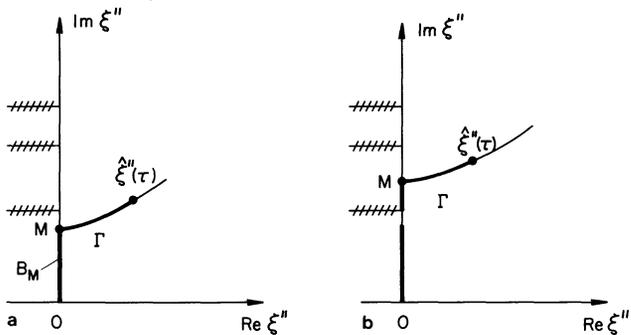


Fig. 4a, b. Cagniard's contour of integration (Γ in ξ'' -plane, with branch cuts in the second quadrant, extends into the first quadrant. If $B_M < 0$, contour Γ extends into the fourth quadrant)

represents waves that are transmitted to the surface layer, subsequently transmitted to the intermediate layer 2 and then refracted back to the top layer.

In the following two subsections we derive Eq. (76) explicitly and discuss proper branch points corresponding to early arrivals of head waves.

Arrival times of direct rays

The stationary point of the phase function is determined by solving Eq. (76). To calculate the roots, we must first carry out the differentiation explicitly and then solve this irrational equation numerically. Once the roots $\xi'' = \xi''_M = \pm iB_M$ are found, the value of t at this stationary point is the arrival time of a particular ray,

$$t_M = -g_{j\dots}(\xi''_M). \quad (78)$$

We call t_M the arrival time of a direct ray. By direct ray we mean a ray that is either transmitted or reflected at an interface. The case of a refracted ray, a part of which travels along an interface, is discussed in the next subsection.

In the following we calculate $dt/d\xi''$ for the first two trains of waves in the surface layer.

Differentiation of the phase function, Eq. (49), renders the equation for ξ''_M for the first wave train,

$$\begin{aligned} -i \frac{dg_{jq}}{d\xi''} &= -\frac{1}{\chi} \chi^T \mathbf{x}_S'' - \frac{1}{\zeta''_0} (\zeta''_0{}^T \mathbf{D}_2^{j+1} \mathbf{T}) d' \\ &+ \frac{1}{\zeta''_0} (\zeta''_0{}^T \mathbf{D}_2^{j+1} \mathbf{T}) \frac{1}{\eta'_{j0}} \eta'_{j0}{}^T \mathbf{D}^{q+1} \mathbf{x} = 0, \end{aligned} \quad (79)$$

where the vectors

$$\chi^T = (\chi, -i\xi''), \quad \zeta''_0{}^T = (\zeta''_0, -i\xi'') \quad (80)$$

$$\eta'_{j0}{}^T = (\eta'_{j0}, -i\xi'_j),$$

which are orthogonal to the corresponding differentiated slowness vectors, ξ'' , ξ''_0 , ξ'_{j0} , respectively, have been introduced in the compact matrix notation. Note the scalar factor $\frac{1}{\zeta''_0} (\zeta''_0{}^T \mathbf{D}_2^{j+1} \mathbf{T})$ when differentiating the higher order slowness vector and the transformation

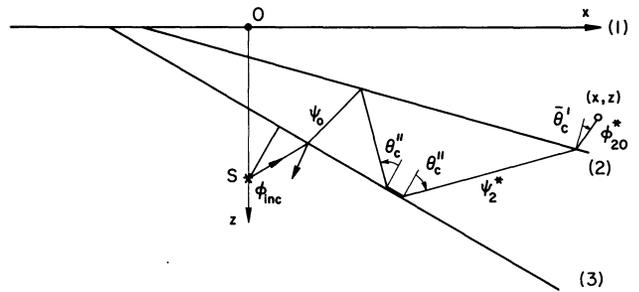


Fig. 5. Refracted ray path ϕ_{20}^* . Refraction of ψ_2^* at fast bottom to interface (3). (No counterpart corresponding to a head wave travelling up-dip is observed by an up-dip receiver point under the conditions assumed in the figure)

$$i \zeta'_j = \zeta_0{}^T \mathbf{D}_2^{j+1} \mathbf{T}: \quad (81)$$

$$\begin{aligned} -i \frac{dg_{jqpn}}{d\xi''} &= -\frac{1}{\chi} \chi^T \mathbf{x}_S'' - \frac{1}{\zeta''_0} (\zeta''_0{}^T \mathbf{D}_2^{j+1} \mathbf{T}) d' \\ &+ \frac{1}{\zeta''_0} (\zeta''_0{}^T \mathbf{D}_2^{j+1} \mathbf{T}) \frac{1}{\eta'_{j0}} (\eta'_{j0}{}^T \mathbf{D}^{q+1} \mathbf{T}) d' \\ &- \frac{1}{\zeta''_0} (\zeta''_0{}^T \mathbf{D}_2^{j+1} \mathbf{T}) \frac{1}{\eta'_{j0}} (\eta'_{j0}{}^T \mathbf{D}^{q+1} \mathbf{T}) \frac{1}{\zeta'_{jq0}} \\ &\cdot (\zeta'_{jq0}{}^T \mathbf{D}_2^{p+1} \mathbf{T}) d' \\ &+ \frac{1}{\zeta''_0} (\zeta''_0{}^T \mathbf{D}_2^{j+1} \mathbf{T}) \frac{1}{\eta'_{j0}} (\eta'_{j0}{}^T \mathbf{D}^{q+1} \mathbf{T}) \frac{1}{\zeta'_{jq0}} \\ &\cdot (\zeta'_{jq0}{}^T \mathbf{D}_2^{p+1} \mathbf{T}) \frac{1}{\eta'_{jqp0}} \eta'_{jqp0}{}^T \mathbf{D}^{q+1} \mathbf{x} = 0, \end{aligned} \quad (82)$$

where the additional higher order vectors

$$\zeta'_{jq0}{}^T = (\zeta'_{jq0}, -i\xi'_j) \quad (83)$$

and

$$\eta'_{jqp0}{}^T = (\eta'_{jqp0}, -i\xi'_{jqp}),$$

which are orthogonal to ξ'_{jq0} and ξ'_{jqp0} , respectively, enter through differentiation. Slownesses are related to source ray slowness by

$$\begin{aligned} i \zeta'_{jq} &= \zeta_0{}^T \mathbf{D}^{q+1} \mathbf{T} = (\zeta_0{}^T \mathbf{D}_2^{j+1} \mathbf{T}, \eta'_{j0}) \mathbf{D}^{q+1} \mathbf{T}, \\ i \zeta'_{jqp} &= \zeta'_{jq0}{}^T \mathbf{D}_2^{p+1} \mathbf{T} \\ &= [(\zeta_0{}^T \mathbf{D}_2^{j+1} \mathbf{T}, \eta'_{j0}) \mathbf{D}^{q+1} \mathbf{T}, \zeta'_{jq0}] \mathbf{D}_2^{p+1} \mathbf{T}. \end{aligned} \quad (84)$$

Arrival times of refracted rays

We assume layers with increasing slowness, $b < a_2 < a$. Since refraction occurs only during critical reflection at a fast bottom, the lowest order refracted ray is ψ_2^* , see Fig. 5, when the once-transmitted ray ψ_0 after reflection at interface (2), ψ_1 , is incident under critical angle condition, $\theta'_c = \sin^{-1}(b/a_2)$, at interface (3). A portion of the refracted ray coincides with interface (3) and the last two segments to an observation point in layer 1 are under critical angle condition, θ'_c , and $\theta'_c = \sin^{-1}(a_2/a)$ is changed to θ'_c due to the divergence through the angle β .

We calculate early arrival times of head waves from the branch points

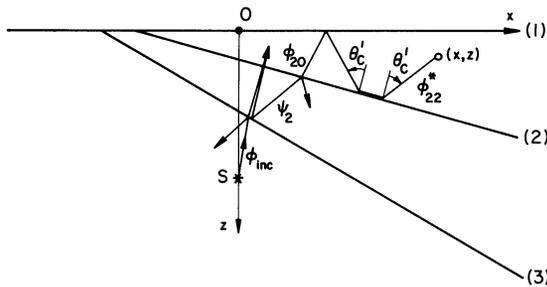


Fig. 6. Refracted ray path ϕ_{22}^* . Refraction of ϕ_{22} at fast bottom to interface (2). First train of waves in surface layer

$$\xi_j'' = \zeta_j \pm ib, \quad \chi_j'' = 0. \quad (85)$$

The location in the ξ'' -plane is calculated by solving the equations

$$\begin{aligned} \xi_j''^T &= \xi_0''^T \mathbf{D}_2^j = -b(1, -\kappa_2), \\ \kappa_2 &= \left(\frac{a_2^2}{b^2} - 1\right)^{1/2}, \quad \text{real} \quad j(\text{even}). \end{aligned} \quad (86)$$

Hence,

$$\xi_0''^T = -b(\zeta_j \pm 1, -\kappa_2)(\mathbf{D}_2^j)^T \quad j(\text{even}) \quad (87)$$

determines the branch point which may fall below the stationary point. The minus sign (-) applies to the ray $j=1$ only when incident under critical angle condition and ψ_2^* renders a head wave propagating up-dip in the intermediate layer 2.

Head waves of the first train of waves propagating in the surface layer are given by rays ϕ_{jq}^* refracted at interface (2), Fig. 6.

The location of the branch points

$$\xi_{jq}' = \zeta_j \pm ia_2, \quad \chi_{jq}' = 0 \quad (88)$$

in the ξ'' -plane are calculated by successive solution of the equations

$$\xi_{jq}''^T = \xi_{j0}''^T \mathbf{D}^q = -a_2(1, -\kappa), \quad \kappa = \left(\frac{a}{a_2} - 1\right)^{1/2}, \quad \text{real}, \quad (89)$$

$$\xi_j''^T = \xi_0''^T \mathbf{D}_2^{j+1} \quad (90)$$

to be at

$$\begin{aligned} \xi_{j0}''^T &= -a_2(1, -\kappa)(\mathbf{D}^q)^T, \quad i\xi_j' = -a_2(1, -\kappa)(\mathbf{D}^q)^T \mathbf{T}, \\ \zeta_j' &= (a_2^2 + \xi_j'^2)^{1/2}, \quad \xi_0''^T = \xi_j'(\mathbf{D}_2^{j+1})^T, \quad i\xi_j'' = \xi_j'^T (\mathbf{D}_2^{j+1})^T \mathbf{T}. \end{aligned} \quad (91)$$

Head waves in the second train of waves may travel in the intermediate layer 2, the branch points being

$$\xi_{jq}'' = \zeta_j \pm ib, \quad \chi_{jq}'' = 0, \quad (92)$$

or in the surface layer. Refraction of the ray ϕ_{jqpn}^* at interface (2) is given by the branch points

$$\xi_{jqpn}' = \zeta_j \pm ia_2, \quad \chi_{jqpn}' = 0. \quad (93)$$

The locations in the ξ'' -plane are found by successive solutions of proper transformation equations: For even j we have

$$\xi_j'' = \zeta_j' [\cos(j+1)\beta - \sqrt{(ia_2/\zeta_j')^2 - 1} \sin(j+1)\beta], \quad (94)$$

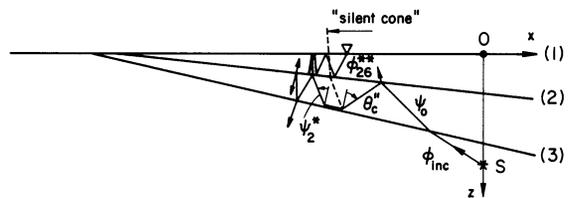


Fig. 7. Refracted ray path ϕ_{26}^{**} corresponding to backscattering of an up-dip travelling head wave ψ_2^* . Note the steepening effect in the direct ray path shown in the surface layer 1

where, in the case of refraction at interface (3),

$$\begin{aligned} \xi_j' &= ib [\zeta_j \pm \cos p\beta - \sqrt{(a_2/b)^2 - 1} \sin p\beta] \{ \cos(q+1)\alpha \\ &\quad - \sqrt{[a/a_3(\zeta_j \pm \cos p\beta - \sqrt{(a_2/b)^2 - 1} \sin p\beta)]^2 - 1} \sin(q+1)\alpha \}. \end{aligned} \quad (95)$$

In the case of refraction at interface (2) the local slowness to be inserted in Eq. (94) becomes

$$\xi_j' = \xi_{jq}' [\cos(q+1)\alpha - \sqrt{(ia/\xi_{jq}')^2 - 1} \sin(q+1)\alpha], \quad (96)$$

where, with p (odd)

$$\xi_{jq}' = \xi_{jqp}' [\cos(p+1)\beta - \sqrt{(ia_2/\xi_{jqp}')^2 - 1} \sin(p+1)\beta] \quad (97)$$

and with n (even)

$$\xi_{jqp}' = ia_2 [\zeta_j \pm \cos n\alpha - \sqrt{(a/a_2)^2 - 1} \sin n\alpha]. \quad (98)$$

All radicals are assumed real and the branch points located on the imaginary ξ'' -axis may fall below the stationary point M , indicating the early arrival of the head wave. The arrival time is calculated by substituting ξ'' of the branch point into Eq. (70).

Observation stations receiving those head waves with a last segment under critical angle condition must be located outside the "silent cone" defined by the conditions

$$\frac{dt}{db} > 0, \quad \frac{dt}{da_2} > 0. \quad (99)$$

Up-dip travelling head waves, corresponding to branch points like $\xi_1'' = -ib$ or $\xi_{j1}' = -ia_2$, are subject to backscattering due to the steepening effect of the wedge. A delayed signal, corresponding to a ray path with incident segment under critical angle condition followed by a ray portion coinciding with the proper interface and a subsequent direct ray path to the observational point, is also received within the silent cone, see Fig. 7.

Numerical results: One dipping layer

We omit the intermediate layer and take the limit $h_2 \rightarrow 0$ and $\beta \rightarrow 0$ in Eq. (45). The ray integrals of displacement V become

$$\bar{V}_q(s) = \bar{F}(s) \int_{-\infty}^{\infty} S_0(\xi') \Pi_{0q} e^{s g_{0q}(\xi')} d\xi'. \quad (100)$$

Π_{0q} is given by Eq. (43) and $S_0 = ST^{(2)}(\xi')$. The phase function after q reverberations is given in the limit of

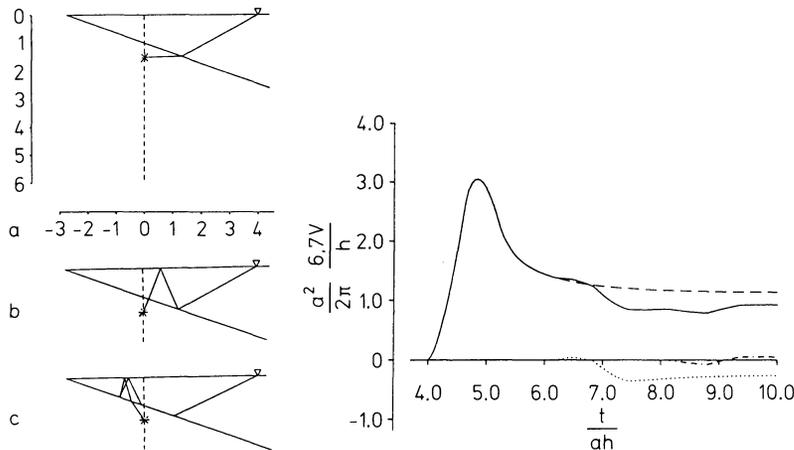


Fig. 8a-c. Dipping surface layer, $\alpha=20^\circ$. Source location in (fast) bedrock at $z_0=1.5h$. Observational point at free surface, epicentral distance $x=4h$. Triangular source time function: $2\Delta=2.4$. Arrival times: **a** source ray-3.995, **b** ray 2-6.1828, **c** ray 4*-8.0847

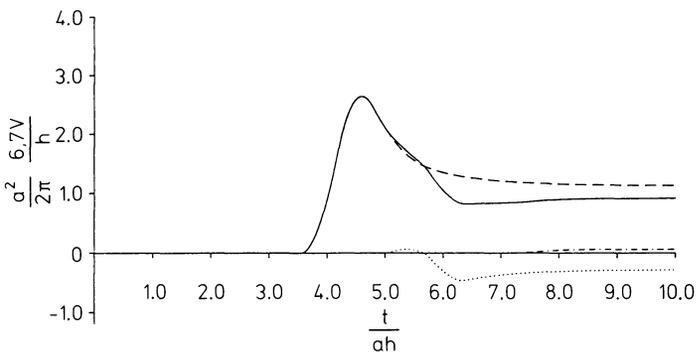


Fig. 9. Parallel surface layer, $\alpha=0^\circ$. Source location in (fast) bedrock at $z_0=1.5h$. Observational point at free surface, epicentral distance $x=|4h|$. Triangular source time function: $2\Delta=2.4$

Eq. (49)

$$g_{0q}(\xi') = -\xi'^T \mathbf{D} \mathbf{x}_0 + \xi_0'^T \mathbf{D}^{q+1} \mathbf{x}, \quad \mathbf{x}_0^T = (x_0, z_0) \quad (101)$$

with the newly defined slowness vectors

$$\xi'^T = (i\xi', \xi'), \quad \xi_0'^T = (i\xi', \eta_0'), \quad \xi' = (b^2 + \xi'^2)^{1/2}, \quad \eta_0' = (a^2 + \xi'^2)^{1/2}. \quad (102)$$

For convenience of integrations in the ξ -plane, we calculate the response to a source time function of triangular shape of duration 2Δ , with equal rise and fall time Δ , when the line source is located at $x=0$, $z_0=1.5h$. Observational points are situated on the free surface at epicentral distances $x=4h$, $\pm h$, respectively.

Material properties of the layer and the underlying bedrock are $a=c^{-1}=1$, $a_2=c_2^{-1}=1/\sqrt{2}$, $\rho_2/\rho=1$. Dipping angle $\alpha=20^\circ$ ($\alpha=0^\circ$ in the parallel layer case). Individual and summed ray integrals are shown in Figs. 8 and 9. The pulse recorded in the parallel layer case of Fig. 9 is much smoother in the less peaked source ray at early times as well as in the tail at later times. Thus Fig. 8, indicating higher accelerations in the seismogram on top of an inclined layer with arrival of a head wave, contributes to the understanding of high damage records of the Skopje (1963) earthquake for structures above sloping parts of the interface, reported by Porceski (1969). Note that only a finite number of rays arrive at a fixed receiver in the dipping layer case.

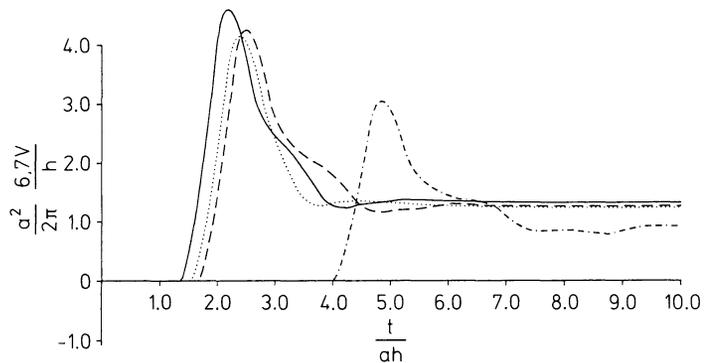


Fig. 10. Signals observed at the free surface of a dipping layer, $\alpha=20^\circ$. Epicentral distances are: $x=0$ (epicentre) —, $x=-h$..., $x=h$ ---, $x=4h$ - · - · - ·. Source location in bedrock, $z_0=1.5h$

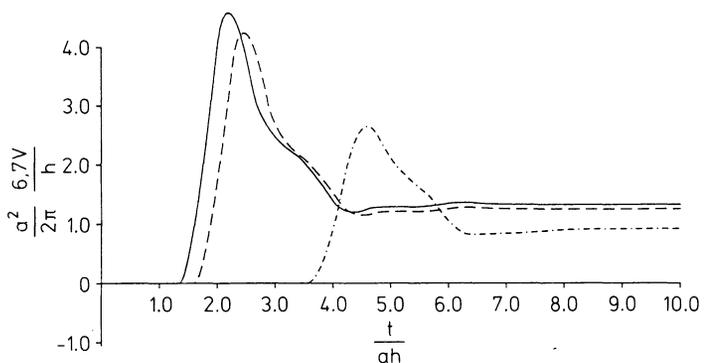


Fig. 11. Signals observed at the free surface of a parallel layer, $\alpha=0^\circ$. Epicentral distances are: $x=0$ (epicentre) —, $x=|h|$ ---, $x=|4h|$ - · - · - ·. Source location in bedrock, $z_0=1.5h$

There are two direct rays and a returning and refracted ray shown in Fig. 8a-c, respectively. Within the recording time of Fig. 9 there are also three rays observed, but they are all direct rays. Individual ray signals arriving at later times in the parallel layer case are all weakened by transmission of energy through interface (2). Figures 10 and 11 clearly show the inclination effect and the asymmetry in the synthetic seismograms recorded at $x=\pm h$. Amplification effects and comparatively higher accelerations are much more pronounced at larger down-dip distances of the observation points. Note the critical angle of incidence, at the interface to the fast bedrock, to be 45° .

Conclusions

Cylindrical waves emitted by a non-stationary line source located in bedrock and transmitted through two dipping layers several times up and down, with an arbitrary number of subsequent reflections within those layers considered between transmission, are described in closed form by generalized ray integrals. Through proper transformations of local slowness vectors by rotational matrices, the ray integrals are put in a form suitable for the application of the Cagniard technique. Thus, inversion of the Laplace-transformed solution is the same as in the parallel layer case; hence, becomes a standard procedure. Contrary to the parallel layer case, the number of ray integrals received at a fixed point of observation becomes finite. Generalization of the analytical part of the solution to a multi-layered configuration is rather simple, as long as the source remains in the underlying bedrock.

Divergence effects of a single dipping layer on synthetic seismograms, observed at receivers at the free surface of the half-space, are studied by comparison of individual and summed ray signals to those of the parallel layer case; equal thickness of the layers is understood at source location. Integration in the source ray slowness plane is performed by Gaussian quadrature in both cases, thus the standard procedure of the parallel layer case is directly applicable. A triangular source time shape function was chosen to save computer time on the convolution integrals.

The number of ray integrals to be evaluated in the multi-layered case is dramatically increased together with computer time required for numerical integration. Thus, numerical results had to be restricted to one layer only. The variation of locations of the observational point on top of the half-space and the amplification and increased unevenness of the records observed give a quantitative interpretation of damage results reported on structures above dipping parts of an interface in the Skopje (1963) earthquake, where a single layer has been reported.

Acknowledgements. The authors gratefully acknowledge support of this research by a grant from the National Science Foundation (CEE8206758). Part of the research of Franz Ziegler was performed while he was appointed Visiting Professor in the Department of Theoretical and Applied Mechanics at Cornell University during the summers of 1983 and 1984.

References

- Aki, K., Richards, P.G.: Quantitative seismology, Theory and methods. Vols. 1 and 2. San Francisco: W.H. Freeman and Company 1980
- Cagniard, L.: Reflection and refraction of progressive seismic waves. (Translated and revised by E.A. Flinn and C.H. Dix.) New York: McGraw-Hill Book Company 1962
- Drimmel, J., Gangl, G., Gutdeutsch, R., König, M.: Model seismic experiments for the interpretation of macroseismic data observed in the Eastern Alps (in German). *Z. Geophys.* **39**, 21–39, 1973
- Hong, T.L., Helmberger, D.V.: Generalized ray theory for dipping structure. *Bull. Seismol. Soc. Am.* **67**, 995–1008, 1977
- Hudson, J.A.: *SH*-Waves in a wedge-shaped medium. *Geophys. J.R. Astron. Soc.* **7**, 517–546, 1963
- Ishii, H., Ellis, R.M.: Head and reflected waves from an *SH* line source in a dipping layer overlaying an elastic medium. *J. Phys. Earth* **18**, 1–17, 1970a
- Ishii, H., Ellis, R.M.: Multiple reflection of plane *SH*-waves by a dipping layer. *Bull. Seismol. Soc. Am.* **60**, 15–28, 1970b
- Kennett, B.L.N.: Seismic waves in a stratified half-space-II, Theoretical Seismograms. *Geophys. J.R. Astron. Soc.* **61**, 1–10, 1980
- Müller, G.: Theoretical seismograms for some types of point-sources in layered media.
Part I. *Z. Geophys.* **34**, 15–35, 1968a
Part II. *Z. Geophys.* **34**, 147–162, 1968b
Part III. *Z. Geophys.* **35**, 347–371, 1969
- Pao, Y.H., Gajewski, R.R.: The generalized ray-theory and transient responses of layered elastic solids. In: *Physical Acoustics*, **13**, W.P. Mason and R.N. Thurston, eds.: pp. 184–265. New York: Academic Press 1977
- Pao, Y.H., Ziegler, F.: Transient *SH*-waves in a wedge-shaped layer. *Geophys. J.R. Astron. Soc.* **71**, 57–77, 1982
- Porceski, A.: The ground effect of the Skopje July 26, 1963 earthquake. *Bull. Seismol. Soc. Am.* **59**, 1, 1969
- Wiedmann, S.: Interpretation eines Alpenprofils unter Berücksichtigung der Beugung mit Hilfe seismischer Modellierungsmethoden. Dissertation, University of Vienna, 1983
- Wiggins, R.A., Helmberger, D.V.: Synthetic seismogram computation by expansion in generalized rays. *Geophys. J.R. Astron. Soc.* **37**, 73–90, 1974
- Ziegler, F., Pao, Y.H.: Theory of generalized ray for *SH*-waves in dipping layers. *Wave Motion* **7**, 1–24, 1985

Received October 10, 1984; Revised version February 7, 1985
Accepted February 12, 1985

The formulae for the calculation of the Fresnel zones or volumes

B. Gelchinsky

Department of Geophysics and Planetary Sciences, Tel-Aviv University, Ramat-Aviv, Tel Aviv 69978, Israel

Abstract. The symmetrized invariant formulae for the calculation of Fresnel zones or volumes are derived. It is assumed that an inhomogeneous medium with curvilinear interfaces is located between the source and/or the receiver and along the central ray within the Fresnel zone or volume. In the vicinity of the zone centre, the medium is considered locally homogeneous.

The formula for the leading term of the field of a wave scattered by a bent body immersed in the above-mentioned medium is obtained by the Kirchhoff approximation. With the help of this formula and the expressions for the Fresnel radii for a particular case, the formulae for the Fresnel zones in the general case considered are obtained on the basis of the reciprocity relation. The formulae for the Fresnel zones are used to obtain the expressions for the Fresnel volumes.

The physical consequences of the derived formulae with respect to the validity of the ray formulae and the resolution of seismic methods etc. are discussed.

Key words: Area essential for reflection (propagation) – Symmetrized invariant formulae – Validity conditions – Resolution.

Introduction

The Fresnel diffraction theory has occupied a central position in optics and in the theory of wave propagation in general since 1818, when a well-known Fresnel memoir appeared. In 1882, Kirchhoff gave the Fresnel diffraction theory a rigorous mathematical foundation; since that time the explanation of diffraction and wave propagation has been based essentially on the Fresnel-Kirchhoff theory. The concept of Fresnel zones plays an important role in this theory and is continually being developed and generalized.

This problem has been examined in many books and articles and it is impossible to review them all here. We shall mention only the works of Al'pert et al. (1953), Bertoni et al. (1971) and Kravtsov and Orlov (1980) in which special attention is paid to the consideration of regions essential to the formation of fields of reflected and transmitted waves¹. This problem was

investigated in depth in the book by Kravtsov and Orlov (1980). The following points connected with the Fresnel volume are considered on a heuristic basis: an area of ray localization, a finite thickness of physical ray, an area of applicability and resolution of the ray method.

In the seismic literature, a certain amount of attention is paid to the question of computation of the Fresnel zones and their connection with the resolution of the seismic method (see, for example, Hagedoorn, 1959; Hilterman, 1970; Sheriff, 1980; Sheriff and Geldart, 1982; Kleyn, 1983). However, only the simplest cases are considered while, in practice media of a rather complicated structure are generally encountered. However, as far as we know, the formulae for computation of the Fresnel zones and volumes for the case of a sufficiently complicated structure have not been given, although many formulae for the Fresnel zones for various particular cases are presented in the literature (Tatarsky, 1967; Flatte, 1979; Kravtsov and Orlov, 1980).

The aim of the present paper is to derive symmetrized invariant formulae for the computation of Fresnel zones and volumes for media of complex structure². In this paper it is assumed that an inhomogeneous medium with curvilinear interfaces is located between the sources and/or the receiver and the centre of the zone or volume. There is one restriction: in the vicinity of the zone (or volume) centre the medium is considered locally homogeneous.

In order to show how the notion of the Fresnel zones appears in the Fresnel-Kirchhoff theory, we first consider the problem of the scattering of a wave on a body of arbitrary shape. This consideration is also the basis for determining the Fresnel volume. Since the techniques of evaluation of integrals obtained in the Fresnel-Kirchhoff theory are well known (Keller, 1957; Bleistein and Handelsman, 1975; Born and Wolf, 1980; Felsen and Marcuvitz, 1973), the computational scheme with some improvements concerning the smooth continuation of a surface beyond the body contour (Gelchinsky, 1982a) is presented in a very brief form.

In conclusion, some physical consequences of the

¹ Let us note, by the way, that there is no conventional terminology for an area essential for wave propagation. Bertoni et al. (1971) call it the 3-D Fresnel zone and Kravtsov and Orlov (1980), the Fresnel volume. We use the terminology of the latter

² A formula has an invariant form if the quantities included in it do not depend on the choice of the coordinate system. We say that a formula is written in a symmetrized form if the reciprocity principle follows *explicitly* from the written form

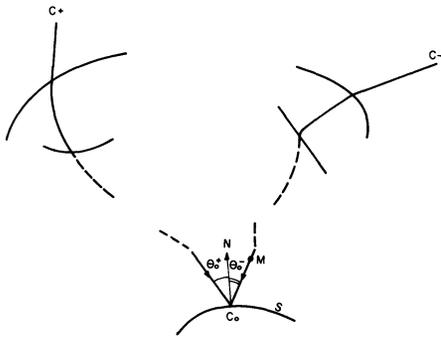


Fig. 1. Ray path for considered model (a 2-D ray scheme is shown for simplicity). C^+ is the point of emission; C^- is the point of observation; S is the scattering surface; θ_0^+ and θ_0^- are angles of incidence in positive and negative directions; C_0 is the specular point

derived formulae are discussed. We try to present this discussion so that readers unfamiliar with the derivation of the formulae obtained can, at least, understand their consequences.

Derivation of a formula for the field of a scattered wave in the Kirchhoff approximation

Let a time harmonic wave with frequency ω fall on the surface S of a body and let the time dependence $\exp(-i\omega t)$ be ignored. The field of the scattered wave $U(M)$ at the point M can be determined by the Green formula (often known as the Kirchhoff formula in the case of a scalar wave equation):

$$U(M) = \iint \left\{ U(C) \frac{\partial G(M, C)}{\partial N} - G(M, C) \frac{\partial U(C)}{\partial N} \right\} dS(C), \quad (1)$$

where $G(M, C)$ is the Green function with the source at point M and the receiver at point C on S , and $\partial/\partial N$ denotes differentiation with respect to the normal \mathbf{N} to the surface S .

Since the shape of the contour limiting the surface S , and also the type of the point source, does not influence the parameters of the Fresnel zone or volume, we limit ourselves to the consideration of scattering of a wave excited by a point source in the form of the δ -function, at a general curvilinear surface having the form of a bent rectangle.

In the following, we will use the reciprocity relation, changing the source and the receiver at the fixed points C^+ and C^- . In this way the wave motion in the positive direction (the path $C^+ \dots C_0 \dots C^-$) as well as in the negative one (the path $C^- \dots C_0 \dots C^+$) will be considered (Fig. 1). When the wave scattering in the positive (negative) direction is treated, the field of the incident wave will be denoted as $U_0(C^+, C)$ or $U_0^+(C)$ [$U_0(C^-, C)$ or $U_0^-(C)$], where the point C is located on the surface S . Under the given conditions, the Green function $G(C^{(v)}, C)$ [$v = +$ or $-$] and the incident field $U_0^{(v)}(C)$ are equal. The leading part of these fields can be written in the form:

$$U_0^{(v)}(C_0) = U_0(C^{(v)}, C) = G(C^{(v)}, C) = I_0^{(v)}(C) \exp\{i\omega\tau_0^{(v)}(C)\}, \quad (v = + \text{ or } -), \quad (2)$$

where $I_0^{(v)}(C)$ is the amplitude and $\tau_0^{(v)}(C)$ is the time of propagation (eiconal) of the incident wave from the source at the point $C^{(v)}$ to the point C .

It is assumed that the front ($\tau^{(v)} = \text{constant}$) of the wave moving in the v -th direction is of arbitrary shape. This means that an inhomogeneous medium with curved interfaces could exist between the source (or the receiver) at the point $C^{(v)}$ and the point of observation, C . In the vicinity of the point C on the scattering surface S , the medium is considered to be homogeneous.

It is known (Alekseev and Gelchinsky, 1959; Červený and Ravindra, 1971) that, in the Kirchhoff approximation, the amplitude of the scattered wave on the surface S at the point C is determined by the relation:

$$I^{(v)}(C) = \begin{cases} K(\theta_0, \omega) I_0^{(v)}(C) & \text{in the lit area} \\ 0 & \text{in the shadow} \end{cases} \quad (3)$$

where $K(\theta_0, \omega)$ is the coefficient of reflection (transmission) depending on the angle of incidence θ_0 and the frequency ω . The leading part of the scattered field $U(M)$ at the point M in the vicinity of the surface S may be represented by the formula:

$$U^{(v)}(M) = I^{(v)}(C) \left\{ \frac{d\Sigma(M)}{d\Sigma(C)} \right\}^{\frac{1}{2}} \exp\left\{i\omega \left[\tau^{(v)}(C) + \frac{\Delta l}{v} \right]\right\}, \quad (4)$$

where $\Delta l = CM$ is the ray path between the point C and the nearby point M , v is the propagation velocity of the scattered wave and $\left\{ \frac{d\Sigma(M)}{d\Sigma(C)} \right\}^{\frac{1}{2}}$ is the geometrical spreading function of the scattered wave.

If we now consider the scattering of a wave moving in the positive direction³ and substitute the expressions of the field $U(C^+, C)$ and the Green function $G(C^-, C)$ and of its derivatives according to formulae (2)–(4) in Eq. (1), we obtain the following integral:

$$U(C^+, C^-) = \iint_S F(C) \exp\{i\omega\tau(C)\} dS, \quad (5)$$

where

$$F(C) = \frac{-i\omega}{4} I_0^{(+)}(C) I_0^{(-)}(C) K\{\theta^-(C)\}, \quad (6)$$

$$\tau(C) = \tau(C^+, C, C^-) = \tau_0^{(+)}(C_0) + \tau_0^{(-)}(C_0). \quad (7)$$

Since the function $F(C)$ can usually be considered to be a slowly varying function⁴, the approximate value of the integral (5) can be obtained by the well-known method of stationary phase (MSP) (Keller, 1957; Felsen and Marcuvitz, 1973; Bleistein and Handelsman, 1975). The results of computations can be presented in the form (Gelchinsky, 1982a):

$$U(C^+, C^-) = U_{\text{ray}}(C^+, C^-) W(C^+, C^-), \quad (8)$$

³ If a wave moves in the positive direction, the source is at the point C^+ and the point of observation coincides with the receiver C^- .

⁴ We shall later recall some of the physical conditions under which the function F can be considered as a slowly varying one.

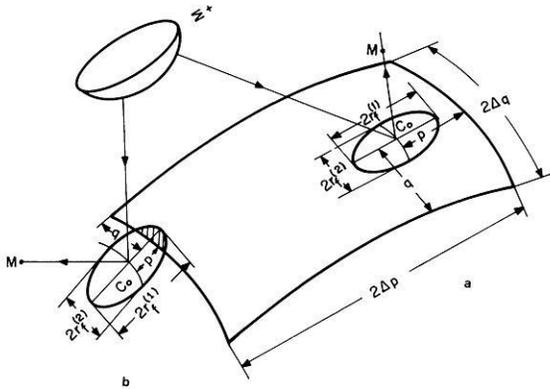


Fig. 2a and b. The Fresnel zone on the curved rectangle, S : **a** Real point of reflection (lit area). **b** Fictitious point of reflection (shadow part of half-shadow area). $r_f^{(1)}$ and $r_f^{(2)}$ are half-axes (radii) of the curved Fresnel zone; p and q are distances from C_0 to the nearest edge of the rectangle; $2\Delta p$ and $2\Delta q$ are the length and width of the rectangle, respectively; Σ^+ is the front of the wave incident in the positive direction

where U_{ray} is equal to the leading part of the reflected wavefield calculated according to the formulae of the ray method (Alekseev and Gelchinsky, 1959; Červený and Ravindra, 1971) as if the surface S were unbounded.

W is the so-called weakening function which takes account of the influence of the restricted size of the surface S on the scattered field. When S is a bent rectangle, the function W can be presented in the form of the product of two Fresnel integrals (Gelchinsky, 1982a):

$$W = \left\{ (2i)^{\frac{1}{2}} \int_{\xi_1}^{\xi_2} \exp(i\pi X^2/2) dX \right\} \cdot \left\{ (2i)^{\frac{1}{2}} \int_{\eta_1}^{\eta_2} \exp(i\pi X^2/2) dX \right\}, \quad (9)$$

where:

$$\xi_1 = \frac{2^{\frac{1}{2}} p}{r_f^{(1)}}, \quad \xi_2 = \frac{2^{\frac{1}{2}} (2\Delta p - p)}{r_f^{(1)}},$$

$$\eta_1 = \frac{2^{\frac{1}{2}} q}{r_f^{(2)}}, \quad \eta_2 = \frac{2^{\frac{1}{2}} (2\Delta q - q)}{r_f^{(2)}}. \quad (10)$$

The arguments, (10), of each of the integrals (9) are dimensionless ratios of certain distances, $p(q)$ and $2\Delta p - p$ ($2\Delta q - q$), and of a certain characteristic size, $r_f^{(1)}$ ($r_f^{(2)}$). The geometrical sense of these quantities is explained in Fig. 2. The quantity $p(q)$ is the distance from the specular point C_0 , computed by the laws of geometrical optics when the positions of the points C^+ and C^- are fixed, to the closest edge of the rectangle. This distance is measured along the surface S parallel to the corresponding side of this rectangle; $2\Delta p - p$ ($2\Delta q - q$) is the distance to the opposite side of the rectangle.

The quantities $r_f^{(1)}$ and $r_f^{(2)}$ are the radii (half-axes) of the Fresnel zones on the surface S with the centre at the specular point C_0 (Fig. 2). The position of C_0

(when the points C^+ and C^- are fixed) is determined by the condition of stationary phase (Snell's law):

$$\frac{\partial \tau}{\partial \xi_1} = \frac{\partial \tau}{\partial \xi_2} = 0 \quad \text{at the point } C_0, \quad (11)$$

where ξ_1 and ξ_2 are curvilinear coordinates on the surface S .

In Fig. 2 two cases are shown: the first when the point C_0 is located in the lit area and the second when it is in the so-called half-shadow, where the point C_0 is located not far from the edge of S at a distance smaller than the respective Fresnel radius. In particular, the point C_0 can be located beyond the body's contour on the so-called smooth continuation of the surface S (a detailed description is given in Gelchinsky, 1982a). This is a fictitious specular point.

The quantity $p_f^{(i)}$ ($i=1, 2$) is determined by the expression:

$$\frac{1}{r_f^{(i)}} = \left\{ \frac{|\alpha + \beta - (-1)^i [(\alpha - \beta)^2 + \gamma^2]^{\frac{1}{2}}|}{2} \right\}^{\frac{1}{2}}, \quad (12)$$

where α , β and γ are second derivatives of eiconal τ

$$a = \frac{\omega}{\pi} \frac{\partial^2 \tau}{\partial \xi_1^2}, \quad \beta = \frac{\omega}{\pi} \frac{\partial^2 \tau}{\partial \xi_2^2},$$

$$\gamma = \frac{2\omega}{\pi} \frac{\partial^2 \tau}{\partial \xi_1 \partial \xi_2} \quad \text{at point } C_0 \quad (13)$$

in an orthogonal curvilinear coordinate system where $\xi_1 = \xi_{\parallel}$ and $\xi_2 = \xi_{\perp}$, when the origin is located at the specular point, the tangent to the line ξ_{\parallel} at the point C_0 is in the plane of incidence E_{\parallel} , and the tangent to the line ξ_{\perp} is perpendicular to E_{\parallel} .

It is easy to show for a fixed ray $C^+ \dots C_0 \dots C^-$ using the expression:

$$\tau(C) = \tau(C_0) + \frac{\pi\alpha}{2\omega} \xi_{\parallel}^2 + \frac{\eta\beta}{2\omega} \xi_{\perp}^2 + \frac{\pi\gamma}{\omega} \xi_{\parallel} \xi_{\perp}$$

$$= \tau(C_0) + \frac{\pi}{\omega} \left\{ \frac{p^2}{(r_f^{(1)})^2} + \frac{q^2}{(r_f^{(2)})^2} \right\} \quad (14)$$

that the closed line, the coordinates of which, p^* and q^* , satisfy the expression

$$\tau(C) - \tau(C_0) \equiv \frac{\pi}{\omega} \left[\frac{p_*^2}{(r_f^{(1)})^2} + \frac{q_*^2}{(r_f^{(2)})^2} \right] = \frac{T}{2} \quad (15)$$

where T is the period of wave, determines the boundary of the first Fresnel zone on the surface S . It is easy to see from Eq. (14) that the axes of the Fresnel zone coincide with the axes of the orthogonal curvilinear coordinate system p, q . The angle between the tangents to the lines ξ_{\parallel} and p is determined by the relation:

$$\cos \delta = \left[\frac{1}{2} \left(1 + \frac{\alpha - \beta}{[(\beta - \alpha)^2 + \gamma^2]^{\frac{1}{2}}} \right) \right]^{\frac{1}{2}}. \quad (16)$$

In some cases, it is convenient to introduce the so-called image plane Q , tangent to surface S at the reflection point C_0 (Fig. 3). On this plane the coordinate system x, y is considered where the coordinate line x

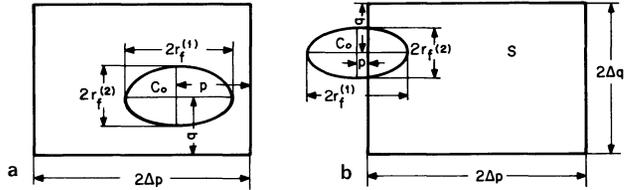


Fig. 3a and b. Images of the curved rectangle S and of the Fresnel zone in the image plane Q : **a** Real point of reflection (lit area). **b** Fictitious point (shadowed area)

(or y) is tangent to the line p (or q) at the point C_0 . On the plane Q the Fresnel zone is an ellipse with centre at the specular point C_0 and the surface S is imaged into a planar rectangle with sides $2\Delta p$ and $2\Delta q$ (see Gelchinsky, 1982a).

The formulae (12) for the Fresnel radii are not convenient for qualitative physical considerations as well as for computations since they include second derivatives, Eq. (13), of the eiconal $\tau(C)$ (in Eq. (7)) which are rather difficult to compute numerically in the case of complex media.

Derivation of invariant symmetrized formulae for second derivatives of the eiconal

In order to obtain the invariant formulae for second derivatives of the time of wave propagation, one can use the various methods developed for derivations of the formulae for the geometrical spreading function or for the curvature of the wavefront (see, for example, Gelchinsky, 1961, 1982; Deschamps, 1972; Hubral, 1980; Hubral and Krey, 1980). However, these methods of derivation are rather cumbersome and, therefore, in our problem we wish to apply the reciprocity relation and the formulae obtained for second derivatives of τ in a more particular case than that considered here. Let us note, by the way, that the reciprocity principle is often applied in the theory of diffraction when the known formulae for the field of a wave moving in one direction are used to obtain or to generalize the expressions for a wave moving in the opposite direction.

Later on the formulae for the second derivatives of τ derived in the paper by Gelchinsky (1982a) are used. These formulae are valid when a wave with a front of arbitrary shape is scattered by a curved body and the receiver at point M and the specular point C_0 on the scattering surface are located in a homogeneous medium (Fig. 2). The formulae are also applicable to the case of converted waves (the velocities of the incident and reflected waves are not equal), as well as to the cases of reflection or refraction.

The following relations:

$$\begin{aligned} \alpha &= g_0^{\frac{1}{2}} \left\{ \frac{\cos^2 \theta_0^+}{\lambda^+ r_{\parallel}^+} + \frac{\cos^2 \theta_0^-}{\lambda^- l} + \frac{1}{R_{\parallel}} \left(\frac{\cos \theta_0^+}{\lambda^+} \pm \frac{\cos \theta_0^-}{\lambda^-} \right) \right\}, \\ \beta &= g_0^{\frac{1}{2}} \left\{ \frac{1}{\lambda^+ r_{\perp}^+} + \frac{1}{\lambda^- l} + \frac{1}{R_{\perp}} \left(\frac{\cos \theta_0^+}{\lambda^+} \pm \frac{\cos \theta_0^-}{\lambda^-} \right) \right\}, \\ \gamma &= g_0^{\frac{1}{2}} \left\{ \left(\frac{\cos \theta_0^+ \sin 2\phi^+}{\lambda^+} \right) \left(\frac{1}{r_1^+} - \frac{1}{r_2^+} \right) \right. \\ &\quad \left. + \sin 2\Phi \left(\frac{1}{R_1} - \frac{1}{R_2} \right) \left(\frac{\cos \theta_0^+}{\lambda^+} - \frac{\cos \theta_0^-}{\lambda^-} \right) \right\} \end{aligned} \quad (17)$$

were derived in the above-mentioned paper (see Eq. (28) in Gelchinsky, 1982a) for the second derivatives of the eiconal τ in Eq. (13).

Besides the notation introduced earlier, the following notation is used: $\lambda^{(v)}$ ($v = +$ or $-$) is the wavelength of the wave incident in the v -direction; $r_1^{(v)}$ and $r_2^{(v)}$ (R_1 and R_2) are the curvature radii of the principal normal sections of the wavefront $\Sigma^{(v)}$ (of the surface S); $r_{\parallel}^{(v)}$ and $r_{\perp}^{(v)}$ (R_{\parallel} and R_{\perp}) are the curvature radii of the normal sections of the wavefront $\Sigma^{(v)}$ (of the surface S) corresponding to the planes E_{\parallel} and E_{\perp} ($E_{\perp} \perp E_{\parallel}$); $\phi^{(v)}$ (Φ) is the angle between the plane of incidence E_{\parallel} and the first principal normal section of the front $\Sigma^{(v)}$ (of the surface S).

The curvature radius $R_j(r_j^{(v)})$ ($j = 1, 2$ or \parallel or \perp) is considered to be positive if the normal section of the surface S (of the front $\Sigma^{(v)}$) is a convex curve from the side of the incident wave. The quantity l is the distance between the point C_0 and point M :

$$l = CM = \tau^-(C_0) v^-. \quad (18)$$

The quantity g_0 is the determinant value for the metric tensor of the surface S in the special orthogonal curvilinear coordinate system p, q , determined by the following local condition at the point C_0 :

$$g_0^{\frac{1}{2}} = \left(\frac{\partial \mathbf{r}(p, q)}{\partial p} \right)^2 = \left(\frac{\partial \mathbf{r}(p, q)}{\partial q} \right)^2, \quad (19)$$

where $\mathbf{r}(p, q)$ is the radius vector to the specular point C_0 .

To generalize the obtained formulae to the case where there are intermediate surfaces between the reflection point C_0 and the point of observation C^- , we consider the formulae for the field, Eq. (8), from the reciprocity principle standpoint (or of the symmetry with respect to the direction of the wave propagation). It is easy to see that the formulae (8) and (9) and the position of the reflection point C_0 are symmetric in this sense. Only the formulae (17) for the second derivatives of the eiconal are not symmetric. Let us try to symmetrize them.

We begin by considering α . The formula for α contains the values characterizing waves incident in both the positive and negative directions. For example, θ_0^+ (θ_0^-) and λ^+ (λ^-) are the angle of incidence and the corresponding wavelength for propagation in the positive (negative) direction. At the same time, the quantities r_{\parallel}^+ and l are similar but not identical characteristics, as both are curvature radii. The difference between them is easily explained as follows: the front Σ^+ of the wave incident on S in the positive direction is not, generally speaking, spherical while the front of the wave incident on S from the point M is spherical because formulae (17) were derived for this case. To generalize the expression for α in the more general case where the front Σ^- is of arbitrary shape, we have to substitute the curvature radius r_{\parallel}^- of the normal section of Σ^- in the plane of incidence E_{\parallel} instead of the quantity l .

In addition, we change the rule for the sign of the curvature radii r_j^- and R_j^- ($j = 1, 2, +$ or $-$). The curvature radius \tilde{R}_j^- or \tilde{r}_j^- is considered to be positive

if the normal section of the surface S or of the front Σ^- is a convex curve from the side of the incident wave propagating in the negative direction. For example, if we consider the transmitted wave and $R_j^+ > 0$, then $R_j^- < 0$. Thus, according to the new rule for the sign of the curvature radii of the surface S and of the front Σ^- , there is only one sign in the corresponding parentheses in formulae (17).

Now the expression for α takes a symmetrized form. If we consider the formula for β in Eq. (17), it is easy to make an analogous generalization by substituting the curvature radius r_{\perp}^- of the normal section (in the plane E_{\perp}) of the front Σ^- instead of the quantity l .

Considering the expression of γ in Eq. (17), we observe that its nonsymmetry is determined by the fact that, from the negative side, the incident front Σ^- has a difference of the principal curvatures equal to zero. The generalization for the general case is easily carried out by addition to the formula for γ of the term analogous to the first item in the expression for γ . This additional term corresponds to the non-spherical wave incident on the negative side of the surface in the general case.

We can now write the symmetrized formulae for the second derivatives of τ :

$$\begin{aligned} \alpha &= g_0^{\frac{1}{2}} \left(\frac{\cos^2 \theta_0^+}{\lambda^+ r_{\parallel}^+} + \frac{\cos^2 \theta_0^-}{\lambda^- r_{\parallel}^-} + \frac{\cos \theta_0^+}{\lambda^+ R_{\parallel}^+} + \frac{\cos \theta_0^-}{\lambda^- R_{\parallel}^-} \right), \\ \beta &= g_0^{\frac{1}{2}} \left(\frac{1}{\lambda^+ r_{\perp}^+} + \frac{1}{\lambda^- r_{\perp}^-} + \frac{\cos \theta_0^+}{\lambda^+ R_{\perp}^+} + \frac{\cos \theta_0^-}{\lambda^- R_{\perp}^-} \right), \\ \gamma &= g_0^{\frac{1}{2}} \left\{ \frac{\cos \theta_0^+ \sin 2\phi^+}{\lambda^+} \left(\frac{1}{r_1^+} - \frac{1}{r_2^-} \right) \right. \\ &\quad + \frac{\cos \theta_0^- \sin 2\phi^-}{\lambda^-} \left(\frac{1}{r_1^-} - \frac{1}{r_2^-} \right) \\ &\quad \left. + \sin 2\Phi \left[\frac{\cos \theta_0^+}{\lambda^+} \left(\frac{1}{R_1^+} - \frac{1}{R_2^+} \right) + \frac{\cos \theta_0^-}{\lambda^-} \left(\frac{1}{R_1^-} - \frac{1}{R_2^-} \right) \right] \right\}. \end{aligned} \quad (20)$$

Thus we obtain formulae (12) and (20) for the radii of the Fresnel zone in the general case when the incident fronts Σ^+ and Σ^- are of arbitrary shape.

The formulae for the Fresnel volume

Let the position of the source and of the receiver at points M_+ and M_- be given and the ray path M_+M_- calculated (Fig. 4). The following procedure is then used to find the Fresnel volume surrounding the centre ray M_+M_- .

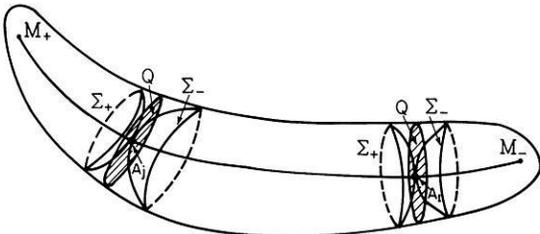


Fig. 4. Plot showing the Fresnel zone construction: Σ^+ (Σ^-) is the front arriving at the point A_j ($j=1, 2, \dots$) from source M_+ (receiver M_-); Q is the cross-section of the Fresnel volume

The fronts $\Sigma^+(A_j)$ and $\Sigma^-(A_j)$ arriving from the points M_+ and M_- to the point A_j are computed for the series of successive points $A_1, A_2, \dots, A_j, \dots$ on the ray. We then assume that the front $\Sigma^+(A_j)$ [and $\Sigma^-(A_j)$] coincides with the surface S . In this case the sizes of the curved Fresnel zones on the wavefronts $\Sigma^+(A_j)$ and $\Sigma^-(A_j)$ and of the Fresnel ellipse on the image planes $Q(A_j)$ can be determined with the help of the formulae derived above. The surface of the Fresnel volume is obtained as an envelope of the Fresnel zones (or of the ellipses) calculated on the series of points A_j ($j=1, 2, \dots$) along the central ray M_+M_- . In the case where the surface S coincides with the incident front [for example $\Sigma^+(A_j)$], the following relations hold at the point A_j :

$$\begin{aligned} \theta_0^+ &= \theta_0^- = \phi^+ = \Phi = 0, & \lambda^+ &= \lambda^- = \lambda, \\ r_1^+ &= r_{\parallel}^+ = R_1^+ = R_{\parallel}^+, & r_2^+ &= r_{\perp}^+ = R_2^+ = R_{\perp}^+. \end{aligned} \quad (21)$$

Taking into account the relation (21), we can rewrite the expression (20) in the form

$$\begin{aligned} \alpha &= \frac{g_0^{\frac{1}{2}}}{\lambda} \left(\frac{1}{r_1^+} + \frac{1}{r_{\parallel}^-} \right), & \beta &= \frac{g_0^{\frac{1}{2}}}{\lambda} \left(\frac{1}{r_2^+} + \frac{1}{r_{\perp}^-} \right), \\ \gamma &= \frac{g_0^{\frac{1}{2}}}{\lambda} \sin 2\Delta\phi \left(\frac{1}{r_1^-} - \frac{1}{r_2^-} \right), \end{aligned} \quad (22)$$

where $\Delta\phi$ is the angle between the two first principal normal planes of the fronts $\Sigma^+(A_j)$ and $\Sigma^-(A_j)$.

If we now substitute the values of α , β and γ from the relations (22) in the expression (12) for the Fresnel radii and use Euler's formula

$$\frac{1}{r_{\parallel}} = \frac{\cos^2 \phi}{r_1} + \frac{\sin^2 \phi}{r_2}, \quad (23)$$

we obtain the following equations:

$$1/r_f^{(j)} = |[c - (-1)^j d]^{\frac{1}{2}}|, \quad (24)$$

where

$$\begin{aligned} c &= 2g_0^{\frac{1}{2}}(h_+ + h_-)/\lambda, \\ d &= \{g_0(\Delta K_+^2 + \Delta K_-^2 + 2\Delta K_- \Delta K_+ \cos^2 2\Delta\phi)/\lambda^2\}^{\frac{1}{2}}, \end{aligned} \quad (25)$$

$$h_v = \frac{1}{2} \left(\frac{1}{r_1^{(v)}} + \frac{1}{r_2^{(v)}} \right),$$

$$\Delta K_v = \frac{1}{r_1^{(v)}} - \frac{1}{r_2^{(v)}}, \quad (v = + \text{ or } -).$$

The symmetrized invariant expressions (24) and (25) determine the radii of the curved Fresnel zones on the fronts Σ^+ and Σ^- or of the Fresnel ellipse in the normal cross-section of the Fresnel volume (in this case $g_0=1$) at the point A_j (Fig. 4).

Some physical consequences

Now we shall consider some implications connected with the formulae obtained. If the following inequalities hold:

$$\Delta p/r_f^{(1)} \gg 1, \quad \Delta q/r_f^{(2)} \gg 1, \quad (26)$$

$$p/r_f^{(1)} \gg 1, \quad q/r_f^{(2)} \gg 1, \quad (27)$$

where, as previously, Δp and Δq are the body sizes, p and q are the distances of the specular point C_0 from the corresponding body edges, $r_f^{(1)}$ and $r_f^{(2)}$ are the radii (semi-axes) of the Fresnel zone (Figs. 2 and 3), then the weakening function, Eq. (9), in Eq. (8) is:

$$W \simeq 1. \quad (28)$$

The relation (28) follows from the asymptotic formula for the Fresnel integrals in Eq. (9) (Abramovitz and Stegun, 1970). The equality of the weakening function W to unity means that the scattering by the body surface S is the "pure" reflection (refraction) occurring according to the laws of the ray method. The impact of the body edges (or of the diffraction effect) is then negligible, so that the reflection (refraction) takes place in accordance with the ray method if:

a) the scattering body is large-scaled, i.e. its sizes are large as compared to the Fresnel zone [conditions (26)];

b) the source and receiver (points C^+ and C^- in Fig. 1) are located in the lit area, i.e. the corresponding specular point (point C_0 in Fig. 2) is far from the body edges or from the boundary of the geometrical shadow [condition (27)].

From both physical and practical points of view, the inequalities (26) and (27), which follow from the conditions of validity of the asymptotic formulae for the Fresnel integrals, are, however, too strict. According to the well-known Fresnel explanation, the leading part of the wavefield at some point is determined by the first Fresnel zone as the contributions of the following even and odd zones extinguish each other. This physical interpretation of weak impact of the following Fresnel zones on the wavefield could easily be explained by the properties of an integral with a rapidly varying integrand, such as type (5). Therefore, the practical conditions for the pure reflection (refraction) can be written in the form:

$$\Delta p \gtrsim r_f^{(1)}, \quad \Delta q \gtrsim r_f^{(2)}, \quad (29)$$

$$p \gtrsim r_f^{(1)}, \quad q \gtrsim r_f^{(2)}. \quad (30)$$

Conditions (29) and (30) are necessary, but they are not sufficient to provide the pure reflection. In addition, it is necessary that the factor F in the integrand of integral (5) be a slowly varying function. The conditions which provide this property of factor F can be written in different forms (Felsen and Marcuvitz, 1973; Bleistein and Handlesman, 1975). We will use the following approximate condition of validity of the method of stationary phase:

$$r_f^{(\gamma)} \frac{\partial \ln F(\xi_1, \xi_2)}{\partial \xi_\gamma} \ll 1, \quad (\gamma = 1, 2), \quad (31)$$

where F is the integrand of integral (5) without an exponential factor, $\xi_1 = p$ and $\xi_2 = q$ are the curvilinear coordinates on the body surface S or on the fronts $\Sigma^+(A_j)$ and $\Sigma^-(A_j)$ (Figs. 2–4).

In the case of the plane Q tangent to S (Figs. 2 and

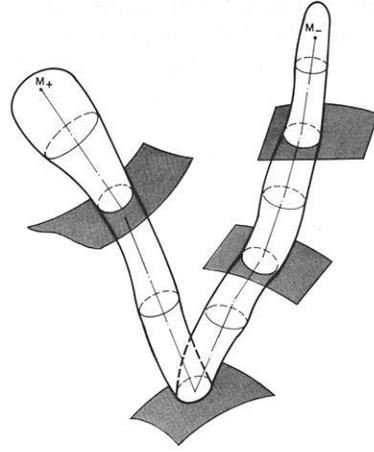


Fig. 5. The Fresnel volume for the reflected wave transmitted through three intermediate interfaces

3) or to Σ^+ and Σ^- (Fig. 4), the coordinate system ξ_1, ξ_2 could be replaced by the Cartesian:

$$\xi_1 = x, \quad \xi_2 = y. \quad (32)$$

The conditions of Eq. (31) impose some restrictions on the speed of variation of the wavefield and of the medium parameters. We can rewrite them in the form:

$$r_f^{(j)} \frac{\partial \ln I}{\partial \xi_j} \ll 1, \quad r_f^{(j)} \frac{\partial \ln K}{\partial \xi_j} \ll 1, \quad (33)$$

$$r_f^{(j)} \frac{\partial \ln v}{\partial \xi_j} \ll 1 \quad (j = 1, 2),$$

where I is the wave amplitude, K is the coefficient of reflection, v is the velocity of wave propagation.

These conditions, Eq. (33), of validity of the method of stationary phase also prove to be too strict from the practical (physical) point of view. The comparison of data obtained by calculation according to asymptotic formulae (in particular in the Kirchhoff approximation) and according to the exact numerical or analytical formulae, or by physical modelling, shows that the asymptotic formulae give a fairly good approximation when the conditions

$$r_f^{(j)} \frac{\partial \ln I}{\partial \xi_j} \gtrsim 1, \quad r_f^{(j)} \frac{\partial \ln K}{\partial \xi_j} \gtrsim 1, \quad (34)$$

$$r_f^{(j)} \frac{\partial \ln v}{\partial \xi_j} \gtrsim 1 \quad (j = 1, 2)$$

are met (Vainstein, 1957; Felsen and Marcuvitz, 1973; James, 1974; Zahradnik, 1977; Borovikov and Kinber, 1978; Gelchinsky and Karaev, 1980).

This consideration together with conditions (29) and (30) form the basis of the assertion that the Fresnel volume (zone) is the area essential for propagation (reflection). In other words, the Fresnel volume (zone) is the domain in which the wavefield coming from the source M_+ to the receiver M_- is formed (Figs. 2 and 4). From this fact in particular, it follows that the resolution of the seismic method is determined by the sizes of the Fresnel volume and the Fresnel zones surrounding the ray $M_+ M_-$ (Fig. 5).

Let us now consider the structure of the formulae (12), (22), (24) and (25) for the radii of the Fresnel zone and volume. The important peculiarity of the formulae obtained is their locally invariant form, i.e. that all the quantities included in them are characteristics of the medium and fronts at the centre of the zone (point C_0 or A_j on Figs. 2–4) and are independent of the choice of coordinate system.

The formulae (20) and (25) include the value g_0 . This quantity is the determinant value for the metric tensor of the scattering surface S [or the fronts $\Sigma^+(A_j)$ or $\Sigma^-(A_j)$] in the special coordinate system (p, q) at the specular point C_0 (Fig. 2) [or at the considered point A_j on the centre ray M_+M_- – Fig. 4]. The directions of the coordinate lines at the origin C_0 (or A_j) are tangent to the corresponding Fresnel areas (Fig. 2); owing to the local conditions, Eq. (18), this coordinate system can be called the quasi-Cartesian. Thus, the orthogonal curvilinear coordinate system (p, q) is determined by the orientation of the Fresnel zone on the surface S (or Σ^+ or Σ^-) and by the geometry of S (or Σ^+ or Σ^-).

The formulae (20) and (25) for the Fresnel radii have a symmetrized form, i.e. they do not change if the source and the receiver exchange places at the fixed points A^+ and A^- . These formulae include three types of terms: the first depends on the geometrical characteristic of the front Σ^+ , on its velocity and on the orientation of the plane of incidence E_{\parallel} ; the second depends on the characteristics of the Σ^- , v and E_{\parallel} ; and the third on the geometrical characteristics of the scattering (reflecting) surface and on E_{\parallel} .

The formulae obtained are essentially simplified in particular cases. For example, if the plane of incidence, E_{\parallel} , coincides with the principal normal sections of the fronts Σ^+ , Σ^- and the scattering surface S at the specular point C_0 , then the radii of the Fresnel zone are determined by the expressions:

$$\frac{1}{r_f^{(1)}} = g_0^{\frac{1}{2}} \left| \left(\frac{\cos^2 \theta^+}{\lambda^+ r_1^+} + \frac{\cos^2 \theta^-}{\lambda^- r_1^-} + \frac{\cos \theta^+}{\lambda^+ R_1^+} + \frac{\cos \theta^-}{\lambda^- R_1^-} \right) \right|^{\frac{1}{2}} \quad (35)$$

$$\frac{1}{r_f^{(2)}} = g_0^{\frac{1}{2}} \left| \left(\frac{1}{\lambda^+ r_2^+} + \frac{1}{\lambda^- r_2^-} + \frac{1}{\lambda^+ R_2^+} + \frac{1}{\lambda^- R_2^-} \right) \right|^{\frac{1}{2}},$$

where

$$r_1^{(v)} = r_{\parallel}^{(v)}, \quad r_2^{(v)} = r_{\perp}^{(v)},$$

$$R_1^{(v)} = R_{\parallel}^{(v)}, \quad R_2^{(v)} = R_{\perp}^{(v)}, \quad (v = + \text{ or } -). \quad (36)$$

The analogous formulae for the Fresnel volume (Fig. 4), in the case where the angle $\Delta\phi$ between the first principal normal section of the fronts $\Sigma^+(A_j)$ and $\Sigma^-(A_j)$ is equal to zero, take the form:

$$\frac{1}{r_f^{(1)}} = \left| \frac{1}{r_1^+} + \frac{1}{r_2^+} \right|, \quad \frac{1}{r_f^{(2)}} = \left| \frac{1}{r_1^-} + \frac{1}{r_2^-} \right|. \quad (37)$$

It should be noted that the level of difficulty in the calculation of the Fresnel radii $r_f^{(i)}$ ($i=1, 2$) according to formulae (12), (20), (24) and (25), is of the same order as that for the computations of the radii of curvature of the fronts Σ^+ and Σ^- . The procedure for the calculation of the curvature of wavefronts is described in various

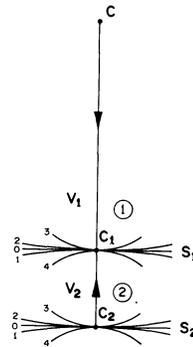


Fig. 6. Sketches of the cross-sections of the considered models in the plane of incidence E_{\parallel} : $v_1=2,500$ m/s; $v_2=4,000$ m/s; $CC_1=1,500$ m; $C_1C_2=500$ m; $R_1^{(i)}(N)$ is the N -th version ($N=0, 1, 2, 3, 4$) of the principal radius of curvature of the interface S_i ($i=1, 2$) in the plane $E_1=E_{\parallel}$; $R_1^{(i)}(0)=\infty$; $R_1^{(i)}(1)=4,000$ m; $R_1^{(i)}(2)=-4,000$ m; $R_1^{(i)}(3)=500$ m; $R_1^{(i)}(4)=-500$ m

works (see, for example, Gelchinsky, 1958, 1961; Červený and Ravindra, 1971; Deschamps, 1972; Shah, 1973; Hubral, 1980). Particularly useful in this sense is the work by Hubral and Krey (1980) in which these procedures are treated systematically in the context of solving the general problem of seismic prospecting.

We now present the numerical illustration of the results obtained. The model under consideration is the two layers on the half-space (Fig. 6). The principal normal planes E_1 to both interfaces S_1 and S_2 coincide. The 25 versions of the model corresponding to the five values of the radius of curvature, $R_1^{(i)}$, of the principal normal section to each interface S_i ($i=1, 2$) in the plane E_1 were considered (Fig. 6). It was also assumed that the radius of curvature, $R_2^{(i)}$ of each interface in the plane E_2 ($E_2 \perp E_1$) was the same for all versions of the model and was equal to infinity.

The Fresnel zones and volumes for the frequency $\nu=40$ Hz and for the central normal ray of waves reflected from S_2 were calculated. The thickness CC_1 or C_1C_2 of each layer in the normal direction does not change, therefore the zero time and the average velocity for the wave reflected from S_2 remain the same in each version considered. In the situation under consideration, the plane of incidence E_{\parallel} coincides with the principal normal plane E_1 for S_1 and S_2 , so that formulae (35) and (37) for the Fresnel zone and volume are valid.

Sizes computed for the half-axes, $r_f^{(i)}$, of the Fresnel ellipses in E_{\parallel} for the interfaces S_1 and S_2 are presented in Tables 1 and 2.

Table 1. The radii of the Fresnel ellipses on plane Q tangent to the interface S_1 in the plane E_{\parallel}

$R_2^{(2)}$ (m)	$R_1^{(1)}$ (m)				
	∞	4,000	-4,000	500	-500
∞	250	242	269	183	580
4,000	210	204	248	174	769
-4,000	298	263	312	206	306
500	201	193	208	165	291
-500	128	126	131	102	170

Table 2. The radii of the Fresnel ellipses on plane Q tangent to the interface S_2 in the plane E_{\parallel}

$R_1^{(2)}$ (m)	$R_1^{(1)}$ (m)				
	∞	4,000	-4,000	500	-500
∞	268	258	285	216	587
4,000	231	224	240	194	690
-4,000	335	312	368	246	357
500	136	134	138	116	164
-500	208	201	190	236	159

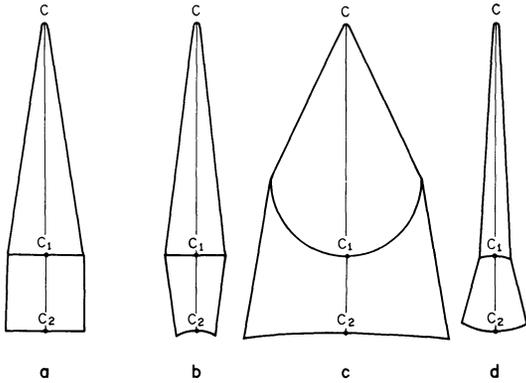


Fig. 7a-d. Examples of the cross-sections of the Fresnel volumes in E_{\parallel} for some of the models shown in Fig. 6: **a** Cross-section for model with $R_1^{(1)}(0) = R_1^{(2)}(0) = \infty$; **b** Cross-section for model with $R_1^{(1)}(0) = \infty$; $R_1^{(2)}(3) = 500$ m; **c** Cross-section for model with $R_1^{(1)}(4) = -500$ m; $R_1^{(2)}(2) = 4,000$ m; **d** Cross-section for model with $R_1^{(1)}(3) = 500$ m; $R_1^{(2)}(4) = -500$ m

In all versions the Fresnel radii $r_f^{(2)}$ in the plane $E_2 = E_1$ are equal to 250 m for S_1 and 268 m for S_2 .

The four examples of the calculated cross-section of Fresnel volumes in the plane E_{\parallel} are shown in Fig. 7. The calculated data show that the Fresnel zones and volumes can be essentially different for waves with a fixed central ray in the models with fixed values of interval velocities and time of wave propagation along the ray. The essential changes in the Fresnel zones and volumes can take place when the ray path crosses the interface with a large curvature. In the theory of wave propagation, it is accepted that the presence of inhomogeneities with large curvature (or with large gradients) along the ray results in essential decreases in the Fresnel zone (Tatarsky, 1967; Flatte, 1979). The data presented show that the intersection of the central ray with the surface of large curvature could lead to an increase or decrease in the Fresnel zones and volumes as compared to the case of smooth interfaces. The results obtained can be explained as the effects of strong focusing or defocusing of rays intersecting the interfaces with large curvature – for example, the Fresnel volume in Fig. 7c is essentially larger than that in Fig. 7a.

Such a decrease in the Fresnel zone is caused by strong defocusing of rays transmitted through the first surface S_1 with large curvature ($R_1^{(1)} = -500$). It is useful to note that the essential changes considered in the Fresnel zones and volumes are not isolated effects, but are also accompanied by strong variations in the kinematic and dynamic properties of the wavefield. In particular, the RMS velocities are also altered in these cases, although the zero time and average velocity remain constant.

In conclusion, it should be noted that the Fresnel radii are also important characteristics in cases where scattering by a body differs essentially from reflection (refraction). Generally speaking, this problem is the subject of special consideration and we wish only to point out here that in some of these cases [e.g. when the reflection (refraction) properties change rapidly over the length of the Fresnel radii, $r_f^{(i)}$], the complex parameter, $\rho_f^{(i)}$, characterizing the variation of a wavefield could be introduced (Gelchinsky, 1982b). This parameter is called the Fresnel parameter: its imaginary part is equal to the corresponding Fresnel radii, $r_f^{(i)}$, and its real part characterizes the speed of variation of the reflection properties. The behaviour and resolution of the wavefield depends on the relation between the imaginary part and the real part of the Fresnel parameter.

Acknowledgements. The author wishes to express his gratitude to Dr. Peter Hubral and Dr. Morris Podolak who read the manuscript and made valuable comments.

References

- Abramovitz, M., Stegun, I.: Handbook of mathematical functions with formulae, graphs and mathematical tables. Washington D.C., National Bureau of Standards, Applied Mathematics Series 55, 1970
- Alekseev, A.S., Gelchinsky, B.: On the ray method of computation of a wave field for inhomogeneous media with curved interfaces. In: Problems in the dynamic theory of seismic wave propagation (in Russian) **3**, 102–160, Leningrad University Press 1959
- Al'pert, Yal, Ginzburg, V.L., Fainberg, E.H.: Radio wave propagation (in Russian). Moscow Fizmatgiz, 1953
- Bertonei, H.L., Felsen, L.B., Hessel, A.: Local properties of radiation in lossy media. IEEE Trans. AP-19, N-3, 226–238, 1971
- Bleistein, N., Handelsman, R.: Asymptotic expansions of integrals. New York: Holt, Rinehard and Wiston 1975
- Born, M., Wolf, E.: Principles of optics. Oxford: Pergamon Press 1980
- Borovikov, V.A., Kinber, E.E.: Geometrical theory of diffraction (in Russian). Moscow Svjaz, 1978
- Červený, V., Ravindra, R.: Theory of seismic head waves. University of Toronto Press, 1971
- Deschamps, G.H.: Ray techniques in electromagnetics. Proceedings of the IEEE, **60**, 1022–1035, 1972
- Felsen, L.B., Marcuvitz, N.: Radiation and scattering of waves. New Jersey: Prentice-Hall 1973
- Flatte, S.M. (Ed.): Sound transmission through a fluctuating ocean. Cambridge University Press, 1979
- Gelchinsky, B.: Calculation of the spreading function of rays in layered homogeneous isotropic media. In: Materials of quantitative investigations of dynamics of seismic waves (in Russian), Vol. 3, Petrashen, G.I., ed.: pp. 295–305. Leningrad University Press, 1958
- Gelchinsky, B.: An expression for the spreading function. In: Problems in dynamic theory of seismic wave propagation (in Russian), Vol. 5, Petrashen, G.I., ed.: pp. 47–53. Leningrad University Press, 1961
- Gelchinsky, B.: Scattering of waves by a quasi-thin body of arbitrary shape. Geophys. J.R. Astron. Soc. **71**, 425–453, 1982a

- Gelchinsky, B.: Formulae for field of wave scattered by an inhomogeneous body and regions essential for reflection and propagation. Technical Program of 52nd Annual Meeting of SEG, 173-174, Dallas, 1982b
- Gelchinsky, B., Karaev, N.: Theoretical and model investigation waves scattered by quasithin bodies of arbitrary shape. *Ann. Geophys.* **36**, 509-518, 1980
- Hagedoorn, T.G.: The process of seismic reflection interpretation. *Geophys. Prospecting* **2**, 85-127, 1959
- Hilterman, F.T.: Three-dimensional seismic modeling. *Geophysics* **35**, 1020-1037, 1970
- Hubral, P.: Wavefront curvatures in three-dimensional laterally inhomogeneous media with curved interfaces. *Geophysics* **45**, 905-913, 1980
- Hubral, P., Krey, T.H.: Interval velocities from seismic reflection time measurements. Society of Exploration Geophysicists, Tulsa, 1980
- James, G.L.: Geometrical theory of diffraction. Peter Peregrinus, Stevenage, England, 1974
- Keller, J.B.: Diffraction by an aperture. *Appl. Phys.* **28**, 426-445, 1957
- Kleyn, A.H.: Seismic reflection interpretation. Applied Science, New York, 1983
- Kravtsov, Y.A., Orlov, Y.A.: The geometrical optics of inhomogeneous media (in Russian). Moscow: Nauka Press 1980
- Shah, P.M.: Use of wavefront curvature to relate seismic data with subsurface parameters. *Geophysics* **38**, 812-825, 1973
- Sheriff, R.E.: Nomogram for Fresnel-zone calculation. *Geophysics* **45**, 968-972, 1980
- Sheriff, R.E., Geldart, L.P.: Exploration seismology, Vol. 1. Cambridge University Press, 1982
- Tatarsky, V.I.: Wave propagation in a turbulent atmosphere (in Russian). Moscow: Nauka Press 1967. [Translated into English: The effects of the turbulent atmosphere on wave propagation. Israel Program for Scientific Translation, Jerusalem, 1971]
- Vainstein, L.A.: Electromagnetic waves (in Russian). Moscow Sovetskoe Radio, 1957
- Zahradnik, T.U.: SH-waves in the vicinity of rectangular impenetrable wedge. *Ann. Geophys.* **33**, 201-207, 1977

Received July 18, 1984; Revised version February 18, 1985

Accepted March 25, 1985

Stromfunktion des erdmagnetischen Hauptfeldes in der Quellschicht an der Kern/Mantel-Grenze

J. Meyer, J.-H. Hufen und M. Siebert

Institut für Geophysik der Universität Göttingen, Herzberger Landstr. 180, Postfach 2341, D-3400 Göttingen, Bundesrepublik Deutschland

Zusammenfassung. Das an der Erdoberfläche beobachtete Magnetfeld kann durch einen bekannten Formalismus im Rahmen der Potentialtheorie zurückgeführt werden auf eine äquivalente Stromverteilung für eine beliebig vorgegebene Kugel­fläche im Innern der Erde. Eine reale Bedeutung gewinnt ein solches Stromsystem, wenn es bezogen wird auf die Quellschicht nahe der Oberfläche des Erdkerns, in 3000 km Tiefe, deren Annahme nahegelegt wird durch das räumliche Spektrum der Energiedichte des Feldes. Auf der Grundlage neuerer Magsat-Ergebnisse (J.C. Cain's Feldmodell M051782) wird die Stromfunktion in dieser Tiefe global berechnet und die Stromverteilung graphisch dargestellt. Hierbei werden die Koeffizienten einer Kugelfunktionsentwicklung bis zu dem Grad und der Ordnung 12 berücksichtigt, sowohl unter Einbeziehung als auch bei Fortfall des Dipolanteils. Als der Realität am nächsten kommend wird dasjenige Stromsystem betrachtet, bei dem die äquatoriale Komponente des Dipols voll, die axiale Komponente aber nur zu rund einem Drittel einbezogen wird, gemäß einer Extrapolation des räumlichen Spektrums. (Der größere Teil des axialen Dipols wird vermutlich durch einen anderen Mechanismus erzeugt, der nicht direkt bzw. nicht allein in Verbindung steht mit Vorgängen innerhalb der Quellschicht an der Erdkernoberfläche.) Zum Vergleich wird die globale Verteilung des erdmagnetischen Hauptfeldes für den gleichen Bereich des räumlichen Spektrums herangezogen. Die Korrespondenz zur Stromverteilung wird besonders deutlich, wenn auch das Magnetfeld auf die Tiefe der Quellschicht reduziert wird. Dabei ist die Horizontalkomponente eng verknüpft mit der Stromdichte, d.h. mit dem Gradienten der Stromfunktion, die Vertikalkomponente mit der Stromfunktion selbst.

Current function of the geomagnetic main field in the source layer at the core/mantle boundary

Abstract. The geomagnetic field observed at the surface of the earth can formally be produced, through a well-known procedure, by an equivalent current distribution on any spherical surface inside the earth. A real physical significance can be attributed to the current system if it is referred to the source layer near the surface of

the earth's core, at a depth of 3000 km, as inferred from the spatial energy density spectrum of the field. Based on the latest Magsat results (J.C. Cain's field model M051782), the current function at this depth has been calculated and plotted incorporating terms up to order and degree 12, both with and without the dipole constituent. The current system regarded as the most realistic one includes the equatorial and, by extrapolation from the spatial spectrum, about one third of the axial dipole is presumably caused by another mechanism not directly connected with the revealed source layer.) For comparison, the global distribution of the main field is shown likewise. The equivalence is particularly apparent if the field is also reduced to the depth of the source layer. Its horizontal component is closely related to the surface density of the current, i.e. the gradient of the current function, and its vertical component to the current function itself.

Key words: Geomagnetic main field – Current function – Source layer – Spherical harmonic analysis

Einleitung

Seit der grundlegenden Arbeit von Gauß (1839), in der er festgestellt hat, daß die Quellen des Erdmagnetfeldes im wesentlichen ihren Sitz im Erdinnern haben, sind mannigfache Anstrengungen unternommen worden, den genauen Sitz dieser Quellen näher zu erschließen. Die Auswertung umfangreicher globaler Vermessungen im Verein mit theoretischen Erwägungen mündeten in die Erkenntnis, daß der Hauptteil des erdmagnetischen Innenfeldes dem flüssigen Erdkern entstammt und nur ein kleinerer Teil von magnetischen Mineralien in der Kruste herrührt, während der Erdmantel praktisch als quellenfrei angesehen werden kann. Die lange Zeit übliche Zweiteilung des Innenfeldes in Dipolanteil und Restfeld ließ eine weitergehende Trennung in Kern- und Krustenanteil jedoch nicht zu.

Ansätze zu einer solchen Trennung wurden erst gewonnen mit den erheblichen Fortschritten in der Vermessung des Feldes mittels Satelliten-Magnetometer – insbesondere durch die erste globale Vektorvermessung beim Magsat-Projekt 1979/80 – und der Beschreibung der globalen Feldverteilung durch das räumliche Spek-

trum der mittleren Energiedichte. Abgeleitet wird das räumliche Spektrum aus den Ergebnissen einer Kugelfunktionsanalyse der Meßdaten. Von Cain et al. (1984) ist eine derartige Analyse durchgeführt worden bis zu Termen vom Grad $n=29$. Das entsprechende räumliche Spektrum zeigt eine Aufteilung in zwei quasi-lineare Äste unterschiedlicher Neigung, von denen der eine (der „langwellige“ Spektralast) dem Kernfeld und der andere (der „kurzwellige“ Spektralast) dem Krustenfeld zugeschrieben werden kann. Der Übergang zwischen beiden liegt bei etwa $n=13$. Ein Vergleich mit analogen Ergebnissen für ein globales Krustenmodell-Feld zeigt, daß im Bereich niedrigerer Grade ein etwaiger Krustenanteil in allen Fällen klein ist gegenüber dem jeweiligen Kernfeldanteil (Meyer et al., 1983). Dadurch ermöglicht sich eine spezielle Untersuchung des Kernfeldes und seiner Quellen bereits nach dem pauschalen Abzug des Krustenanteils im räumlichen Spektrum des Innenfeldes.

Ungeachtet des besonderen Erzeugungsmechanismus kann angenommen werden, daß das aus dem Erdkern stammende Magnetfeld in Verbindung steht mit dort fließenden elektrischen Strömen. Eine Zurückführung des an der Erdoberfläche beobachteten Feldes auf eine äquivalente Stromverteilung ist in eindeutiger Weise möglich, wenn die Ströme beschränkt sind auf eine beliebige Kugelfläche (Chapman und Bartels, 1940). So hat bereits Vestine et al. (1947) für verschiedene Kugelradien äquivalente Stromverteilungen berechnet und dabei die Entwicklungskoeffizienten des Feldes bis zum Grad $n=6$ berücksichtigt. Eine realistische Bedeutung gewinnt ein solches Stromsystem aber erst, wenn unabhängige Indizien dafür vorliegen, daß die Ströme tatsächlich in einer bestimmten, einheitlichen Tiefe fließen. Der Befund, daß das räumliche Spektrum des Kernfeldes in einer Tiefe von rund 3000 km quasi „weiß“ ist, läßt in der Tat auf die Existenz einer relativ dünnen Stromschicht knapp unterhalb der Erdoberfläche schließen und rechtfertigt zugleich, ebendiese Schicht als Quellschicht zu bezeichnen (Meyer et al., 1983).

Die nachfolgenden Untersuchungen behandeln die in dieser Quellschicht fließenden Ströme im Vergleich zu der entsprechenden Verteilung des auf das gleiche Bezugsniveau reduzierten Feldes, unter Einbeziehung des vollen Kernfeldanteils, d.h. aller Terme bis einschließlich $n=12$. Da Strom und Feld sich gegenseitig bedingen, erhält man ein anschauliches Bild von der Säkularvariation des Feldes mit der Vorstellung von zeitlichen Änderungen der dargestellten Stromverteilung.

Berechnung der Stromfunktion

Sieht man von luftelektrischen Strömen ab, so läßt sich das Erdmagnetfeld \mathbf{B} überall außerhalb der Quellen darstellen als negativer Gradient eines Potentials V ,

$$\mathbf{B} = -\text{grad } V. \quad (1)$$

Dieses Potential wird in bekannter Weise an der Erdoberfläche (Erdradius $R_E = 6371$ km) entwickelt in eine doppelte Reihe von Kugelflächenfunktionen vom Grad n und der Ordnung m (Normierung nach A. Schmidt),

$$V = \sum_{n=1}^{\infty} V_n = R_E \sum_{n=1}^{\infty} \sum_{m=0}^n (g_n^m \cos m\lambda + h_n^m \sin m\lambda) P_n^m(\theta) \quad (2)$$

mit

$$V_n = R_E \sum_{m=0}^n (g_n^m \cos m\lambda + h_n^m \sin m\lambda) P_n^m(\theta) \quad (3)$$

($\lambda =$ geographische Länge, $\theta =$ Poldistanz).

Im folgenden wird angenommen, daß der äußere Anteil des Feldes durch eine Voranalyse bereits in Abzug gebracht worden ist. Demgemäß sind unter den Entwicklungskoeffizienten g_n^m und h_n^m in Gl. (2) und (3) allein die Koeffizienten des inneren Anteils zu verstehen.

Der Potentialanteil n -ten Grades, V_n , entspricht dem Feld eines bestimmten Multipoles (2^n -Pol) im Erdmittelpunkt. Die Verteilung der dieses Feld erzeugenden Flächenströme auf einer zur Erdoberfläche konzentrischen Kugel vom Radius $a_c < R_E$, d.h. im Innern der Erde, wird beschrieben durch die Stromfunktion $J_n(\theta, \lambda)$. Zwischen J_n und V_n besteht die Beziehung ($\mu_0 =$ Permeabilität des Vakuums)

$$J_n(\theta, \lambda) = \frac{1}{\mu_0} \frac{2n+1}{n} \left(\frac{R_E}{a_c}\right)^{n+1} V_n = \frac{1}{\mu_0} \frac{2n+1}{n} \left(\frac{R_E}{a_c}\right)^{n+1} \times R_E \sum_{m=0}^n (g_n^m \cos m\lambda + h_n^m \sin m\lambda) P_n^m(\theta). \quad (4)$$

Über die Herleitung s. Chapman und Bartels (1940). Allgemein ist die Stromfunktion so definiert, daß der Strom in Richtung der Linien gleichen J -Wertes fließt und die Differenz ΔJ zweier Isolinien die Stärke des zwischen ihnen fließenden Stromes angibt. Aus dem Abstand der Isolinien von J ist somit unmittelbar die Flächenstromdichte ersichtlich. Dabei ist das Vorzeichen so gewählt, daß der Strom im Uhrzeigersinn um ein relatives Minimum von J fließt. Eine bei dieser Definition zunächst noch zulässige willkürliche additive Konstante verschwindet durch die Festsetzung, daß die Stromfunktion im Mittel über die Kugel gleich Null sein soll.

Die Stromfunktion ist wie das Potential des Magnetfeldes eine skalare Ortsfunktion, bei der die Anteile mit unterschiedlichem n summiert werden können. Für eine weitestgehende Einbeziehung aller Bestandteile des Kernfeldes wird die Summe bis $n=12$ erstreckt, dem höchsten Multipolterm im räumlichen Spektrum des Innenfeldes, der noch ganz dem Kernfeld zugeschrieben werden darf. Im Prinzip erfolgt die Berechnung der gesamten Stromfunktion des Hauptfeldes damit in der Form

$$J(\theta, \lambda) = \sum_{n=1}^{12} J_n(\theta, \lambda) = \frac{R_E}{\mu_0} \sum_{n=1}^{12} \frac{2n+1}{n} \left(\frac{R_E}{a_c}\right)^{n+1} \times \sum_{m=0}^n (g_n^m \cos m\lambda + h_n^m \sin m\lambda) P_n^m(\theta). \quad (5)$$

Globale Stromverteilung in der Quellschicht

Die Auswertung der räumlichen Spektren von sechs neueren Feldmodellen, die unter Verwendung von Magsat-Daten abgeleitet worden sind, hat eine mittlere

Quellschicht-Tiefe von 147 km unterhalb der seismisch erschlossenen Kern/Mantel-Grenze ergeben (Meyer et al., 1983). In Anbetracht eines mittleren Fehlers von ± 50 km genügt es, für die Tiefe unterhalb der Erdoberfläche einen gerundeten Wert von 3,000 km anzusetzen, entsprechend einem Radius der Bezugskugel von $a_c = R_E - 3000$ km = 3371 km. Außerdem können die Ergebnisse dann direkt mit denen von Vestine et al. (1947) verglichen werden. Genauer gesagt, sind die auf diese Kugel bezogenen Flächenströme eine Projektion aller in der relativ dünnen Quellschicht nahe der Kern/Mantel-Grenze fließenden Ströme auf die zentrale Fläche dieser Schicht.

Nach Gl. (5) wurden die Werte der Stromfunktion in der Quellschicht-Tiefe für ein $1^\circ \times 1^\circ$ -Gitternetz über die ganze Erde berechnet, basierend auf den Koeffizienten des Feldmodells M051782 von Cain et al. (1984). Die Abb. 1 zeigt das Ergebnis in Form von Isolinien, die zugleich Stromlinien sind, unter Einbeziehung des vollen Dipolanteils. Linien mit positiven Werten von J sind durchgezogen, solche mit negativen Werten gestrichelt. Zwischen je zwei benachbarten Linien fließt ein Strom von 2×10^8 A. Zur Verdeutlichung ist die Richtung des Stromes durch Pfeile gekennzeichnet. Das Überwiegen des Dipolanteils im Hauptfeld äußert sich in der globalen Vorzugsrichtung der Ströme von Ost nach West. Darüber hinaus spiegelt sich jedoch nur wenig von der großräumigen Struktur des Oberflächenfeldes in der Stromverteilung wider. Dies ändert sich erst, wenn auch das Magnetfeld auf die Quellschicht-Tiefe reduziert wird (s.u.).

Die besondere Rolle, die das Dipolfeld ($n=1$) im räumlichen Spektrum des Feldes spielt, deutet auf einen Erzeugungsmechanismus hin, der – anders als bei den Feldanteilen mit $n>1$ – nicht nur eine relativ dünne Schicht nahe der Erdoberfläche betrifft, sondern offenbar größere Teile des Kernes einbezieht, wenn nicht gar den Kern als Ganzes. Es liegt deshalb nahe, eine äquivalente Stromverteilung in analoger Weise auch für das Magnetfeld nach Abzug des Dipolanteils zu berechnen. Dies geschieht, indem von vornherein in der Gl. (5) nur Koeffizienten mit $n \geq 2$ berücksichtigt werden. In der Tat zeigt das entsprechende Bild der Isolinien (Abb. 2) keinen bevorzugten zonalen Anteil mehr. Die Spitzenwerte der Stromfunktion in höheren Breiten sind auf weniger als die Hälfte ihres Betrages zurückgegangen und positive wie negative Werte gleichermaßen über beide Hemisphären verteilt. Überhaupt ist das Gesamtbild deutlich differenzierter geworden; eine Feststellung, die formal darin begründet liegt, daß ohne den Dipol die Anteile der höheren Multipole klarer hervortreten.

In Anbetracht des quasi-weißen Feld-Spektrums in der Quellschicht-Tiefe im Bereich der höheren Harmonischen ist jedoch nicht anzunehmen, daß die globale Stromverteilung in dieser Tiefe gänzlich ohne Dipolanteil bleibt. Als der Realität am nächsten kommend wird dasjenige Stromsystem betrachtet, bei dem der Anteil ersten Grades mit Hilfe der spektralen Regressionsgeraden von $n=2$ bis 12 extrapoliert wird. Hierbei wird die äquatoriale Komponente voll, die axiale Komponente aber nur zu einem Bruchteil einbezogen, der so bestimmt wird, daß der gesamte Term $n=1$ auf der Regressionsgeraden liegt. Das nach wie vor negative Vorzeichen

des betreffenden Koeffizienten gewährleistet weiterhin eine ausgeglichene Bilanz der Energiedichte bei der Säkularvariation des Quellschicht-Feldes (vgl. Meyer, 1985a, b).

Mit einer wiederum aus sechs neueren Feldmodellen abgeleiteten mittleren Regressionsgeraden ergibt sich für den extrapolierten Koeffizienten g_1^0 ein Wert von -11741 nT. Dies ist gut ein Drittel des mittleren Wertes von -30107 nT aus den verwendeten Feldmodellen direkt. Die damit berechnete Stromverteilung in der Quellschicht zeigt Abb. 3. Das Bild ähnelt dem der Abb. 2. Doch kommen gerade die großen Stromwirbel unter Nordamerika und Zentralasien bei der partiellen Berücksichtigung des Dipolanteils in Abb. 3 besser zum Ausdruck. Auffällig ist, daß die beiden entgegengesetzten Stromwirbel bei Südafrika in allen drei Darstellungen deutlich in Erscheinung treten.

Zusammenhang mit den Komponenten des Erdmagnetfeldes

Die Stromverteilung in der Quellschicht-Tiefe, wie sie in den Abb. 1-3 dargestellt ist, ist eingeführt als Quellfunktion des an der Erdoberfläche beobachteten Feldes. Um den Zusammenhang beider zu demonstrieren, wird das Potential V_n für den Feldanteil n -ten Grades im Abstand r vom Erdmittelpunkt ausgedrückt durch den entsprechenden Anteil J_n der Stromfunktion (vgl. Gl. (4)):

$$V_n(r, \theta, \lambda) = \mu_0 \frac{n}{2n+1} \left(\frac{a_c}{r}\right)^{n+1} J_n(a_c, \theta, \lambda), \quad r \geq a_c. \quad (6)$$

Die Komponenten der Kraftflußdichte an der Erdoberfläche sind

$$\begin{aligned} X_n(R_E, \theta, \lambda) &= \frac{1}{R_E} \left. \frac{\partial V_n}{\partial \theta} \right|_{r=R_E} \\ &= \mu_0 \frac{n}{2n+1} \left(\frac{a_c}{R_E}\right)^{n+1} \frac{1}{R_E} \frac{\partial J_n}{\partial \theta}, \end{aligned} \quad (7a)$$

$$\begin{aligned} Y_n(R_E, \theta, \lambda) &= -\frac{1}{R_E \sin \theta} \left. \frac{\partial V_n}{\partial \lambda} \right|_{r=R_E} \\ &= -\mu_0 \frac{n}{2n+1} \left(\frac{a_c}{R_E}\right)^{n+1} \frac{1}{R_E \sin \theta} \frac{\partial J_n}{\partial \lambda}, \end{aligned} \quad (7b)$$

$$\begin{aligned} Z_n(R_E, \theta, \lambda) &= \left. \frac{\partial V_n}{\partial r} \right|_{r=R_E} \\ &= -\mu_0 \frac{n}{2n+1} \left(\frac{a_c}{R_E}\right)^{n+1} \frac{n+1}{R_E} J_n. \end{aligned} \quad (7c)$$

Damit ergibt sich für die Energiedichte der Horizontalkomponente dieses Feldanteils

$$\begin{aligned} W_{H_n}(R_E, \theta, \lambda) &= \frac{1}{2\mu_0} H_n^2 = \frac{1}{2\mu_0} (X_n^2 + Y_n^2) \\ &= \frac{\mu_0}{2} \left(\frac{n}{2n+1}\right)^2 \left(\frac{a_c}{R_E}\right)^{2(n+1)} \\ &\quad \times \frac{1}{R_E^2} \left[\left(\frac{\partial J_n}{\partial \theta}\right)^2 + \frac{1}{\sin^2 \theta} \left(\frac{\partial J_n}{\partial \lambda}\right)^2 \right]. \end{aligned} \quad (8)$$

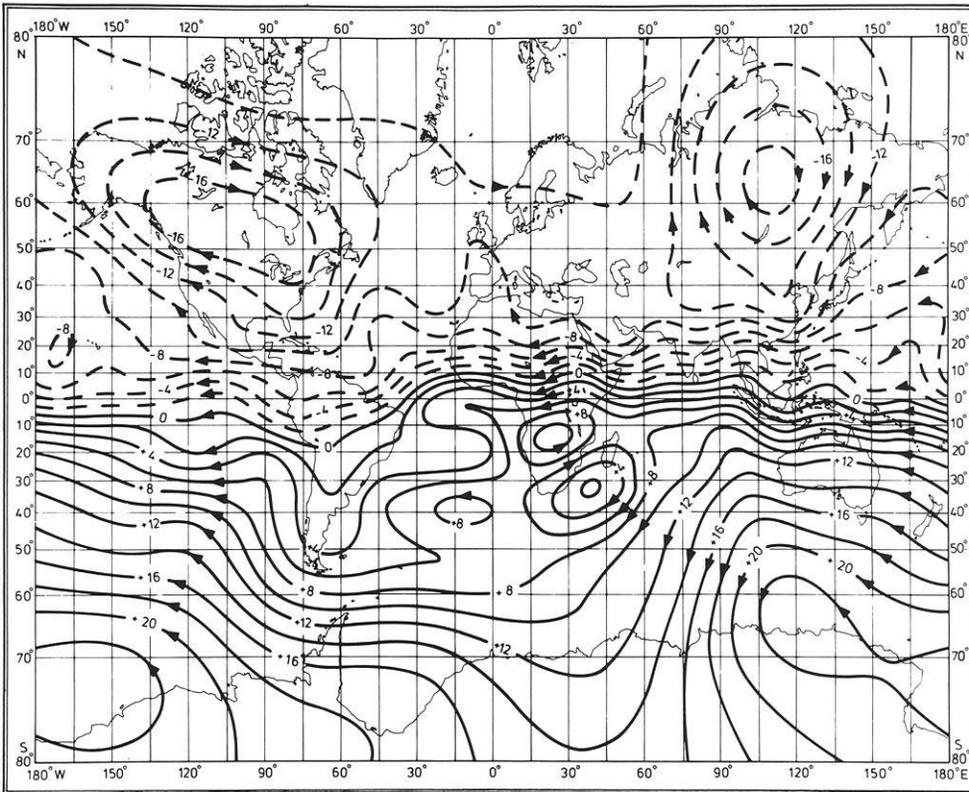


Abb. 1. Isolinien der Stromfunktion für das gesamte Hauptfeld ($n=1, \dots, 12$; Epoche 1980) in der Quellschicht-Tiefe von 3000 km. Einheit 10^8 A. Die Richtung des Stromes ist durch Pfeile gekennzeichnet

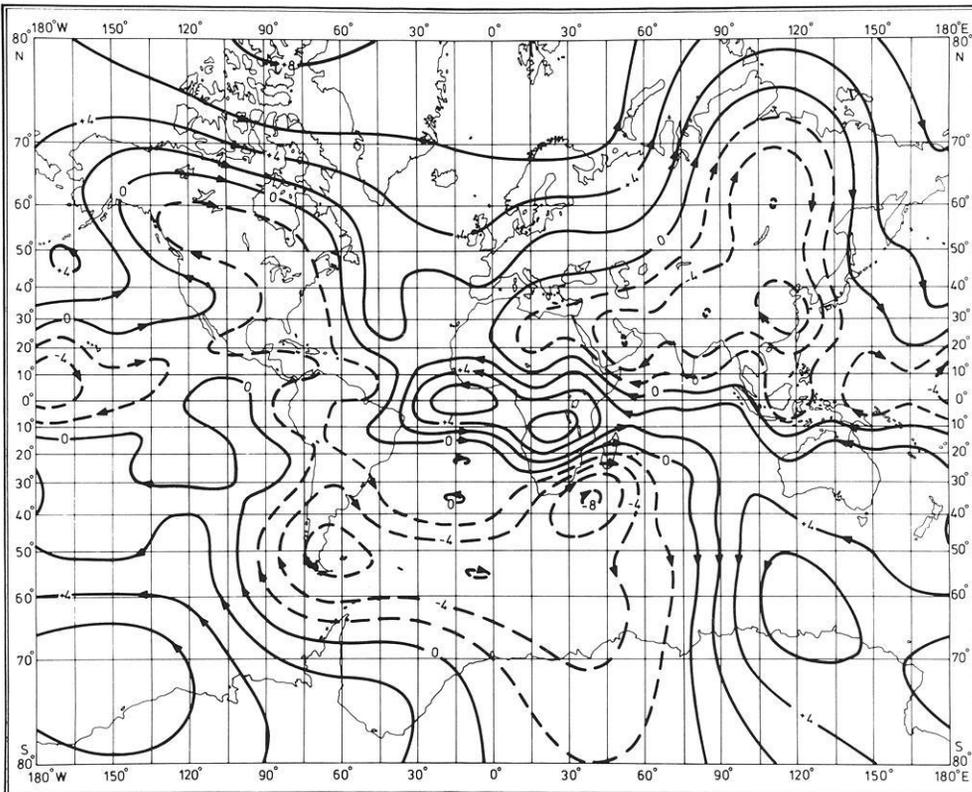


Abb. 2. Isolinien der Stromfunktion für das Hauptfeld ohne Dipolanteil ($n=2, \dots, 12$; Epoche 1980). Einheit 10^8 A. Ansonsten wie Abb. 1

Andererseits ist die zugehörige Flächenstromdichte an der Kugelfläche mit dem Radius a_c (\mathbf{e}_r = Einheitsvektor in radialer Richtung)

$$\mathbf{j}_n = -\mathbf{e}_r \times \text{grad } J_n \quad (9)$$

mit

$$j_n^2 = j_{\theta n}^2 + j_{\lambda n}^2 = \frac{1}{a_c^2 \sin^2 \theta} \left(\frac{\partial J_n}{\partial \lambda} \right)^2 + \frac{1}{a_c^2} \left(\frac{\partial J_n}{\partial \theta} \right)^2. \quad (10)$$

Daraus folgt als Beziehung zwischen dem Betrag der Flächenstromdichte j_n in der Quellschicht und der Energiedichte der Horizontalkomponente H_n des Ma-

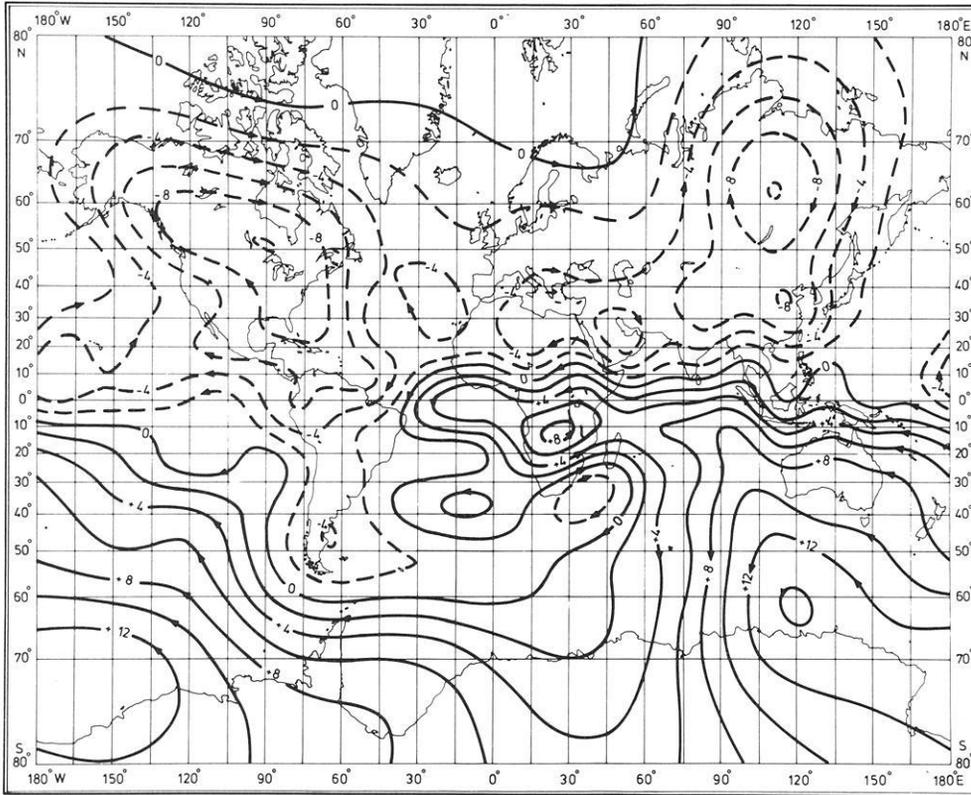


Abb. 3. Isolinen der Stromfunktion bis $n=12$ in der Quellschicht-Tiefe von 3000 km für das Hauptfeld mit extrapoliertem Dipolanteil (39% des axialen Dipolmomentes; Epoche 1980). Einheit 10^8 A

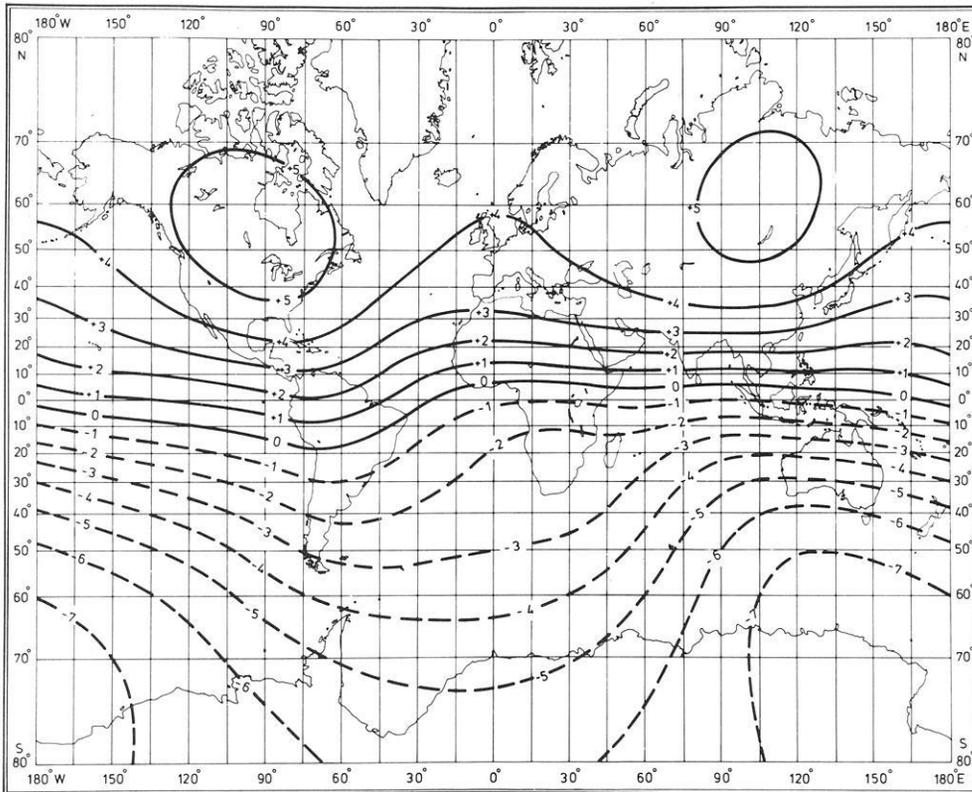


Abb. 4. Isodynamen der Vertikalkomponente des Hauptfeldes an der Erdoberfläche ($n=1, \dots, 12$; Epoche 1980). Einheit 10^4 nT

gnetfeldes an der Erdoberfläche:

$$W_{H_n}(R_E, \theta, \lambda) = \frac{\mu_0}{2} \left(\frac{n}{2n+1} \right)^2 \left(\frac{a_c}{R_E} \right)^{2(n+2)} j_n^2(a_c, \theta, \lambda). \quad (11)$$

Flächenstromdichte und Horizontalkomponente des Magnetfeldes bedingen sich gegenseitig. H_n ist am größ-

ten, wo auch j_n am größten ist, und umgekehrt. Dabei ist die Richtung von H_n stets senkrecht zu den Isolinen der Stromfunktion, positiv nach abnehmenden J_n -Werten hin.

In ähnlicher Weise bedingen sich Stromfunktion J_n und Vertikalkomponente Z_n des Magnetfeldes gegensei-

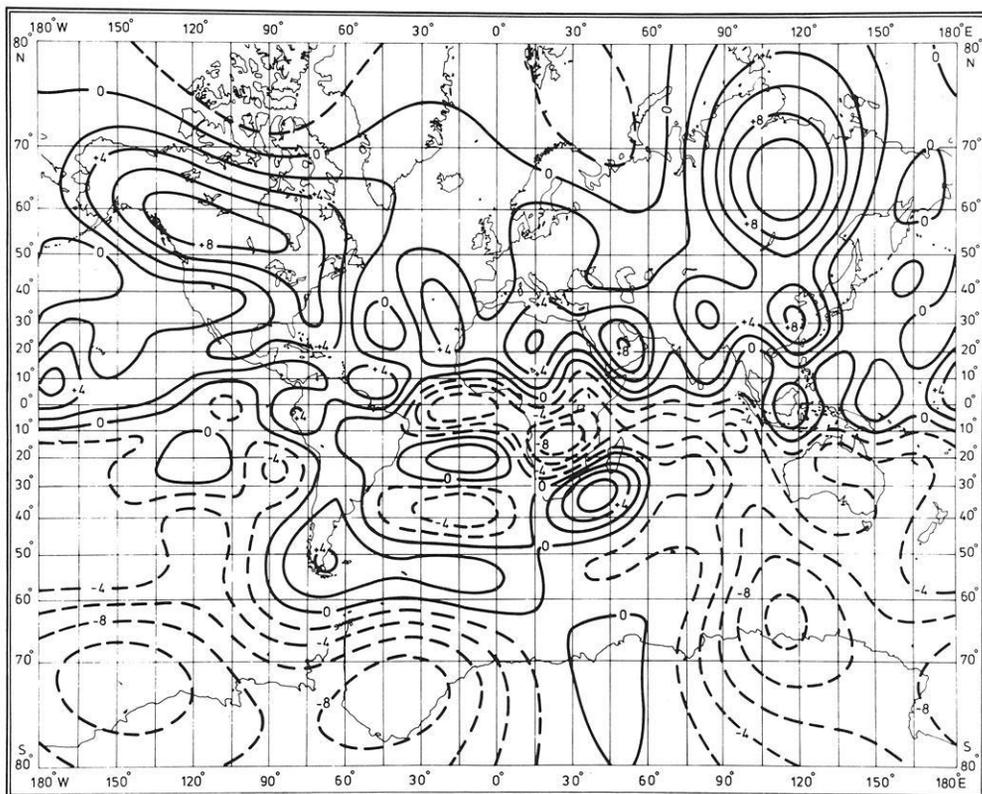


Abb. 5. Isodynamen der Vertikalkomponente des Hauptfeldes in der Quellschicht-Tiefe von 3000 km ($n=1, \dots, 12$; Epoche 1980) Einheit 10^5 nT

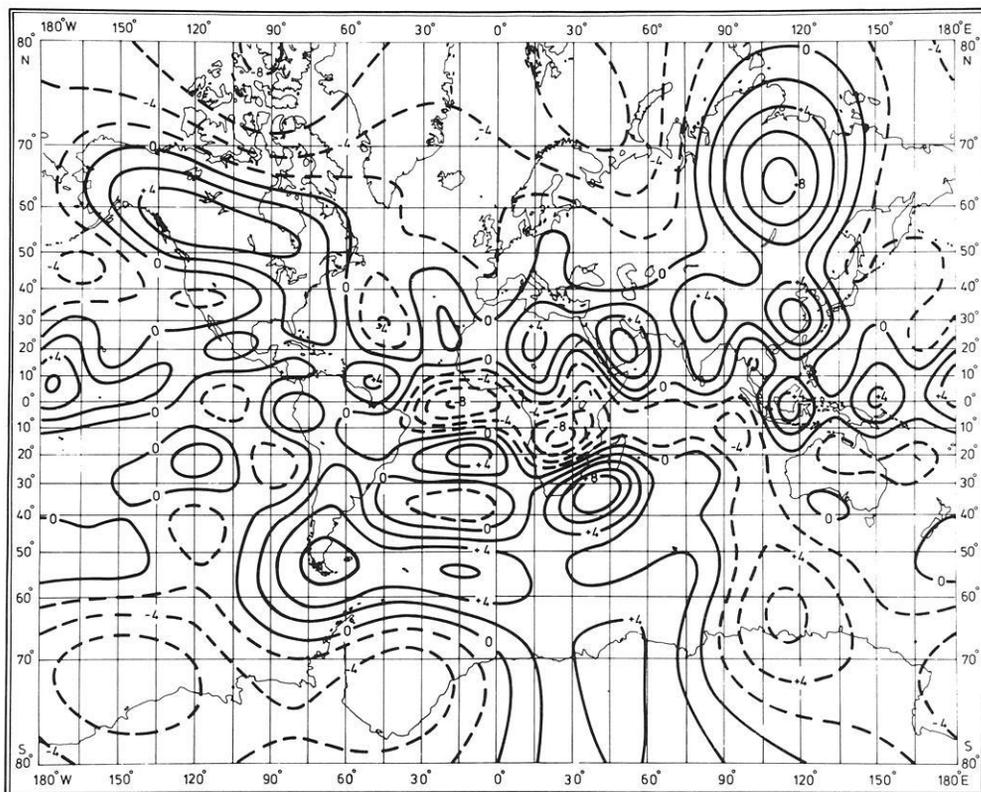


Abb. 6. Isodynamen der Vertikalkomponente des Hauptfeldes in der Quellschicht-Tiefe von 3000 km nach Abzug des Dipolanteils ($n=2, \dots, 12$; Epoche 1980). Einheit 10^5 nT

tig; vgl. Gl. (7c). An Stellen mit minimalem J_n ist Z_n maximal und umgekehrt. Für die betreffende Energiedichte der vertikalen Feldkomponente gilt

$$W_{Z_n}(R_E, \theta, \lambda) = \frac{1}{2\mu_0} Z_n^2$$

$$= \frac{\mu_0}{2} \left(\frac{n}{2n+1} \right)^2 \left(\frac{a_c}{R_E} \right)^{2(n+1)} \left(\frac{n+1}{R_E} \right)^2 J_n^2(a_c, \theta, \lambda). \quad (12)$$

Die Energiedichte des gesamten Feldanteils vom Grad n an der Erdoberfläche wird somit bestimmt sowohl

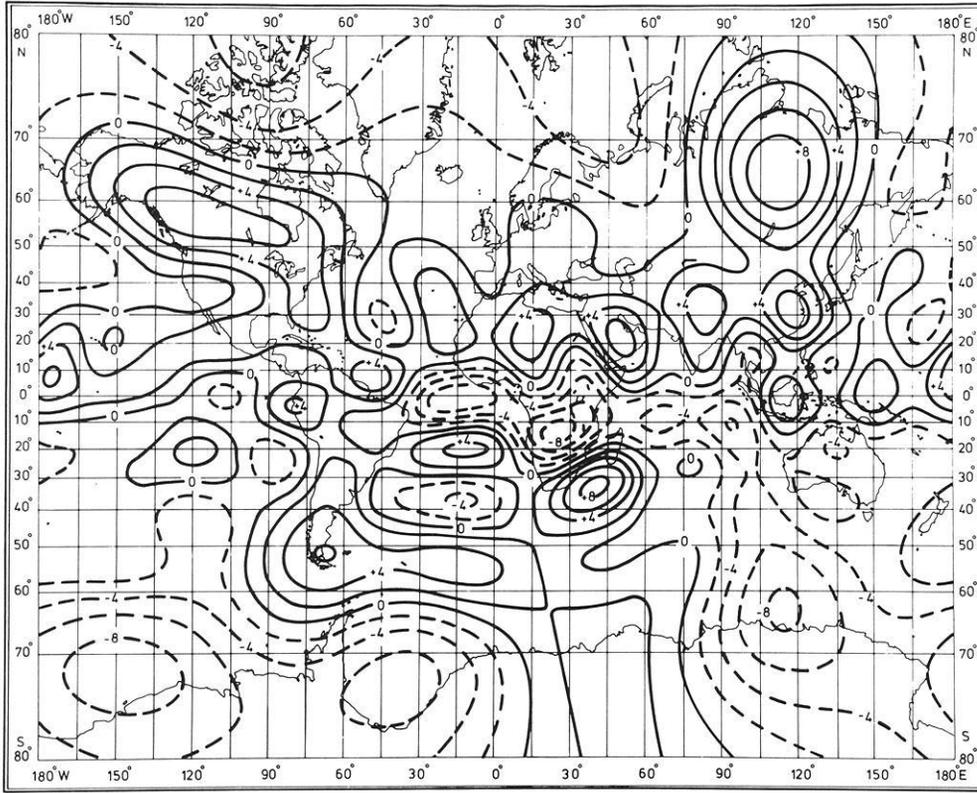


Abb. 7. Isodynamen der Vertikalkomponente des Hauptfeldes bis $n=12$ mit extrapoliertem Dipolanteil (39% des axialen Dipolmomentes; Epoche 1980) in der Quellschicht-Tiefe von 3000 km. Einheit 10^5 nT

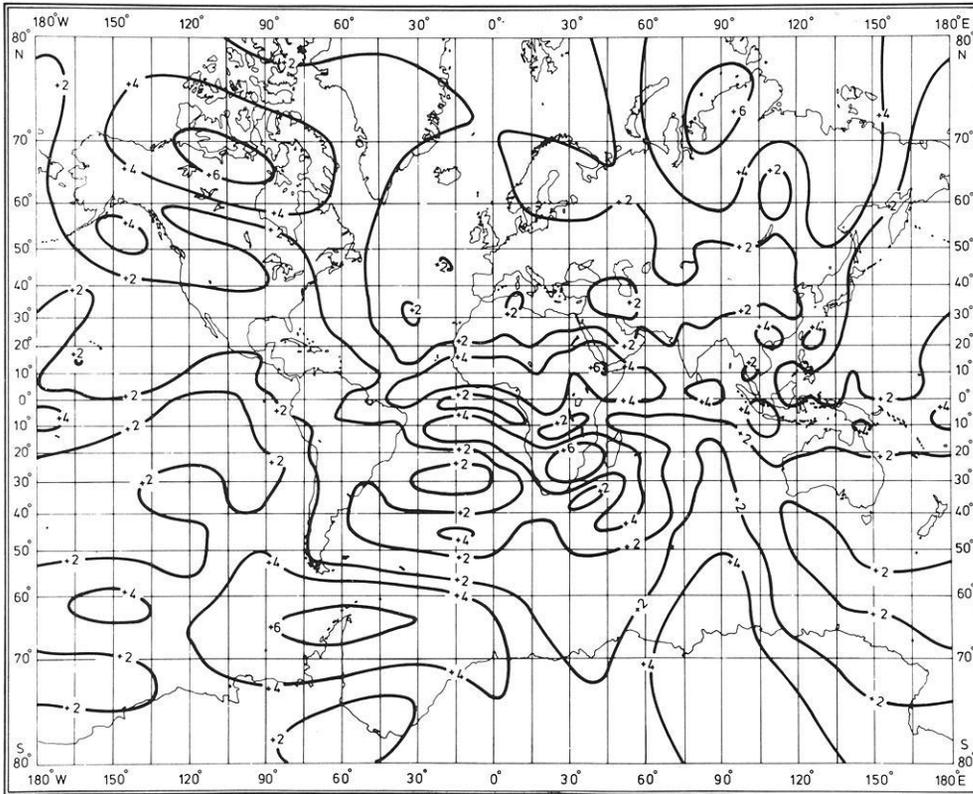


Abb. 8. Isodynamen der Horizontalkomponente des Hauptfeldes bis $n=12$ mit extrapoliertem Dipolanteil (39% des axialen Dipolmomentes; Epoche 1980) in der Quellschicht-Tiefe von 3000 km. Einheit 10^5 nT

durch den absoluten Wert der Stromfunktion am Quellschicht-Punkt mit den gleichen Koordinaten, als auch durch den dortigen Gradienten:

$$W_n(R_E, \theta, \lambda) = \frac{\mu_0}{2} \left(\frac{n}{2n+1} \right)^2 \left(\frac{a_c}{R_E} \right)^{2(n+2)}$$

$$\times \left[j_n^2(a_c, \theta, \lambda) + \left(\frac{n+1}{a_c} \right)^2 J_n(a_c, \theta, \lambda) \right]. \quad (13)$$

Es sei jedoch darauf hingewiesen, daß die aufgezeigten Zusammenhänge in strenger Form nur für Verteilungen eines bestimmten, wenngleich beliebigen Grades

n gelten. Sie gelten nicht mehr streng für die Summen von Verteilungen mit verschiedenem n oder gar für beliebige Verteilungen. Der Grund liegt in den unterschiedlichen, von n abhängigen Gewichtungsfaktoren zu den Ableitungen der Stromfunktion in der Formel für die Energiedichte des Feldes, beispielsweise der Horizontalkomponente:

$$W_H(R_E, \theta, \lambda) = \frac{\mu_0}{2R_E^2} \left\{ \left[\sum_n \frac{n}{2n+1} \left(\frac{a_c}{R_E} \right)^{n+1} \frac{\partial J_n}{\partial \theta} \right]^2 + \frac{1}{\sin^2 \theta} \left[\sum_n \frac{n}{2n+1} \left(\frac{a_c}{R_E} \right)^{n+1} \frac{\partial J_n}{\partial \lambda} \right]^2 \right\}. \quad (14)$$

Für die Flächenstromdichte dagegen ergibt sich

$$j^2(a_c, \theta, \lambda) = \frac{1}{a_c^2} \left\{ \left[\sum_n \frac{\partial J_n}{\partial \theta} \right]^2 + \frac{1}{\sin^2 \theta} \left[\sum_n \frac{\partial J_n}{\partial \lambda} \right]^2 \right\}. \quad (15)$$

Zwischen den entsprechenden Verteilungen des Gesamtfeldes an der Erdoberfläche und der zugehörigen Stromfunktion in der Quellschicht-Tiefe ist deshalb keineswegs eine Kongruenz, sondern lediglich eine Korrelation zu erwarten. In der Tat ist die Stromfunktion sowohl mit der Vertikal- als auch mit der Horizontalkomponente des Magnetfeldes eng korreliert.

Die Abb. 4 zeigt die Isodynamen der Z -Komponente des Hauptfeldes an der Erdoberfläche für den gleichen spektralen Bereich ($n=1, \dots, 12$) wie bei der Stromfunktion in Abb. 1. Die Dominanz des Dipolfeldes, die im wesentlichen bedingt ist durch die immer stärker werdende radiale Abnahme der einzelnen Spektraltermen mit wachsendem n , erschwert allerdings den direkten Vergleich mit der Stromverteilung in der Quellschicht. Nur die großen Stromwirbel unter Nordamerika und Zentralasien sowie am Rande der Antarktis sind in der Feldverteilung wiederzuerkennen.

Ein besserer Vergleich ist möglich, wenn auch das Magnetfeld reduziert wird auf die Tiefe der Quellschicht, in der alle Terme annähernd gleiches Gewicht haben (Abb. 5). Hier im gemeinsamen Bezugsniveau ist die Ähnlichkeit der Strom- und der Feldverteilung unverkennbar. Im Uhrzeigersinn umflossene Stromwirbel sind verbunden mit Zentren positiver Z -Komponente, entgegen dem Uhrzeigersinn umflossene Stromwirbel mit Zentren negativer Z -Komponente. Darüber hinaus finden sich hohe Werte von Z an Stellen mit starker Änderung der Stromdichte, d.h. des Gradienten von J .

Noch enger wird die Korrelation zwischen der J - und der Z -Verteilung, wenn man in beiden Fällen den Dipolterm ($n=1$) fortläßt (s. Abb. 6 im Vergleich zu Abb. 2). Dabei ändert der Abzug des Dipolanteils – anders als bei der Stromverteilung – die allgemeine Struktur der vertikalen Feldkomponente in niederen Breiten nur wenig. Er äußert sich hauptsächlich in einer verminderten Stärke der großen Zentren in höheren Breiten. Entsprechend gering ist die Änderung der Feldverteilung, wenn man vom Anteil ersten Grades gerade soviel einbezieht, wie sich durch Extrapolation des räumlichen Spektrums ergibt. Die Abb. 7 zeigt das so bestimmte Magnetfeld in der Quellschicht-Tiefe, das qualitativ mit der Stromverteilung in Abb. 3 verknüpft ist. Es dürfte ein realistisches Bild sein vom Feld der ge-

samten, unmittelbar in der Quellschicht fließenden Ströme.

Für die Horizontalkomponente H ist das Bild insofern etwas verwickelter, als diese nicht direkt proportional ist der Stromfunktion J , sondern verknüpft ist mit deren Gradienten (s. Gl. (11) bzw. (14)). Ermöglicht wird der Vergleich auch hier wieder erst nach Reduktion der Feldverteilung auf die Quellschicht-Tiefe von 3000 km. Der Kürze halber ist lediglich das Hauptfeld (bis einschließlich $n=12$) mit dem für die Quellschicht extrapolierten Dipolanteil dargestellt (Abb. 8 im Vergleich zu Abb. 3). Entsprechend den Erwartungen, liegen die Zentren maximaler Horizontalkomponente an den Stellen größter Stromdichte, während die Zentren der Stromwirbel sich im allgemeinen in minimalen H -Werten widerspiegeln. Besonders deutlich treten diese Verhältnisse hervor bei den beiden Stromwirbeln nahe Südafrika.

Diskussion

Das bisher rein formal bestimmbare Stromsystem zur Erzeugung des Erdmagnetfeldes kann den Anspruch weitestgehender Realität erheben, wenn es bezogen wird auf die Quellschicht nahe der Erdoberfläche, in 3000 km Tiefe, deren Annahme nahegelegt wird durch das räumliche Spektrum der mittleren Energiedichte des Feldes. Die besondere Bedeutung dieser Stromverteilung erwächst aus dem Sachverhalt, daß sie unmittelbar und in eindeutiger Weise bestimmt wird durch die gegebene Feldverteilung. Außer einem quellenfreien Erdmantel sind keinerlei zusätzliche Annahmen bzw. einschränkende theoretische Vorstellungen erforderlich. Daraus folgt eine direkte Verknüpfung der abgeleiteten Stromverteilung sowohl mit der Vertikal- als auch mit der Horizontalkomponente des beobachteten Feldes.

Der Darstellung von Details der Stromverteilung ist nur insofern eine Grenze gesetzt, als im räumlichen Spektrum ab $n=15$ der Krustenanteil des Feldes das Kernfeld verdeckt. Umgekehrt kann im unteren Spektralbereich, bis $n=12$, ein etwaiger Feldanteil aus der Kruste vernachlässigt werden, wie schon durch Extrapolation des eigentlichen Spektralastes für das Krustenfeld leicht abzuschätzen ist. In der Tat führt eine gehäufte Berücksichtigung des Krustenanteils im Bereich niedrigerer Grade, unter Verwendung eines globalen Krustenmodell-Feldes (Meyer et al., 1983), zu Korrekturen der Stromverteilung, die im allgemeinen innerhalb der Zeichengenauigkeit liegen und auf die deshalb hier verzichtet worden ist.

Eine Sonderstellung innerhalb des Kernfeldes kommt lediglich dem Dipolterm ($n=1$) zu, der von dem allgemeinen linearen Abfall des spektralen Hauptfeldastes in signifikanter Weise abweicht. Dies deutet darauf hin, daß das Dipolfeld und die mit ihm verknüpften Ströme wiederum aus zwei ursächlich getrennten Anteilen bestehen: einem ersten Anteil, der zusammen mit den höheren Multipolfeldern der erschlossenen Quellschicht entstammt, und einem zweiten Anteil, dem Hauptdipol, dessen Erzeugungsmechanismus vermutlich größere Teile des Erdkerns oder gar den Kern als Ganzes betrifft. Letzterer hat paläomagnetischen Untersuchungen zufolge offenbar eine achsenparallele Rich-

tung. Damit wird es möglich, durch Extrapolation des räumlichen Spektrums die Ströme für den aus der Quellschicht stammenden Dipolanteil voll mit in die Betrachtungen einzubeziehen. Das Bild der Abb. 3 stellt die gesamten Ströme dar, die ursächlich der Quellschicht nahe der Erdoberfläche zuzuordnen sind.

Eine gesonderte Behandlung von Quellschicht-Anteil des Dipolfeldes und Hauptdipol ist weiter gerechtfertigt durch das unterschiedliche Verhalten beider bezüglich der zeitlichen Änderung. Das räumliche Spektrum der Säkularvariation, bei dem – anders als bei dem räumlichen Spektrum des Feldes selbst – der Dipolterm voll mit in die lineare Regression einbezogen werden kann, läßt darauf schließen, daß die Ursache der gegenwärtig beobachteten Säkularvariation zur Gänze in den Änderungen der Quellschicht-Ströme zu sehen ist (vgl. Meyer, 1985a, b). Über die Verteilung der mit dem Feld des Hauptdipols verknüpften Ströme, die hierbei quasi als stationär anzusehen sind, kann vorerst nichts Genaueres gesagt werden. Überhaupt können diese Ströme wohl nur erschlossen werden in Verbindung mit Untersuchungen über den Mechanismus ihrer Erzeugung, der zugleich ihre langfristige Änderung, einschließlich ihrem Verhalten bei der Feldumkehr, einbezieht. So muß bis auf weiteres offenbleiben, wie groß der Anteil des Hauptdipols an den Strömen in der Quellschicht ist.

Die Abb. 3 zeigt die Stromverteilung, deren strukturelle Änderung unmittelbar im Zusammenhang steht mit der beobachteten Säkularvariation. Während die Säkularvariation des Feldes als räumliche Vektorverteilung nur unvollkommen in Form von Komponenten-Profilen graphisch dargestellt werden kann, läßt sie sich im Prinzip auf diese Weise durch eine „ebene“ Kugelflächen-Verteilung vollständig beschreiben.

Danksagung. Wir danken der Deutschen Forschungsgemeinschaft für die Förderung der vorliegenden Untersuchungen. Die numerischen Rechnungen wurden durchgeführt auf der Sperry 1100/83 der Gesellschaft für Wissenschaftliche Datenverarbeitung Göttingen (GWDG), bei der auch die Vorlagen für die Abbildungen angefertigt wurden. Herrn Dr. J.C. Cain vom U.S. Geological Survey in Denver/Colorado danken wir für die vorzeitige Überlassung der Koeffizienten des Feldmodells M 051782.

Literaturverzeichnis

- Cain, J.C., Schmitz, D.R., Muth, L.: Small-scale features in the earth's magnetic field observed by Magsat. *J. Geophys. Res.* **89**, 1070–1076, 1984
- Chapman, S., Bartels, J.: *Geomagnetism*, Vol. II. Oxford: Clarendon 1940
- Gauss, C.F.: Allgemeine Theorie des Erdmagnetismus. In: Resultate aus den Beobachtungen des magnetischen Vereins im Jahre 1838, hrsg. v. C.F. Gauss und W. Weber: pp. 1–57. Leipzig: Verlag d. Weidmannschen Buchhdlg. 1839
- Meyer, J.: Secular variation of magnetic mean energy density at the source-layer depth. *Phys. Earth Planet. Int.* **39**, 1985a (im Druck)
- Meyer, J.: The decrease of the geomagnetic dipole field as part of the general secular variation. *J. Geomag. Geoelectr.* **37**, 153–158, 1985b
- Meyer, J., Hufen, J.-H., Siebert, M., Hahn, A.: Investigations of the internal geomagnetic field by means of a global model of the earth's crust. *J. Geophys.* **52**, 71–84, 1983
- Vestine, E.H., Laporte, L., Lange, I., Scott, W.E.: *The geomagnetic field. Its description and analysis.* Publ. no. 580, Carnegie Institution of Washington 1947

Eingegangen 13. März 1985

Angenommen 6. Mai 1985

Measurement of magnetic susceptibility anisotropy in Buntsandstein deposits from southern Germany

Thomas Schultz-Krutisch¹ and Friedrich Heller²

¹ Institut für Geologie, Universität Würzburg, Pleicherwall 1, D-8700 Würzburg, Federal Republic of Germany

² Institut für Geophysik, ETH Zürich, CH-8093 Zürich, Switzerland

Abstract. The anisotropy of magnetic low-field susceptibility in the Triassic Plattensandstein formation (Upper Buntsandstein) from northern Bavaria has a typical sedimentary fabric. The anisotropy ellipsoids are strongly oblate with minimum susceptibility axes normal to sedimentary bedding. The directions of the maximum susceptibility axes are consistent with the NNE-NE-trending general sediment transport direction that is derived from geological field observations of cross-bedding structures in the sandstones. However, the very small intensity differences between maximum and intermediate susceptibility require extremely sensitive measurement techniques. Comparative measurements were made with a spinner magnetometer, a cryogenic magnetometer and a susceptibility bridge. Directionally, the most consistent results were obtained with the spinner magnetometer after it was stabilized by means of a low-pass active filter. The directional consistency of the anisotropy principal axes can be improved further by annealing the sandstones at 750° C in air. During this treatment a strongly magnetic, low-coercivity mineral phase – probably magnetite – is formed which enhances the degree of magnetic anisotropy as well as the bulk susceptibility. Low temperature measurements indicate that, in the natural unheated state, paramagnetic minerals contribute substantially to the low-field susceptibility of the sandstones at room temperature.

Key words: Rock magnetism – Low-field susceptibility – Anisotropy – Triassic red beds – Southern Germany

Geological introduction

For the past 50 years cross-bedding analysis has been one of the sedimentologist's tools to delineate transport directions in sandstones (cf. Wurster, 1958). Cross-bedding structures are generated by continual development of sediment fore-casts. In this process, sand bars in fluvial deposits progress in the current direction when sediment is deposited on the lee side of the sand bars so that the sediment strata dip very gently along the current direction. This direction can be derived from field measurements of the azimuth and dip of the bedding planes.

The Plattensandstein member of the Upper Buntsandstein formation in northwestern Bavaria is a typical cross-bedded fluvial sediment in which palaeocurrent direc-

tions have been studied previously (Vossmerbäumer et al., 1979; Teyssen and Vossmerbäumer, 1980). The sandstone sequence is about 30–40 m thick and consists of red, partly violet, well-stratified and well-sorted fine sandstones. In a few zones the sediments are not oxidized and therefore green coloured. Silt and clay layers of up to 15-cm thickness are interbedded with the sandstones. Horizontal bedding, platy and especially trough-like cross-bedding structures are the dominant types of stratification. The trough channels have concave basal planes which cut across silt- or clay-horizons.

In an outcrop at Dietenhan, near the town of Würzburg, where the uppermost 12–15 m of the Plattensandstein have been quarried, the channels are up to 4 m wide and 1 m thick (Fig. 1). In the lower part they are bedded parallel to the slightly inclined base, in the upper part the bedding sometimes changes to the horizontal. Some channels are filled asymmetrically. Multiple cross-cutting of different channels indicates positional changes of the river. Some channels were developed at the same time, probably due to a braided river system (Allen, 1965) during the Plattensandstein sedimentation. Additional sedimentary textures are oscillation ripple marks (with irregularly divided ridges), flute casts, current crescents and sporadically load casts.

The measured cross-bedding values (Fig. 1) have been taken from all accessible cross-bedded strata at Dietenhan. In order to resolve variable strike and dip of the cross-bedded layers, each layer was measured several times, if possible. The direction of sediment transport was calculated from field measurements of the dip azimuth of the foresets following the method developed by Wurster (1958). The observations of the dip azimuth on both flanks of the cross-bedded layers result in a bimodal distribution (Fig. 2b). The vectorial mean value of the bedding poles dips steeply towards SSW with a circular standard deviation (ψ 63) of 10°. From this value, the local palaeocurrent direction pointing towards N18°E is derived. The dip of the cross-bedded layers is always low, especially towards the core of the channels and averages about 11° (Fig. 2a). This value is typical for trough-like cross-bedding lamination.

The small dip angles and the often recognized horizontal bedding (Fig. 1) in the Plattensandstein formation make field measurements with a compass tedious, inaccurate or in some situations impossible. Thus the determination of the resulting mean current direction (Fig. 2) is strongly influenced by the quality of the measuring points as well as by their distribution which may be limited due to outcrop

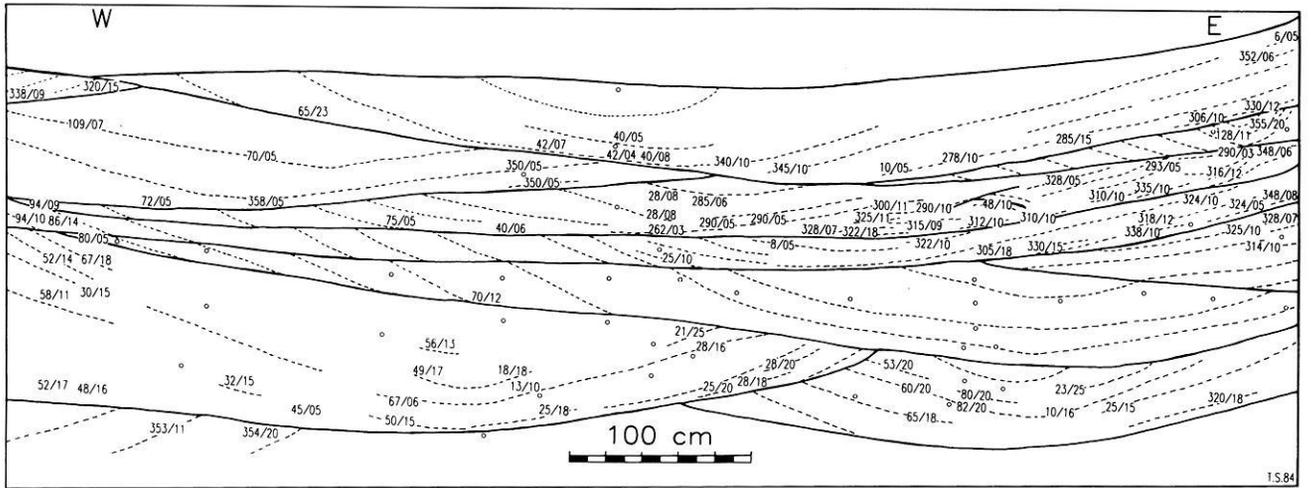


Fig. 1. Lithostratigraphic section of the Plattensandstein member at Diethan. *Solid lines*: boundaries of cross-bedded bodies; *dashed lines*: traces of bedding planes; *circles*: drill holes; *numbers* denote azimuth and dip of cross-bedding

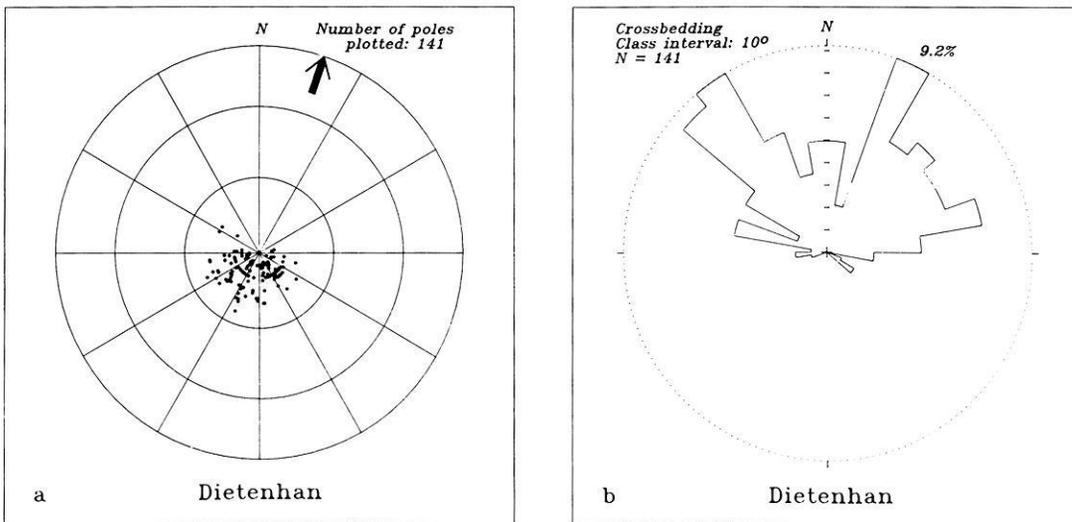


Fig. 2. **a** Orientation of cross-lamination foreset poles on equal area projection (lower hemisphere). The angular standard deviation (ψ_{63}) of the mean pole amounts to 10° . — *Arrow* represents resulting current direction to the NNE. **b** Rose diagram of the dip azimuth of cross-lamination foresets. Bimodal distribution is caused by the E and W dipping cross-bedding structures (cf. Fig. 1)

conditions. Even worse, it may be erroneous due to subjective selection criteria.

Undeformed sedimentary rocks have a magnetic fabric which is due to forces acting during and after deposition. Therefore, ancient current directions may be determined precisely by the measurement of magnetic susceptibility anisotropy (Rees, 1965). This paper describes magnetic measurement and analysis techniques applicable to the weak magnetic fabrics of red sandstones. We have drilled 40 oriented mini cores in the various sedimentary structures at Diethan (Fig. 1). In addition, two other nearby outcrops have been sampled, but most of the results presented in this paper refer to samples from the main quarry at Diethan. The cores have a diameter of 2.54 cm and are up to 10 cm long so that up to four specimens of 2.25-cm length can be cut from the same core. The length-to-diameter ratio has been chosen in order to avoid anisotropies arising from the shape of the specimens (Scriba and Heller, 1978).

Magnetic mineralogy of the Plattensandstein

The magnetic mineralogy of red sandstones is usually dominated by haematite which occurs either as specularite or as pigmentary grain aggregations (Turner, 1980). Microscopic observations of Plattensandstein polished sections indicate specular haematite to be the major ferromagnetic constituent with negligible additional amounts of ilmeno-haematite (with ilmenite exsolutions) and of unexsolved ilmenite.

The ferromagnetic minerals of the sandstones may also be identified from the analysis of the coercivity and blocking temperature characteristics of isothermal remanent magnetization (Dunlop, 1972). The shape of IRM acquisition curves and the fact that IRM is far from saturation in a 1 T field, indicates a high coercivity mineral as the main carrier of remanence (Fig. 3, Type I). Low coercivity phases are largely absent. Stepwise thermal demagnetiza-

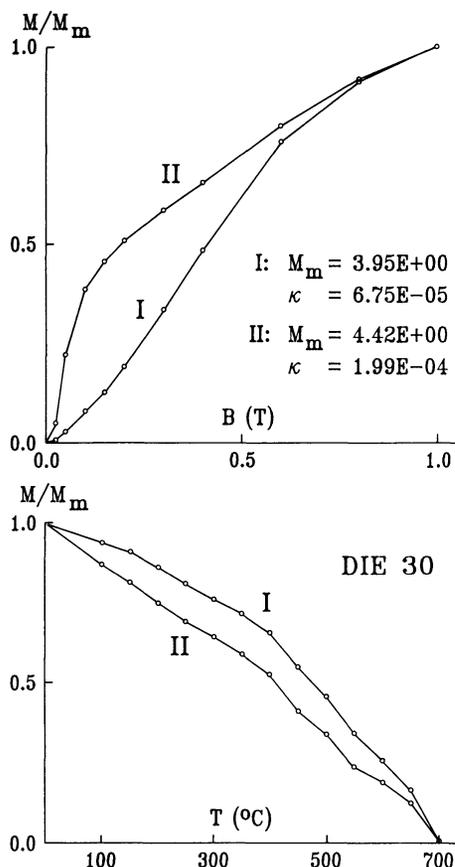


Fig. 3. IRM acquisition and subsequent progressive thermal demagnetization of sandstone samples in the natural state (I) and after annealing at 750° C (II). M_m denotes maximum IRM [A/m], k is the low-field bulk susceptibility [SI-units/cm³] before thermal demagnetization

tion of IRM shows maximum unblocking temperatures above 650° C which, together with the high coercivity, indicate haematite as the main ferromagnetic mineral.

When the sandstones are annealed in air at 750° C for one hour, the magnetic properties are altered drastically (Fig. 3, Type II). A low coercivity phase with maximum unblocking temperatures around 550° C is created. Following Stephenson (1967), who performed similar heating experiments with Old Red sandstones, we tentatively ascribe this phase to magnetite. The magnetite is responsible for the strongly increased bulk susceptibility (Fig. 3) after heating. The maximum IRM intensities at room temperature, however, have nearly the same values whether heated or not. At liquid nitrogen temperature, the IRM intensities of the heated samples are up to twice as strong as the room temperature IRM. This suggests that most of the magnetite formed during annealing is superparamagnetic at room temperature. For unheated sandstones, IRMs at liquid nitrogen temperature and at room temperature are not appreciably different, which indicates an absence of superparamagnetic mineral phases and, therefore, negligible superparamagnetic contributions to the low-field susceptibility of the sandstones in the natural state.

The source of the low-field susceptibility in red sediments is often ambiguous (Collinson, 1965). In addition to ferromagnetic sources, paramagnetic and diamagnetic components may be important. Collinson (1968) and Shive

et al. (1984) have shown by means of high-field experiments that the induced magnetization of red beds can be controlled by substantial amounts of paramagnetic material such as iron-rich clays, phyllosilicates and ilmenite.

The temperature dependence of the low-field susceptibility may give some information about the magnetic state of the minerals causing the measured susceptibility signal. The following technique was used to measure the susceptibility at variable temperature between liquid nitrogen and room temperature. A small-sized (4 cm³) cylindrical sandstone sample was put into a styrofoam box and cooled down to liquid nitrogen temperature. Then the sample was taken out of the liquid nitrogen and allowed to warm up slowly. Its susceptibility was measured at regular time intervals with a KLY-1 susceptibility bridge. After 75 min the sample reached room temperature. The temperature calibration was performed independently in the same manner with the sample outside the susceptibility bridge, since the susceptibility of any thermocouple would completely overshadow the weak susceptibility signal of a sandstone sample.

The temperature dependence of low-field susceptibility and its reciprocal have been plotted in Fig. 4 for two unheated sandstones. The samples are dominated by paramagnetic mineral contributions which lead to a linear increase of the reciprocal susceptibility between 77 K and 250 K. Towards room temperature, non-paramagnetic mineral phases become of greater importance resulting in a non-linear reciprocal susceptibility curve. However, the extrapolated linear regression segments intercept the temperature axis at negative values. These negative Néel points, at first inspection, are suggestive of antiferromagnetic ilmenite which has a Néel temperature around -210° C. An ilmenite content (bulk susceptibility value from Bleil and Petersen 1982) of more than 1% by volume would be needed for the observed signal. However, such an amount of ilmenite is incompatible with the microscopic evidence. Microscopic examination showed that biotite, clinocllore and traces of ilmenite are present in these Plattensandstein samples. We prefer an interpretation which assigns the paramagnetic susceptibility to the biotite and clinocllore. Since the basal planes of the biotite crystals lie within the bedding planes, they may indeed be responsible for a "detritic" anisotropy.

If the amount of ilmenite, as suggested by the optical examination, is negligible and if the ferromagnetic susceptibility can be represented as temperature-independent constant between liquid nitrogen and room temperature, then an approximate estimate can be made (Table 1) about the contributions of diamagnetic, paramagnetic and ferromagnetic minerals to the measured susceptibility. Since diamagnetic minerals such as quartz and feldspar make up nearly 100% of the sandstones, a constant diamagnetic value can be added to the measured bulk susceptibility. A rough estimate of the ferromagnetic susceptibility is then obtained by shifting the signal until the regression line hits the origin of the temperature axis.

In both samples the paramagnetic minerals predominate over the other two contributions (Table 1). The ferromagnetic susceptibility of sample DIE38A is about 45% higher than that of the sample WES02B. It is interesting to note that the same relation holds for the IRM intensities (Table 1) and thus confirms the derivation of a higher content of ferromagnetic minerals, i.e. of haematite, in sample DIE38A.

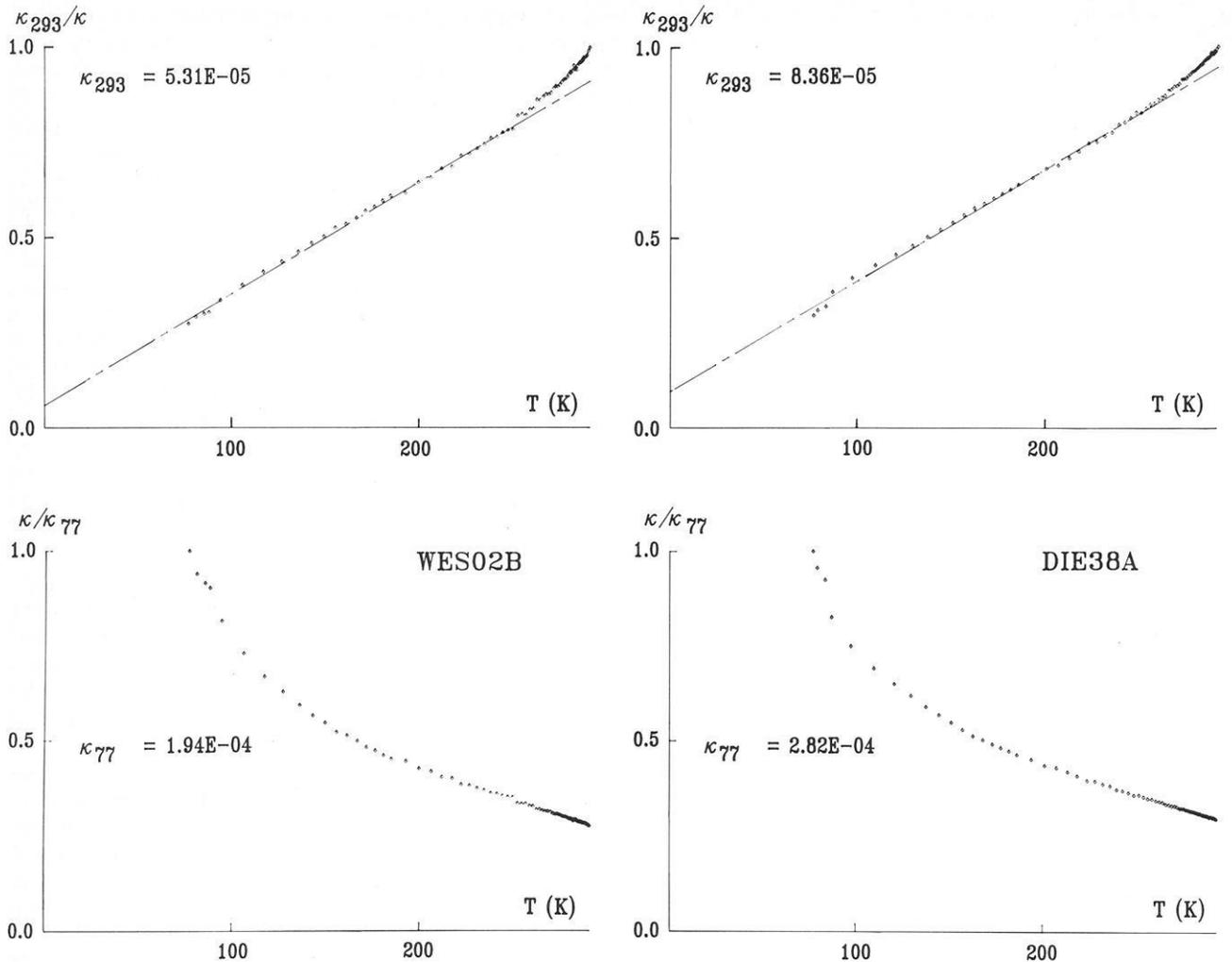


Fig. 4. Temperature dependence of low-field susceptibility and reciprocal of low-field susceptibility for two unheated red sandstone samples. Regression lines were computed between 77 K and 250 K. Values in SI-units/cm³

Table 1. Estimates of paramagnetic, diamagnetic and ferromagnetic susceptibility contributions and IRM intensity (acquired in a 1 Tesla field) of two unheated sandstone samples

Sample	Susceptibility [SI-units/cc]				IRM [A/m]
	Measured	Paramagnetic	Diamagnetic	Ferromagnetic	
DIE38A	8.36E-05	5.69E-05	-1.53E-05	4.20E-05	2.02
WES02B	5.31E-05	4.09E-05	-1.53E-05	2.83E-05	1.41

Magnetic susceptibility anisotropy

The magnetization (\mathbf{M}_i) of a sample which is placed in a magnetic field \mathbf{H}_i can be characterized by

$$\mathbf{M}_i = k_{ij} \mathbf{H}_i$$

with the susceptibility k_{ij} being a tensor of the second order, which can be described geometrically as an ellipsoid with the three principal axes k_{\max} , k_{int} and k_{\min} . The degree of anisotropy is often expressed by the ratios of the intensities of the principal axes,

$$P_1 = k_{\max}/k_{\min},$$

$$P_2 = k_{\max}/k_{\text{int}},$$

$$P_3 = k_{\text{int}}/k_{\min}.$$

Ising (1942) established that the minimum axes of the anisotropy ellipsoid in Swedish varved clays are oriented perpendicular to the bedding planes. Various natural and laboratory produced depositional sedimentary fabrics were investigated by Rees (1961, 1965, 1968), Rees et al. (1968), Hamilton (1967) and Hamilton and Rees (1971) measuring magnetic susceptibility anisotropy. Generally, the orientation of magnetic particles depends on the earth's gravity field, the strength of the depositing current, the dip of the bedding plane, the shape of a sediment particle and the strength and direction of the geomagnetic field (Hamilton et al., 1968). The laboratory experiments of these authors proved that gravity and hydrodynamic forces are much more efficient than the influence of the geomagnetic field for the particle settling in fine grained sandstones. They

Table 2. Anisotropy data obtained by different measuring techniques

Instrument	k_{\max}				k_{\min}				P_1	P_2	Susc.	N
	Az	Dip	ψ_{63}	k	Az	Dip	ψ_{63}	k				
Digico without filter	0.6	-1.7	34.9	5.4	107.7	87.7	13.7	34.8	1.068	1.014	6.309	69
Digico with filter	46.8	1.0	28.6	8.0	185.8	89.0	9.1	79.7	1.050	1.009	6.335	69
ScT	21.6	3.4	41.5	3.8	223.0	85.3	8.7	86.7	1.055	1.007	6.211	69
KLY-1 manual	54.2	0.8	48.5	2.8	338.1	83.8	7.8	107.3	1.054	1.008	5.315	8
KLY-1 on-line	27.8	4.9	33.3	5.9	356.6	84.7	5.4	221.4	1.050	1.006	5.128	8
Digico with filter	39.8	2.7	24.9	10.6	352.8	87.6	6.4	162.0	1.044	1.006	6.119	8
Digico unheated	42.0	1.3	21.8	13.8	248.5	88.9	9.8	68.5	1.050	1.009	6.583	39
Digico heated	38.5	0.6	18.9	18.4	200.3	89.2	7.6	113.5	1.192	1.025	19.280	42

Azimuth (Az) and dip of k_{\max} and k_{\min} mean directions with circular standard deviation (ψ_{63}) and Fisher (1953) precision parameter (k). Mean volume susceptibilities (Susc.) in SI-units/ 10^{-5} from N samples

produce preferential alignment of grain short axes normal to bedding and grain long axes parallel to current flow.

The relation between palaeocurrent directions and susceptibility anisotropy in different types of undeformed sandstones has been investigated by several authors (Crimes and Oldershaw, 1967; Galehouse, 1968; Rad, 1970; Hrouda and Janak, 1971; Hamilton and Rees, 1971; Argenton et al., 1975; Van den Ende, 1975; Channell et al., 1979). Hrouda and Janak (1971) found that P_2 factors were always very low ($1.0 < P_2 < 1.01$) in red sandstones. The maximum anisotropy directions were often poorly defined, but generally were oriented parallel to sedimentological features such as cross-bedding foresets, ripple marks or flute casts (Rad, 1970). According to Hrouda and Janak (1971), the P_3 factors are always higher than P_2 ($P_3 > 1.01$), representing an anisotropy to be expected for a sedimentary or compaction texture. On the other hand, in Permian red sandstones where the magnetic fabric was thought to be caused by the alignment of platy haematite crystals, Van den Ende (1975) found that the maximum axes of susceptibility were oriented perpendicular to the current direction markers and great circle distributions of these axes could often not be interpreted.

The differences amongst these observations and the poor definition of the orientation of maximum anisotropy axes may result partly from inaccurate measuring methods. Therefore, when measuring the Plattensandstein magnetic fabric, several magnetometers and different measurement techniques were applied and tested.

Different techniques for magnetic anisotropy measurement of Plattensandstein samples

1. The Digico spinner magnetometer measures relative susceptibility differences in three orthogonal planes of a sample. To get the absolute anisotropy values, the axial bulk susceptibility is measured with a susceptibility bridge (KLY-1) and combined with the anisotropy values. Both instruments use alternating current methods. Two series of measurements were performed with the spinner magnetometer (effective applied field: 0.246 mT at 10 kHz). Initially, the cylindrical sandstone samples were measured with the instrument in its commercial configuration. In a second

measuring series, an accessory Krohn-Hite low-pass active filter (amplification: 20 dB; cut-off frequency: 16 kHz) was adapted to the instrument to improve its sensitivity and stability.

2. To determine magnetic anisotropy with the KLY-1 susceptibility bridge, the bulk susceptibility of cube-shaped samples is measured in 15 different positions according to Jelinek's (1973, 1977) measuring scheme. Again two measurement series were made. The first series was based on manual balancing of the bridge using its potentiometers, whereas the second series utilized the unbalanced output voltage of the bridge which then was connected on-line with the laboratory computer. In the latter configuration the instrumental noise is reduced appreciably by signal stacking and instrumental drift can be easily compensated for.

3. The ScT cryogenic magnetometer can be used to measure absolute susceptibilities by trapping a constant magnetic field in the instrument (applied field used in this study: 0.045 mT) during cooling through the superconducting critical point (Scriba and Heller, 1978). The total magnetization is recorded for nine positions, 45° apart, in each of three orthogonal planes. The signals are processed on-line so that the anisotropy ellipsoid can be calculated after subtraction of the remanence signal and correction for instrumental drift.

4. Finally, a set of 44 sandstone samples was measured after annealing at a temperature of 750°C , which had been observed to increase the low-field bulk susceptibility.

The results of the different measuring techniques are presented in Figs. 5–7 and Table 2. Figure 5 shows data from 69 sandstone samples measured (a) with the unmodified spinner magnetometer, (b) with this instrument utilizing the low-pass active filter and (c) with the cryogenic magnetometer. The susceptibility minima cluster very well around a direction nearly normal to bedding, as expected for a sedimentary fabric. When the mean direction of each principal anisotropy axis is calculated independently (although this is mathematically not strictly correct), small circular standard deviations (Table 2) result for the minimum axes. This is because the oblateness of the anisotropy ellipsoids is always strong with a 5%–7% degree of anisotropy between maximum and minimum susceptibility (P_1 factors in Table 2). The ratio of maximum to intermediate anisotropy axes (P_2), however, only rarely exceeds 1%. The

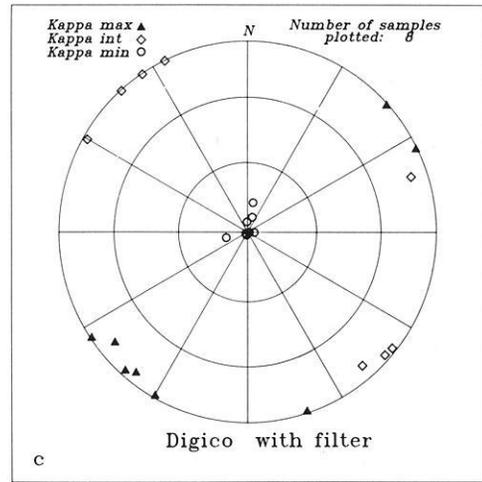
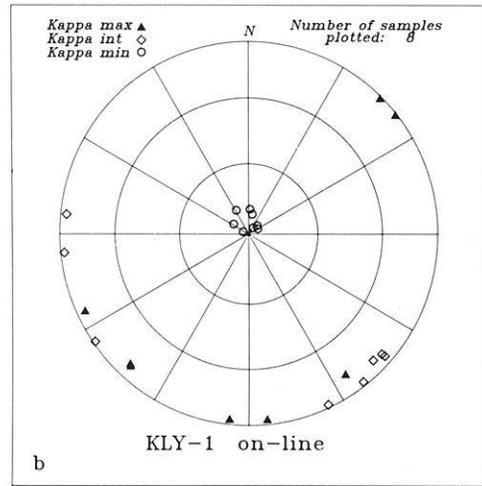
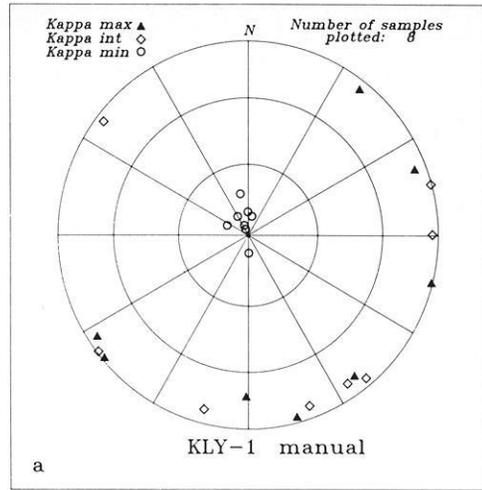
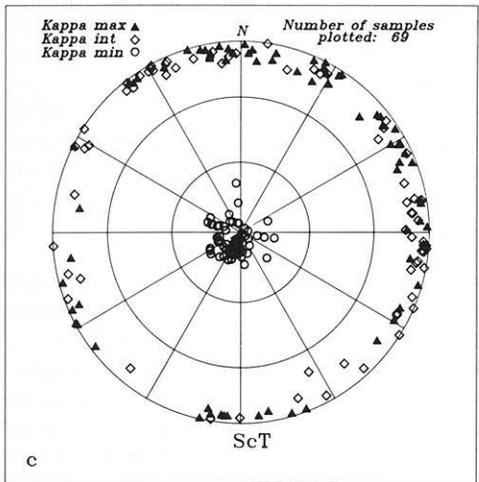
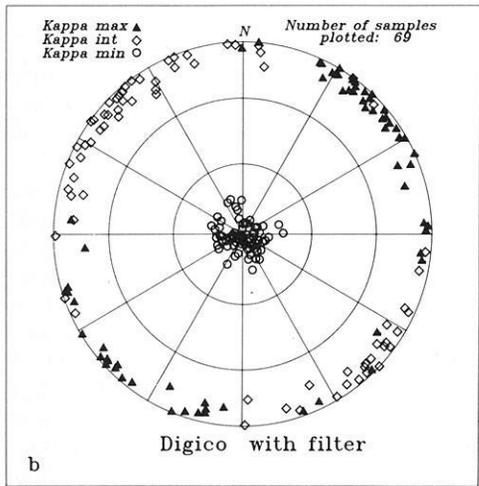
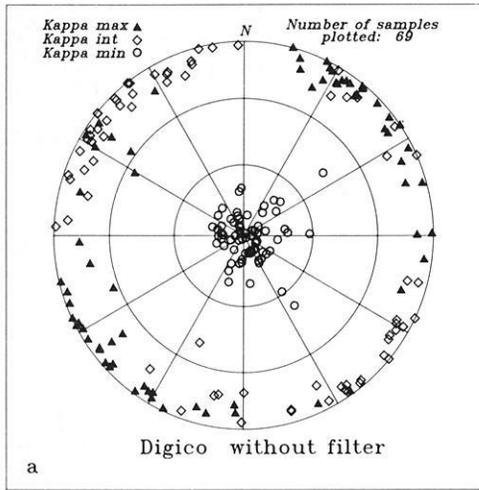


Fig. 5a-c. Anisotropy data measured **a** with the spinner magnetometer in its commercial configuration, **b** stabilized with a low-pass active filter and **c** with the cryogenic magnetometer. Most consistent data are obtained with the stabilized spinner magnetometer

Fig. 6a-c. Comparison of AC bridge and spinner magnetometer anisotropy results. **a** "manual", **b** "on-line" bridge measurements and **c** stabilized spinner magnetometer

directional scatter in the horizontal plane of the unfiltered Digico and the ScT measurements is high and the maximum axes are distributed along a great circle, with a slight preferred clustering about a NE-SW axis. Measurements with these techniques cannot be used to determine current direc-

tions in the sandstones with sufficient precision. Implementation of the low-pass active filter in the Digico spinner magnetometer lowers the instrumental noise level by a factor 10-100. This results in a distinct reduction of the scatter of the maximum susceptibility directions. Now they group

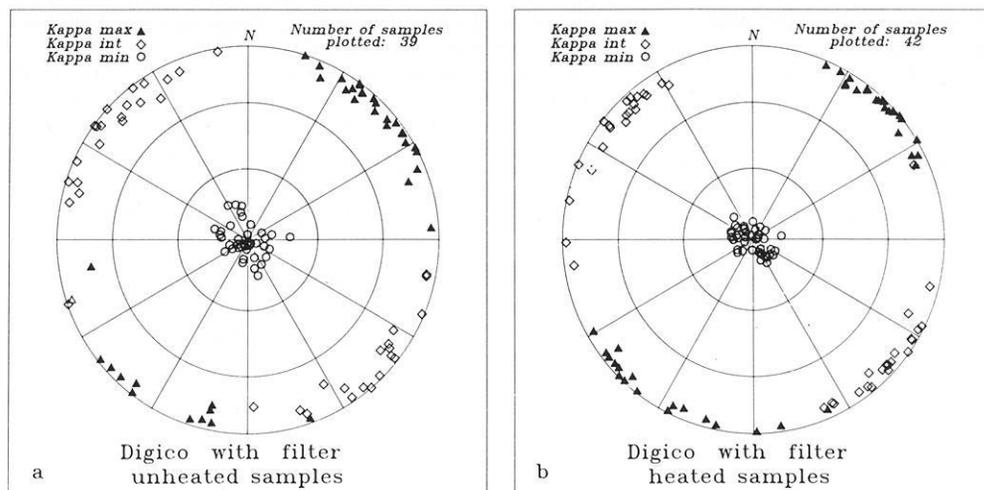


Fig. 7a, b. Anisotropy of a unheated and b heated sandstone samples measured with the stabilized spinner magnetometer

closely around an axis with mean azimuth of N47°E (S47°W).

Ten cube-shaped samples were measured on the susceptibility bridge and with the improved spinner magnetometer (Fig. 6). On both instruments two samples were discarded because the 95% confidence angle of the sample measurement of at least one axis (maximum or intermediate) exceeded the critical value for non-randomness (Jelinek, 1977). Even with this restriction, the manually adjusted bridge measurements of the maximum axes are highly scattered. When the bridge measurements are monitored on-line with the computer, more consistent data result which, however, are still of inferior quality compared to the actively filtered spinner magnetometer results. The minimum anisotropy directions are again well defined.

The most consistent anisotropy data were obtained on the stabilized spinner magnetometer from a set of 39 unheated samples and a sister set of 42 heated samples (Fig. 7) which gave individual measurements acceptable at the 95% confidence limit (Jelinek, 1977). In both data sets the directions of all three principal axes group well. The sedimentary fabric with very steeply inclined minimum directions is confirmed and the maximum axes group rather tightly around a sub-horizontal NE-SW direction (Table 2). An up-current imbrication of the maximum axes was noted by Crimes and Oldershaw (1967) but cannot be discerned in the Plattensandstein. On the contrary, there are maximum axes clusters on both the NE and SW quadrants of the lower hemisphere projection, but with a tendency (2:1) to dip towards the NE along the main current direction suggesting a down-current imbrication of these axes. This is also reflected by the very steeply dipping individual minimum anisotropy axes which form a subvertical mean direction (Table 2). The minimum axes follow an elliptical distribution the long axis of which is oriented NW-SE. This distribution is very similar to that of the cross-bedding poles (Fig. 2a) and is caused by the symmetric cross-bedding stratification of the trough fillings (see Fig. 1). The magnetic low-field anisotropy of the Plattensandstein formation in this manner, reflects the original depositional fabric.

The heat treatment improves the quality of the anisotropy measurements. The mean susceptibility increases by about 300% and the degree of anisotropy rises sharply to 20% (Table 2). The resulting directional distribution of

the maximum and intermediate anisotropy axes is very similar to that of the unheated samples but less scattered (Fig. 7b; Table 2). This improvement, however, can be achieved only with anisotropy meters using an alternating current method. The cryogenic magnetometer method, using a locked-in constant magnetic field, deteriorates because of a strong remanence viscosity originating from the superparamagnetic magnetite produced during the heating.

Discussion and conclusion

The Plattensandstein sediments in the outcrop investigated have a clearly developed magnetic anisotropy. The minimum axes lie perpendicular to the bedding planes of the cross-bedded layers, thus indicating a sedimentary detritic fabric rather than a post-sedimentary compaction fabric. The differences between maximum and intermediate susceptibility axes are very small (below 1%) for the samples in the natural state. Their orientation can be resolved best by accurate measurements with the modified spinner magnetometer. This technique is also a very quick procedure since one sample can be measured within less than 1 min. Even higher accuracy is achieved by using heated sandstone samples which contain magnetite produced at high temperature. The magnetite fabric obviously mimics the natural pre-heating fabric, but can be measured more accurately due to the largely increased bulk susceptibility and its anisotropy.

Field-measurements of foresets yield a bimodal distribution (Fig. 2) which is caused by the symmetric filling from both sides of the cross-bedded structures. The resulting vector mean represents the main palaeocurrent direction. It points towards the NNE with an azimuth of $18.5^\circ \pm 5.1^\circ$ (circular standard error) and deviates by about 20° from the mean of the maximum susceptibility axes of the heated sandstones which is aligned parallel to a NE-SW axis with an azimuth of $38.5 \pm 3.0^\circ$ or $218.5^\circ \pm 3.0^\circ$ (Fig. 7b). The azimuthal difference between these directions is significant, since their error cones do not overlap each other.

The magnetic determination of the palaeocurrent axial direction is more accurate than the result of the geological field observations. Since the mean of the susceptibility minima, however, is subvertical, up-stream or down-stream direction of the palaeocurrent cannot be discerned from the

present anisotropy measurements alone. Thus a combination of both geological field measurements, which resolve the trend of the transport direction, and precise measurements of the anisotropy principal directions gives the best estimate of the palaeocurrent directions in the sandstones of the Plattensandstein formation in southern Germany.

Acknowledgements. We thank H. Vossmerbäumer and K. Ernstson for critically reading the manuscript, W. Lowrie and R. Freeman for invaluable suggestions for improvement of the manuscript, M. Grieder for his help with the installation of the low pass filter and W. Gruber for sample preparation.

References

- Allen, J.R.L.: A review of the origin and characteristics of recent alluvial sediments. *Sedimentology* **5**, 89–191, 1965
- Argenton, H., Bobier, C., Polveche, J.: La mesure de l'anisotropie de susceptibilité magnétique dans les flysches: application à la recherche des directions des paleocourants. *Sediment. Geol.* **14**, 149–167, 1975
- Beil, U., Petersen, N.: Magnetic properties of minerals. In: *Physical properties of rocks*. G. Angenheister ed. pp 308–365 Springer: Berlin-Heidelberg-New York, 1982
- Channell, J., Heller, F., Stuijvenberg, J. van: Magnetic susceptibility anisotropy as an indicator of sedimentary fabric in the Gurnigel Flysch. *Eclog. Geol. Helv.* **73**, 781–787, 1979
- Collinson, D.W.: The origin of remanent magnetization and initial susceptibility of certain red sandstones. – *Geophys. J. R. Astron. Soc.* **9**, 203–217, 1965
- Collinson, D.W.: Ferrous and ferric iron in red sediments and their magnetic properties. *Geophys. J. R. Astron. Soc.* **16**, 531–542, 1968
- Crimes, T.P., Oldershaw, M.A.: Paleocurrent determinations by magnetic fabric measurements of the Cambrian rocks of St. Tudwal's Peninsula, North Wales. *Geol. J.* **5**, 217–232, 1967
- Dunlop, D.J.: Magnetic mineralogy of unheated and heated red sediments by coercivity spectrum analysis. *Geophys. J. R. Astron. Soc.* **27**, 37–55, 1972
- Fisher, R.A.: Dispersion on a sphere. *Proc. Roy. Soc. London, A* **217**, 295–305, 1953
- Galehouse, J.S.: A test of the method using the anisotropy of magnetic susceptibility as a paleocurrent indicator. *Geol. Soc. Am. Bull.* **79**, 387–390, 1968
- Hamilton, N.: The effect of magnetic and hydrodynamic control on the susceptibility anisotropy of redeposited silt. *J. Geol.* **75**, 738–743, 1967
- Hamilton, N., Rees, A.I.: The anisotropy of magnetic susceptibility of the Franciscan rocks of the Diablo Range, Central California. *Geol. Rdsch.* **60**, 1103–1124, 1971
- Hamilton, N., Owens, W.H., Rees, A.I.: Laboratory experiments on the production of grain orientation in shearing sand. *J. Geol.* **76**, 465–472, 1968
- Hrouda, F., Janak, F.: A study of the hematite fabric of some red sandstones on the basis of their magnetic susceptibility anisotropy. *Sediment. Geol.* **6**, 187–199, 1971
- Ising, G.: On the magnetic properties of varved clay. *Ark. Math. Astr. Phys.* **29A**, 1–37, 1942
- Jelinek, V.: Precision a.c. bridge set for measuring magnetic susceptibility of rocks and its anisotropy. *Studia Geophys. Geod.* **17**, 36–47, 1973
- Jelinek, V.: The statistical theory of measuring anisotropy of magnetic susceptibility of rocks and its application. *Geofyzika Brno*, 88p, 1977
- Rad, U. von: Comparison between "magnetic" and sedimentary fabric in graded and cross-laminated sand layers, Southern California. *Geol. Rdsch.* **60**, 331–354, 1970
- Rees, A.I.: The effect of water currents on the magnetic remanence and anisotropy of susceptibility of some sediments. *Geophys. J. R. Astron. Soc.* **5**, 251, 1961
- Rees, A.I.: The use of the anisotropy of magnetic susceptibility in the estimation of sedimentary fabric. *Sedimentology* **4**, 257–271, 1965
- Rees, A.I.: The production of preferred orientation in a concentrated dispersion of elongated and flattened grains. *J. Geol.* **76**, 457–465, 1968
- Rees, A.I., Rad, U. von, Shepard, F.P.: Magnetic fabric of sediments from the La Jolla submarine canyon and fan. *Marine Geology* **6**, 145–178, 1968
- Scriba, H., Heller, F.: Measurements of anisotropy of magnetic susceptibility using inductive magnetometers. *J. Geophys.* **44**, 341–352, 1978
- Shive, P.N., Steiner, M.B., Huycke, D.T.: Magnetostratigraphy, paleomagnetism, and remanence acquisition in the Triassic Chugwater formation of Wyoming. *J. Geophys. Res.* **89**, 1801–1815, 1984
- Stephenson, A.: The effect of heat treatment on the magnetic properties of Old Red Sandstone. *Geophys. J. R. Astron. Soc.* **13**, 425–440, 1967
- Teyssen, T., Vossmerbäumer, H.: Schrägschichtungsanalyse am Beispiel des Buntsandsteins in Nordbayern. *N. Jb. Geol. Paläont. Mh.* **10**, 620–642, 1980
- Turner, P.: Continental red beds. *Developments in Sedimentology*, **29**, pp 1–562. Elsevier: Amsterdam, 1980
- Van den Ende, C.: On the origin of anisotropy of magnetic susceptibility in the Permian red beds from the western parts of the Dôme de Barrot (S. France). In: *Progress in geodynamics*, Borradaile, G.J. et al. eds. *Proc. Nat. Symp. Geodynamics*. pp 176–189, 1975
- Vossmerbäumer, H., Achtnich, T., Kistner, A., Priebe, H., Rongitsch, A., Rückert, E., Schultz, T., Teyssen, T.: Schüttungsrichtungen im Plattensandstein (Trias, Oberer Buntsandstein) Frankens. *Geol. Bl. NO-Bayern*, **29**, 50–60, 1979
- Wurster, P.: Geometrie und Geologie von Kreuzschichtungskörpern. *Geol. Rdsch.* **47**, 322–359, 1958

Received December 17, 1984; Revised February 4, 1985

Accepted February 4, 1985

Deep drilling through the accreting plate boundary of Asal, Southern Afar: palaeomagnetism and magnetic properties of basaltic cores

P.Y. Galibert, N. Bonhommet

C.N.R.S. (LP 4661) and Laboratoire de Géophysique Interne, Université de Rennes 1, F-35042 Rennes Cedex, France

Abstract. Palaeomagnetic and rock magnetic investigations have been carried out on unoriented samples from two deep vertical boreholes in the axial valley of the Asal Rift (Republic of Djibouti). The deepest samples (1280 m), submitted to temperatures of 200° C and long-term hydrothermal alteration, show low intensity values of magnetization ($0.6 \pm 0.2 \text{ Am}^{-1}$), in contrast to normal values of oceanic layer 2A, and pure magnetite as magnetization carrier. Thermal demagnetization, low-temperature saturated magnetization, high-temperature spinner experiments and optical examination show that: (1) There is evidence of thermal overprinting which did create a secondary component in the present day field direction; this direction is thermally locked-in during extraction from the borehole. (2) The use of this magnetization to orientate the core together with vectorial analysis of thermal demagnetization diagrams of the NRM reveals a high blocking temperature primary component with reverse polarity. (3) This contrasts with all previous surface palaeomagnetic surveys in the Asal Rift where only normal polarity directions have been found; it agrees, however, with an age of about 1 MY for the early stage of rifting at the present place as suggested from propagating rift theory in this area.

Key words: Magnetic properties – Basalt – Asal Rift – Hydrothermal alteration – Multivectorial magnetization

Introduction

The magnetic structure of the oceanic crust has been well explained by Vine and Matthews (1963) through the interpretation of marine magnetic anomalies. Dredged samples and submarine surveys have enabled us to state in more detail the magnetic structure of the crust, allowing us to give some constraints on the surface layers. Intensive studies have been carried out in some local areas of the mid-Atlantic Ridge at 45° N (Brooke et al., 1970; Irving, 1970; Ade-Hall et al., 1973) and the Famous area (Prévot et al., 1976). They could not, however, offer data needed for the vertical distribution of magnetic properties. With the next phases of crustal sampling (IPOD and DSDP programmes) data were obtained up to 600 m depth in basaltic layers 2A and 2B. A direct test of the Vine and Matthews hypothesis was possible (Johnson and Merrill, 1978) but increased complexity has arisen from the results (Hall, 1976), in par-

ticular when polarity inversion appeared in a vertical drilled section of the oceanic crust (Ryall et al., 1977; Johnson and Merrill, 1978). Magnetic properties of deep intrusive layers of the oceanic crust have been obtained from comparison with geothermal deep drilling in Iceland (Kristjansson and Watkins, 1977), obducted ophiolite suites (Stern and Elton, 1980; Banerjee, 1980) or sea floor dredgeholes (Kent et al., 1978).

These comparisons, however, may not always be justified as Iceland might not have true oceanic crust (Gibson, 1979) and there is evidence that most ophiolite suites have undergone metamorphism during obduction. Dunlop and Prévot (1982) have emphasized this difficulty. Their new data on magnetic properties on drilled submarine rocks from legs 30, 37 and 45 led them to propose a new magnetic layering of the oceanic crust where magnetite is the main magnetic carrier of all layers. They even show that deep crust and upper mantle are possible candidates to oceanic magnetic anomalies, an exciting hypothesis which has since been challenged.

In the search for the sources of magnetic anomalies over the Republic of Djibouti territory, a region where pseudo-oceanic crust has long since been recognized (Barberi and Varet, 1974; Lepine et al., 1972; Needham et al., 1976; Ruegg, 1975), aeromagnetic mapping was done by CNRS, France (Courtilot et al., 1980) and paleomagnetic surveys undertaken (Galibert et al., 1980; Courtilot et al., 1984). We report here results of magnetic properties of samples from two ancient deep geothermal holes in that same area.

During a geothermal survey in the Republic of Djibouti the "Bureau de Recherches Géologiques et Minières" (BRGM) drilled two boreholes close to Lake Asal in a volcano-tectonic axis known as the Asal Rift (Varet, 1978). This tectonic area is a unique example (together with Iceland) of an emerged spreading segment at the West of the Gulf of Aden near the Red Sea. The Asal Rift is at the westernmost segment of the Carlsberg ridge (Laughton, 1966), making a transition (Fig. 1a) between the oceanic domain of the Tadjourah gulf and the Afar depression (Lepine et al., 1972; Ruegg, 1975; Richard, 1979). Though morphological and structural evidence shows a similarity to a low-rate spreading ridge (Needham et al., 1976), the magmatism (Stieltjes, 1973; Richard, 1979) is transitional, the tholeiitic trend appearing in most recent times. The rifting process studied during the last volcano-tectonic crisis in 1978 (Ruegg et al., 1979), first compared to the one previously studied in Iceland (Bjornsson et al., 1979), is considered as a typical example of a propagating rift (Courtilot

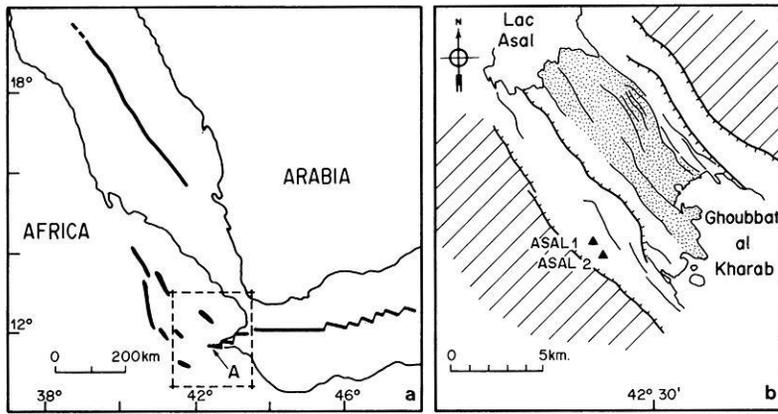


Fig. 1. a Sketch map of the Afar area with main tectonic trends: *black*: known spreading axis and active oceanic type volcanic ranges. *A* indicates the site of the Asal rift. b Locations of the two boreholes in the Asal rift. *Shaded area*: Dahla and stratoid basaltic trapps (8–1 M.Y. old). *White*: central zone, containing the Inner Floor (*dotted area*), after Needham et al. (1976)

et al., 1980; Courtillot et al., 1984). The affinity of the Asal Rift with an oceanic ridge should result in an oceanic-type magnetic structure of the crust and magnetic properties of samples from deep layers might help to elaborate magnetic layering models for this area and be useful for a comparison with true oceanic crust.

Geological setting of boreholes and sample description

The two boreholes named ASAL1 and ASAL2, were drilled by BRGM during the 1975 survey for detecting steam fields in this area. They lie inside the rift near its SW limit (Fig. 1b) and are both on the same tectonic block of the "Central zone" as described by Needham et al., 1976) by analogy with a rift valley of an oceanic ridge. During drilling operations the alkaline basaltic Stratoid and Dahla series which are, respectively, 1–4 MY and 4–8 MY old (Barberi et al., 1975; Richard, 1979) and outcrop only outside the rift, were encountered beneath the present tholeiitic basalt. In situ temperature measurements for ASAL 2 yielded values above 220°C at a depth of 1260 m, indicating a strong temperature gradient at this site; a petrographic study of the two borehole cuttings (BRGM, 1975) indicates that samples have undergone increasing metamorphism with depth from zeolite to greenschist facies with chlorite and epidote below 1200 m. It is important to point out that this metamorphism is not a simple deep sea water circulation mechanism, as for surface layers of the oceanic crust (Bohlke et al., 1981), but results from hydrothermal alteration at temperatures of about 300°C.

Table 1 lists the different standard specimens obtained by drilling from fragments provided by BRGM and recovered in three different cores: (1) Core 1C2 (ASAL 1) is a rhyolitic flow with strong silica recrystallization; substitution of pyroxene by chlorite reported by BRGM (BRGM, 1975) was not observed in our samples. Coring was done from 450 to 453 m (90% recovered); we used fragments 1, 5, 11, 15, 18 and 21 spread along the 3 m of coring to get 16 specimens. (2) Core 2C1 (ASAL 2) is from an alkaline basalt with low chlorite content. Coring was done from 980 to 981.5 m (60% recovered); we used fragments 1 and 2 to get 6 specimens. (3) Core 2C2 (ASAL 2) is also an alkaline basalt having extensive chloritization, secondary silica calcite and epidote; it has a characteristic greenschist facies metamorphism. Coring was done from 1281 to 1284 m (60% recovered); we used fragments 1, 2 and 3 spread along the distance of coring to get 10 specimens. A total of 32 specimens were available for our study.

Table 1. Palaeomagnetic results. 1C2. 1a stands for Asal 1 borehole, coring C2, fragment n° 1 (BRGM) palaeomagnetic standard specimen *a*. *I*, inclination. *J*, intensity of magnetization in Am^{-1} . *H* and *T*: range of AF peak field and temperature giving the cleaned direction

Specimen number	N.R.M.		Inclination after cleaning		
	I_{arc}	J_{NRM}	$H_{\text{(mT)}}$	$T_{\text{(°C)}}$	I_{arc}
Asal 1					
1C2 1a	(82)	(0.71)	12–40		(82)
1b	(80)	(0.22)		125°–375°	(81)
5a	42	0.02	20–35		13
5b	18	0.03		450°	15
5c	13	0.02			
11a	1	0.03	30–50		15
b	5	0.03		150°–450°	14
c	12	0.04			
d	19	0.03			
15a	45	0.02	30–50		0
b	49	0.03			
18a	31	0.03	10–35		10
b	34	0.03			
21a	51	0.04	15–40		–
b	50	0.03		300°–400°	7
c	63	0.03			
$N=5, J_{\text{NRM}}=0.03 \pm 0.01 \text{ Am}^{-1}, I_m=9^\circ \pm 6^\circ$					
Asal 2					
2C1 1a	7	2.00	50–60		18
b	10	1.98		broken at 450°	9
2a	18	2.47	40–70		19
b	16	2.34		broken at 450°	18
c	18	2.43		broken at 375°	19
d	19	–	50–70		22
$N=2, J_{\text{NRM}}=2.2 \pm 0.3 \text{ Am}^{-1}, I_m=17^\circ \pm 4^\circ$					
2C2 1a	50	0.34	30–70		34
b	45	0.56		500°–700°	36
c	44	0.37			
2a	40	0.50	15–70		38
b	41	0.65		600°–700°	36
c	42	0.95		450°–580°	36
d	43	0.81			
3a	33	0.55	10–70		32
b	31	1.25		400°–600°	30
c	46	0.48			30
$N=3, J_{\text{NRM}}=0.64 \pm 0.2 \text{ Am}^{-1}, I_m=34^\circ \pm 3^\circ$					

Natural remanent magnetization and viscous remanent magnetization

Stability of the natural remanent magnetization (NRM) has been studied by the "storage test" using a 1-week period: samples are stored for 1 week with their vertical axes upward in the geomagnetic field and for another week with their vertical axes downward, in a position antiparallel to the first one. Subtraction and addition of the two NRMs measured after the two storages yield, respectively, the VRM (viscous remanence) acquired in 1 week in the laboratory and the NRM free of viscous component at this time scale (Dunlop and Prévot, 1982; Thellier and Thellier, 1944). The ratio VRM/TRM yields a mean viscosity of 12% for the rhyolites in ASAL 1 and lower than 5% for the two basaltic flows in ASAL 2 showing that short-term VRM is small or negligible for the two basaltic flows but quite important for the rhyolitic flow. NRM intensities for all samples are given in Table 1; they show consistent results within the same fragment and in the same core except for fragment 1C2-1 which has disproportionately high NRM intensity and abnormal inclination, AF demagnetization of specimen 1C2-1a has shown the typical behaviour of an IRM component which might have been acquired during drilling, a process previously suggested by different authors (Dunlop and Prévot, 1982).

Rejecting these anomalous samples, we obtain mean intensities of 0.03 ± 0.01 , 2.2 ± 0.3 and $0.64 \pm 0.20 \text{ Am}^{-1}$, respectively, for cores 1C2, 2C1 and 2C2. Putting aside the rhyolite 1C2, which is not a major constituent of the crust in the Asal area, the last two results from the basaltic flows can be compared with data obtained in two previous surface samplings made on the Asal Rift, keeping in mind the limited number of samples we have. The mean NRM intensities of the two sets available are 2.7 Am^{-1} (Harrison et al., 1977) and 5.6 Am^{-1} (Galibert et al., 1980); we interpret the discrepancy between these surveys by a variable proportion of highly magnetized subaquatic flows and less magnetized aerial flows in the two samplings. Such differences in intensity have been pointed out by Prévot and Grommé (1975). The value of $2.2 \pm 0.3 \text{ Am}^{-1}$ found at 980 m in ASAL 2 for core 2C1 agrees well with these figures, which are all characteristic of layer 2A. On the contrary, core 2C2 at 1280 m with $0.6 \pm 0.2 \text{ Am}^{-1}$ has a much too low intensity. Two hypotheses may be considered to explain this result: (1) The variability of petrographic type. (2) The effect of high temperature in this area. If we consider the mean value 3.6 Am^{-1} (with bounds of 2.1 and 14.7 for the standard deviation) computed by Prévot and Grommé (1975) for young basaltic flows of variable type, our data is significantly lower.

Similar studies in Iceland (Kristjansson, 1972; Wood and Gibson, 1976) did not yield a clear answer, but the prime importance of hydrothermal circulation in lowering the proportion of magnetic minerals has been pointed out (Ade-Hall et al., 1971; Kristjansson, 1972). However, we must not forget the low number of samples we have worked with.

Alternating field and thermal demagnetization

Table 1 gives the absolute value I of the NRM inclination (we know the vertical direction of each fragment of cylindrical shape but not which face is up or down). For core 1C2 (ASAL 1) we get variable values, rather high, relative

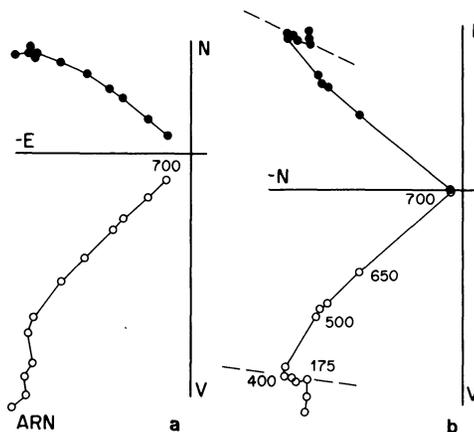


Fig. 2a, b. Zijderveld (1967) diagrams. **a** AF cleaning of sample 2C2.1a; **b** Thermal treatment of sample 2C2.1b. Full symbols: horizontal projection; open symbols: vertical projection. Dotted lines: B component

to the expected dipole inclination (22°) at the latitude of Asal. However, except for fragment 1 which might have acquired an IRM during drilling, after AF and thermal demagnetization the mean inclination is lowered to $9^\circ \pm 6^\circ$, a value compatible with secular variation if we consider it as an instantaneous record of the field direction from one flow unit. For core 2C1 a stable direction is defined for each specimen with coercive force above 40 mT and high unblocking temperatures giving a mean inclination, after cleaning, of $17^\circ \pm 4^\circ$. The decrease of magnetization is small during heating; this corresponds to a stable direction up to 400° C after which most samples break up, preventing higher temperature steps being done.

Core 2C2 shows very different behaviour during both types of demagnetization. Thermal cleaning allows the separation of three components (Fig. 2b) while, on the other hand, AF cleaning (Fig. 2a) shows only one direction corresponding to the one with high unblocking temperatures. Table 1 gives the inclination of the high-temperature/high-field component, defined in the range 10–70 mT or unblocking temperatures greater than 400° C . The corresponding mean inclination of this component is about 34° . The characteristics of the three components of magnetization are the following: (1) Component A with unblocking temperatures greater than 400° C and a mean inclination of about 34° . (2) Component B with shallow inclination of about 4° and unblocking temperatures in the range $175^\circ\text{--}400^\circ \text{ C}$. (3) Component C with steep inclination, defined by unblocking temperatures lower than 175° C (Fig. 2b) which probably represents the viscosity acquired during 1 year of storage in the laboratory field. The B component is sometimes poorly defined when the two blocking-temperature spectra of B and A do overlap; nevertheless, the result is that B is roughly antiparallel to A (Fig. 2b). We shall come back to this result later on.

Rock magnetism

Opaque minerals were identified by microscope observation of polished sections and thermomagnetic analysis. J_s - T curves were obtained with a vertical type balance in a field of about 0.1 T and samples heated in vacuum. Acquisition of isothermal remanence magnetization (IRM) in fields up to 1.2 T, observation of transition temperature using a Di-

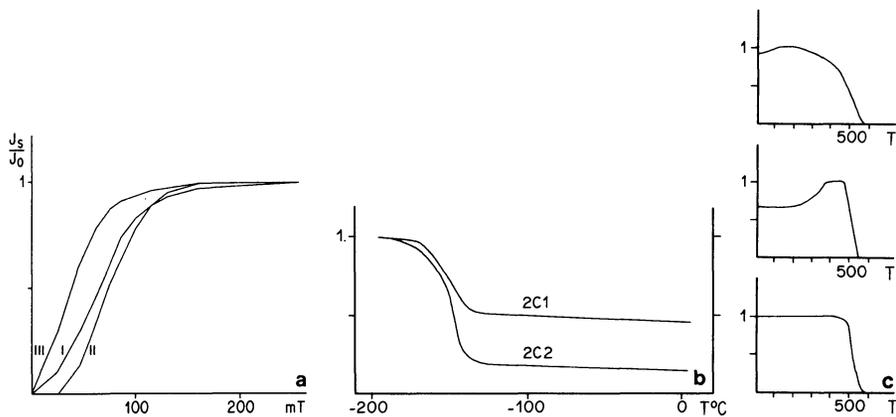


Fig. 3. a SIRM acquisition for core 1C2 (I), 2C1 (II) and 2C2 (III). b SIRM behaviour at low temperatures for core 2C1 and 2C2. c J_s - T curves for cores 1C2 (top) 2C1 (middle) and 2C2 (bottom)

gic high/low temperature spinner and X-ray diffractometry were also used to better estimate the magnetic carriers.

Opagques for cores 1C2 and 2C1 are abundant and have a mean size of $10 \mu m$. The grains consist of primary titanomagnetite (TTM) with scarce exsolved lamellae of ilmenite ("trellis" intergrowths). Early stages of maghemitization can be observed, the minerals having only suffered a slight alteration at the time of emplacement. The J_s - T curve for 2C1 shows a particular behaviour with enhancement of magnetization below the Curie point (Fig. 3c); the external field, too low for saturation, explains the maximum, a characteristic of the Hopkinson effect (Dunlop, 1974) which may be an indication of abundance of mono-domain grains. The reversible curve indicates no or minor chemical transformation during heating and a Curie point near pure MT. IRM acquisition curves (Fig. 3a) show that 2C1 remanence in a field of 50 mT is only 20% of the saturation remanence (SIRM) which is, however, always obtained in fields lower than 200 mT.

The behaviour of SIRM at low temperature has been studied following the method given by Nagata et al. (1964) for the separation of the magnetite and hematite phases. After the sample was magnetized, both at room and liquid-air temperature ($-160^{\circ}C$ in our experience), continuous recording of the variation of magnetization shows that a transition at $-150^{\circ}C$ exists during warming up to $20^{\circ}C$ (Fig. 3b). This is characteristic of almost Ti-free MT. The relatively small decrease of SIRM at $-150^{\circ}C$, however, might indicate abundant SD grains. The absence of a transition at $-20^{\circ}C$ on the other hand is evidence of lack of hematite at least in coarse grains. X-ray determination of the lattice parameter has been carried out on one sample from core 2C1 giving $a=8.35 \text{ \AA}$, a value characteristic of maghemite or substitute TTM.

Core 2C2 is totally different; opaques are scarce and have a mean size of a few μm . The grains are interstitial in small veins and consist of secondary magnetite. X-ray determination of the lattice parameter yields an unlikely low and poorly defined value of $a=8.27 \text{ \AA}$. Saturation remanence is obtained in a field below 150 mT compatible with MT/TTM grains. Hematite is never apparent in 2C2 although unblocking temperatures greater than $580^{\circ}C$ have been observed in one sample (Fig. 2).

To summarize, MT with low Ti content appears to be the major constituent of the magnetic minerals in all sam-

ples. Hematite was not observed directly, but thermal demagnetization occasionally exhibited $T_b > 600^{\circ}C$. For 2C2, secondary MT crystallized during hydrothermal alteration, a temperature of $300^{\circ}C$ being high enough for this process to take place (Cann, 1979). The origin of MT observed in 1C2 and 2C1, however, at a depth where temperature was much lower is not understood. The samples do not show any alteration and there is no indication for deuteric oxidation. The MT looks like primary MT.

Interpretation of the results of core 2C2

Rock magnetic studies show the presence of secondary magnetite in 2C2 as the only visible magnetic mineral. The effect of hydrothermal alteration on TRM may explain the result observed during thermal demagnetization: if all primary minerals have broken down as indicated by microscope observation, the original TRM has been replaced by CTRM (component A) with grains which may be locked at temperatures of the order of $230^{\circ}C$. If, on the contrary, primary magnetic minerals still subsist, but are small enough and not visible (for instance, inclusion of MT into silicates would protect them from alteration), part of TRM subsist in component A. In both cases the B component is a PTRM carried by secondary grains which remain unblocked at $230^{\circ}C$. The acquisition of magnetization is the following: (1) Acquisition of primary TRM (component A). (2) Stepwise burial of the flow with evolution of the Asal Rift; in situ temperature increases slowly, up to $230^{\circ}C$ with hydrothermal activity. This prolonged effect possibly destroys the primary magnetic minerals and secondary magnetic grains grow, some of them staying unblocked owing to that temperature. (3) Acquisition of TRM by the unblocked grains during uplift of the core and cooling in the present magnetic field ($I=5^{\circ}$), yielding the shallow inclination B component.

Thermal overprinting has been theoretically studied (Pullaiah et al., 1975) for synthetic minerals (magnetite and hematite); the corresponding diagrams allow us an estimate of the duration of heating, assuming his amplitude and the blocking temperature for laboratory heating times. Whole rock applications are critical as they rely on the assumption of the NRM to be TRM with constant mineralogy through time (either magnetite or hematite in the cases studied by Pullaiah) but have already been attempted (Schwarz, 1977; Van der Voo et al., 1978; Buchan and Schwarz, 1980). The simple magnetic mineralogy of the

Table 2. Characteristic direction of magnetization for core 2C2 computed after undirect orientation of the samples (see text)

Sample number	Declination	Inclination
2C2 1b	169.9	-35.8
2C2 1c	212.3	-35.5
2C2 1d	162.1	-32.2
2C2 2b	221.1	-36.3
2C2 2c	173.2	-35.7
2C2 2d	183.5	-36.7
2C2 3b	136.4	-29.9
Mean direction $N=7$	179.1	-37.7 $k=12$ $\alpha_{95}=15.4$

Asal samples allows us to use the magnetite diagram (Pul-laiiah et al., 1975); blocking temperatures given by thermal cleaning are about 400° C (disappearance of the B component) for roughly 5 min heating. These two parameters, carried back onto the diagram, yield a heating time of about 1 MY at 230° C (in situ temperature).

If we assume a constant temperature through time and no evolution of the magnetic grains which is hardly probable, the residence time we found of 1 M.Y., which gives the age of the hydrothermal circulation responsible for the alteration, is coeval with the Asal Rift formation. An age of 1 M.Y. for the Asal Rift is in agreement with other results deduced from tectonic and microtectonic studies (Arthaud et al., 1980). The B component, whose presence has been explained, is therefore parallel to the present magnetic field at Asal ($D=0^\circ$, $I=5^\circ$); we can now fully orientate the samples (Watts and Van der Voo, 1976). We have only to apply a vertical axis rotation to make the apparent B declination equal to zero. Table 2 gives the directions of A components obtained for seven samples and the mean direction (179/-38°). This result must be used carefully, owing to multiple errors caused by core rotation during uplift or overlap of blocking temperature spectra leading to bad determination of B , for example. Nevertheless, the A component undoubtedly represents a reversed direction of the earth's magnetic field.

Conclusions

The core 2C2 from the deepest borehole (ASAL 2) in the inner floor of the Asal Rift, taken at a depth of 1280 m, has been submitted to in situ temperature of about 220° C. This process results in a multivectorial magnetization for core 2C2 as PTRM was locked in the present field direction during extraction of the core from the borehole, as shown by the direction of component B revealed by thermal demagnetization.

The resulting intensity of magnetization of core 2C2 ($0.6 \pm 0.2 \text{ Am}^{-1}$) is anomalously low compared to oceanic basalts from layer 2A. Although limited to a small number of samples available from core 2C2, this result would suggest a negligible contribution of deep sources to aeromagnetic anomalies that have been mapped in this area. We suggest that surface palaeomagnetic survey data might be used in inversion models at least for the inner floor of the Asal Rift. In fact, correlation of palaeomagnetic polarity with the sign of the aeromagnetic anomalies over the older (1-4 MA) area between Lakes Asal and Abbe are in favour

of a simple vertical structure of the magnetic crust in the area (Courtillot et al., 1984).

Use of the B component of magnetization, parallel to the present field, to fully orientate the samples shows that the primary magnetization of basalt from core 2C2 has a reversed direction. This contrasts with all surface flows in the rift which have normal polarity (Galibert et al., 1980; Harrison et al., 1977) corresponding to recent Brunhes-age volcanic activity as illustrated by the last volcanic crisis of 1978 (Ruegg et al., 1979). A reverse direction, on the contrary, fits with a pre-Brunhes age of the last phase of rifting at that place, when rifting propagated from the Ghoubbat straights to the present Asal area, as suggested by Courtillot et al., (1984).

Acknowledgements. This research was suggested to us by J. Varet from B.R.G.M. and funded by a grant from D.G.R.S.T., N° 76,7.1326.

References

- Ade-Hall, J.M., Palmer, H.C., Hubbard, T.P.: The magnetic and opaque petrological response of basalts to regional hydrothermal alteration. *Geophys. J. R. Astron. Soc.* **24**, 137-174, 1971
- Ade-Hall, J.M., Aumento F., Ryall, P.J.C., Gerstein, R.E., Brooke, J., McKeon, D.L.: The Mid Atlantic Ridge near 45° N, magnetic results from basalts drill cores from the median valley. *Can. J. Earth Sci.* **10**, 676-696, 1973
- Arthaud, F., Choukroune, P., Robineau, B.: Tectonique, microtectonique et évolution structurale du golfe de Tadjourah et du sud de la dépression Afar. *Bull. Soc. Geol. France*, (7) **XXII**, 6, 901-908, 1980
- Banerjee, S.K.: Magnetism of the oceanic crust: evidence from ophiolite complexes. *J. Geophys. Res.* **85**, 3557-3566, 1980
- Barberi, F., Varet, J.: Nature of the Afar crust: A discussion. In: Afar Depression of Ethiopia, A. Pilger and A. Rosler eds.: pp 375-378, Stuttgart 1974
- Barberi, F., Ferrara, G., Santocroce, R., Varet, J.: Structural evolution of Afar triple junction. In: Afar Depression of Ethiopia, A. Pilger and A. Rosler eds.: pp 38-54, Stuttgart 1975
- Bjornsson, A., Johnsen, G., Sigurdsson, S., Thoberfsson, Tryggvasson, E.: Rifting of the plate boundary in North Iceland 1975-1978. *J. Geophys. Res.* **84**, 3029-3038, 1979
- Bohlke, J., Honorez, J., Honorez-Guerstein, N., Muehlenbachs, K., Petersen, N.: Heterogeneous alteration of the upper oceanic crust: Correlation of rock chemistry, magnetic properties, and O isotope ratios with alteration patterns in basalts from site 3968, DSDP. *J. Geophys. Res.* **86**, 7935-7950, 1981
- B.R.G.M.: Forages géothermiques dans la région du lac Asal. Drilling Report. Orléans, 1975
- Brooke, J., Irving, E., Park, J.V.: The Mid Atlantic Ridge at 45° N. XIII. Magnetic properties of basalt bore core. *Can. J. Earth Sci.* **7**, 1515-1527, 1970
- Buchan, K.L., Schwarz, E.J.: Remanent magnetization in the contact zone between Columbia plateau flows and feeder dykes: Evidence for ground water at the time of intrusion. *J. Geophys. Res.* **85**, 1888-1897, 1980
- Cann, J.R.: Metamorphism in the oceanic crust. In: Deep drilling results in the Atlantic ocean: Ocean crust. Talwani M., Harrison C.G.A., Hayes, D.E. Washington: A.G.U., 1979
- Courtillot, V., Galdeano, A., Le Mouel, J.L.: Propagation of an accreting plate boundary: A discussion of new aeromagnetic data in the gulf of Tadjourah and southern Afar. *Earth Plan. Sci. Lett.* **47**, 144-160, 1980
- Courtillot, V., Achache, J., Landre, F., Bonhommet, N., Galibert, P.Y., Montigny, R., Feraud, G.: Episodic spreading and rift propagation: new paleomagnetic and geochronologic data from the Afar nascent passive margin. *J. Geophys. Res.* **89**, 3315-3333, 1984

- Dunlop, D.J.: Thermal enhancement of magnetic susceptibility. *J. Geophys.* **40**, 439–451, 1974
- Dunlop, D.J., Prévot, M.: Magnetic properties and opaque mineralogy of drilled submarine intrusive rocks. *Geophys. J. R. Astron. Soc.* **69**, 763–803, 1982
- Galibert, P.Y., Sichler, B., Smith, B., Bonhommet, N.: Paléomagnétisme en zone d'accrétion: le cas de l'Afar. *Bull. Soc. Geol. France (7) XXII*, 6, 881–890, 1980
- Gibson, I.L.: Crust of oceanic affinity in Iceland. *Nature* **281**, 347–351, 1979
- Hall, J.M.: Major problems regarding the magnetization of oceanic crustal layer 2. *J. Geophys. Res.* **81**, 4224–4230, 1976
- Harrison, C., Stieltjes, L., Tarasiewicz, E.: Paleomagnetism of samples from the axial zone of the Afar depression. *Earth Planet. Sci. Lett.* **34**, 273–283, 1977
- Irving, E.: The Mid Atlantic Ridge at 45° N. XIV. Oxidation and magnetic properties of basalt; review and discussion. *Can. J. Earth. Sci.* **7**, 1528–1538, 1970
- Johnson, H.P., Merrill, R.T.: A direct test of the Vine-Matthews hypothesis. *Earth Planet. Sci. Lett.* **40**, 263–269, 1978
- Kent, D.V., Honorez, B.M., Odyke, N.D., Fox, P.J.: Magnetic properties of dredged oceanic gabbros and the source of marine magnetic anomalies. *Geophys. J. R. Astron. Soc.* **55**, 513–537, 1978
- Kristjansson, L.: On the thickness of the magnetic crustal layer in SW Iceland. *Earth Planet. Sci. Lett.* **16**, 237–244, 1972
- Kristjansson, L., Watkins, N.D.: Magnetic studies of basalt fragments recovered by deep drilling in Iceland, and the “magnetic layer” concept. *Earth Planet. Sci. Lett.* **34**, 365–374, 1977
- Laughton, A.S.: The Gulf of Aden in relation with the Red Sea and the Afar depression of Ethiopia, the world rift system. (UMC symposium, Ottawa, 1965). *Geol. Surv. Pap. Can.* **66**, 78–97, 1966
- Lepine, J.C., Ruegg, C., Steinmetz, L.: Seismic profiles in the Djibouti area. *Tectonophysics* **15**, 59–64, 1972
- Nagata, T., Kobayashi, K., Fuller, M.D.: Identification of magnetite and hematite in rocks by magnetic observations at low temperatures. *J. Geophys. Res.* **69**, 2111–2120, 1964
- Needham, H.D., Choukroune, P., Cheminee, J.L., Le Pichon, X., Francheteau, J., Tapponier, P.: The accreting plate boundary: Ardoukoba rift (North East Africa) and the oceanic rift valley. *Earth Planet. Sci. Lett.* **28**, 439–453, 1976
- Prévot, M., Grommé, S.: Intensity of magnetization of subaerial and submarine basalts and its possible change with time. *Geophys. J. R. Astron. Soc.* **40**, 207–224, 1975
- Prévot, M., Lecaille, A., Francheteau, J., Hebikian, R.: Magnetic inclination from the Mid Atlantic Ridge near 37° N. *Nature* **259**, 649–653, 1976
- Pullaiah, M., Irving, E., Buchan, K.L., Dunlop, D.J.: Magnetization changes caused by burial uplift. *Earth Planet. Sci. Lett.* **28**, 133–143, 1975
- Richard, O.: Etude de la transition dorsale océanique-rift émerge: Le golf de Tadjourah (République de Djibouti), Thèse 3ème cycle; Orsay, 1979
- Ruegg, J.C.: Main results about the crustal and upper mantle structure of the Djibouti region (T.F.A.I.). In: Afar Depression of Ethiopia, A. Pilger and A. Rosler eds.: pp 120–134 Stuttgart 1975
- Ruegg, J.C., Lepine, J.C., Tarantola, A., Kasser, M.: Geodetic measurements of rifting associated with seismo-volcanics crisis in Afar. *Geophys. Res. Lett.* **11**, 817–820, 1979
- Ryall, P.J., Hall, J.M., Clark, J., Milligan, T.: Magnetization of oceanic crustal layer 2; results and thoughts after DSDP leg 37. *Can. J. Earth Sci.* **14**, 684–706, 1977
- Schwarz, E.J.: Depth of burial from remanent magnetization: the Sudbury eruptive at the time of diabase intrusion (1250 my). *Can J. Earth Sci.* **14**, 82–88, 1977
- Stern, C., Elton, D.: Vertical variation in the effects of hydrothermal metamorphism in Chilean ophiolites: their implications for ocean floor metamorphism. *Tectonophysics* **55**, 179–213, 1980
- Stieltjes, L.: L'axe volcano-tectonique d'Asal. Thèse 3ème cycle, Orsay, 1973
- Thellier, E., Thellier, O.: Recherches géomagnétiques sur des coulées volcaniques d'Auvergne. *Ann. Geophys.* **1**, 37–52, 1944
- Van der Voo, R., Henry, S.G., Pollack, H.N.: On the significance and utilisation of secondary magnetizations in red beds. *Phys. Earth Planet. Int.* **16**, 12–19, 1978
- Varet, J.: Geologic map of central and southern Afar. CNRS-CNR 1978
- Vine, F.J., Matthews, D.H.: Magnetic anomalies over oceanic ridges. *Nature* **199**, 947–949, 1963
- Watts, D., Van der Voo, R.: Further paleomagnetic results from Michigan basin deep borehole (abstract T1). *EOS* **57**, 760, 1976
- Wood, D., Gibson, I.: The relationship between depth of burial and mean intensity of magnetization for basalts from eastern Iceland. *Geophys. J. R. Astron. Soc.* **6**, 497–498, 1976
- Zijderveld, J.D.A.: Demagnetization of rocks: analysis of results. In: *Methods in paleomagnetism*. D.W. Collinson, K.M. Creer, S.K. Runcorn, eds.: pp 254–286. Amsterdam: Elsevier 1967

Received October 29, 1984; Revised March 22, 1985

Accepted March 28, 1985

Observation of kinetic Alfvén waves excited at substorm onset

N. Klöcker¹, H. Lühr¹, A. Korth² and P. Robert³

¹ Institut für Geophysik und Meteorologie der Technischen Universität Braunschweig, FRG

² Max-Planck-Institut für Aeronomie, Katlenburg-Lindau, FRG

³ Centre de Recherches en Physique de l'Environnement Terrestre et Planétaire, Issy-les-Moulineaux, France

Abstract. Ground-based observations of locally confined, very intense, drifting current systems by the EIS-CAT magnetometer cross in correlation with GEOS-2 measurements will be explained in terms of kinetic Alfvén waves. Particle and magnetic flux measurements on GEOS-2 indicate an excitation of the waves at the inner edge of the earthward-drifting plasma sheet by resonance mode conversion from hydromagnetic surface waves. The collapsing tail-like field configuration itself is identified as the surface wave. The comparison of theoretically deduced quantities with observational results reveals a satisfactory agreement between observations and theory.

Key words: Kinetic Alfvén waves – Resonance mode conversion – Expanding plasma sheet – Magnetosphere-ionosphere coupling

Introduction

Recent ground observations of very intense magnetic pulsations with periods near the lower border of the Pi2 frequency spectrum, in combination with riometer and STARE data, suggest the interpretation as ionospheric reflections of kinetic Alfvén waves. A detailed description of the event is published by Lühr et al. (1984) which will be referred to in the following as Paper 1.

The proof of kinetic Alfvén waves is of special interest in the actual discussion about high-latitude Pi2 pulsations (Samson, 1982; Pashin et al., 1982; Baumjohann and Glaßmeier, 1984). Kinetic Alfvén waves are a special type of shear mode Alfvén waves, which always play an important role in the transmission of information along magnetic field lines. Shear mode Alfvén waves are generated by any change of the electric potential distribution in the magnetosphere-ionosphere system (i.e. any change in the magnetospheric convection pattern) and are always accompanied by field-aligned currents (Mallinckrodt and Carlson, 1978). The characteristic features of kinetic Alfvén waves are their large perpendicular wavenumber and their parallel elec-

tric field. Electrons can be accelerated by the electric field up to energies of a few keV. Narrow auroral structures and intense particle precipitation, which can be produced by these waves, are typical for the substorm breakup phase and Pi2 pulsations, which occur in conjunction with breakup phenomena (see Samson, 1982 and references therein).

The propagation of shear mode Alfvén waves was theoretically studied by Fejer and Lee (1967) and Fejer and Kan (1969). On this basis and on Hasegawa's (1977) treatment of kinetic Alfvén waves, several attempts have been made to study the dynamics of coupling processes between the magnetosphere and the ionosphere. Goertz and Boswell (1979) analysed the effects of propagation and reflection of kinetic Alfvén waves at the ionosphere. Lysak and Carlson (1981) and Lysak and Dum (1983) extended this work by including the effect of turbulence, leading to a partial decoupling of the ionosphere from the magnetosphere. A possible source for kinetic Alfvén waves was provided by Hasegawa (1976) and further discussed by Goertz (1983). Hasegawa has shown that a large-scale incompressible MHD surface wave can be converted to small-scale Alfvén waves.

During the first mentioned pulsation event the European Geostationary Satellite GEOS-2 was in a position magnetically conjugate to the EISCAT magnetometer chain (Paper 1). Particle and magnetic field fluctuations measured on GEOS in conjunction with the ground-based observations will be presented in this paper. They can satisfactorily be explained in terms of the theoretical frame outlined by Hasegawa (1976) and Goertz (1983). At first the GEOS-2 observations will be introduced followed by a detailed comparison of observed and theoretically deduced parameters.

Observations on GEOS-2

Energetic particle flux

On the day of interest – 2 November 1982 – the geostationary satellite GEOS-2 was located at a geographic longitude of 32.2° E and a geomagnetic latitude of about 3° S.

The particle measurements onboard GEOS-2 were made with the MPAe charged-particle spectrometer (Korth and Wilken, 1978; Korth et al., 1978) which

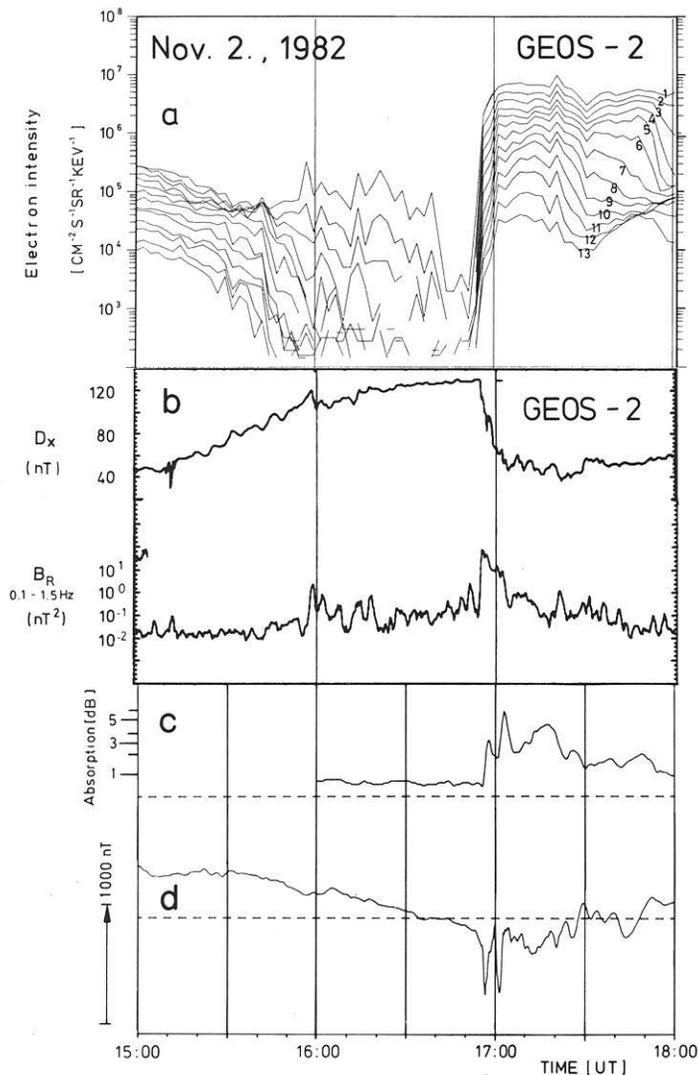


Fig. 1 a-d. Synchronous observations made onboard the geostationary satellite GEOS-2 and on ground in northern Scandinavia in the same local time sector.

a Differential electron intensities in 13 energy channels. The data were averaged over 2.5 min. The band widths in keV of the different labelled channels are as follows: 1: 24-30; 2: 30-38; 3: 38-45; 4: 45-52; 5: 52-61; 6: 61-71; 7: 71-83; 8: 83-96; 9: 96-114; 10: 114-133; 11: 133-154; 12: 154-180; 13: 180-214. Around 1710 UT these channels can be identified in the given sequence with the highest intensity for the lowest energy band.

b The equatorial component of the magnetic field D_x measured on board GEOS-2 with the search coil magnetometer and the integrated power in the ULF range from 0.1-1.5 Hz in B_R , the right-handed polarized component in the equatorial plane.

c and d Ground-based observations of the cosmic radio noise absorption at the Finnish station Rovaniemi (c), indicating an unusual enhancement of the ionospheric plasma density at 1656 UT, and the geographic north component of the EISCAT magnetometer station Muonio (d). Simultaneously with the onset of the very strong magnetic pulsation, all other observations show rapid variations

detected ions and electrons. For this study only electron data are shown. Electrons were measured simultaneously at energies >22 keV as well as in 13 energy channels between 24 and 213 keV.

The electron intensities in the 13 energy channels are presented in Fig. 1a. The electrons were detected perpendicular ($90^\circ \pm 5^\circ$) to the S/C spin axis (pitch angle calculations could not be carried out because the fluxgate magnetometer failed in 1979). From 1500 UT we observe a steady decrease of the intensities in all energy channels. At 1654 UT a sudden energy dispersionless increase by 2-3 orders of magnitude is recorded. About 10 min later the maximum intensity is reached.

The decrease in the electron intensity can be explained by a change in the local magnetic field topology at geostationary orbit. From Fig. 1b it can be demonstrated that after 1500 UT the magnetic field (measured by the search coil magnetometer) becomes more and more tail-like until substorm onset at 1654 UT. The onset is associated with a strong injection of energetic electrons and ions (Fig. 2) at the geostationary orbit and with a rapid return to a dipole-like magnetic field configuration. This behaviour is well

known from the work of Baker and co-workers (e.g. Baker et al., 1982). In addition, Fig. 3 shows count rates of electrons with energies >22 keV for a longer time period than in Fig. 2. Fluctuations in the period range between 250 and 500 s can clearly be identified.

Magnetic flux variations

Figure 1b shows the equatorial projection D_x of the magnetic field and the integrated power in the frequency range 0.1-1.5 Hz in B_R , the right-handed polarized component in the equatorial plane. The ULF fluxmeter and the deconvolution method used to obtain an equivalent dc magnetic measurement are described by Robert et al. (1984). Around 1500 UT the increasing D_x indicates a developing tail-like magnetic field configuration in the same way as discussed by Shepherd et al. (1980) and Baker et al. (1982). During this time GEOS-2 entered the nightside magnetosphere where, during the growth phase of a substorm, cross-tail currents have often been observed in the near earth region. At 1654 UT this configuration starts collapsing, lasting until 1703 UT when the undisturbed dipolar configuration was reached again. The additional field value in this

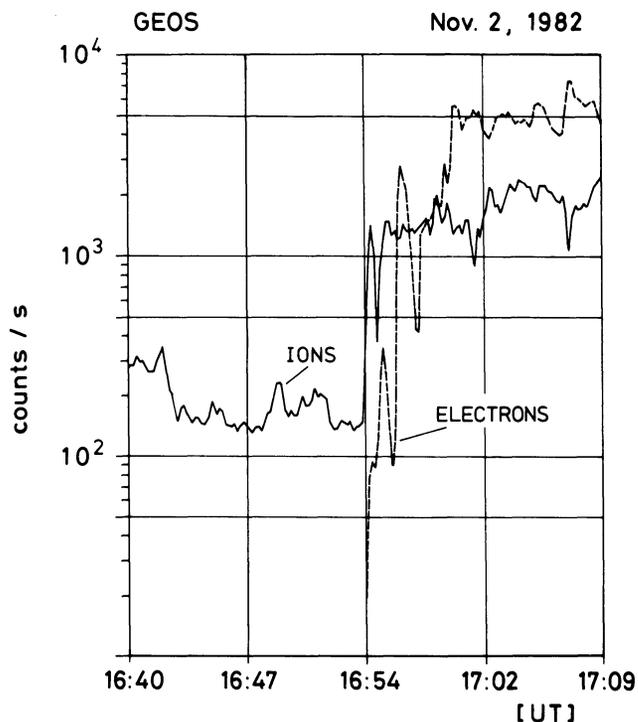


Fig. 2. This shows the detected electrons as well as the ions in the spin plane of GEOS-2 integrated over all energy channels, i.e. from 24 to 214 keV for the electrons and from 35 to 403 keV for the ions. The data were averaged over 11 s yielding a much better time resolution than in Fig. 1a. The rise time for ions is about 30 s. The electrons increase stepwise with peak intensities after 75 s, 150 s, and 350 s

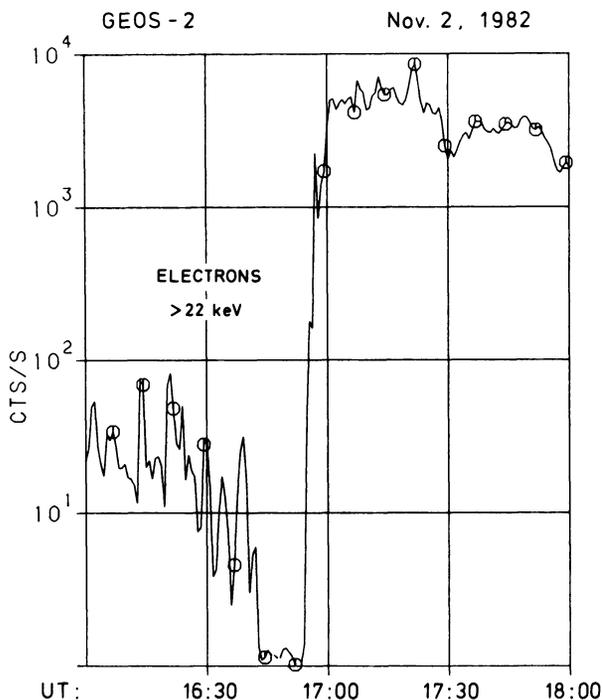


Fig. 3. Count rates of electrons in the spin plane integrated above an energy level of 22 keV. The data were averaged over 45 s. Pulsations in the period range 250–500 s before and after substorm onset are present

component amounted to 80 nT. At the beginning of the collapse the field changed by -0.2 nT/s.

During the transition from a tail-like towards a dipolar configuration, strong ULF emissions with short irregular pulsations (SIPs) occurred which are typical of the substorm onset (Shepherd et al., 1980). The SIPs are interpreted by Robert et al. (1984) as the signature of current structures parallel to the magnetic field passing by the spacecraft with a high velocity. In their analysis they obtain an average drift velocity of 70 km/s, an average current density of 8×10^{-8} A/m² and an average dimension of 200 km. In our case, pulsations are observed in the D_x component with a period of 265 ± 60 s (Fig. 1b).

For completeness, E -field measurements would be highly desirable. Unfortunately, the electron beam experiment on GEOS-2 failed around June 1979 and since the data received from the electric field experiment (Pedersen et al., 1978) need such serious corrections due to photo-electrons and the wake effect (O. Bauer, personal communication), we decided not to include them.

Summary of ground observations

Figure 1 shows the simultaneously recorded ground observations of the cosmic radio noise absorption (Fig. 1c) at the station Rovaniemi and the geographic north component of the magnetic field at Muonio (Fig. 1d). A detailed description of these observations is contained in Paper 1. For completeness, the main characteristics are listed in the following:

- The amplitudes of the magnetic pulsations reached values of up to 1000 nT in the horizontal component.
- The period of the pulsations was 340 ± 50 s.
- The magnetic perturbations on the ground originated from almost two-dimensional ionospheric current systems with a width of the order of 20 km.
- The current bands drifted like parallel wavefronts in a southwest direction with a velocity $v_d = 2.3 \pm 0.3$ km/s. They have been observed down to latitudes corresponding to $L < 3.3$.
- The distance between the wavefronts was more than a factor 10 larger than the width of the current bands, so that the magnetic perturbations could be observed on the ground.
- There was a nearly one-to-one correspondence between the drifting wavefronts and the locally confined enhancement in ionization.
- The simultaneous onset of a substorm was indicated by mid-latitude Pi 2 activity.
- The current bands we observed were on the poleward side of the well-developed Harang discontinuity.

Excitation of kinetic Alfvén waves

A theory by which most of the observed features can be explained coherently has been provided by Hasegawa (1976) and was further discussed by Goertz (1983). This theory combines the creation of narrow structures like auroral arcs and small-scale ionospheric current bands by kinetic Alfvén waves with the possible

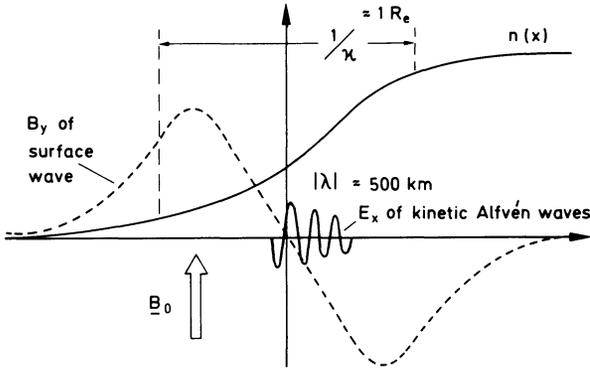


Fig. 4. Profile of the surface wave, the density variation and the kinetic Alfvén wave (after Hasegawa, 1976)

source of these waves in the magnetosphere. The basic mechanism is the resonant mode conversion of large-scale MHD surface waves to kinetic Alfvén waves.

At boundaries within the magnetosphere like the plasmopause, the inner edge of the plasma sheet and the plasma sheet boundary layer, surface waves can be excited by macroscopic MHD instabilities, e.g. the Kelvin-Helmholtz instability. The dispersion relation for a surface wave along the plasma-vacuum interface is given by (Hasegawa, 1976):

$$\omega_s = \sqrt{2} k_{\parallel} v_A,$$

where $v_A = B_0 / \sqrt{\mu_0 n_i m_i}$ is the Alfvén velocity and k_{\parallel} is the wavenumber parallel to the ambient magnetic field. The index s denotes surface wave parameters.

If the plasma density varies smoothly between zero at $x \rightarrow -\infty$ and n_0 at $x \rightarrow \infty$ (Fig. 4) the eigenmode of local shear Alfvén waves can be in resonance with the surface wave and, therefore, can extract energy from the surface wave. The coupling occurs where the local Alfvén velocity $v_A(x)$ is equal to $\sqrt{2} v_A(x \rightarrow \infty)$.

The sharp density gradient in the keV electrons and ions (Fig. 2) observed on GEOS is regarded as the inner edge of the expanding plasma sheet. The ratio of magnetic to kinetic energy density β is of the order of $0.3 \gg m_e/m_i$. We are therefore dealing with the warm plasma case. Then the ion gyroradius ρ_i becomes a critical length for small-scale perturbations. Under simplified conditions, i.e. the plasma density is taken locally constant, near the resonance point the dispersion relation of the Alfvén waves has the form [Hasegawa, 1976; Eq. (23)]:

$$\omega^2 = k_{\parallel}^2 v_A^2 (1 + k_x^2 \rho^2); \quad (1)$$

with

$$\rho^2 = \left(\frac{3}{4} + \frac{T_e}{T_i} \right) \rho_i^2. \quad (2)$$

This expression implies a propagation of the wave across the ambient magnetic field. (T_e , T_i are the temperatures of the electron and ion gas, respectively, and ρ_i is the ion gyroradius.)

The solution of the wave equation for the perpen-

dicular electric field derived by Hasegawa [his Eq. (31)] is given by

$$E_x = -\frac{\pi^{1/2} E_0}{(\kappa \bar{\rho})^{2/3}} \left(\frac{\delta}{x} \right)^{1/4} \exp \left\{ i \left[\frac{2}{3} \left(\frac{x}{\delta} \right)^{3/2} + \frac{\pi}{4} \right] \right\} + \frac{E_0}{\kappa x} \quad \text{for } x > 0. \quad (3)$$

In this equation E_0 is the amplitude of the surface electric wave, κ is the inverse scale length of the plasma gradient, which is taken to be linear. $\bar{\rho}$ is similar to ρ in Eq. (2), but differs by a factor multiplied to the temperature ratio. This factor has a value between 0 and 1 depending on the plasma density fraction $n(x)/n(x \rightarrow \infty)$ at the resonance point. $\delta = (\bar{\rho}^2 / \kappa)^{1/3}$ is also a scale length.

The first term in Eq. (3) represents the kinetic Alfvén wave and the second term the MHD surface wave. The wave equation solution for the low-density side of the surface ($x < 0$) contains only the surface wave, i.e. under given conditions of a warm plasma the Alfvén wave propagates only into the higher-density side of the resonance point. The perpendicular wavelength of the kinetic Alfvén wave is given by the exponent of Eq. (3)

$$\lambda_x = 0.7 \delta. \quad (4)$$

A sketch of the spatial distribution of the plasma density, the surface wave fields and E_x of the kinetic Alfvén wave is shown in Fig. 4.

The amplitude E_x of the kinetic Alfvén wave near the resonance point can be expressed in terms of the y component of the surface wave magnetic field B_{sy} if we replace E_0 in Eq. (3) by $v_A B_{sy}$ and consider the region where $\kappa \cdot x$ is of the order 1. Then we get

$$|E_x| = \frac{v_A}{\sqrt{\kappa \rho}} B_{sy}, \quad (5)$$

and hence the magnitude of the perpendicular potential drop ϕ becomes

$$|\phi| = \frac{1}{k_x} \frac{v_A}{\sqrt{\kappa \rho}} B_{sy}. \quad (6)$$

Goertz (1983) derived the ratio of the parallel potential drop ψ to ϕ in the case of warm plasma

$$\frac{\psi}{\phi} \approx 2 \left[\kappa \rho_i \cdot \frac{\left(\frac{T_e}{T_i} \right)^3}{\left(\frac{3}{4} + \frac{T_e}{T_i} \right)} \right]^{2/3}. \quad (7)$$

Finally, the equation for the parallel current density j_z , given by Hasegawa (1976), has the form

$$j_z = \frac{k_x}{\mu_0} \frac{T_i}{T_e} \frac{B_{sy}}{\sqrt{\kappa \rho_i}}. \quad (8)$$

The results from these equations will now be confronted with observational results.

Comparison of observational and theoretical results

First of all we have to identify the MHD surface wave to which the Alfvén waves can couple and which can provide enough energy. The properties of the surface wave are not too restricted by the theory. Hasegawa (1976) has emphasized that the coupling does not depend on the frequency of the surface wave. We only have to look for a large-scale $\partial B/\partial t$. This is given by the reconfiguration of the dipole field occurring simultaneously with the crossing of the plasma sheet boundary.

The excess magnetic field energy near the inner edge of the plasma sheet, which has partly been converted to kinetic Alfvén waves, can be estimated by the equation

$$E_B = d_{\perp} \cdot d_{\parallel} \frac{(\Delta B)^2}{2\mu_0}.$$

E_B is magnetic energy per length along the plasma sheet boundary. d_{\perp} is the thickness of the boundary in x taken roughly as $1R_e$ (earth radius), which can be considered as a lower limit derived from the decrease in D_x (Fig. 1b), and an assumed drift velocity of the convecting plasma sheet of about 70 km/s. d_{\parallel} is the extent of the tail-like field configuration perpendicular to the equatorial plane. With $d_{\parallel} \approx 5R_e$ and $\Delta B = 80$ nT we get:

$$E_B \approx 5 \times 10^5 \text{ J/m}.$$

This value has to be compared with the energy density of the excited waves.

For mode conversion we have to consider the y component of the large-scale magnetic field perturbation. This component is generated either by the cross-tail currents themselves which are responsible for the tail-like field distortion or by the field-aligned currents in the plasma sheet boundary layer, by which the interrupted equatorial currents are closed via the ionosphere. In both cases we have to take the field component along the plasma sheet boundary. For the following computations we assume $B_{sy} = 10$ nT, which seems to be a conservative estimate.

The temperatures T_e , T_i are taken to be equal, hence we get $\rho = 1.3\rho_i$ in Eq. (2). The uncertainties in the actual values justify the assumption $\bar{\rho} \approx \rho_i$. A 20-keV proton has a $\rho_i \approx 200$ km in a magnetic field of flux density $B = 100$ nT.

For the estimation of the scale length $1/\kappa$ of the plasma gradient we need the earthward-directed drift velocity of the expanding plasma sheet. On the ground we have observed equatorward-directed velocities around 2.3 km/s. Projected on the equatorial plane, taking into account the divergence of the field lines in the meridional plane by a factor of 30 for a $L = 6.6$ field line, we get $v_d = 70$ km/s. Together with the rise time of the warm plasma density on GEOS (see Fig. 2) of about 30 s for the ions and 130 s for the stepwise increasing electron flux, we achieve a scale length $1/\kappa = 1R_e$ as a mean value ($1/\kappa = 0.3R_e$, if we regard only the warm ions). Hence from Eq. (4) we get the perpendicular wavelength of the kinetic Alfvén wave in the equatorial plane, $\lambda_x \approx 430$ km (≈ 300 km). If we project

these values down along the field lines, we have

$$\lambda_x \approx 15 \text{ km} \quad (\approx 10 \text{ km})$$

at the ionospheric altitude. This is in good agreement with the observational result of 20 km for the width of the ionospheric current band.

Next we compare the amplitude of the wave electric field, given by Eq. (5), with our observations. With the above values for κ , ρ and B_{sy} , and taking $v_A = 1,000$ km/s, as a typical value for the equatorial plane, we get $|E_x| \approx 60$ mV/m near the source region of the Alfvén waves. Fields of this order of magnitude have been published by Aggson et al. (1983). Again, projected down to the ionospheric level, the electric field of the downward-travelling Alfvén wave has an amplitude

$$E_x \approx 1.8 \text{ V/m}.$$

From ground observations we deduced $E_x \approx 2$ V/m (see Paper 1). The potential ϕ is then easily computed in accordance with Eq. (6) by multiplying E_x with $\lambda_x/2\pi = 1/k_x \approx 3 \times 10^3$ m:

$$\phi \approx 5 \text{ kV}.$$

And hence, from Eq. (7), we have for the parallel potential drop over the wavelength of the guided Alfvén waves

$$\psi \approx 1 \text{ kV}.$$

The surface wave, which is mode converted to the kinetic Alfvén wave, does not change in time harmonically but varies with nearly constant $\partial B/\partial t$. For an observer moving with the inner edge of the plasma sheet nothing happens in time. For this reason, trapping by the wave and direct acceleration by the parallel potential may account for the energization and precipitation of electrons. The estimated value of ψ is somewhat small but of the right order of magnitude. Typical energy values for auroral electrons are between 2 and 3 keV.

Landau damping may also be considered as a possible wave-particle interaction for electrons which move with nearly the phase velocity of the Alfvén wave. Since v_A grows from about 1,000 km/s near the equator up to a maximum of 10,000 km/s at a height of $1-2R_e$, thermal electrons can be accelerated up to a few ten keV. The efficiency of this mechanism depends on the fraction of resonant electrons in the plasma sheet and a detrapping of the electrons when the wave approaches the ionosphere with decreasing phase speed.

The field-aligned current density associated with the kinetic Alfvén wave can be estimated by Eq. (8). Near the equator we obtain $j_z \approx 0.6 \mu\text{A/m}^2$. At the ionospheric altitude the projected value comes out as

$$j_z \approx 200 \mu\text{A/m}^2.$$

This agrees with the observational result.

Finally, we compare the energy carried by the kinetic Alfvén wave with the released magnetic energy E_B . The integrated energy flux per length along the wavefront is approximately given by:

$$E_w = \lambda_x \Delta t \frac{E_1^2}{v_A \mu_0}.$$

Δt is the time period of the wave to travel along the ambient magnetic field over one wavelength. $\Delta t = \lambda_z / v_A = \lambda_x / v_d \approx 10$ s. We obtain

$$E_w \approx 1.5 \times 10^4 \text{ J/m}$$

near the equatorial plane. Hence the wave energy in a wavefront in one hemisphere makes up 3% of the total available free magnetic energy.

Discussion

The agreement between observational and theoretical results is striking. It is even so good that one may forget the estimative character of the theoretical quantities. Therefore, we wish to point out that only the order of magnitude of the quantities is relevant for their comparison.

The good agreement confirms our opinion that the pulsation event of 2 November 1982 can be explained in terms of kinetic Alfvén waves generated along the inner edge of the plasma sheet. This explanation is also supported by the fact that the pulsation occurred only on the poleward side of the plasma sheet boundary layer (see Paper 1). According to Eq. (3), the kinetic Alfvén waves propagate into the dense region when the warm plasma case applies. The observed drift velocity on ground then results from the difference between the earthward-directed drift of the warm plasma sheet plasma and the roughly oppositely directed propagation of the kinetic Alfvén waves. The direction of the drift velocity on the ground is determined by the relative magnitude of both velocities concerned. For similar events during that winter (not discussed here), poleward-directed drifts of the wavefronts have also been observed.

After the studies of Lysak and Carlson (1981) and Lysak and Dum (1983) the question arises whether our interpretation of the observations may be in conflict with their theoretical results, i.e. that damping processes must be taken into account. They would cause a reflection of the Alfvén waves at heights above the ionosphere. Up to an altitude of about $4R_e$, where the cold plasma approximation is valid, kinetic Alfvén waves can be affected by anomalous resistivity due to microscopic turbulences (Lysak and Carlson, 1981). The turbulence is generated when the drift velocity, $v_D = j_z / ne$, exceeds v_c , the critical drift for instability. Hence, the reflection of Alfvén waves at turbulent regions in the model of Lysak and Dum (1983) becomes more effective with increasing current or decreasing scale size. At altitudes around 6,000 km where v_D reaches its maximum, we get, for a plasma density $n = 5 \times 10^7 \text{ m}^{-3}$ and $j_z = 30 \mu\text{A/m}^2$, $v_D \approx 4,000 \text{ km/s}$. For $T_i = 100 \text{ eV}$ we have $v_c \approx 2,000 \text{ km/s}$. These calculated values do not allow a clear decision as to whether we have to expect a partial decoupling of the ionosphere from the magnetosphere or not. In situ measurements in this region would allow for more precise statements. The existence of a turbulent region would also cause a broadening of the wavefront by a factor of two or

more. But this effect is too small to lead to significant discrepancies in our comparison.

In the same altitude region wave energy can be converted to particle energy by Landau damping (Fejer and Kan, 1969; Klöcker, 1982). For this wave-particle interaction a sufficient amount of resonating electrons has to be present in the magnetic flux tube. The resonant growth rate also depends on the scale size of the Alfvén waves perpendicular to the magnetic field. The existence of precipitated electrons during the event of 2 November can be interpreted as an indication for such a coupling process. But for final conclusions we would need more information, such as the spectral energy distribution of the auroral electrons.

The next step in this study of kinetic Alfvén waves will be to look for more pulsation events in the evening sector occurring concurrently with the onset of substorms in order to identify general characteristics of these phenomena and in order to get more reliable quantities for a comparison with theory.

Acknowledgements. We thank Dr. H. and A. Ranta from Sodankylä Geophysical Observatory for providing the riometer recording. The EISCAT Magnetometer Cross is a joint enterprise of the Finnish Meteorological Institute, the Geophysical Observatory of Sodankylä and the Technical University of Braunschweig. The German part has been funded by grants of the Deutsche Forschungsgemeinschaft.

References

- Aggson, T.L., Heppner, J.P., Maynard, N.C.: Observations of large magnetospheric electric fields during the onset phase of a substorm. *J. Geophys. Res.* **88**, 3981–3990, 1983
- Baker, D.N., Hones, jr., E.W., Belian, R.D., Higbie, P.R., Lepping, R.P., Stauning, P.: Multiple-spacecraft and correlated riometer study of magnetospheric substorm phenomena. *J. Geophys. Res.* **87**, 6121–6136, 1982
- Baumjohann, W., Glaßmeier, K.-H.: The transient response mechanism and Pi 2 pulsations at substorm onset. *Planet. Space Sci.* **31**, 1984 (in press)
- Fejer, J.A., Kan, J.R.: A guiding centre Vlasov equation and its application to Alfvén waves. *J. Plasma Phys.* **3**, 331–351, 1969
- Fejer, J.A., Lee, K.F.: Guided propagation of Alfvén waves in the magnetosphere. *J. Plasma Phys.* **1**, 387–406, 1967
- Goertz, C.K.: Nonstationary coupling between the magnetosphere and ionosphere. Sixth ESA-Sympos. European rocket balloon programme. (ESA SP-183), 221–228, 1983
- Goertz, C.K., Boswell, R.W.: Magnetosphere-ionosphere coupling. *J. Geophys. Res.* **84**, 7239–7246, 1979
- Hasegawa, A.: Particle acceleration by MHD surface wave and formation of aurora. *J. Geophys. Res.* **81**, 5083–5090, 1976
- Hasegawa, A.: Kinetic properties of Alfvén waves. *Proc. Indian Acad. Sci.* **86A**, 151–174, 1977
- Klöcker, N.: Observation of guided ULF-waves correlated with auroral particle precipitation theoretically explained by negative Landau damping. *J. Geophys.* **51**, 119–128, 1982
- Korth, A., Wilken, B.: New magnetic electron spectrometer with directional sensitivity for a satellite application. *Rev. Sci. Instrum.* **49**, 1435–1440, 1978
- Korth, A., Kremser, G., Wilken, B.: Observations of substorm associated particle-flux variations at $6 \leq L \leq 8$ with GEOS-1. *Space Sci. Rev.* **22**, 501–509, 1978
- Lühr, H., Klöcker, N., Thürey, S.: Ground-based observations

- of a very intense substorm-related pulsation event. *J. Geophys. Res.* **55**, 41–53, 1984
- Lysak, R.L., Carlson, C.W.: The effect of microscopic turbulence on magnetosphere-ionosphere coupling. *Geophys. Res. Lett.* **8**, 269–272, 1981
- Lysak, R.L., Dum, C.T.: Dynamics of magnetosphere-ionosphere coupling including turbulent transport. *J. Geophys. Res.* **88**, 365–380, 1983
- Mallinckrodt, A.J., Carlson, C.W.: Relations between transverse electric fields and field-aligned currents. *J. Geophys. Res.* **83**, 1426–1432, 1978
- Pashin, A.B., Glaßmeier, K.H., Baumjohann, W., Raspopov, O.M., Yahnin, A.G., Opgenoorth, H.J., Pellinen, R.J.: Pi2 magnetic pulsations, auroral break-ups, and the substorm current wedge: a case study. *J. Geophys. Res.* **51**, 223–233, 1982
- Pedersen, A., Grard, R., Knott, K., Jones, D., Gonfalone, A.: Measurements of quasi-static electric fields between 3 and 7 earth radii on GEOS-1. *Space Sci. Rev.* **22**, 333–346, 1978
- Robert, P., Gendrin, R., Perraut, S., Roux, A., Pedersen, A.: GEOS-2 identification of rapidly moving current structures in the equatorial outer magnetosphere during substorms. *J. Geophys. Res.* **89**, 819–840, 1984
- Samson, J.C.: Pi2 pulsations: high latitude results. *Planet. Space Sci.* **30**, 1239–1247, 1982
- Shepherd, G.G., Boström, R., Derblom, H., Fälthammar, C.-G., Gendrin, R., Kaila, K., Korth, A., Pedersen, A., Pellinen, R., Wrenn, G.: Plasma and field signatures of poleward propagating auroral precipitation observed at the foot of the GEOS-2 field line. *J. Geophys. Res.* **85**, 4587–4601, 1980

Received December 14, 1984; Revised February 28, 1985
Accepted March 28, 1985

*Short communication***Velocity-viscosity correlation in convection cells
with non-uniform viscosity**

U.R. Christensen

Institut für Geowissenschaften, Universität Mainz
and Max-Planck-Institut für Chemie*, Saarstraße 23, D-6500 Mainz, Federal Republic of Germany

Abstract. The correlation between the logarithm of local viscosity η and local velocity v ($\log |v| \cong a - b \log \eta$) in numerical models of variable viscosity convection is studied. Nineteen selected cases of 2-D stationary convection with different viscosity laws including non-Newtonian rheology, different Rayleigh numbers and different aspect ratios are studied. The quality of correlation is very different, ranging from non-existent to very good. The coefficient b is found to be below 0.5 in almost all cases. Heuristic arguments on the structure of mantle convection which are based on assumptions as $|v| \propto \eta^{-1}$ therefore appear unfounded.

Key words: Convection – Viscosity**Introduction**

The rheology of mantle rock is strongly temperature and pressure dependent. Possibly it is also non-Newtonian or stress dependent. Some attempts have been made to determine the influence of non-constant viscosity on the structure of mantle convection by solving the equations of motion and energy transport (e.g. Torrance and Turcotte, 1971; Houston and DeBremaecker, 1975; Schmeling and Jacoby, 1981). However, these efforts have been sporadic and did not attack the problem in a very systematic way. Recently I have studied the heat transport properties of variable viscosity convection in a large number of model cases (Christensen, 1984a, 1985). These models provide the data basis for the present investigation.

The lack of exact systematic solutions for the variable viscosity convection problem has often been bypassed by the use of scaling analysis, heuristic arguments and intuition. For example, it appears reasonable to assume a higher flow velocity v in parts of the convection cell where the viscosity η is lower than in those where η is high. It has been speculated that the thermal state of the mantle may be closer to isoviscous than to adiabatic because the advective heat transport should be less efficient in regions of high viscosity. In order to study this issue by scaling analysis, Fowler (1983) as-

sumed that local velocity and local viscosity correlate like

$$|v| \propto \eta^{-1}. \quad (1)$$

Such a relation, if it could be confirmed, would have important consequences. A viscosity increase by two orders of magnitude from the upper to the lower mantle (Hager, 1984) would reduce the velocity from $O(1 \text{ cm/year})$ to $O(0.1 \text{ mm/year})$ in the lower mantle. Such low velocity would significantly influence ideas about mixing of mantle heterogeneities and tapping of geochemical 'reservoirs'. Although Eq. (1) may appear intuitively appealing, the arguments to support it are rather weak. Given a certain stress level, it is not the velocity but its spatial derivatives which are related to the viscosity. Furthermore, there is no a priori reason to assume the same level of stress in regions of different viscosity within the cell. In a previous comment to Fowler's paper (Christensen, 1983), I have presented some qualitative arguments as to why I expect the relation between v and η to be much weaker than Eq. (1) assumes. However, it appears useful to determine the correlation between local velocity and local viscosity systematically from numerical solutions of variable viscosity convection.

Results

The finite element solutions which are used to study the correlation are described in more detail elsewhere (Christensen, 1984a, 1985). They are for two-dimensional steady state convection in rectangular boxes. In most case studies the boundaries are stress-free, the temperature difference from top to bottom is fixed and there are no internal heat sources. Besides the aspect ratio l , a Rayleigh number Ra_0 based on the viscosity at the top boundary and two rheological parameters (θ and ζ) describing temperature and pressure dependence determine the state of convection:

$$Ra_0 = \frac{\alpha g \rho \Delta T h^3}{\kappa \eta_0} \quad (2)$$

$$\eta = \eta_0 \exp(-\theta \bar{T} + \zeta \bar{z}). \quad (3)$$

α stands for the coefficient of thermal expansion, g for the gravitational acceleration, ρ for the density, ΔT is

* Mailing address

Table 1. Model parameters, Nusselt-numbers and correlation parameters

Model	Ra_0	l	$e^{\theta-\zeta}$	e^ζ	n	Nu	a	b	r
1	1,000	1	1,000	1	1	4.49	1.024	0.383	-0.719
2	10,000	1	1,000	1	1	10.06	1.909	0.289	-0.506
3	100,000	1	1,000	1	1	25.58	2.770	0.219	-0.279
4	1,250	1	64,000	1	1	6.64	0.760	0.518	-0.886
5	3,750	1	64,000	1	1	9.15	1.202	0.488	-0.829
6	1,897	1	250,000	1	1	8.46	0.908	0.519	-0.878
7	10,000	2	1,000	1	1	9.11	2.315	0.019	-0.032
8	10,000	3	1,000	1	1	8.46	2.336	0.000	0.000
9	100,000	2	1,000	1	1	20.46	2.960	0.044	-0.049
10 ^a	10,000	1	1,000	1	1	4.73	0.749	0.680	-0.734
11	10,000	1	1,000	64	1	9.31	2.241	0.086	-0.244
12	10,000	1	1,000	1,000	1	8.28	1.942	0.100	-0.354
13	50,000	1	250	1,000	1	10.19	1.920	0.163	-0.606
14	50,000	2	250	1,000	1	8.43	1.936	0.207	-0.723
15	500	1	10 ⁷	1	3	27.10	2.642	0.129	-0.245
16	500	2	10 ⁷	1	3	16.64	2.726	0.043	-0.077
17	1,000	1	10 ⁵	10 ⁵	3	10.98	1.734	0.133	-0.371
18 ^b	1.5 × 10 ⁷	1	7.2 × 10 ⁻⁵	2 × 10 ⁷	1	3.44	2.714	0.460	-0.922
19 ^c	343,600	1	1	1,000	1	9.10	1.925	0.029	-0.113

^a Rigid top and bottom boundary

^b Purely internal heating, zero bottom heat flux. The Nusselt number is obtained by dividing the 'conductive reference temperature' at the bottom by the actual mean bottom temperature

^c 3/4 internal heating, 1/4 bottom heat flux. For Nu see ^b

the temperature difference across the convective layer of height h , κ is the thermal diffusivity and η_0 the viscosity at the upper boundary. \bar{T} is the temperature normalized to zero on top and one at the bottom, \bar{z} is the vertical coordinate normalized in the same way. In some cases a non-Newtonian third-power law rheology is used with an effective viscosity given by

$$\eta = \eta_0 \left(\frac{s}{\eta_0 \kappa / h^2} \right)^{-2} \exp(-\theta \bar{T} + \zeta \bar{z}), \quad (4)$$

where s is the second invariant of the deviatoric stress tensor. Details of the numerical method are described in Christensen (1984b). Convergence tests indicate an accuracy of the solutions concerning the Nusselt number of better than 1% and concerning the local velocity of at least better than 10%.

The correlation is studied by 'sampling' the local viscosity and velocity at N random points within the convection cell, where N is typically 500. A correlation of the form

$$|v| \propto \eta^{-b} \quad (5)$$

or

$$\log |v| = a - b \log \eta \quad (6)$$

is assumed. The optimal parameters a and b in Eq. (6) are determined by a least-squares fit and the usual correlation coefficient

$$r = \frac{\sum \log \eta \log |v| - \frac{1}{N} \sum \log \eta \sum \log |v|}{\left\{ \left[\sum (\log \eta)^2 - \frac{1}{N} (\sum \log \eta)^2 \right] \left[\sum \log |v|^2 - \frac{1}{N} (\sum \log |v|)^2 \right] \right\}^{1/2}} \quad (7)$$

is taken as a measure of the quality of the fit. One complication must be considered. At some points, like the corners of the cell or the centre of the circulation, the velocity is zero while the viscosity may have arbitrary values. The velocity vanishes because of geometrical reasons and not because of high viscosity. To include the vicinity of these points may obscure an intrinsic correlation between η and $|v|$. At these points $v \rightarrow 0$, whereas either the strain rate $\dot{\epsilon}$ or the vorticity ω or both remain large. When, on the other hand, $|v|$ becomes small due to high local viscosity, both $\dot{\epsilon}$ and ω should likewise be small. To get rid of the stagnation points and their immediate vicinity, I thus rejected all points where

$$\frac{v^2}{\dot{\epsilon}^2 + \omega^2} < K^2 \quad (8)$$

with K of the order of 0.04. Typically 2%-5% of the samples were rejected. Varying K within reasonable limits had only a small effect on the values of a , b and r , but taking it as zero significantly deteriorated the correlation in some cases.

In Table 1 the results for all 19 models are listed. Instead of θ and ζ the values $e^{\theta-\zeta}$ and e^ζ are given. They are, respectively, the actual ratio of top to bottom viscosity due to the combined effects of temperature and pressure dependence, and the hypothetical increase due to the pressure effect alone. For $n=3$ (non-Newtonian power-law creep with stress exponent 3), the actual viscosity difference is much smaller than $e^{\theta-\zeta}$ because of the moderating influence of the stress dependence (Christensen, 1984b).

The quality of the correlation differs strongly from case to case; r values between 0.00 and -0.92 are found. The coefficient b varies in the range 0.00-0.68. In Fig. 1 some typical examples of correlation diagrams

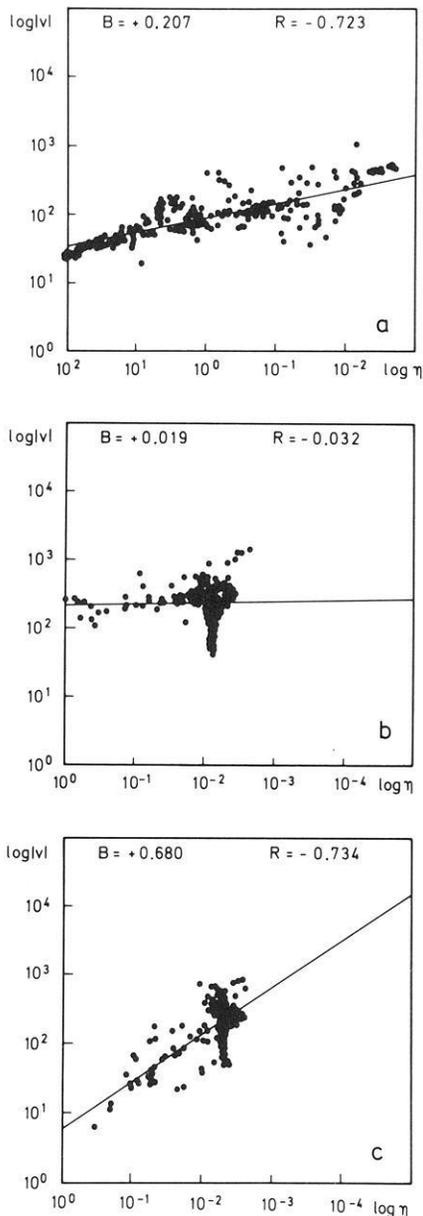


Fig. 1 a-c. Correlation diagrams for three convection models. Of 500 points which have been used to determine the correlation, only 200 are plotted. **a** Model number 14 with strong pressure effect on the rheology, good correlation but low b value. **b** Model number 7 with simply temperature-dependent viscosity, aspect ratio 2, no correlation is found. **c** Model number 10 with rigid boundaries and temperature-dependent viscosity, highest b value of all models

are shown. The correlation is good ($|r| > 0.6$) for strong enough temperature dependence of viscosity ($e^{\theta} > 10^4$), where a sluggishly moving surface layer is formed. The b values are of the order of 0.5 in these cases (models 4–6 in Table 1). The correlation is also good with strong pressure influence (models 13 and 14), but here b is only of the order 0.2. The highest value of $b = 0.68$ is obtained for a model with rigid boundaries (model 10). Due to the boundary condition, a stagnant high-viscosity layer is produced near the surface. Given the same viscosity contrast, this layer moves relatively fast with a free upper boundary (leading to lower b). The

best correlation ($|r| = 0.922$) is obtained in a model with purely internal heating and extremely strong pressure influence on the flow law (model 18, $b = 0.46$). In many cases, however, the correlation is quite poor ($|r| < 0.4$) and the b values are small.

The influence of changing the aspect ratio is not clear. With purely temperature-dependent Newtonian or non-Newtonian rheology, the correlation deteriorates when increasing l (models 2, 7 and 8 and models 15 and 16). However, with additional pressure influence there is even a slight improvement on increasing l (models 13 and 14). With non-Newtonian rheology the correlation seems slightly weaker than in equivalent cases of Newtonian rheology (I consider those Newtonian cases equivalent which have a similar Nusselt number and about half the θ and ζ values of the non-Newtonian model, cf. Christensen, 1984b).

Different ways of determining the correlation do not lead to qualitatively different results. If, instead of random sampling, the horizontal and vertical boundary layers are sampled four times more frequently than the centre of the cell, the correlation coefficient improves by typically 0.1–0.2, but b increases by no more than 0.05. If horizontally averaged velocities and viscosities are correlated, the result is comparable to that of point-wise sampling.

Discussion

Although only a restricted number of cases was studied, a wide range of possible parameter combinations was covered. I found that there is either no good correlation between local velocity and viscosity or, if there is a good correlation, the coefficient b is much less than unity. Thus it seems that the assumption $v \propto \eta^{-1}$ must be abandoned for variable viscosity convection and that speculations based on it are misleading. The dependence between v and η is strongest when the surface layer forms a stagnant lid. However, the Earth's surface plates are actually moving and the viscosity in the deep mantle is probably not high enough to produce stagnant regions. The parameter b is then expected to be less than 0.5. Recently, the idea that the viscosity of the lower mantle is significantly higher than that of the upper mantle has found more support (Hager, 1984; Christensen, 1984c). The result of the present study suggests that even with a viscosity of $O(10^{24}$ Poise), two orders of magnitude higher than in the upper mantle, the velocity would probably be more than one-tenth of the typical plate velocity. The overturn time of whole mantle convection would then be of the order of 1 billion years.

References

- Christensen, U.: Comments on "On the thermal state of the earth's mantle" by A.C. Fowler. *J. Geophys.* **53**, 201–202, 1983
- Christensen, U.: Heat transport by variable viscosity convection and implications for the Earth's thermal evolution. *Phys. Earth Planet. Inter.* **35**, 264–282, 1984a
- Christensen, U.: Convection with pressure- und temperature-dependent non-Newtonian rheology. *Geophys. J. R. Astron. Soc.* **77**, 343–384, 1984b

- Christensen, U.: Instability of a hot boundary layer and initiation of thermo-chemical plumes. *Ann. Geophys.* **2**, 311–320, 1984c
- Christensen, U.: Heat transport by variable viscosity convection II: Pressure influence, non-Newtonian rheology, and decaying heat sources. *Phys. Earth Planet. Inter.* **37**, 183–205, 1985
- Fowler, A.C.: On the thermal state of the Earth's mantle. *J. Geophys.* **53**, 42–51, 1983
- Hager, B.H.: Subducted slabs and the geoid: constraints on mantle rheology and flow. *J. Geophys. Res.* **89**, 6003–6016, 1984
- Houston, M.H., DeBremaecker, J.-Cl.: Numerical models of convection in the upper mantle. *J. Geophys. Res.* **80**, 742–751, 1975
- Schmeling, H., Jacoby, W.R.: On modelling the lithosphere in mantle convection with non-linear rheology. *J. Geophys.* **50**, 89–100, 1981
- Torrance, K.E., Turcotte, D.L.: Thermal convection with large viscosity variations. *J. Fluid Mech.* **47**, 113–125, 1971

Received January 28, 1985; Revised version March 19, 1985

Accepted March 20, 1985

The influence of dirt bands and faults on the propagation of Love seam waves

C. Kerner and L. Dresen

Institut für Geophysik, Ruhr-Universität Bochum, PO Box 102148, D-4630 Bochum 1, Federal Republic of Germany

Abstract. Discontinuities in a coal seam can be detected by in-seam seismic surveys. However, a detailed understanding of wave propagation in coal seams interlayered with dirt bands can provide a more efficient use of this technique. In addition, improvements in this method are possible if one knows the effect of the fault parameters on the field data.

This paper presents the results from a numerical simulation of the propagation of Love seam waves in two-dimensional heterogeneous geological structures. The frequency content of the source signal and the profiles in the centre of the coal layers correspond to current field techniques. Finite difference methods are used.

The effect of modes higher than the first mode on the propagation of waves in seams which are interlayered with dirt bands was studied. The Airy phase of the second mode contributes significantly to the seam wave signal if the dirt band lies near the centre of the seam. Interference with the first mode produces seam wave signals with irregular shapes and amplitudes. In practice, these irregularities might be responsible for poor processing results.

Furthermore, fault characteristics were investigated by means of the reflectivity and transmissivity curves. The throw, the dip angle of the fault plane and the impedance in a fault zone were varied and are discussed with respect to fault detection. Comparison of the curves exhibits ambiguities which might be reduced by supplementing the field data by broad-band recordings and data acquisition on profile points in the surrounding rocks.

Key words: Finite difference method – Love seam waves – Wave guide – Dirt band – Coal seam discontinuities – Faults

Introduction

Mechanized longwall coal mining is economical only for exploiting regions of hard coal in which the seam is not disturbed by extensive dirt bands or faults. In fact, these disturbances occur frequently in European coal fields. Their detection provides an estimate of the productivity of the coal field during mine planning. So, a capital-intensive interruption of the coal production can be avoided.

A powerful geophysical detection method is in-seam seismology (Arnetzl and Klinge, 1982; Brentrup, 1979; Buchanan, 1983; Dresen et al., 1985; Millahn, 1980; Rüter and Schepers, 1979). As coal has lower body-wave velocities compared with those of the cheek – i.e. the rock above and below the seam – channel waves are generated by a seismic source inserted into the coal (Krey, 1963). These channel waves are guided within the seam (seam waves) and are recorded at distances as far as 1.5 km from the source (Arnetzl, 1971). Until now, underground surveys have concentrated on Love seam waves.

The in-seam seismic method is well established. However, the manifold effects of the layering of a seam (i.e. dirt band embeddings), or the layering of the cheek (i.e. root clay layers), for example, can lead to an erroneous interpretation of the underground data. On the other hand, more detailed information on the characteristics of a discontinuity (fault, erosion, fold, etc.) or the geometry (fault offset, length of the eroded region in the seam, etc.) is required.

A better understanding of the physical effects of the geological disturbances on the transmission and reflection of seam waves is one way to improve data interpretation regarding target recognition, location and identification. For this purpose, seismic modelling with a layer matrix method was introduced by Räder et al. (1985). Asten et al. (1984) and Edwards et al. (1985) used a finite element method for their studies on seam waves. In this paper, similar to Korn and Stöckl (1982), we have numerically simulated Love seam wave propagation with a finite difference (FD) method. Amplitude and dispersion curves are compared with theoretical values calculated with the layer matrix method. Two topics are studied:

- the effect of rock and/or soil embeddings in the seam (dirt bed or dirt band) on Love seam wave propagation
- the effect of fault characteristics on the reflection and transmission of Love seam waves.

Layered seams have previously been investigated by Räder et al. (1985). From dispersion curves and amplitude-depth distributions these authors concluded that, for dirt band embeddings, wave groups of higher modes will occur in addition to the Airy phase of the first mode. Here, we investigate the effect of interfering modes. Vertical faults have been analysed by Korn and Stöckl (1982). We extend their work and study dipping fault planes and fault zones of finite width which interrupt the continuity of the coal layer.

Methods of computation

Finite difference method

In Cartesian coordinates, the two-dimensional equation of motion for an *SH* wave propagating in heterogeneous media is

$$\frac{\partial^2 v}{\partial t^2} = 1/\rho \left[\frac{\partial}{\partial x} \left(\mu \frac{\partial v}{\partial x} \right) + \frac{\partial}{\partial z} \left(\mu \frac{\partial v}{\partial z} \right) \right], \quad (1)$$

where x and z are the two components of the radius vector, $v(x, z, t)$ is the displacement, $\rho(x, z)$ is the density and $\mu(x, z)$ is the shear modulus.

This equation can be solved numerically by applying a finite difference method: discretization of the medium by a rectangular grid of equal spacing Δx and Δz and time discretization by time steps Δt yield a sampling of the continuous functions v , μ and ρ . The following notation is used:

$$\begin{aligned} v(m \cdot \Delta x, n \cdot \Delta z, l \cdot \Delta t) &= v_{m,n}^l, \\ \rho(m \cdot \Delta x, n \cdot \Delta z) &= \rho_{m,n} \quad (2) \\ \mu(m \cdot \Delta x, n \cdot \Delta z) &= \mu_{m,n}. \end{aligned}$$

Standard central differences with a truncation error of second-order (Boore, 1970) are employed to approximate the partial differentials in the equation of motion. The following recursion formula for the calculation of $v_{m,n}^{l+1}$, the displacement in the grid point m, n at the time $l+1$, is obtained (Boore, 1972):

$$\begin{aligned} v_{m,n}^{l+1} &= 2v_{m,n}^l - v_{m,n}^{l-1} + [\Delta t^2 / \Delta x^2 [\mu_E (v_{m+1,n}^l - v_{m,n}^l) \\ &\quad - \mu_W (v_{m,n}^l - v_{m-1,n}^l)] \\ &\quad + \Delta t^2 / \Delta z^2 [\mu_S (v_{m,n+1}^l - v_{m,n}^l) \\ &\quad - \mu_N (v_{m,n}^l - v_{m,n-1}^l)]] / \rho_{m,n}, \end{aligned} \quad (3)$$

where:

$$\begin{aligned} \mu_E &= 1/2(\mu_{m+1,n} + \mu_{m,n}) & \mu_N &= 1/2(\mu_{m,n} + \mu_{m,n-1}) \\ \mu_W &= 1/2(\mu_{m-1,n} + \mu_{m,n}) & \mu_S &= 1/2(\mu_{m,n} + \mu_{m,n+1}). \end{aligned}$$

At the start of the recursion, in accordance with an arbitrary source signal, the displacements are prescribed at a given source point.

Numerical stability of the explicit scheme in Eq. (3) is guaranteed if the relation $\Delta t = h_{\min} / (\beta_{\max} \cdot \sqrt{2})$ between the time step Δt , the minimum grid spacing $h_{\min} = \text{MIN}(\Delta x, \Delta z)$ and the maximum shear wave velocity in the model $\beta_{\max} = \text{MAX}[(\mu_{m,n} / \rho_{m,n})^{-1/2}]$, is used. Artificial reflections from the edges of the model are sufficiently suppressed by applying the boundary conditions suggested by Reynolds (1978). Grid dispersion is reduced if the smallest predominant wavelength λ_H , defined by the upper frequency of half the maximum amplitude in the source spectrum, corresponds to ten or more grid spacings $h_{\max} = \text{MAX}(\Delta x, \Delta z)$ (Alford et al., 1974).

A large computational effort is needed for FD simulations even on fast machines. This limits the use of the method to problems where the wave propagation over distances of only some few wavelengths is to be investigated. In particular, seam waves can be studied by FD as these waves propagate in the coal and the neighbouring rock. Thus, only a relatively small part of the geological structure has

to be modelled. Nevertheless, 250 grid points in the x -direction and 100 grid points in the z -direction are necessary for even a simple structure. This means that at each time step $250 \times 100 = 25,000$ 'new' displacements have to be calculated. If the wave needs 2,000 time steps to travel through the model, the above recursion formula must be solved 50 million times: CPU-times of around half an hour are required for the computation of such a model on a conventional computer like the Cyber 175.

A remarkable improvement in the speed of the FD program by a factor of 35 was obtained on a Cyber 205 vector machine. Computations on the Cyber 205 demand a coding of the algorithm in 'vector-FORTRAN'. It is the non-recursive portion of the FD algorithm, i.e. the iteration over the grid points, that is performed by parallel vector operations (Kerner, 1985).

Amplitude and dispersion analysis

An analysis of amplitudes and phase-velocity dispersion of the Love seam waves was carried out. In the case of a fault, the maximum displacement amplitudes of the reflected and transmitted waves (A_R, A_T) are determined. The values are normalized by the amplitude of the incident seam wave. Their representation as a function of the parameters of the fault – e.g. the throw – gives concise information about the effect of the parameters on the detectability of the fault.

More detailed information is drawn from a spectral amplitude analysis of the reflected and transmitted seam waves. We analysed the waves recorded in the centre of the seam on both sides of the fault and normalized the spectra by the amplitude spectrum of the incident Love seam wave. Thus, we remove the influence of the source signal and the filtering effects of the layering. Korn and Stöckl (1982) called the normalized spectrum of the reflected wave $S_R(f)$ the 'reflectivity' and that of the transmitted wave $S_T(f)$ the 'transmissivity', where f is the frequency. In the case of a seam interlayered with a dirt band, we are interested in the mode identification of the waves propagating in each of the layers. This is done by analysing the phase velocities and comparing the results with theoretical dispersion curves. A phase difference method was used for the dispersion analysis (Dziewonski and Hales, 1972).

From the amplitude-depth curves $V(f, z)$, the transfer properties of the layering with respect to seam waves recorded in the coal are deduced. In a sequence rock-coal-rock, the ratio of the signal amplitudes recorded within the seam to those recorded in the rock grows with increasing frequency. This implies a high-pass filtering effect in the coal. Two quantities are used to describe the filtering effect.

For each frequency, the kinetic energy density averaged over one period can be determined for a layer of thickness d and density ρ_l with the formula

$$E(f) = -1/4(2\pi f)^2 \rho_l \int_{-d/2}^{+d/2} V^2(f, z) dz. \quad (4)$$

Normalization of the energy density in the layer by the total energy density in all layers, gives values of the 'relative energy'

$$E_R(f) = \rho_l \int_{-d/2}^{+d/2} V^2(f, z) dz \bigg/ \int_{-\infty}^{+\infty} \rho(z) V^2(f, z) dz. \quad (5)$$

For each frequency, the energy transport in any given layer is estimated with respect to the total energy: $E_R=1$ means that the total energy transport is restricted to the layer (e.g. the coal layer), $E_R=0$ indicates that the total energy is transported outside the layer (Dresen and Freystätter, 1976). The relative energy has been determined from amplitude-depth distributions calculated with the layer matrix method.

Instead of the relative energy, the 'relative amplitudes' can be evaluated from the formula

$$V_R(f) = \frac{\int_{-d/2}^{+d/2} V(f, z) dz}{\int_{-\infty}^{+\infty} V(f, z) dz} \quad (6)$$

Relative amplitudes are determined from amplitude-depth distributions extracted from synthetic seismograms. They provide similar information to that given by the relative energies. They were calculated in the case of seismogram analysis to avoid the use of information concerning the layer densities which are not known in real situations.

Models

Figure 1 shows sketches of the models under investigation. They represent vertical sections through earth layers and through tectonic structures. The characteristics and parameters of each model are specified in Table 1. Mean values of shear-wave velocities and densities in different rock and coal layers known from field surveys are assumed in all models (Table 2). As usual in the in-seam seismic technique, the source is positioned in the middle of the thickest coal layer.

Four models are chosen to examine the influence of dirt bands (Fig. 1a). The seam, consisting of two coal layers (c, c_s) and one dirt band (d), is embedded in a homogeneous rock material (r). Since, in nature, the dirt band material is often the same as the rock material, we choose the same velocities and densities in both the rock and dirt bands. In models Db(1) and Db(2) (Table 1), the dirt band separates the seam into two coal layers with equal thickness. In view of the results of Räder et al. (1985), we choose the thickness of the dirt band so that it significantly affects the wave propagation. This means that the dirt band thickness relative to the seam thickness has to take on values in the range from 0.05 to 0.33 (Kerner, 1984). Räder et al. (1985) pointed out that, outside this range, the dirt band is either so thin that it scarcely affects the seam wave or the dirt band causes a separation between the coal layers so large that the seam wave is guided mainly in the source-containing coal layer. In models Dp(1) and Dp(2), the position of the dirt band is varied, yielding two characteristic ratios of coal layer thicknesses.

Two types of fault models are designed for studying the effect of various fault parameters on the seam wave propagation. A fault where the degree of throw is several times greater than the seam thickness was realized by terminating the seam at the fault (seam end, Fig. 1b). Investigation of this fault type is basic because there are no effects from that part of the seam behind the fault. In the second model type, the throw of the fault is less than twice the seam thickness. Hence, the part of the coal layer behind the fault is modelled as well (Fig. 1c-e).

In models Ez(1)–Ez(4), the fault plane is vertical. A fault zone was constructed and the impedance of the mate-

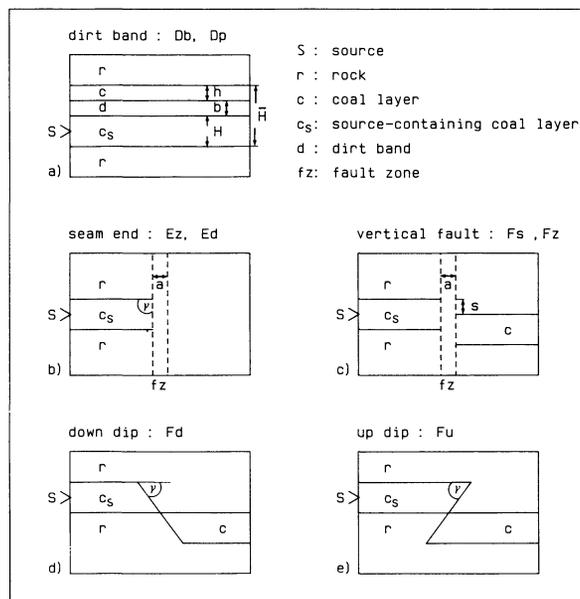


Fig. 1a-e. Sketches of the models. H : thickness of the source-containing coal layer, h : thickness of the coal layer without source in the dirt band models, b : thickness of the dirt band, \bar{H} : thickness of the total seam in the dirt band models, s : throw, a : width of the fault zone, γ : dip angle of the fault plane

rial in the fault zone was varied. Choosing the same elastic parameters in the fault zone as in the rock, in the case of model Ez(1), means that there is no fault zone at all. In model Ez(2), the elastic parameters are chosen so that the impedance in the fault zone is lower than in the coal and rock. In reality, this can be interpreted as disaggregation of the rock material in the zone by fracturing. Choosing an impedance in the fault zone of Ez(4) higher than the impedances in the coal and rock, we model a consolidation of the material in the fault zone. In model Ez(3), an intermediate value is chosen for the impedance in the fault zone. In all the models Ez, the thickness of the fault zone is about 10% of the seam thickness, which means a two-way travel distance of less than a quarter wavelength for waves propagating horizontally through the zone. In models Ed(1)–Ed(3), three different values γ less than 90° are chosen for the dip angle of the oblique fault plane.

In models Fs(1)–Fs(5), the value of the throw relative to the seam thickness is varied. In models Fz, a fault zone is constructed using the same impedances as in the case of models Ez. For all Fz models, the throw is about half the seam thickness. While the fault plane is vertical in models Fs and Fz, in models Fd and Fu the fault plane dips. Models Fd contain down-dipping faults and models Fu contain up-dipping faults. The same angles as in the seam end models Ed are chosen. A constant throw of one seam thickness is used.

The FD grid consists of 250 points in the x -direction and 100 points in z -direction. In the fault models, the thickness of the coal layer is represented by 9 grid points. A grid spacing of 0.4 m in the x - and z -directions is chosen.

The number of grid points for the seam has to be increased to model an interlayering of the coal. We used 15 grid points in the vertical direction for the source-containing coal layer in all dirt band models except Dp(1). In the latter model, 11 grid points represent the seam. Grid

Table 1. Model parameters

Model type	Characteristic	Variable parameters	Constant parameters
Db	Variation of the thickness of the dirt band	1: $b/\bar{H}=0.14$ 2: $b/\bar{H}=0.27$	$h/H=1$
Dp	Variation of the position of the dirt band	1: $h/H=0.82$ 2: $h/H=0.33$	$b/\bar{H}=0.2$
Ez	Variation of the impedance in a fault zone	1: $I_{fz}=6.0 \cdot 10^5 \text{ g}/(\text{cm}^2\text{s})$ 2: $I_{fz}=0.9 \cdot 10^5 \text{ g}/(\text{cm}^2\text{s})$ 3: $I_{fz}=3.5 \cdot 10^5 \text{ g}/(\text{cm}^2\text{s})$ 4: $I_{fz}=8.0 \cdot 10^5 \text{ g}/(\text{cm}^2\text{s})$	$\gamma=90^\circ$ $a/H=0.11$ $s/H \gg 1$
Ed	Variation of the dip angle of the fault plane	1: $\gamma=27^\circ$ 2: $\gamma=45^\circ$ 3: $\gamma=63^\circ$	$a/H=0$ $s/H \gg 1$
Fs	Variation of the offset at the fault	1: $s/H=0.33$ 2: $s/H=0.56$ 3: $s/H=0.78$ 4: $s/H=1.00$ 5: $s/H=1.67$	$a/H=0$ $\gamma=90^\circ$
Fz	Variation of the impedance in a fault zone	1: $I_{fz}=0.9 \cdot 10^5 \text{ g}/(\text{cm}^2\text{s})$ 2: $I_{fz}=3.5 \cdot 10^5 \text{ g}/(\text{cm}^2\text{s})$ 3: $I_{fz}=6.0 \cdot 10^5 \text{ g}/(\text{cm}^2\text{s})$ 4: $I_{fz}=8.0 \cdot 10^5 \text{ g}/(\text{cm}^2\text{s})$	$\gamma=90^\circ$ $s/H=0.56$ $a/H=0.11$
Fd	Down-dipping fault: variation of the dip angle of the fault plane	1: $\gamma=27^\circ$ 2: $\gamma=45^\circ$ 3: $\gamma=63^\circ$	$s/H=1.00$
Fu	Up-dipping fault: variation of the dip angle of the fault plane	1: $\gamma=27^\circ$ 2: $\gamma=45^\circ$ 3: $\gamma=63^\circ$	$s/H=1.00$

Table 2. Elastic parameters

	Shear-wave velocity (km/s)	Density (g/cm^3)	Impedance [$10^5 \text{ g}/(\text{cm}^2\text{s})$]
Rock (dirt band)	2.3	2.6	6.0
Coal	1.2	1.4	1.7

spacings of 0.2 m in the z -direction and 0.4 m in the x direction are chosen.

We used Küpper's wavelets (Küpper, 1958) as source signals, with two amplitude extrema. The duration of the wavelet is chosen such that the spectral amplitudes of the generated seam waves are higher than 5% of the maximum spectral amplitude in a range from 200 to 1400 Hz·m on the $f \cdot H$ scale (f : frequency; H : thickness of the source-containing coal layer. We use this scale, instead of the frequency scale, to make the results independent of the thicknesses of the layers chosen in the models). The maximum spectral amplitude occurs at 800 Hz·m (fault models), or close to this value (dirt band models). At this value, the group-velocity dispersion curve of the first mode takes on its minimum value. So, the generated seam wave contains the Airy phase of the first mode which is of special interest in the field survey due to its low amplitude decrease with distance.

For all models except Dp(1), the numerical errors from grid dispersion are less than 5% for the amplitudes and less than 3.5% for the phase velocities in the $f \cdot H$ -range

below 1400 Hz·m. The same holds for the model Dp(1), below 1100 Hz·m. Interference of the signals with residual artificial reflections from the edges of the model can cause errors of a maximum of 5% in the phase-velocity analysis.

Seams interlayered with dirt bands

Bands of different thicknesses

The seismogram sections displayed in Fig. 2 are calculated for profiles running vertically through the layer sequences in models Db(1) and Db(2). The layer interfaces are marked on the vertical axis.

The seismogram examples show that one part of the seam wave, occurring in both sections in the time interval from about 36 to 90 ms, is guided within the total seam. Amplitudes of this part of the seam wave are high in both the coal layers and in the dirt band. Another part of the seam wave, occurring in the time interval from about 90 to 108 ms, exhibits high amplitudes only within the coal layers while the amplitudes of the phases recorded within the dirt band are nearly zero. For this part of the seam wave, only the coal layers act as wave guides.

In addition to this information, which was also obtained from the dispersion curves and the amplitude-depth distributions (Räder et al., 1985), the seismograms recorded in the coal show that both parts of the seam wave occur simultaneously, forming an extremely long seam wave signal. The two wave parts do not form a continuous wavelet as would be expected if both parts belong to the same (first) mode, but there are gaps with low amplitudes at times of

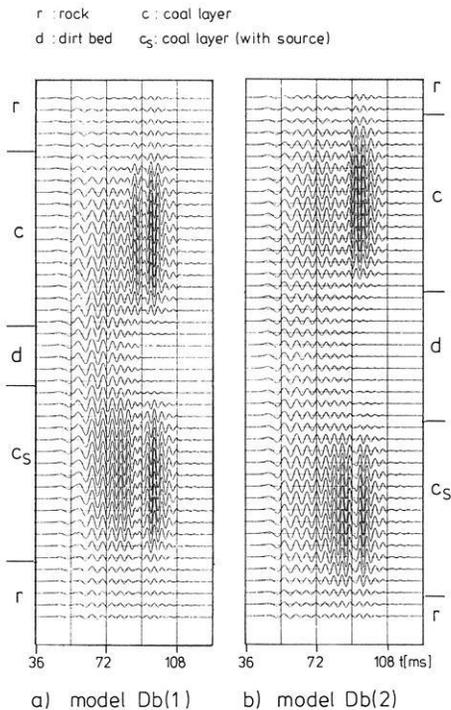


Fig. 2a and b. Seismogram sections calculated for vertical profiles in the models Db. Distance of the profiles from the source position: $x/H=30.7$

about 90 ms (Fig. 2a: refer to the signals in c and c_s; Fig. 2b: refer to the signals in c_s).

Phase-velocity analysis was carried out to determine the modes with which the seam wave phases have to be associated. Seismogram sections calculated for horizontal profiles positioned in the centre of either of the two coal layers were analysed. For the models Db, the results together with the phase- and group-velocity dispersion curves of the first and the second mode are displayed in Fig. 3.

Both graphs (Fig. 3a and b) show that the low-frequency part of the seam wave propagating in the total seam belongs to the first mode. The results agree with the theoretical curve in the $f \cdot H$ -range from 300 to 650 Hz·m, where an indentation characterizes the group-velocity dispersion curves of the first mode. However, the high frequency parts of the seam wave propagating within the coal layers in Db(1) and Db(2) belong to the Airy phase of the second mode in the $f \cdot H$ -range from 650 to 900 Hz·m. Above 900 Hz·m the phase velocities from the analysis correlate with the phase-velocity dispersion curves of both the first and the second mode.

The following should make plausible why the change from the first to the next higher mode occurs. From the amplitude-depth distributions (Räder et al., 1985), it is obvious that the source position in the centre of the coal layers is optimal for an excitation of either the first or the second mode in an $f \cdot H$ -range from about 600 to 1200 Hz·m. Maximum amplitudes occur in this position. Kerner (1984) demonstrated that sharp group-velocity minima, i.e. steep descents in the vicinity of minima, yield a sharp impulse-like Airy phase signal with large amplitudes, while smooth minima yield less prominent Airy phase signals. Hence, it is obvious that in combination with optimal source conditions for generating both modes, the preferential excitation of

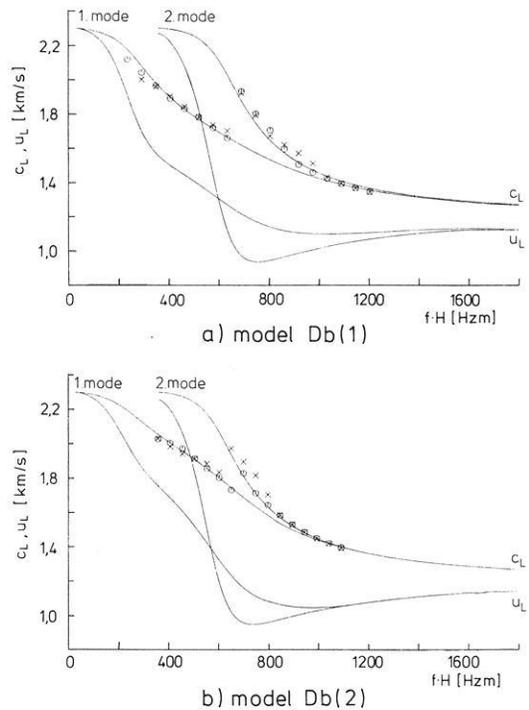


Fig. 3a and b. Phase(c_L)- and group(u_L)-velocity dispersion curves of the first and second mode and results from the phase-velocity analysis of the seam waves propagating in the source-containing coal layer (crosses) and in the coal layer without source (circles)

the second mode is caused by the more pronounced group-velocity minimum of this mode, as apparent in Fig. 3.

For model Db(1), the relative energy versus $f \cdot H$ (Fig. 4) is plotted. The curves derived from the theoretical amplitude-depth distributions were computed for the first and second mode. They describe the portions of energy guided in each of the coal layers. As in the case of the simple layer sequence rock-coal-rock, the general trend of the curves indicates the high-pass filtering effect on the amplitudes of the seam wave phases guided in the coal layers as well as in the case of the layering with dirt band. The results from the analysis (Fig. 3a) agree with the information obtained from the energy distributions (Fig. 4). In the $f \cdot H$ -range below 650 Hz·m, where the first mode dominates, the energy of the first mode is much higher than the energy of the second mode. In the range from 650 to 900 Hz·m, the curve of the relative energy determined for the second mode lies above the one calculated for the first mode. Above 900 Hz·m, both curves approach the maximal value of relative energy guided in each of the two coal layers ($E_R=0.5$). This means that there is no preferential excitation of either of the two modes. Similar energy curves were obtained for model Db(2).

Curves of the relative amplitude distribution were calculated for the models Db from amplitudes extracted from the seismogram sections presented in Fig. 2. In contrast to the curves of relative energy, the curves for the relative amplitude exhibit effects which result from an overlapping of phases of the first and the second mode. The curves of relative amplitude depend specifically on the distance of the vertical profile from the source and the source position. The curves in Fig. 5 describe the dependence on $f \cdot H$ of the relative amplitudes of seam wave phases guided in the two coal layers. There are significant effects on the

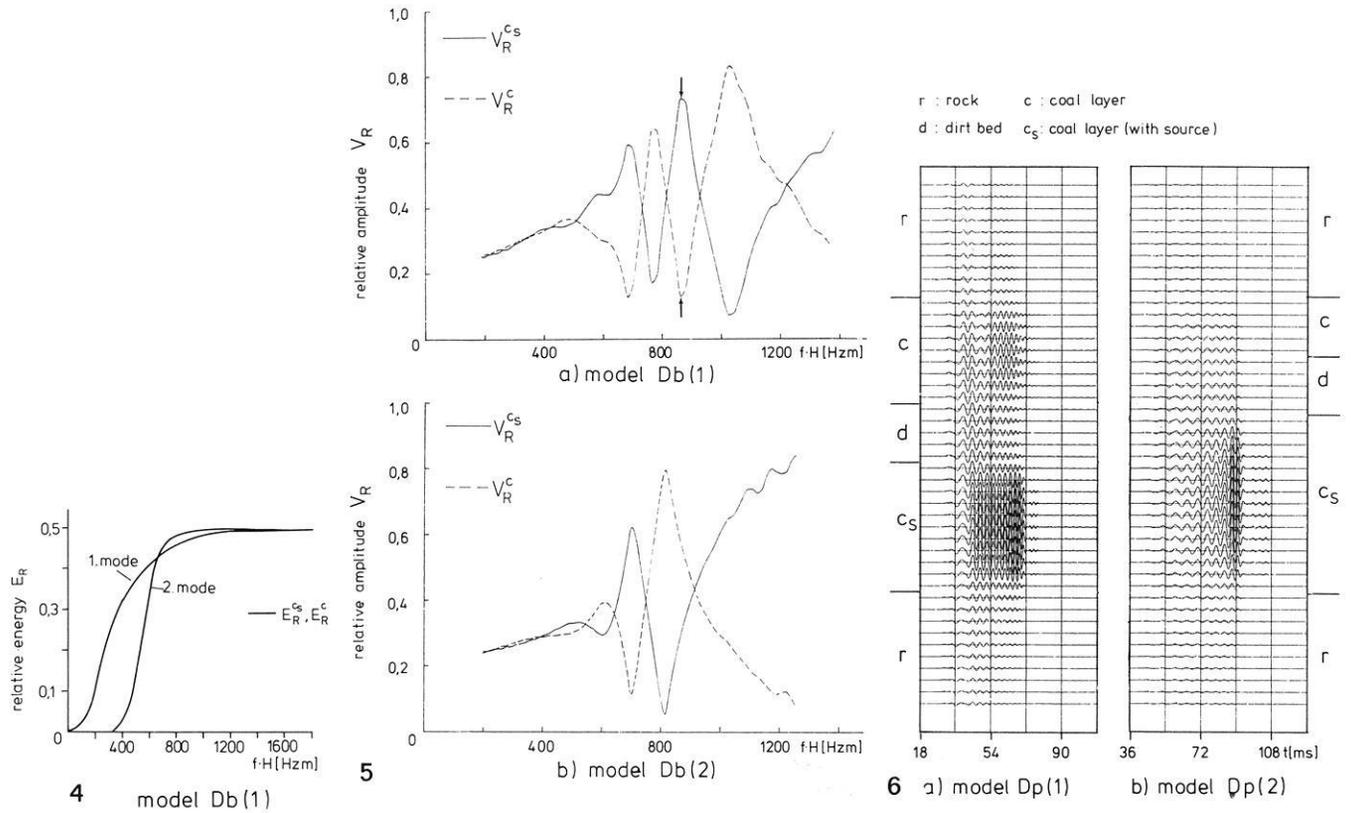


Fig. 4. Curves of the relative energy E_R for waves guided in the coal layers $c_s(E_R^{c_s})$ and $c(E_R^c)$ of the model Db(1). The curves are calculated for the first and second mode

Fig. 5a and b. Curves of the relative amplitude V_R for waves guided in the coal layers $c_s(V_R^{c_s})$ and $c(V_R^c)$ of the models Db

Fig. 6a and b. Seismogram sections calculated for vertical profiles in the models Dp. Distance of the profiles from the source position: $x/H=30.7$

amplitudes in the $f \cdot H$ -range from 600 to 1200 Hz·m in the case of model Db(1) (Fig. 5a) and from 600 to 900 Hz·m in the case of model Db(2) (Fig. 5b). The oscillations in this range can be explained by constructive and destructive interference of seam wave phases of the first and second mode. The effect of the interference depends on the phase difference between waves of the first and the second mode due to different phase velocities and the signs of the amplitudes in the amplitude-depth distributions. For example, in Fig. 5a the maximum (arrow) at 850 Hz·m in the curve describing the relative amplitude in the source-containing coal layer (solid line) occurs because the interfering first and second mode are in phase (Fig. 6 in Räder et al., 1985). The equivalent minimum (arrow) at 850 Hz·m in the curve describing the relative amplitude in the other coal layer (dashed line) occurs because the two modes are 180° out of phase (Fig. 6 in Räder et al., 1985).

Summary 1

It was found that besides the seam wave propagating in the wave guide 'total seam' (which is associated with the first mode), it is not, as usually assumed in practice, the Airy phase of the first mode which propagates in the wave guides 'coal layers', but the Airy phase of the second mode. Interference of waves belonging to the two modes causes pronounced minima and maxima in the curves of relative

amplitude. No significant effects which depend on the thickness of the dirt band are found.

Position of dirt band within the seam

The seismogram sections in Fig. 6 calculated for vertical profiles in the models Dp show that the seam wave propagation is significantly influenced by the position of the dirt band within the seam. In the case of Dp(1) (Fig. 6a), the two wave portions guided in the wave guides 'total seam' and 'coal layers' can be distinguished, but the amplitudes of the signals recorded in the coal layer without source are smaller than those recorded in the source-containing coal layer. In the case of Dp(2) (Fig. 6b), seam wave propagation is concentrated in the source-containing coal layer.

The dispersion curves and the results from the velocity analysis for the models Dp are depicted in Fig. 7. In the case of Dp(2) (Fig. 7b), only the first mode is excited with significant amplitudes. All phase velocities of the seam waves propagating in the source-containing coal layer, and most of the phase velocities of the seam waves propagating in the coal layer without source, coincide with the dispersion curve of the first mode. The amplitude of the Airy-phase signal is evanescent in the coal layer without source. Therefore, the high-frequency wave part which can be seen at the beginning of the seam wave signal (Fig. 6b at 54 ms)

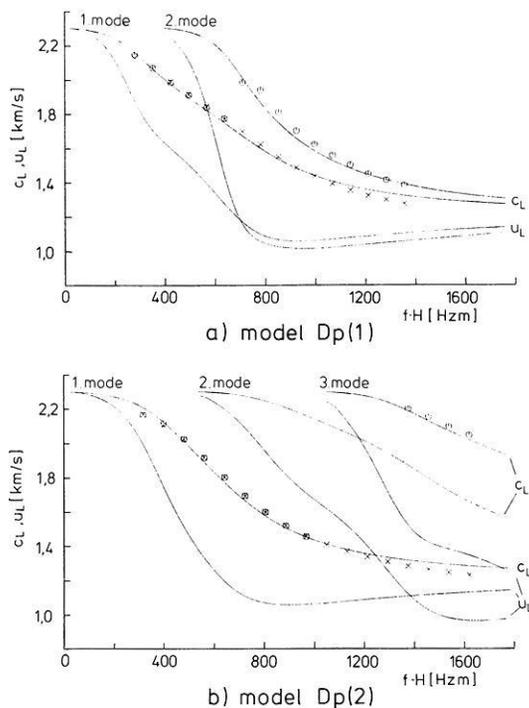


Fig. 7a and b. Phase (c_L)- and group (u_L)-velocity dispersion curves of the first and second mode and results from the phase-velocity analysis of the seam waves propagating in the source-containing coal layer (crosses) and in the coal layer without source (circles)

was identified by the velocity analysis as belonging to the third mode (Fig. 7b).

In the case of model Dp(1), the results obtained for the seam wave propagating in the coal layer without source again show that the first mode is excited in the $f \cdot H$ -range where the group-velocity dispersion curve has an indentation and that the Airy phase of the second mode is excited. In contrast to this finding, all phase velocities of the seam wave propagating in the source-containing coal layer coincide with the phase-velocity dispersion curve of the first mode.

The curves of relative energy in Fig. 8 explain the above results. The curve calculated for the first mode of the seam wave propagating in the source-containing coal layer (solid line in Fig. 8a) has the previously mentioned high-pass filter characteristic with values approaching the maximum value above 800 Hz·m. This means that nearly all of the wave energy of the first mode is guided in the source-containing coal layer. The relative energy curve belonging to the seam wave of the first mode propagating in the coal layer without source (dashed line in Fig. 8a) indicates that only in the $f \cdot H$ -range around 400 Hz·m, where the indentation in the group-velocity dispersion curve occurs, is a small energy portion of the first mode guided within this coal layer. The results for the second mode are just the opposite (Fig. 8b). Only a small portion of wave energy of the low-frequency part of the second mode is guided within the source-containing coal layer, while the second mode is completely trapped within the coal layer without source if this mode is excited in the $f \cdot H$ -range above 1000 Hz·m. As energy transport is confined to one coal layer, the effects of interference are not as pronounced as in the case of models Db. Only around 650 Hz·m do the curves of relative ampli-

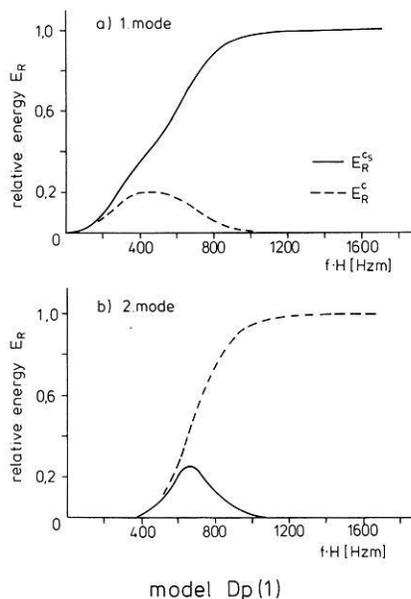


Fig. 8a and b. Curves of the relative energy E_R for waves guided in the coal layers c_s (E_R^{cs}) and c (E_R^c) of the model Dp(1)

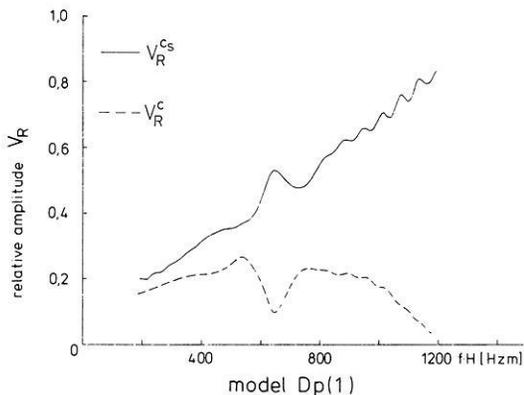


Fig. 9. Curves of the relative amplitude V_R for waves guided in the coal layers c_s (V_R^{cs}) and c (V_R^c) of the model Dp(1)

tude in Fig. 9 show weak oscillations resulting from overlapping of waves of the first and second mode.

Summary 2

The influence of higher modes decreases if the dirt band is not positioned in the centre of the seam. From a comparison of the group-velocity dispersion curves of all dirt band models investigated, it is concluded that for any given layer sequence the significance of higher modes can be deduced from the positions of the group-velocity minima on the $f \cdot H$ scale relative to each other and from the extent to which they are pronounced. Curves of relative energy are useful in this prognosis.

Seams interrupted by faults

Seam end

In Fig. 10, curves of the reflectivity calculated for models Ez are presented. The normalized maximum signal amplitudes of the reflected seam wave signals are listed in a table.

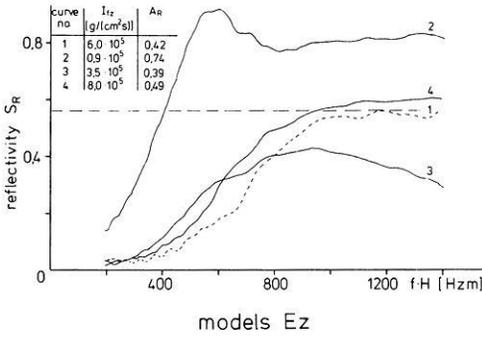


Fig. 10. Reflectivity S_R versus $f \cdot H$ calculated for the seam end models E_z containing fault zones. *Table:* Maximum signal amplitudes A_R of the reflected seam wave

These values correspond to the reflectivity values of the Airy phase of about $800 \text{ Hz} \cdot \text{m}$.

The dashed curve in Fig. 10, calculated for model $E_z(1)$ without a fault zone (Fig. 1), serves as a reference curve. From an approximation of this curve by an analytically calculated curve, Korn and Stöckl (1982) deduced that in the case of a simple vertical truncation of the seam (seam end) the reflectivity is mainly determined by two factors:

- The faults acts as a high-pass filter on the reflected seam wave. Accordingly, the reflectivity curve exhibits the characteristics of a high-pass filter curve.
- In the pass-band of the filter ($> 1000 \text{ Hz} \cdot \text{m}$), the reflectivity approaches the reflection coefficient for normal incidence of the interface coal-rock. This value is equal to 0.56 in our models (horizontal line).

Comparison of the reflectivity curve (2) with the dashed curve shows that a zone of fractured rock material (disaggregation) at the fault produces a significant increase in the amplitudes of the reflected seam wave signal. For high frequencies ($f \cdot H > 800 \text{ Hz} \cdot \text{m}$), the reflectivity is about 0.8. A zone of consolidated rock material also produces an increase in the amplitudes, but the effect is less prominent. Above $800 \text{ Hz} \cdot \text{m}$, the reflectivity curve (4) lies about 10% above the reference curve and approaches values of about 0.6. For an intermediate impedance, the reflectivity curve (3) lies below the reference curve for $f \cdot H > 800 \text{ Hz} \cdot \text{m}$. A maximum value of 0.43 occurs at $950 \text{ Hz} \cdot \text{m}$. Below $800 \text{ Hz} \cdot \text{m}$, all of the solid curves lie above the reference curve. Thus, there is a reduction of the high-pass filter effect of the fault if there is a fault zone. However, this effect is significant only for fractured zones.

To extend the results obtained by Korn and Stöckl (1982) to fault zones, we assume that at high frequencies the reflectivity is mainly determined by the reflection of the seam wave propagating in the coal layer. Therefore, the reflectivity values should approach the reflection coefficient for a plane SH wave reflected from the two interfaces of the fault zone, the coal-fault zone interface and the fault zone-rock interface. Furthermore, it seems reasonable to suppose that the reduction of the high-pass filter effect of the fault is caused by the reflection of the low-frequency seam wave propagating in the rock. For these waves, the reflection coefficients for an SH wave reflected from the interfaces rock-fault zone and fault zone-rock are relevant.

We have calculated reflection coefficients for a thin layer surrounded by two half-spaces with the layer matrix method (Fertig, 1982). The half-spaces are taken to have the

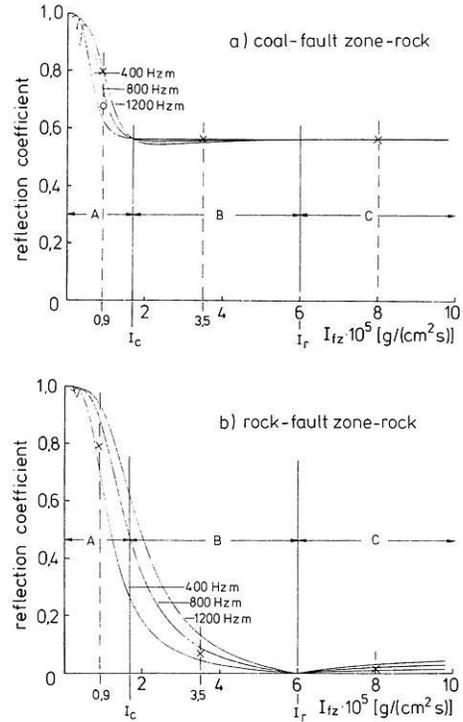


Fig. 11a and b. Reflection coefficient of the SH wave vertically incident on a thin layer (fault zone) versus the impedance of the layer I_z (I_c : impedance of the coal, I_r : impedance of the rock)

elastic parameters of the coal and the rock, respectively, and the thin layer is taken to have the thickness of the fault zone. Figure 11 shows the results for the sequences coal-fault zone-rock (Fig. 11a) and rock-fault zone-rock (Fig. 11b) as functions of the impedance of the fault zone. The reflection coefficient resulting from interference of the reflections from the two interfaces of the thin layer is complex-valued and depends on frequency. We calculated the amount of the reflection coefficient for three frequencies, 400, 800 and $1200 \text{ Hz} \cdot \text{m}$. In the further interpretation we omit the relative minima occurring for small impedance values. These minima are due to destructive interference of the double reflection.

Figure 11a shows that only in the range A, for impedances smaller than those of the coal, does the reflection coefficient of the SH wave reflected from the fault zone (thin layer) exceed the reflection coefficient of 0.56 for a reflection at the interface coal-rock. The reflection coefficient is nearly constant and approaches this value in the ranges B and C, for impedances between those of the rock and coal, and impedances larger than that of the rock, respectively. In addition, Fig. 11a reveals that only for the disaggregation is there a distinct dependence of the reflection coefficient on the frequency. Comparison of the reflection coefficients (crosses in Fig. 11a) with the reflectivity values above $1000 \text{ Hz} \cdot \text{m}$ (Fig. 10) shows that indeed the reflectivity in this $f \cdot H$ -range is mainly determined by the reflection coefficient in the disaggregation case and the consolidation case. For the intermediate impedance, the reflection coefficients can not explain the decrease of the reflectivity values above $1000 \text{ Hz} \cdot \text{m}$. Although Korn and Stöckl (1982) concluded that diffractions generated at the corners of the seam end are unimportant for the reflectivity, we

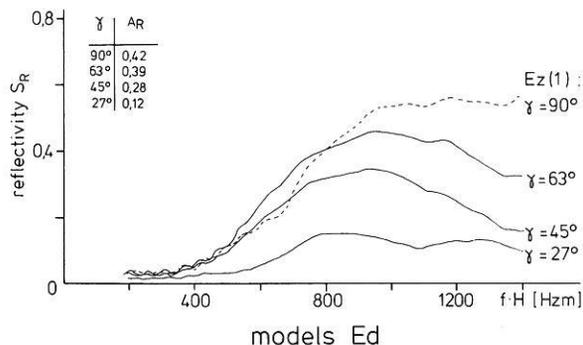


Fig. 12. Reflectivity S_R versus $f \cdot H$ calculated for the models Ed containing truncated seams (seam end) with dipping fault planes
Table: Maximum signal amplitudes A_R of the reflected seam wave

think that in the case of the fault zone, especially in the range B, diffraction phenomena are responsible for details, such as the one mentioned, in the reflectivity curves.

From amplitude calculations (Kerner, 1984), we know that the effects from the reflected seam wave portions propagating in the rock are strongest in the $f \cdot H$ -range near 600 Hz·m. In this range, the seam wave exhibits amplitudes of comparable size in both coal and rock. Figure 1 b shows that for $I_{fz} = 0.9 \times 10^5$ g/(cm²s), the reflection coefficient is about 0.8 at 600 Hz·m (cross). As the wave portion in the seam is also reflected with similar reflection coefficients ($\cong 0.7$, circle in Fig. 11 a), this explains the behaviour of the reflectivity curve from 300 to 800 Hz·m (Fig. 10) in the disaggregation case. At 600 Hz·m, a maximum value even occurs.

For $I_{fz} = 3.5 \times 10^5$ g/(cm²s) and 8.0×10^5 g/(cm²s), the reflection coefficients are 0.07 and 0.01, respectively (crosses). These small values cannot completely explain the actual reflectivity values which, for example, are nearly twice as large as the value in the reference curve at 600 Hz·m. In this low $f \cdot H$ -range, phase differences between the reflected wave propagating in the rock and the reflected wave propagating in the coal might be responsible for these fine details, besides diffractions. We found phase differences of about $\pi/2$ in the latter two cases. This value changes only little with frequency. For the disaggregation, the phase differences are smaller than $\pi/4$ in the $f \cdot H$ -range around 600 Hz·m.

Figure 12 shows the reflectivity curves calculated for the

models Ed with dipping fault plane. The reflectivity curve determined for the seam end with a vertical fault plane serves as a reference curve. It can be seen that decreasing the dip angle of the oblique fault plane produces a decrease of the maximum value in the reflectivity curve. The diminution of the reflectivity values above 950 Hz·m for $\gamma = 63^\circ$ and $\gamma = 45^\circ$ indicates a reduced high-pass filter effect.

An attempt to explain these reflectivity curves by means of the reflection coefficients, which are dependent on the angle of incidence, failed not only when we considered the reflection of the entire seam wave by analogy with the reflection of a plane SH wave – as done in the case of the seam end with vertical fault plane – but also when we considered the reflection of single phases of the seam waves. This is due to the more complex wave propagation processes at an oblique fault plane as compared to a vertical fault plane.

With the help of the seismogram sections plotted in Fig. 13, we will try to elucidate these processes. For the seam end model with vertical fault plane and for the seam end model with the 27°-dipping fault plane, seismograms were calculated for points lying on a rectangular profile surrounding the seam end. The seismograms are normalized by the maximum amplitude in the sections. The seismograms recorded in the seam are excluded from this normalization because the maximum signal amplitudes of the direct seam waves (D) are about ten times larger than the amplitudes used for the normalization.

In the seismogram sections, it can be seen that the low-frequency portion of the direct seam wave (D), which is incident on the seam end, propagates into the rock region adjacent to the seam end (T). Calculations of the transmissivity show that this transmitted wave is hardly influenced by the obliqueness of the fault plane. In particular, the deficit of reflected wave energy in the high-frequency range above 800 Hz·m, in the case of the oblique fault plane (Fig. 12), cannot be explained by an increased transmission into this region. However, the sections exhibit high-frequency wave phases transmitted into the rock (ellipses). In the case of the vertical fault plane (Fig. 13 a), these wave phases are interpreted as diffractions from the corners at the seam end. The waves exhibit higher amplitudes in the case of the oblique fault plane (Fig. 13 b). This is caused by an additional scattering of refracted waves in all directions, occurring when the high-frequency wave portions are multiply reflected in the region of the fault dip.

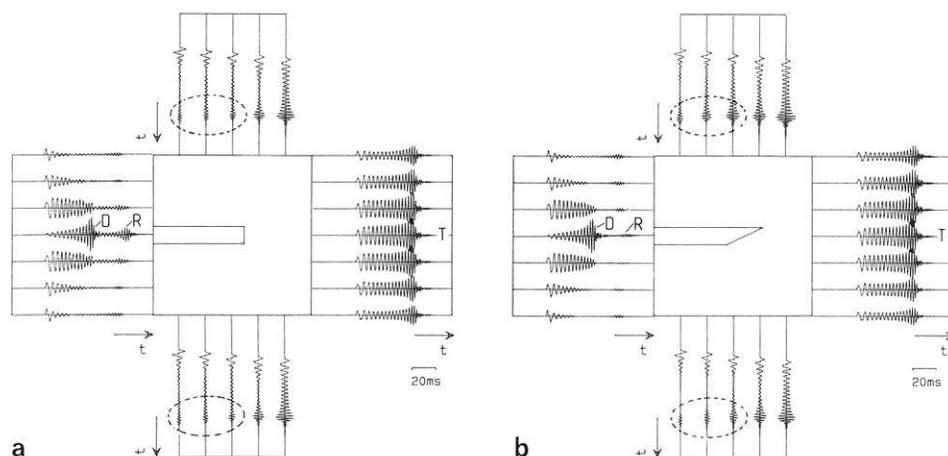


Fig. 13a and b. Seismogram sections calculated for a rectangular profile surrounding the seam end in the models **a** Ez(1) and **b** Ed(1) (D: direct seam wave, R: reflected seam wave, T: transmitted wave, ellipses contain transmitted waves scattered backwards)

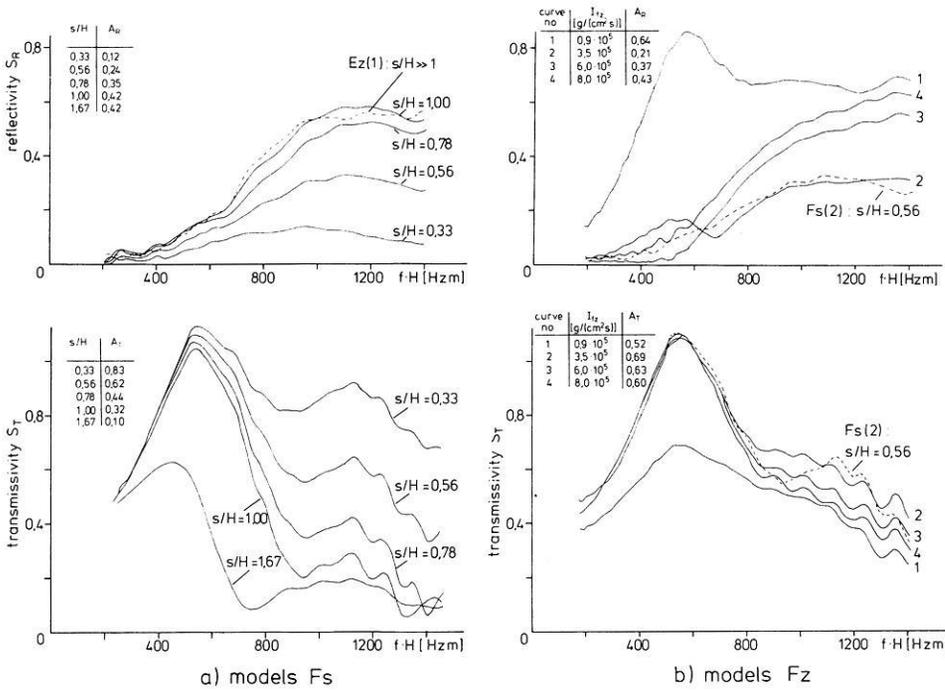


Fig. 14a and b. Reflectivity S_R and transmissivity S_T versus $f \cdot H$ calculated for **a** the models F_s containing faults with differing throws and for **b** the models F_z containing faults with fault zones *Tables*: Maximum signal amplitudes A_R and A_T of the reflected and transmitted seam waves

Summary 3

Only in the case of the disaggregation in the fault zone, is there a significant increase in the amplitude of the reflected seam wave signal compared with the amplitude in the case of the simple fault without a fault zone. The amplitude of the seam wave reflected by a disaggregation zone can be estimated by calculating reflection coefficients.

A dipping fault plane leads to a decrease in the amplitudes of the reflections. The high-frequency portion of the seam wave, especially, is scattered by the dipping fault plane.

Fault throw less/equal one seam thickness

The reflectivity curves for a seam with offset (Fig. 14a, top) exhibit a similar high-pass filter characteristic as does the reflectivity curve in the case of the truncated seam (dashed curve). The throw does not influence the amount of the reflectivity very much below $f \cdot H$ values of $650 \text{ Hz} \cdot \text{m}$. Above this value, the reflectivity becomes significantly lower if the throw decreases below one seam thickness. For example, if the throw is one-third of the seam thickness, the amplitude of the reflection is 12% of that of the incident seam wave (refer to the upper table in Fig. 12a). No significant differences among the curves occur if the throw exceeds one seam thickness.

For the transmitted seam wave, the fault acts as a low-pass filter. Therefore, the transmissivity curves (Fig. 14a, bottom) exhibit a maximum at about $500 \text{ Hz} \cdot \text{m}$. Below $800 \text{ Hz} \cdot \text{m}$, the throw only produces minor differences among the transmissivity curves if the throw is less than or equal to one seam thickness. For the throw exceeding one seam thickness ($s/H = 1.67$), the maximum value of the transmissivity is halved, indicating low amplitudes of the transmitted seam wave for low frequencies. Due to the high-pass filter effect of the layering, the transmissivity values

in the $f \cdot H$ -range from 800 to $1400 \text{ Hz} \cdot \text{m}$ determine the maximum amplitude of the transmitted seam wave. In this range, the curves are characterized by a relative maximum at about $1100 \text{ Hz} \cdot \text{m}$. The transmissivity is significantly dependent on the throw even for offsets less than one seam thickness (refer also to the amplitude values in the table).

Figure 14b shows the effects of the fault zone for a fault with a throw of about half a seam thickness. Inspection of the reflectivity curves (top) shows that all curves (1, 3, 4), except the one calculated for the intermediate impedance in the fault zone (2), lie significantly above the reference curve (dashed curve) calculated for the fault without any fault zone. Thus, both a zone with fractured rock material and a zone with consolidated rock material produce an increase in the amplitude of the reflected seam wave. The comparison of the respective signal amplitudes (table) confirms this result. The effect is explained by the fact that the reflection process at the fault is not restricted to the seam wave portions propagating at the level in the seam where the coal-rock interface occurs (Korn and Stöckl, 1982), as is the case at the simple fault with a throw less than one seam thickness, but is extended to the whole seam wave propagating in both coal and rock. In the case of the intermediate impedance in the fault zone, the reflectivity curve (2) closely follows the reference curve.

Inspection of the transmissivity curves (Fig. 14b, bottom) and a comparison of the amplitudes in the respective table show that there are no significant effects of the fault zone on the transmitted seam wave.

The reflectivity and transmissivity curves in Fig. 15 and the signal amplitudes in the tables show the effects from an oblique fault plane. The reflectivity curves calculated for the models F_d – which contain a down-dipping fault – (Fig. 15a, top) exhibit the same trends as already discussed for the respective seam end models (E_d). Although these reflectivity curves lie slightly (≈ 0.02) above the respective curves (Fig. 12) above $800 \text{ Hz} \cdot \text{m}$, the signal amplitudes of the reflected waves (table) are slightly smaller than

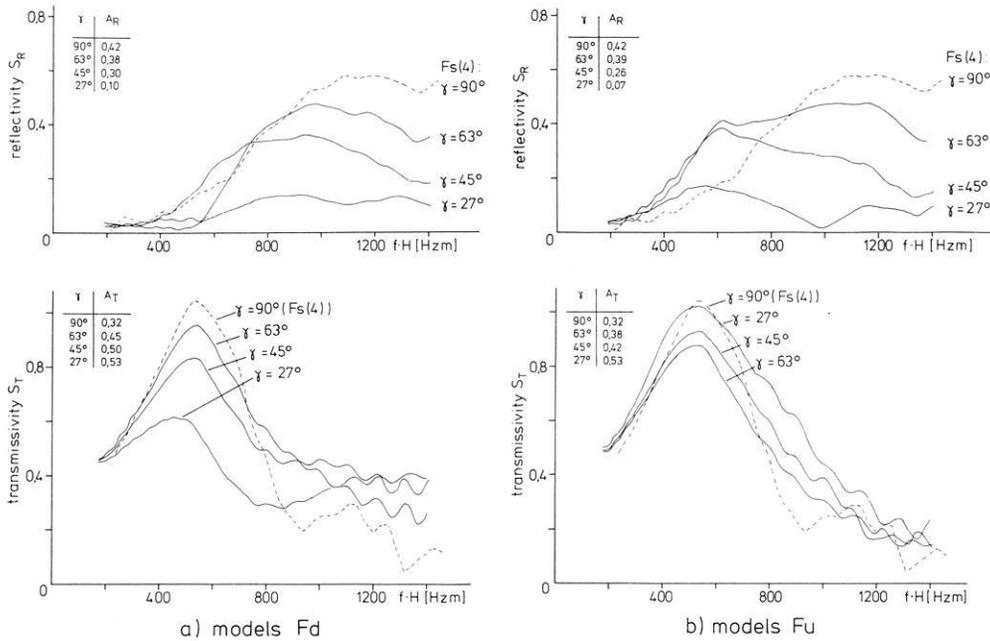


Fig. 15a and b. Reflectivity S_R and transmissivity S_T versus $f \cdot H$ calculated for **a** the models Fd containing faults with down-dipping fault planes and **b** for the models Fu containing faults with up-dipping fault planes
Tables: Maximum signal amplitudes A_R and A_T of the reflected and transmitted seam waves

those obtained for the models Ed. Comparison of the reflectivity curves calculated for the models Fu – which contain an up-dipping fault – (Fig. 15b, top) with the reflectivity curves for the models Fd shows that high reflectivity values occur in the range of low $f \cdot H$ values near 600 Hz·m. Above this value, the curves obtained for models Fd and Fu are similar but the reflectivity values are about 0.06 lower in the latter case. The signal amplitudes (table) are hardly affected by these differences.

The transmissivity curves for models Fu and Fd (Fig. 15a, b, bottom) indicate that the transmission of the high-frequency seam wave portion is favoured by the obliquity of the fault plane. The curves lie above the respective reference curve (dashed curve) for $f \cdot H$ values above 800 Hz·m in the case of models Fd and in the $f \cdot H$ -range from about 750 to 1050 Hz·m in the case of models Fu. The amplitude of the transmitted wave (table) increases if the dip angle decreases. The differences between the transmissivity for the up-dipping and the down-dipping faults become evident when inspecting the curves calculated for dip angle 27°. In the case of the models Fu, low-frequency seam wave portions and high-frequency portions are transmitted with amplitudes which are higher the flatter the dips. In contrast to this, in the case of the models Fd, low-frequency seam wave portions are transmitted with amplitudes which are lower the flatter the dips. The opposite is valid for high-frequency seam wave portions: the flatter the dips, the higher are the amplitudes. From averaging transmissivity values in the $f \cdot H$ -range around 1100 Hz·m, we obtained values of 0.2 for a 27°-up-dipping fault plane and 0.4 for a 27°-down-dipping fault plane. These values indicate that for a high-frequency seam wave signal the difference between the maximum amplitudes of the transmitted waves might be distinct enough to allow for an identification of the dip type at least for flat dip angles.

Some of our fault models are similar to those investigated by Asten et al. (1984). These authors compared finite-element calculations with finite-difference simulations. For models containing vertical faults with differing throws, they proved agreement between the amplitudes in the reflection

and transmission spectra and the normalized amplitudes evaluated by Korn and Stöckl (1982).

We also confirm the results obtained by Asten et al. (1984) (for comments on the conversion of the data, refer to Asten et al., 1984, Section 5). We obtained a value of 0.38 for the normalized amplitude of the seam wave reflected at the 63°-down-dipping fault plane [model Fd(3)], corresponding to an energy value of 14%. For the same fault type, Asten et al. (1984) gave values of about 13% for the reflected energy of the fundamental mode at the period (3.5 ms) where its Airy phase occurs.

Considering the effects of a fault zone with fractured rock, due to differences between our model Ez(2) and the model of Asten et al. (1984), only a rough comparison is possible. Both results indicate that the energy of the reflected seam wave is significantly higher than in the case of a simple fault and that the contribution of low-frequency seam wave phases to the reflection signal is increased.

From their investigations, Asten et al. (1984) deduced that the partition of energy between reflected and transmitted waves and between fundamental and higher modes characterizes the fault. This conclusion is based on a rather complete data set: Asten et al. (1984) calculated total energies of the reflected and transmitted seam waves scattered by the discontinuity in all directions. Furthermore, they considered energy values for the fundamental and higher modes in the $f \cdot H$ -range from 200 to 3,000 Hz·m. Thus, these energy curves enable the discrimination of fault parameters, but they demand that data acquisition is not restricted to profile points in the coal and that broad-band field data are available.

In contrast, we used model data similar to field data: reflectivity and transmissivity curves are deduced from seismograms recorded in the centre of the model seam. Effects from mode conversions are negligible, because the Airy phases of the second and higher modes lie outside the $f \cdot H$ -range in which seam waves are generated. In the field survey there exists a similar limit due to the frequency-band limitation of the equipment.

Although the reflectivity and transmissivity curves differ

significantly they are not particularly appropriate to make a clear-cut conclusion on the fault parameters. For example, no significant differences can be determined between the reflectivity and transmissivity curves for a fault with an offset of half a seam thickness and a fault with the same offset having a fault zone with an impedance of 3.5×10^5 g/(cm²s). Or: the reflectivity curve for a vertical fault with a throw $s/H=0.33$ is similar to the reflectivity curve for a 27°-down-dipping fault. In this latter case, the different curves for the transmissivity would allow the discrimination between the two fault types, if these data are available. On the other hand, if we have information on some of the fault parameters, e.g. the dip and the existence of a fault zone, we can estimate the third parameter, e.g. the throw, from the amplitudes.

Summary 4

Only if the throw is less than one seam thickness does it have an effect on the amplitude of the reflected seam wave: the amplitude decreases with the offset at the fault. The effect is strongest for high-frequency phases. The low-frequency phases of the transmitted seam wave are influenced by the fault only for throws larger than one seam thickness. The high-frequency phases depend on the throw even for offsets less than seam thickness.

For throws less than seam thickness, a fault zone with disaggregated or consolidated rock material leads to an increase in the amplitude of the reflected seam wave due to reflections from the interfaces coal-fault zone-coal.

Transmission of high-frequency seam waves is favoured by a down-dipping fault: there is an increase in the amplitudes of the transmitted seam wave with decreasing dip angle.

Discrimination of fault parameters on the basis of reflectivity and transmissivity curves is not reliable.

Conclusions

It was found that in seams with a dirt band, the Airy phase of the second mode, as well as Love seam wave phases which are part of the first mode, are excited. Interference of the two modes causes irregularities in the shapes and the amplitudes of the seam wave signals. This can lead to a deterioration of the results of data processing, especially for envelope-stacking and recompression. In extreme cases, two separate wave groups occurring in the reflection seismograms may lead to the false conclusion that two reflectors exist. With knowledge of the geological structure, one is able to calculate dispersion curves and curves of relative energy in order to decide whether higher modes are excited.

Our numerical results correspond to model data sets similar to those obtained from the currently employed field techniques. If the fault parameters can be estimated (e.g. from an opening in the roadway), the reflectivity and transmissivity curves presented and the normalized amplitudes can be exploited for a prognosis on the detectability of the fault.

However, a discrimination of fault parameters on the basis of reflectivity and transmissivity curves alone is not reliable, due to the ambiguity of the curves. Therefore, knowledge of some of the fault characteristics is also necessary in order to extract fault parameters from the field data.

If such information is not available, the field data set must be supplemented, e.g. by measurements along vertical seismic profiles through the sequence rock-coal-rock.

Acknowledgements. This investigation has been carried out in cooperation with Bergbauforschung GmbH Essen, Prakla and Seismos GmbH Hannover, Ruhr-Kohle AG – BAG-Lippe Herne and Westfälische Berggewerkschaftskasse Bochum. Financial support was given by the BMFT and BMWi of the Federal Republic of Germany under the contract 'Flözwellenseismische Vorfelderkundung mit Hilfe digitaler Meßwerterfassung' and 'Einführung der Flözwellenseismik als Hilfsmittel zur Produktivitätssteigerung in den Betrieben des Steinkohlenbergbaus – Innovation III'. Dr. F.K. Brentrup, Dr. U. Klinge and Dr. H. Rüter gave valuable suggestions.

References

- Alford, R.M., Kelly, K.R., Boore, D.M.: Accuracy of finite difference modelling of the acoustic wave equation. *Geophysics* **39**:834–842, 1974
- Arnetzl, H.H.: Seismische Messungen untertage. Tagungsberichte 'Mensch und Maschine im Bergbau' der Gesellschaft Deutscher Metallhütten- und Bergleute, 133–141, 1971
- Arnetzl, H.H., Klinge, U.: Erfahrungen mit der Flözwellenseismik bei der Vorfelderkundung. *Glückauf* **118**, 658–664, 1982
- Asten, M.W., Drake, L.A., Edwards, S.: In-seam seismic Love wave scattering modeled by the finite element method. *Geophys. Prosp.* **32**, 649–661, 1984
- Boore, D.M.: Love waves in nonuniform waveguides: finite difference calculations. *J. Geophys. Res.* **75**, 1512–1527, 1970
- Boore, D.M.: Finite difference methods for seismic wave propagation in heterogeneous materials. In: *Methods in computational physics*, Vol. 11, B. Alder, S. Fernbach, M. Rotenberg, eds. New York: Academic Press 1972
- Brentrup, F.K.: Flözwellenseismische Vorfelderkundung. *Glückauf* **115**, 820–823, 1979
- Buchanan, D.J.: In-seam seismology: A method for detecting faults in coal seams. In: *Developments in geophysical exploration methods*, Vol. 5, A.A. Fitch, ed. London and New York: Applied Science Publishers Ltd., 1983
- Dresen, L., Freystätter, S.: Rayleigh channel waves for the in-seam seismic detection of discontinuities. *J. Geophys.* **42**, 111–129, 1976
- Dresen, L., Kerner, C., Kühbach, B.: The influence of an asymmetry in the sequence 'rock-coal-rock' on the propagation of Rayleigh seam-waves. *Geophys. Prosp.* **33**, 519–539, 1985
- Dziewonski, A.M., Hales A.L.: Numerical analysis of dispersed seismic waves. In: *Methods in computational physics*, Vol. 11, B. Alder, S. Fernbach, M. Rotenberg, eds. New York: Academic Press, 1972
- Edwards, S., Astin, M.W., Drake, L.A.: P-SV wave scattering by coal seam inhomogeneities. *Geophysics* **50**, 214–223, 1985
- Fertig, J.: Ebene Wellen in geschichteten Medien. Beitrag zum Seminar: "Numerische Methoden für seismische Wellenausbreitung und elektromagnetische Induktion", Neustadt, 1982
- Kerner, C.: Untersuchungen an zweidimensionalen analogen und numerischen Modellen zur Transmission und Reflexion von Love- und Rayleigh-Flözwellen. Ph.D. thesis, Berichte des Institutes für Geophysik der Ruhruniversität Bochum, Series A, No. 14, Bochum, 1984
- Kerner, C.: Simulation der Ausbreitung von seismischen Wellen in Wellenkanälen mit Hilfe der Methode der finiten Differenzen. *Conferences on Cyber 205*, Proceedings, H. Ehlich, K. Schlosser, eds., Bochum, 1985
- Korn, M., Stöckl, H.: Reflection and transmission of Love channel waves at coal seam discontinuities computed with a finite difference method. *J. Geophys.* **50**, 171–176, 1982
- Krey, Th.: Channel waves as a tool of applied geophysics in coal mining. *Geophysics* **28**, 701–714, 1963

- Küpper, F.J.: Theoretische Untersuchungen über die Mehrfach-aufstellung von Geophonen. *Geophys. Prosp.* **6**, 194–256, 1958
- Millahn, O.: Flözwellenseismik. *Prakla-Seismos Report 2 + 3*, Hannover, F.R. Germany, 19–30, 1980
- Räder, D., Schott, W., Dresen, L., Rüter, H.: Calculation of dispersion curves and amplitude-depth distributions of Love channel waves in horizontally layered media. *Geophys. Prosp.*, 1985 (in press)
- Reynolds, A.C.: Boundary conditions for the numerical solution of wave propagation problems. *Geophysics* **43**, 1099–1110, 1978
- Rüter, H., Schepers R.: In-seam seismic methods for the detection of discontinuities applied to West Germany coal deposits. In: *Coal Exploration 2*, Argall, G.O., ed.: pp 267–293, Proc. of 2nd Int. Coal Exploration Symposium; San Francisco: Miller-Freeman, 1979

Received November 6, 1984; revised version April 12, 1985

Accepted April 25, 1985

Seismic modelling by methods of the theory of edge waves

K.D. Klem-Musatov and A.M. Aizenberg

Institute of Geology and Geophysics of the Siberian Branch of the Academy of Sciences of the USSR,
Universitetskii prospekt 3, Novosibirsk, 90, USSR

Abstract. This paper deals with the computation of wavefields in 3-D inhomogeneous media containing structural elements such as pinch-outs, vertical and oblique contacts, faults, etc. The approach is based on the theory of edge waves. The total wavefield is considered as the superposition of two parts. The first part is described by the ray method. It has discontinuities because of its shadow boundaries. The second part is a superposition of two types of diffracted waves, caused by the edges and vertices of interfaces. This part smooths the above-mentioned discontinuities so that the total wavefield is continuous. Of special importance is the mathematical form of the amplitudes of diffracted waves, described with unified functions of eikonals. In fact, it allows all additional computations to be considered by finding the eikonals of diffracted waves. A modification of the ray method including diffraction by edges and vertices is described. A generalization of the concept of edge waves for caustic situations is given – the method of superposition of edge/tip waves. The result of such a generalization no longer supplements the geometrical seismic description, but completely replaces it by a new description valid for a broader class of wave phenomena (reflection/refraction, diffraction on edges and vertices, formation of caustics, etc.).

Key words: Diffraction on edges and vertices – Amplitudes of diffracted waves – Superposition of edge/tip waves

Introduction

The present article is an extended version of our report at the Workshop on Seismic Wave Propagation in Laterally Varying Media at Liblice, Czechoslovakia, 1983. Its main points were given in Klem-Musatov and Aizenberg (1984). Here we want to illustrate the main ideas of the theory of edge waves with the simplest examples and show how it is possible to use methods of this theory for seismic modelling.

Of special importance for seismology is the ray method which allows wavefields to be computed efficiently in 3-D inhomogeneous media far from the

source (Babich and Alekseyev, 1958; Karal and Keller, 1959; Červený et al., 1977). However, the method gives only the components of the wavefields connected with the energy flux along the ray tubes but not diffusion through their side walls. If the main part of the wavefield is formed by diffusion, it cannot be described by the ray method. The desire to adapt this method to such situations has resulted in various modifications (Babich and Buldyrev, 1972; Popov, 1981; Červený, 1983; Kennett, 1984).

In the present paper a modification of the ray method for 3-D inhomogeneous block media is considered. The structural elements of interfaces in this type of media have sharp edges, the so-called diffracting edges (for example, the lines of pinch-outs, vertical and oblique contacts of interfaces, faults and so on). The ray method does not give a continuous description of the wavefields in this type of media because of shadow boundaries. The main idea of the present modification is to smooth the discontinuities by diffracted waves, scattered by the edges of interfaces, in such a way that the total wavefield is continuous. From a physical viewpoint, it is the same as adding the diffusion that is not considered by the ray method (Fock, 1965). This principle is well-known in the classical theory of diffraction (Born and Wolf, 1968) and in its modern modifications (Claerbout, 1976; Trorey, 1977; Hilterman, 1982; Fertig and Müller, 1979). However, there were no general formulae to use the above-mentioned idea for improving the ray method.

The very core of the present approach is connected with the so-called boundary layer approximation. It allows us to correct the results of the ray method only within the neighbourhood of the shadow boundaries. It is just this kind of approximation that makes the final formulae general and simple. The simplest way of getting these formulae is shown in this paper. It is based on assumptions concerning the analytical properties of the wavefields, but not the dynamic equations in any case. If a wave velocity is constant, the same formulae can be derived by the parabolic equation method or by an asymptotic analysis of the Kirchhoff integral. The formulae for the edge waves can be derived from a solution of the more general diffraction problem for wedge-shaped structures as well. For details on this subject, see Klem-Musatov (1980, 1981a, b) and Aizenberg (1982).

Note that in the following monochromatic wavefields of angular frequency ω will be considered. The time factor $\exp(-i\omega t)$ with $i^2 = -1$, where t is time, is omitted for convenience. Theoretical seismograms can also be computed in the time domain by the application of the Fourier transform.

Analytic results will be illustrated with theoretical seismograms. Two types of media are chosen for seismic modelling: a two-layered model and a three-layered one with a pinch-out. The top medium always has the following elastic parameters for both types of models: $\rho_1 = 2 \text{ g/cm}^3$, $v_{P1} = 2 \text{ km/s}$, $v_{S1} = 1.25 \text{ km/s}$. The parameters of the bottom are always $\rho_2 = 2.4 \text{ g/cm}^3$, $v_{P2} = 2.5 \text{ km/s}$, $v_{S2} = 1.5 \text{ km/s}$. Parameters of the pinched layer are $\rho_3 = 1.8 \text{ g/cm}^3$, $v_{P3} = 1.75 \text{ km/s}$, $v_{S3} = 1.05 \text{ km/s}$. The observation system is located in the top medium. It contains either a profile of observation points (they are marked by circles in the figures) and a single source of oscillations (it is marked by an asterisk), or a profile of matched sources with observation points (they are marked by crosses). The source of oscillations excites a P wave with a spherical directivity pattern. The shape of the radiated pulse is $f(t) = t \exp(-\beta t) \sin(2\pi t/T)$, where t is time, $0 \leq t \leq 4T$, $\beta = 50 \text{ Hz}$, $T = 0.03 \text{ s}$. On changing to the nonstationary case the Hilbert transformation for narrow-band signals was used. All components of a displacement vector of seismic waves can be computed. In numerical examples (except Fig. 23) we show only the Z -component of PP waves. The X -component is oriented along a profile, the Z -component is oriented upwards. All metrical values are given in kilometres, as time is given in seconds.

Ray method

First of all, let us recollect the basic principles of the ray method. The model of the medium is considered as a combination of domains and interfaces. The functions, describing physical properties within the domains, are continuous and slowly changing. A surface formed by points of discontinuity of any of these functions is called an *interface*. A point of the interface is considered as *regular* if the surface is continuous together with its first and second tangential derivatives. A part of the interface is considered as *regular* if its points are all regular. The ray method allows us to describe only those components of the wavefield that are connected with reflections/transmissions at the regular parts of interfaces. The description has the form of superposition of the single waves

$$f = \sum_m f_m. \quad (1)$$

Let us give the main definitions related to a single wave f_m .

Kinematics

A ray is a space curve, the tangential unit vector \mathbf{e}_m of which complies with the differential equation:

$$\frac{d}{ds} (\mathbf{e}_m / v_m) = \text{grad}(1/v_m) \quad (2)$$

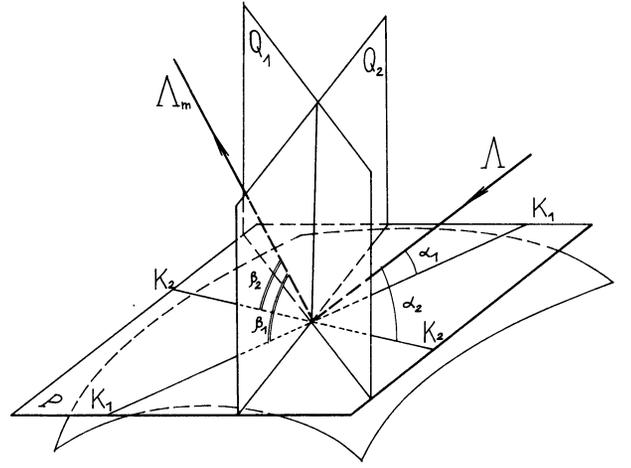


Fig. 1. Interpretation of Snell's law. Explanation in text

where ds is the differential of the arc length and v_m is the wave velocity. This equation determines the ray uniquely if its initial direction is given, and if a connection between the directions of incident and reflected/transmitted rays at the points of the interfaces is also given. The latter is expressed by Snell's law. In this law the geometry of the interface is usually characterized by the position of the normal to the interface. However, in this paper it is more convenient to achieve this by means of a tangential plane to the interface.

Let K_1 and K_2 be tangents to two arbitrary intersecting curves of the interface at the point of incidence (Fig. 1). As a result we get the position of the tangential plane P at any regular point of interface. Let α_1 and α_2 be acute angles between the incident ray A and the lines K_1 and K_2 , respectively. Let β_1 and β_2 be acute angles between the reflected/transmitted ray A_m and the same lines K_1 and K_2 . Let Q_1 and Q_2 be the planes that are normal to the lines K_1 and K_2 at the point of incidence. Then Snell's law can be expressed in the following way (Klem-Musatov, 1980):

1) The incident and the secondary ray lie on different sides of the planes Q_1 and Q_2 ,

2) The directions of the above rays comply with the conditions:

$$\cos \beta_1 / v_m = \cos \alpha_1 / v, \quad \cos \beta_2 / v_m = \cos \alpha_2 / v \quad (3)$$

where v and v_m are the velocities of the incident and reflected/transmitted wave, respectively.

Dynamics

If a set of rays \mathbf{e}_m is a two-parameter set of space curves, it is called a *congruence* (Born and Wolf, 1968). A single wave

$$f_m = \Phi_m \exp(i\omega \tau_m) \quad (4)$$

is connected with a congruence of the rays \mathbf{e}_m . Its eikonal τ_m complies with the differential equation:

$$\text{grad } \tau_m = \mathbf{e}_m / v_m. \quad (5)$$

Equation (4) itself may represent a scalar wave (optics, acoustics) or a vector wave (elastodynamics, electrodynamics). In the first case, the ray amplitude Φ_m is a scalar one. In the second case,

$$\Phi_m = \mathbf{p}_m \varphi_m \quad (6)$$

where \mathbf{p}_m is a unit vector of polarization, and φ_m is a scalar. In an isotropic medium the vector \mathbf{p}_m coincides with the vector \mathbf{e}_m (a longitudinal wave) or is perpendicular to \mathbf{e}_m (a transverse wave).

The scalar amplitude Φ_m (or φ_m) complies with the so-called transport equation

$$2 \text{grad } \tau_m \cdot \text{grad } \Phi_m + B_m \Phi_m = 0 \quad (7)$$

where the coefficient B_m depends on the kind of original accurate equations of optics, acoustics (or elastodynamics, electrodynamics). The solution of Eq. (7) is well known:

$$\Phi_m = \kappa_m L_m^{-1/2}, \quad L_m = \exp \left(\int_0^{\tau_m} v_m^2 B_m d\tau_m \right) \quad (8)$$

where integration must be performed along the ray. The choice of the constant κ_m must comply with the boundary conditions. In fact, κ_m is the product of reflection/transmission coefficients of plane waves. Only the first term of the ray series is shown. As will be seen later, the subsequent approach does not deal with the explicit formulae for the ray amplitude Φ_m .

Edge waves

We extend the theoretical basis using the ideas of the theory of diffraction. Let a certain line be formed by points of discontinuity of an interface or any of its first or second tangential derivatives. It is a common linear element of the regular parts of a single interface or of several interfaces. This type of line is called *an edge*. A point of the edge is considered *regular*, if the corresponding line is continuous together with its first tangential derivative. The edge is considered *smooth* if its points are all regular.

Every single wavefield f_m exists within a connected domain of continuity. This domain is called *the primary illuminated zone*. If the interfaces have edges, there may be a domain in which the wave f_m does not exist (we define it as $f_m \equiv 0$). This type of domain is called *the primary shadow zone* of the wave. The singlyconnected surface dividing these zones is called *the primary shadow boundary*. Let mn be the double number of each primary shadow boundary of the wave f_m . Let Ω_{mn}^+ be a symbol of the primary shadow zone, formed by the mn -th shadow boundary. Let Ω_{mn}^- be the symbol of the primary illuminated zone. The non-caustic shadow boundaries formed by the edges are considered.

Figure 2 shows the simplest example of the above definitions for a wave reflected from a half-plane. Let us mark this wave by $m=1$, and its shadow boundary by $m=1, n=1$. Figure 3a shows the reflected wave. The primary illuminated zone Ω_{11}^- is $x \leq 1.15$ km, the primary shadow zone Ω_{11}^+ is $x > 1.15$ km.

We can see that shortcomings of Eq. (1) appear as dis-

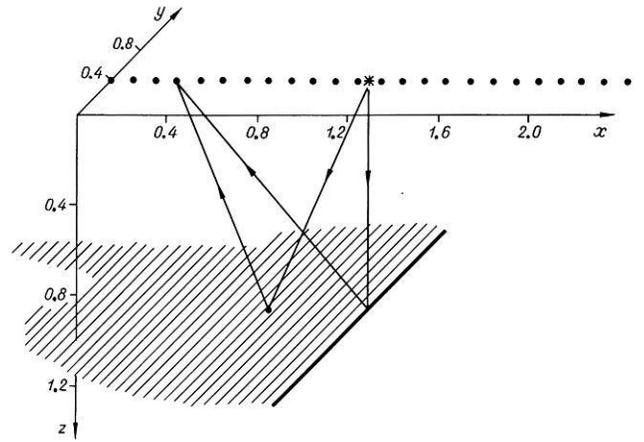


Fig. 2. Model of "half-plane". Reflecting interface coincides with shaded region

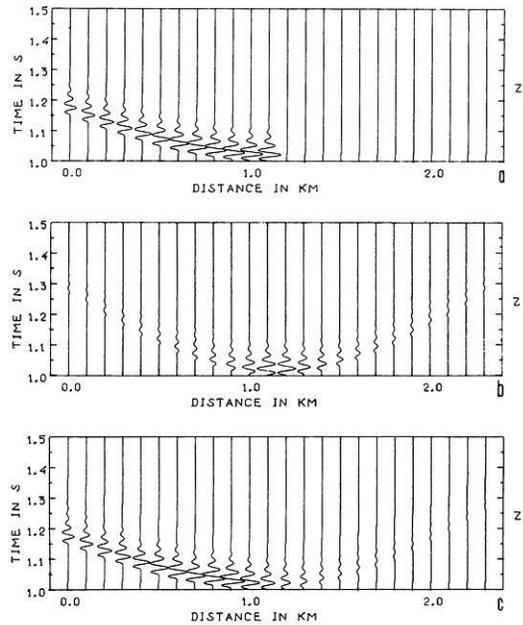


Fig. 3a-c. Theoretical seismograms for model of "half-plane": a reflected wave, b edge wave (twice enlarged), c total field

continuities of the wavefields f_m at the primary shadow boundaries. Let us see how this can be corrected.

Kinematics

We use a formal method to find the directions of the rays, generated at the points of an edge. Let the ray impinge on any regular point of the edge. The direction of a secondary ray must comply with Snell's law, Eq. (3). It is necessary to fix the positions of the pair of lines K_1 and K_2 , i.e. to set the position of the plane P . One of the two lines (for example, K_1) must be the tangent to the edge because it is a common linear element of the interfaces (Fig. 4). However, there are no limitations in choosing the direction of the second line K_2 . That is why any plane, containing the tangent to the edge, may be considered as plane P . Let incident A and secondary A_{mn} rays make the acute angles α and β ,

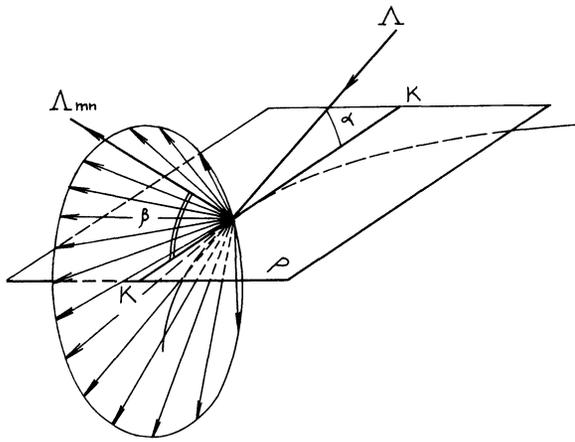


Fig. 4. A cone of diffracted rays. Explanation in text

respectively, with the tangent to the edge. Then Snell's law appears in the following form:

$$\cos \beta / v_m = \cos \alpha / v. \quad (9)$$

Note, the one-parameter set of secondary rays complies with this condition.

The above-mentioned fact is known as *the law of edge diffraction* (Keller, 1962). It reads as follows. Let an incident ray make an acute angle α with the tangent to an edge. The set of secondary rays forms a cone with its vertex at the point of incidence. Its apex angle is 2β , where β and α are connected under Eq. (9). The incident ray and the above-mentioned cone lie on opposite sides of the plane normal to the edge at the point of incidence. Obviously, this law holds true within a small neighbourhood of the point of incidence, in which it is possible to neglect the curvature of the rays.

Take \mathbf{e}_{mn} to be the unit vector of the tangent to the ray. Let this ray comply with Eq. (9) for that edge, which gives the mn -th primary shadow boundary. Then the differential equation:

$$\frac{d}{ds} (\mathbf{e}_{mn} / v_m) = \text{grad}(1/v_m) \quad (10)$$

determines the congruence of *the edge diffracted rays*.

An edge diffraction coefficient

Let the wave

$$f_{mn} = \Phi_{mn} \exp(i\omega\tau_{mn}), \quad \text{grad } \tau_{mn} = \mathbf{e}_{mn} / v_m \quad (11)$$

be connected with the mn -th primary shadow boundary. The latter may be given implicitly by the equation $\tau_{mn} = \tau_m$. For example, in Fig. 2 such a boundary is given by the relation $x = 1.15$ km. The wave, Eq. (11), is called *an edge-diffracted wave*. Now we have come to the description of diffracted waves.

Let τ_{mn} , η , ζ be the ray coordinates of the wave f_{mn} . Here η and ζ give a congruence of the diffracted rays, i.e. every pair of fixed values $\eta = \text{constant}$ and $\zeta = \text{constant}$ gives a single ray. This pair of coordinates may be chosen in many different ways. Let the coordinate surface $\eta = 0$ coincide with the mn -th primary shadow

boundary $\tau_{mn} = \tau_m$, so that the primary shadow zone of the wave f_m coincides with the domain $\eta > 0$. The coordinate surfaces $\zeta = \text{constant}$ may be taken arbitrarily.

In the first place, let us take the case when the amplitude Φ_m of the wave, Eq. (4), is a scalar one. Generalization of the present approach for polarized waves will be given in a following section. However, it is more convenient to use numerical examples for polarized waves before the mentioned generalization.

Let the wave, Eq. (4), be a function of the above-mentioned ray coordinates

$$f_m = \Phi_m(\tau_{mn}, \eta, \zeta) \exp[i\omega\tau_m(\tau_{mn}, \eta, \zeta)]. \quad (12)$$

Then this wavefield in the neighbourhood of its shadow boundary may be represented by the discontinuous function:

$$\begin{aligned} f_m &= f_m(\tau_{mn}, \eta, \zeta) \quad \text{when } \eta < 0, \\ f_m &= 0 \quad \text{when } \eta > 0 \end{aligned} \quad (13)$$

which displays explicitly the shortcomings of the ray method.

Suppose, Eq. (13) represents an analytic function of the variable η and allows us to make an analytic continuation into the complex plane of η for any permissible values τ_{mn} and ζ . Let us find f_{mn} among the piecewise-analytical functions decreasing at infinity ($f_{mn} \rightarrow 0$, when $|\eta| \rightarrow \infty$). Then we may construct the following integral of Cauchy's type

$$f_{mn} = \frac{1}{2\pi i} \int_L f_m(\tau_{mn}, \eta + \alpha, \zeta) \frac{d\alpha}{\alpha} \quad (14)$$

where L is some smooth infinite contour of integration. The specific type of this contour will be given in the following.

The integral in Eq. (14) has the following properties. It is zero when $|\eta| \rightarrow \infty$. It has a discontinuity when $\eta = 0$. However, the superposition of Eqs. (13) and (14) is a continuous and analytic function of η within the neighbourhood of the surface $\eta = 0$. If the function (13) is a solution of some linear differential equation (for example, the wave equation) within the domain $\eta < 0$, then the superposition of Eqs. (13) and (14) complies with the same equation for $\eta < 0$ as well as for $\eta > 0$.

When $\omega \rightarrow \infty$, the asymptotic value of integral (14) can be found by *the method of canonical integrals* (Felsen and Marcuvitz, 1973). Since this value is formed by contributions within a small neighbourhood of the saddle point $\alpha = -\eta$, let us take the standard approximations at this point

$$\begin{aligned} \Phi_m &\approx \Phi_m(\tau_{mn}, 0, \zeta), \\ \tau_m &\approx \tau_m(\tau_{mn}, 0, \zeta) + \eta^2 / 2 \left(\frac{\partial^2 \tau_m}{\partial \eta^2} \right)_{\eta=0} \end{aligned} \quad (15)$$

and use the following relations:

$$\tau_m(\tau_{mn}, 0, \zeta) = \tau_{mn}, \quad \left(\frac{\partial^2 \tau_m}{\partial \eta^2} \right)_{\eta=0} \approx \frac{2}{\eta^2} (\tau_m - \tau_{mn}). \quad (16)$$

Let the contour L within a neighbourhood of the saddle point coincide with *the steepest descent path*

$$\text{Im}[i\omega\tau_m(\tau_{mn}, \eta + \alpha, \zeta)] = \text{Im}[i\omega\tau_m(\tau_{mn}, 0, \zeta)],$$

$$\text{Re}[i\omega\tau_m(\tau_{mn}, \eta + \alpha, \zeta)] < 0.$$

Using τ_m from Eq. (15) gives the following equations for this part of the contour

$$\text{Re}\alpha + \eta = -\text{Im}\alpha \quad \text{for } \tau_{mn} > \tau_m,$$

$$\text{Re}\alpha + \eta = \text{Im}\alpha \quad \text{for } \tau_{mn} < \tau_m.$$

The integral exists, if the contour approaches the points $\text{Im}\alpha = \pm\infty$ within domains $\text{Re}[i\omega\tau_m(\tau_{mn}, \eta + \alpha, \zeta)] < 0$. Using τ_m from Eq. (15) allows us to obtain this condition for $\text{Im}\alpha \rightarrow \infty$ in the following form

$$-\pi < \text{Re}\alpha + \eta < 0 \quad \text{when } \tau_{mn} > \tau_m,$$

$$0 < \text{Re}\alpha + \eta < \pi \quad \text{when } \tau_{mn} < \tau_m,$$

and for $\text{Im}\alpha \rightarrow -\infty$ in the form

$$0 < \text{Re}\alpha + \eta < \pi \quad \text{when } \tau_{mn} > \tau_m,$$

$$-\pi < \text{Re}\alpha + \eta < 0 \quad \text{when } \tau_{mn} < \tau_m.$$

The contours are shown in Fig. 5a for $\tau_{mn} > \tau_m$ and in Fig. 5b for $\tau_{mn} < \tau_m$ (in these figures $q=1$).

Then integral (14) may be written as

$$f_{mn} = s_{mn} \Phi_m W(w_{mn}) \exp(i\omega\tau_{mn}),$$

$$w_{mn} = \sqrt{2\omega(\tau_{mn} - \tau_m)/\pi}, \quad (17)$$

$$s_{mn} = +1 \quad \text{within } \Omega_{mn}^+, \quad s_{mn} = -1 \quad \text{within } \Omega_{mn}^-,$$

$$W(w) = \exp(-i\pi w^2/2) / (2\sqrt{\pi}) \int_{-i\pi w^2/2}^{\infty} t^{-1/2} \exp(-t) dt \quad (18)$$

where W may be regarded as an *edge diffraction coefficient*. If $\tau_{mn} < \tau_m$, we have $w_{mn} = ix$, $x = \sqrt{2\omega(\tau_m - \tau_{mn})/\pi}$, $W(ix) = \overline{W}(x)$, where \overline{W} denotes the complex conjugate of W . In these formulae we may use the analytic continuation of the amplitude Φ_m and the eikonal τ_m into the primary shadow zone by means of any type of parametrization of space.

Figure 3b shows the edge wave computed by Eq. (17).

Note, that the function $W(w)$ can be represented by known special functions

$$W(w) = (2\sqrt{\pi})^{-1} \Psi(1/2, 1/2; z) = (2\sqrt{\pi})^{-1} \exp(z) \Gamma(1/2, z), \\ z = -i\pi w^2/2 \quad (19)$$

where $\Psi(1/2, 1/2; z)$ is a confluent hypergeometric function, and $\Gamma(1/2, z)$ is an incomplete gamma function. If $0 \leq w < \infty$, we have the following approximate formulae:

$$W(w) = W(0) + w/\sqrt{2} \cdot \exp(i3\pi/4) + O(w^2), \\ W(0) = 1/2 \quad \text{when } w \rightarrow 0, \quad (20)$$

$$W(w) = \exp(i\pi/4) / (\sqrt{2}\pi w) + O(w^{-2}) \quad \text{when } w \rightarrow \infty$$

where O is the symbol of asymptotic estimation. Figure 6 shows the graph of the function $W(w)$, Eq. (18).

Now we shall briefly discuss the type of approxima-

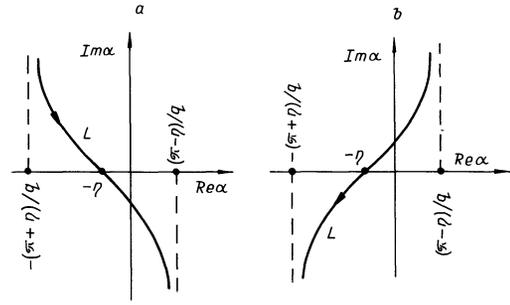


Fig. 5a and b. The contours of integration. Explanation in text

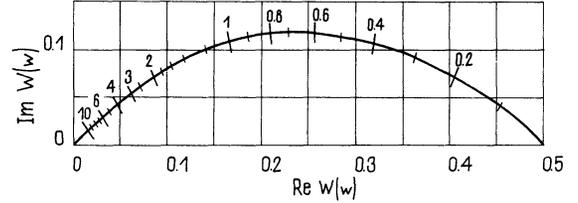


Fig. 6. Function $W(w)$ of the real variable w . Numbers above the curve give the values of w

tion given by Eq. (17). Obviously, the above approach is approximate even in a high-frequency sense, because of the disturbance of the boundary conditions at interfaces by the integral (14). To be more accurate, it would be necessary to add a certain term δf_{mn} to the integral (14) to satisfy the boundary conditions. In principle, this term can be found by using the geometrical theory of diffraction (for example, see Klem-Musatov (1980) for the case $v_m = \text{constant}$). It is essential that such an addition δf_{mn} would have no discontinuity at the primary shadow boundary $\eta=0$. Using the geometrical theory of diffraction (Klem-Musatov, 1980) allows us to estimate the value to the order of $\delta f_{mn} \sim O(\omega^{-1/2})$ except for grazing and critical regions.

Let us see how it would be connected with Eq. (17). According to Eq. (20), the amplitude of the edge wave changes its value from an asymptotic estimation $O(1)$ at $w_{mn}=0$ to $O(\omega^{-1/2})$ at $w_{mn} \rightarrow \infty$. The gradient of this function depends on w_{mn} . Let us write w_{mn} in the form

$$w_{mn} = \sqrt{2N}, \quad N = v_m(\tau_{mn} - \tau_m)/(\lambda/2), \quad \lambda = 2\pi v_m/\omega$$

where N presents the phase difference in the half-period Fresnel zones. If $N=0, 1, 2, 3, 4, 5, \dots$ we have, respectively, $|W|=0.50; 0.17; 0.11; 0.09; 0.08; 0.07; \dots$. We can see that this function changes rapidly when $N < 2 (w_{mn} < 2)$ and slowly when $N > 2$. By the way, the second part ($w_{mn} \rightarrow \infty$) of Eq. (20) is true for $w_{mn} > 2$. The domain of this rapid change forms a neighbourhood of the primary shadow boundary. It is called a *boundary layer*, with a position determined by the inequality $w_{mn} \lesssim 2$. Within the boundary layer an inaccuracy $\delta f_{mn} \sim O(\omega^{-1/2})$ of Eq. (17) may be considered of no importance in comparison with $O(1)$. In Fig. 3 the boundary layer occupies a domain $0.85 \text{ km} \lesssim x \lesssim 1.45 \text{ km}$. Outside the boundary layer the amplitude of the edge wave has the same asymptotic

estimation $O(\omega^{-1/2})$ as δf_{mn} . It is clear that Eq. (17) fails here. More accurate analysis of the integral (14) would not improve the properties of this formula. Thus, Eq. (17) gives a satisfactory result only within the boundary layer. All this is quite enough for correcting the ray method.

Figure 3c shows the total wavefield $f = f_1 + f_{11}$ formed by the superposition of reflected f_1 and edge f_{11} waves.

Equation (17) has one other local property, which allows us to interpret the forthcoming results. According to Eq. (15), a value of Φ_m would be taken at the primary shadow boundary $\Phi_m = \Phi_m(\tau_{mn}, 0, \zeta)$. However, it is possible to consider Φ_m as a function of the free point $\Phi_m = \Phi_m(\tau_{mn}, \eta, \zeta)$ as well because the difference $\Phi_m(\tau_{mn}, \eta, \zeta) - \Phi_m(\tau_{mn}, 0, \zeta)$ is so small in comparison with Φ_m within the boundary layer. The real accuracy of the description of the edge wave is independent of the choice of the above versions. By the way, this is the reason why Φ_m may be continued analytically into the shadow zones.

Interrelation with known physical ideas

We consider briefly, how the above theory matches the concept of Fresnel-Kirchhoff's *secondary sources* and Fock's concept of *transverse diffusion*.

Using Eq. (19) we can represent the superposition of reflected waves, Eq. (13), and edge waves, Eq. (17), within a boundary layer in the form

$$f_m + f_{mn} = f_m F(-s_{mn} w_{mn} \sqrt{\pi/2}), \quad (21)$$

$$F(z) = \pi^{-1/2} \exp(-i\pi/4) \int_{-\infty}^z \exp(ix^2) dx$$

where $F(z)$ is the Fresnel integral. If the wave velocity is constant, the same formula can be derived by an asymptotic analysis of the Kirchhoff integral (Klem-Musatov, 1980; Aizenberg, 1982). Thus Eq. (17) matches the classical ideas of Fresnel-Kirchhoff's theory of diffraction.

Let us show with the simplest example that Eq. (17) also complies with the so-called Fock's parabolic equation of *transverse one-dimensional diffusion* which describes diffusion of the wave energy out of the primary illuminated zone into the primary shadow zone. Let the wave velocity be constant and the wave f_m be plane ($\Phi_m = \text{constant}$) with its wave vector perpendicular to an edge. Let (r, θ, z) be the cylindrical ray coordinates, where r is the distance along a diffracted ray from the edge, θ is the angle between the diffracted ray and the shadow boundary $\theta=0$, z is the distance along the edge. By substitution of Eq. (17) into Helmholtz's equation $(\Delta + k_m^2)f_{mn} = 0$, where $k_m = \omega/v_m$, and neglecting all values within an order less than k_m , we can obtain the well-known equation of transverse diffusion

$$\frac{2ik_m}{\sqrt{r}} \frac{\partial}{\partial r} (\sqrt{r} \Phi_{mn}) + \frac{1}{r^2} \frac{\partial^2}{\partial \theta^2} \Phi_{mn} = 0. \quad (22)$$

Using the following relations within a boundary layer,

$$w_{mn} = \sqrt{2\omega(\tau_{mn} - \tau_m)/\pi} = \sqrt{2k_m r(1 - \cos\theta)/\pi} \approx \theta \sqrt{k_m r/\pi},$$

we can represent Eq. (22) in the form

$$z \frac{d^2}{dz^2} \Phi_{mn} + (1/2 - z) \frac{d}{dz} \Phi_{mn} - 1/2 \Phi_{mn} = 0,$$

$$z = -i\pi w_{mn}^2/2. \quad (23)$$

The solution of this equation is function $W(w)$ in Eq. (19). Thus Eq. (17) describes the phenomenon of transverse diffusion in the form of the edge wave. An analysis shows the following mechanism of this phenomenon. The wave energy flows from the primary illuminated zone through the shadow boundary into the shadow zone along a cone of diffracted rays. There is no energy exchange between neighbouring cones (Klem-Musatov, 1980).

Note, if the wave velocity is constant, Eq. (17) is a first approximation of the more precise description of the edge wave by the successive approximation method in the form of an infinite series (Klem-Musatov, 1980).

Polarization

Now let us take the case when the amplitude of wave (4) is the vector (6). Let $\mathbf{j}_1, \mathbf{j}_2, \mathbf{j}_3$ be the unit vectors of a certain fixed coordinate system (for example, the Cartesian one). Let us decompose the vector (6) on the above basis and represent the wave (4) in the form

$$f_m = \mathbf{p}_m \varphi_m \exp(i\omega\tau_m) = \sum_{q=1}^3 \mathbf{j}_q f_m^{(q)}, \quad (24)$$

$$f_m^{(q)} = \varphi_m^{(q)} \exp(i\omega\tau_m)$$

where the $f_m^{(q)}$ are scalars. We represent the edge wave, Eq. (11), on the same basis:

$$f_{mn} = \sum_{q=1}^3 \mathbf{j}_q f_{mn}^{(q)}, \quad f_{mn}^{(q)} = \varphi_{mn}^{(q)} \exp(i\omega\tau_{mn}) \quad (25)$$

where the $f_{mn}^{(q)}$ are scalars. We use the same approach, Eqs. (12)–(17), for every scalar function $f_m^{(q)}(\tau_{mn}, \eta, \zeta)$ which was used for function (12). It allows us to determine three scalar functions:

$$f_{mn}^{(q)} = s_{mn} \varphi_m^{(q)} W(w_{mn}) \exp(i\omega\tau_{mn}) \quad \text{with } q=1, 2, 3. \quad (26)$$

Inserting Eq. (26) into Eq. (25) gives, once again, Eq. (17), where Φ_m is the vector (6).

This result has to be interpreted. Let \mathbf{p}_{mn} be a unit vector of polarization of the edge wave f_{mn} . In accordance with the general theory this vector must coincide with \mathbf{e}_{mn} (a longitudinal wave) or be perpendicular to \mathbf{e}_{mn} (a transverse wave). But in Eq. (17) the vector \mathbf{p}_{mn} coincides with $\mathbf{p}_m(\tau_{mn}, 0, \zeta)$ which is out of the line of general theory. In other words, the above approach gives an inaccuracy $\delta\mathbf{p} = \mathbf{p}_{mn}(\tau_{mn}, \eta, \zeta) - \mathbf{p}_m(\tau_{mn}, 0, \zeta)$. In fact, the real accuracy of the description of polarization is independent of the choice of any of the versions: $\mathbf{p}_m(\tau_{mn}, 0, \zeta)$, $\mathbf{p}_m(\tau_{mn}, \eta, \zeta)$ or $\mathbf{p}_{mn}(\tau_{mn}, \eta, \zeta)$ because the corresponding $\delta\mathbf{p}$ is of no importance in comparison with \mathbf{p}_{mn} within the boundary layer. That is why the vector Φ_m may be considered a function of the free point $\Phi_m(\tau_{mn}, \eta, \zeta)$ and then continued analytically into the shadow zones.

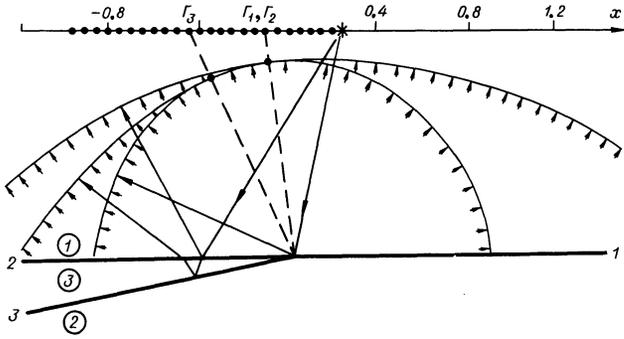


Fig. 7. Model of "pinch-out" and wavefronts in cross-section with projection of the observation system. The profile of observation is perpendicular to the edge. The angle of the pinch-out is 5°

Ray method with consideration of edge diffraction

The present modification of the ray method results in the addition of edge waves, Eq. (17), to the reflected/transmitted wavefield, Eq. (4)

$$f = \sum_m (f_m + \sum_n f_{mn}).$$

This approach can be used outside caustic zones in any inhomogeneous medium with smooth edges of interfaces.

Here we shall show an example of using this approach for the model of a pinch-out (Fig. 7). Inside the pinched layer there are multiple reflections. As a result, there are many reflected waves with sharp shadow boundaries at the profile of observation. Only three of them are shown in Fig. 7. Let us mark them by $m = 1, 2, 3$. We mark their shadow boundaries by m_1 . They are shown with dashed lines (Γ_1, Γ_2 relating to interfaces 1 and 2, and Γ_3 to interfaces 3 with PP -transmission through interface 2). The total wavefield is

$$f = \sum_m (f_m + f_{m_1}).$$

Figure 8 shows the wavefields. Note that all edge waves from the common edge have the same eikonal and form a total diffracted wave. One can see that the addition of edge waves essentially changes the wavefield compared with the ray method wavefield.

The method of superposition of edge waves

Equations (4) and (17) allow us to compute wavefields in 3-D inhomogeneous media with smooth edges of interfaces outside caustic zones. There is an approach, based on the above formulae, which can also be used within the caustic zones if they are formed by the curvature of interfaces (Aizenberg and Klem-Musatov, 1980). Let us describe the idea in brief. The wavefield scattered by a rectifiable interface can be replaced approximately by the field scattered by a piecewise-plane boundary, which is approximated by a sufficiently large number of plane elements.

This approximation guarantees finite values of the terms Φ_m in Eqs. (4) and (17) because they now do not depend on the curvature of the initial interface. The wavefield scattered by the piecewise-plane boundary is

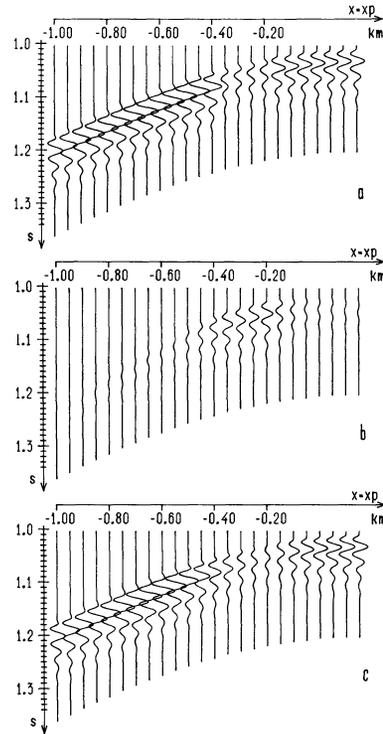


Fig. 8a-c. Theoretical seismograms for model of "pinch-out": **a** reflected waves, **b** edge waves, **c** total field

the superposition of fields scattered by the plane elements composing it.

The simplest version of this approach occurs when the original interface is a cylindrical surface and only single scattering is taken into consideration. Then each element of the piecewise-plane boundary is an infinite band with two rectilinear edges. For a sufficiently small step of approximation, the contributions from the elements, containing reflection/transmission points, are small enough and may be neglected. Then the field scattered by an individual plane element will be the superposition of two edge waves diverging from its edges. The total field can be represented by superposition of just edge waves alone

$$f = \sum_{m=1}^M \Delta f_m, \quad \Delta f_m = \sum_{n=1}^2 f_{mn}, \quad m \neq m_0 \quad (27)$$

where m coincides with the index of an element, n coincides with the index of an edge, m_0 are the indices of elements containing reflection/transmission points.

To justify the above approach, we will take Eq. (27) to the limit of the initial interface, simultaneously letting the step of approximation tend to zero and increasing the number of elements. In this case

$$\lim_{\substack{\max(\Delta l_m) \rightarrow 0 \\ M \rightarrow \infty}} \sum_{m=1}^M \Delta l_m = \int_L dl, \quad (28)$$

$$f = \int_L F(l) \exp[i\omega\tau(l)] dl,$$

$$F(l) = \Phi_0(l) [-i\omega/(4\pi)]^{1/2} \left| \frac{\partial\tau}{\partial l} \right| (\tau - \tau_0)^{-1/2}$$

where $\Phi_0(l)$ and τ_0 are the amplitude and the eikonal of the analytic reflection/transmission for point l , and τ is the eikonal of the scattered wave. Here the analytic reflection/transmission stands for the usual reflection/transmission from/through the plane boundary which is a tangent to the initial interface at point l .

Equation (28) can be regarded as a new asymptotic formulation of the concept of secondary sources. According to this formulation the value of the field at any point of the medium will be the superposition of secondary waves with amplitude $F(l)dl$ diverging from each point of the initial interface. Analysis shows that, unlike the classical Fresnel and Kirchhoff descriptions, the amplitudes of the secondary waves in Eq. (28) are strictly bounded on the contour of integration.

Within the concept of secondary sources obtained we can explain the phenomena of regular reflection/transmission of waves, formation of caustics, edge effects and the phenomena of scattering of waves by interfaces of complex form. It is known that the asymptotic behaviour of integrals of the type in Eq. (28), for large values of ω , is determined by stationary points of the function $\tau(l)$, singular points for the function $F(l)$ and the end points of the contour of integration (Felsen and Marcuvitz, 1973). The asymptotic analysis in the neighbourhood of an isolated stationary point of the first order gives the wavefield, Eq. (28), in the form of a regular wave

$$f = \Phi(l_1) \exp[i\omega\tau(l_1)], \quad (29)$$

$$\Phi(l_1) = \Phi_0(l_1) \left[1 - \frac{\partial^2 \tau_0(l_1)}{\partial l^2} \Big/ \frac{\partial^2 \tau(l_1)}{\partial l^2} \right]^{-1/2}.$$

If the wave velocity is constant, this formula coincides with the formula of the ray method.

The asymptotic analysis of integral (28) in the neighbourhood of an isolated stationary point of second order gives the approximation for a simple caustic

$$\begin{aligned} f &= \Phi(l_2) \text{Ai}(q) \exp[i\omega\tau(l_2)], \\ \Phi(l_2) &= \Phi_0(l_2) (-i\pi/\omega)^{1/2} |q| [\tau(l_2) - \tau_0(l_2)]^{-1/2}, \\ q &= \omega^{2/3} 2^{1/3} \frac{\partial \tau(l_2)}{\partial l} \left[\frac{\partial^3 \tau(l_2)}{\partial l^3} \right]^{-1/3} \end{aligned} \quad (30)$$

where $\text{Ai}(q)$ is the Airy function. The same analysis in the neighbourhood of the irregular points

$$\frac{\partial \tau(l_3+0)}{\partial l} \neq \frac{\partial \tau(l_3-0)}{\partial l}$$

gives the Eqs. (29) and (17). Thus, the above justifies the present approach.

We shall show the potential of this approach for the model "flexure" (Fig. 9). Let the geometrical form of an interface be described by the formula $z = 1 + \delta/\pi \arctan[\alpha(x-0.4)]$ where δ is the difference between the depths of the wings of the flexure. The wavefields are shown in Fig. 10 for $\delta = \lambda_p/4$, where $\lambda_p = 0.06$ km. The top seismogram is for the sloping flexure $\alpha = 5$, the middle one for $\alpha = 50$, and the bottom one for the very steep flexure $\alpha = 500$. It is clearly seen that, for the small curvature of the interface, the wavefield may be obtained by the ray method. If the

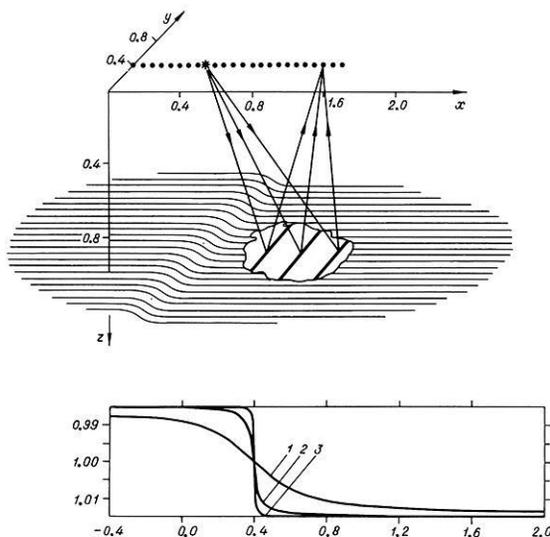


Fig. 9. Model of "flexure" (top) and its cross-section with the reflecting interfaces (bottom). Numbers corresponds to 1: $\alpha = 5$, 2: $\alpha = 50$, 3: $\alpha = 500$. Edge rays, scattering from elements of the interface, are shown

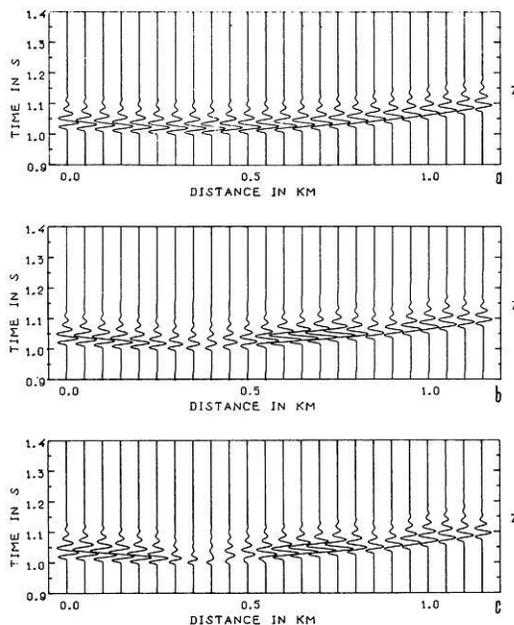


Fig. 10a-c. Theoretical seismograms for model of "flexure": a $\alpha = 5$, b $\alpha = 50$, c $\alpha = 500$

curvature of the interface is around $\alpha = 50$, some kind of caustic phenomenon occurs in the form of a local loss of intensity caused by interference (0.35–0.45 km). If the curvature of the interface is extremely large, $\alpha = 500$, the wavefield coincides with the field for a fault with small throw. In this case there are two reflections from the wings of the fault and two edge waves. The interference of the edge waves forms the field, the character of which depends on the value of the parameter δ .

Tip waves

The point of break (or the end) of a smooth edge is called a tip. The common tip of several edges is a

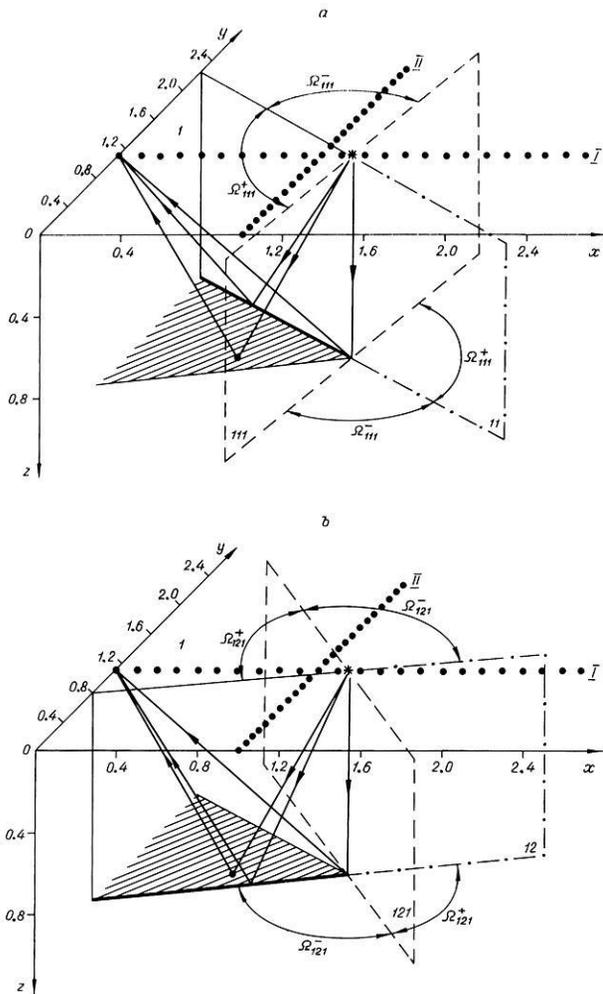


Fig. 11 a and b. Model of "sector". Reflecting interface coincides with shaded region. Further explanation in text

vertex. The sizes of edge wave domains are limited because of the tips. A single edge wave f_{mn} exists within the connected domain coinciding with the corresponding congruence of diffracted rays. This wavefield f_{mn} is continuous everywhere within its domain, with the exception of the primary shadow boundary, $\tau_{mn} = \tau_m$. This type of domain is called the *secondary illuminated zone* of the wave f_{mn} . A domain of absence of the wave ($f_{mn} \equiv 0$) is called the *secondary shadow zone*. A simply connected surface dividing the above zones is called the *secondary shadow boundary*. It looks like the surface of a curvilinear cone whose apex angle complies with the law of edge diffraction. Let mnp be the triple number of each secondary shadow boundary of the edge wave f_{mn} . The non-caustic shadow boundaries formed by tips are considered in the following.

Let us illustrate these definitions by an example of the reflection from an interface "sector". Its location is shown in Fig. 11. According to all the above, we have to describe the wavefield as a superposition of the reflected wave f_1 and two edge waves f_{11} and f_{12} . In Fig. 11 the primary shadow boundaries are marked by the indices 11 and 12 and the secondary shadow boundaries by the indices 111 and 121. The domain of existence of the reflected wave is marked by the index 1

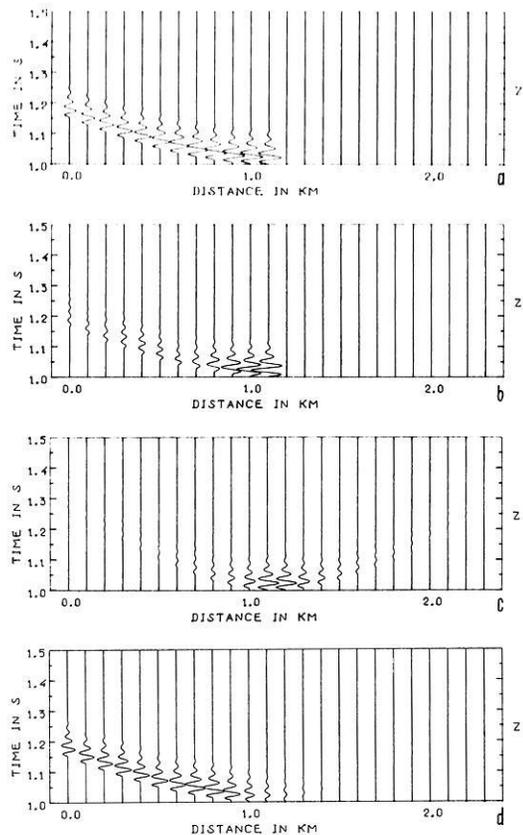


Fig. 12 a-d. Theoretical seismograms for model of "sector" (profile I): a reflected wave, b edge waves (1.5 times enlarged), c tip waves (7.5 times enlarged), d total field

and the domains of the secondary illuminated zones are located on the left of the dashed lines. Figures 12a and 13a show the reflected wave for two profiles marked by I and II. Figures 12b and 13b show the edge wavefields scattered from both edges.

One can see that a shortcoming of Eq. (17) appears as discontinuities of the wavefield f_{mn} at the secondary shadow boundaries. Let us see how this can be corrected.

Kinematics

Let us use a formal method to find the directions of rays arising from a tip. It concerns Snell's law in the form of Eq. (9). However, there are no limitations in choosing the directions of arising rays because the tip is not a linear element of interfaces. Any direction complies with the above-mentioned law formally. This fact is formulated as the *law of tip diffraction* (Keller, 1962). It reads as follows: the incident ray generates rays leaving the tip in all directions.

Let \mathbf{e}_{mnp} be a unit vector of the tangent to a ray. Let this ray comply with the law of tip diffraction at that tip, which gives the mnp -th secondary shadow boundary. Then the differential equation

$$\frac{d}{ds} (\mathbf{e}_{mnp}/v_m) = \text{grad}(1/v_m) \quad (31)$$

determines the congruence of *tip diffracted rays*.

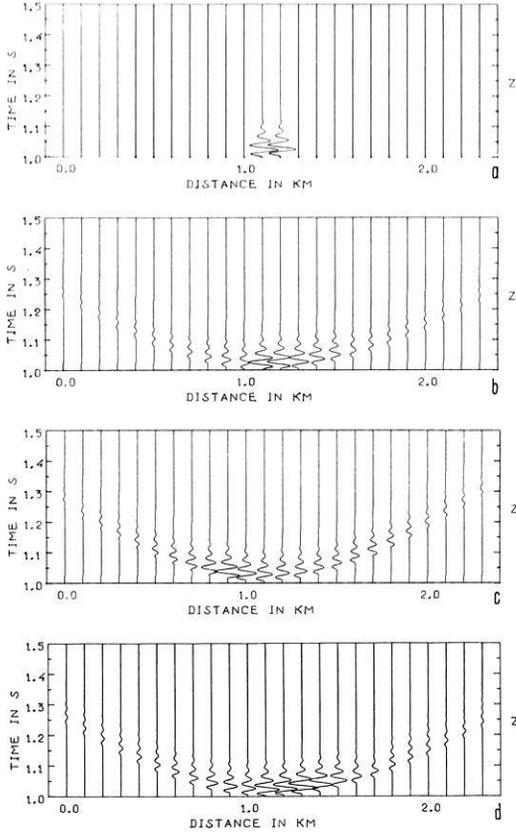


Fig. 13 a–d. Theoretical seismograms for model of “sector” (profile II): **a** reflected wave (twice diminished), **b** edge waves (1.66 times diminished), **c** tip waves (4 times enlarged), **d** total field

A tip diffraction coefficient

Let a wave

$$f_{mnp} = \Phi_{mnp} \exp(i\omega\tau_{mnp}), \quad \text{grad } \tau_{mnp} = \mathbf{e}_{mnp}/v_m \quad (32)$$

be connected with the mnp -th secondary shadow boundary. The latter may be given implicitly by the equation $\tau_{mnp} = \tau_{mn}$. The wave, Eq. (32), is called a *tip diffracted wave*. Let us divide the tip wave domain into separate parts. Suppose, the eikonal τ_{mn} may be continued analytically into the secondary shadow zone. The analytical continuation of the primary shadow boundaries $\tau_{11} = \tau_1$ and $\tau_{12} = \tau_1$ is shown in Fig. 11 by the dash-dotted lines.

Then the primary shadow boundary $\tau_{mn} = \tau_m$ and the secondary shadow boundary $\tau_{mnp} = \tau_{mn}$ divide the domain of the wave f_{mnp} into four parts. These parts are shown in Fig. 14. Let us give them the numbers 1, 2, 3 and 4, going around the line $\tau_{mnp} = \tau_{mn} = \tau_m$ clockwise or counter-clockwise, so that the shortest way from the fourth part to the first would coincide with the shortest way from the primary illuminated zone of the wave f_m to the primary shadow zone through the mn -th primary shadow boundary. The first and third parts have common points only at the line $\tau_{mnp} = \tau_{mn} = \tau_m$. The second and fourth parts have common points at the same line only. Let Ω_{mnp}^- be the symbol of the domain formed by the first and third parts. Let Ω_{mnp}^+ be

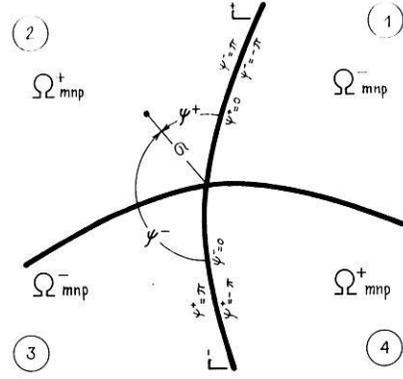


Fig. 14. The domains of continuity of a tip wave. Explanation in text

the symbol of the domain formed by the second and fourth parts. These domains in Fig. 11 for the secondary shadow boundary 111 are denoted by the symbols Ω_{111}^- and Ω_{111}^+ . Let us denote the boundary between the first and second parts by Γ^+ , and between the third and fourth parts by Γ^- . Γ^+ and Γ^- together form the secondary shadow boundary $\tau_{mnp} = \tau_{mn}$.

Let us represent the tip wave, Eq. (32), in the form:

$$f_{mnp} = f^+ + f^-, \quad f^\pm = \Phi^\pm \exp(i\omega\tau_{mnp}). \quad (33)$$

Let the sum $f_{mn} + f^+$ be continuous at the boundary Γ^+ and the sum $f_{mn} + f^-$ be continuous at the boundary Γ^- . Under the above conditions, we can find f^+ and f^- in the same way which is used for finding f_{mn} .

Let τ_{mnp} , ψ^\pm , σ be ray coordinates of the wave f^\pm . Here ψ^\pm and σ give a congruence of the tip diffracted rays, i.e. every pair of fixed values $\psi^\pm = \text{constant}$ and $\sigma = \text{constant}$ gives a single ray. Let ψ^\pm vary in the interval $-\pi \leq \psi^\pm \leq \pi$. We choose ψ^\pm in such a way that the surface $\psi^\pm = 0$ would coincide with the surface Γ^\pm , and the surfaces $\psi^\pm = \pi$ and $\psi^\pm = -\pi$ would coincide with Γ^\mp (Fig. 14). Then

$$\begin{aligned} \psi^\pm &= \pm s_{mnp} |\psi^\pm| & \text{when } |\psi^\pm| < \pi/2, \\ \psi^\pm &= \pm s_{mnp} (|\psi^\pm| - \pi) & \text{when } |\psi^\pm| > \pi/2, \end{aligned} \quad (34)$$

$$\begin{aligned} s_{mnp} &= +1 & \text{within } \Omega_{mnp}^+, \\ s_{mnp} &= -1 & \text{within } \Omega_{mnp}^-. \end{aligned} \quad (35)$$

Let us consider $\sigma = 0$ at the line $\tau_{mnp} = \tau_{mn} = \tau_m$.

In the first place, let us take the case when the amplitude Φ_{mnp} of the wave (11) is a scalar. Let the edge wavefield, Eq. (11), be a function of the above ray coordinates

$$f_{mn} = \Phi_{mn}(\tau_{mnp}, \psi^\pm, \sigma) \exp[i\omega\tau_{mn}(\tau_{mnp}, \psi^\pm, \sigma)]. \quad (36)$$

Then, in the neighbourhood of the secondary shadow boundary, this wavefield may be represented by the discontinuous function

$$\begin{aligned} f_{mn} &= f_{mn}(\tau_{mnp}, \psi^\pm, \sigma) & \text{when } -\pi < \psi^\pm < 0, \\ f_{mn} &= 0 & \text{when } 0 < \psi^\pm < \pi \end{aligned} \quad (37)$$

which displays explicitly the shortcoming of Eq. (17).

Figures 12a and 13a illustrate the discontinuity of the reflected wave. Figure 13b shows the change of sign of two edge waves at the primary shadow boundary 12 ($y=1.06$ km) and at the primary shadow boundary 11 ($y=1.25$ km). In Fig. 12b the discontinuity of the edge wave amplitudes is seen at the secondary shadow boundary $x=1.15$ km.

Suppose Eq. (37) represents an analytical function of the variable ψ^\pm and allows us to make an analytical continuation into the complex plane of ψ^\pm for any permissible values τ_{mnp} and σ . Let us find \tilde{f}^\pm among the piecewise-analytical functions decreasing at infinity ($\tilde{f}^\pm \rightarrow 0$ when $|\psi^\pm| \rightarrow \infty$). Then we may construct the following integral of Cauchy's type

$$\tilde{f}^\pm = \frac{1}{2\pi i} \int_L f_{mn}(\tau_{mnp}, \psi^\pm + \alpha, \sigma) \frac{d\alpha}{\alpha}. \quad (38)$$

The specific type of the contour L will be given in the following. The properties of the similar integral, Eq. (14), have been discussed above. Equation (38) has a discontinuity at $\psi^\pm=0$. However, the superposition of Eqs. (37) and (38) is continuous at this point. The problem is that integral (38) has two extra discontinuities at $\psi^\pm = -\pi$ and $\psi^\pm = \pi$ because of the limited interval $-\pi \leq \psi^\pm \leq \pi$. To eliminate these discontinuities we take the periodic function of ψ^\pm , i.e.

$$f^\pm = \sum_{k=-\infty}^{\infty} \tilde{f}^\pm(\tau_{mnp}, \psi^\pm + 2\pi k, \sigma). \quad (39)$$

Inserting Eq. (38) into Eq. (39) and using the well-known formula

$$\sum_{k=-\infty}^{\infty} (z - 2\pi k)^{-1} = 1/2 \cdot \cot(z/2) \quad (40)$$

we get:

$$f^\pm = \frac{1}{4\pi i} \int_L f_{mn}(\tau_{mnp}, \psi^\pm + \alpha, \sigma) \cot(\alpha/2) d\alpha. \quad (41)$$

Integral (41) has the following properties. It is zero when $|\psi^\pm| \rightarrow \infty$. It has a discontinuity at $\psi^\pm=0$. However, the superposition of Eqs. (37) and (41) is a continuous and analytical function of ψ^\pm in the neighbourhood of the surface $\psi^\pm=0$. If function (37) is a solution of some linear differential equation (for example, the wave equation) within the domain $-\pi < \psi^\pm < 0$, the superposition of Eqs. (37) and (41) complies with the same equation within the whole domain $-\pi \leq \psi^\pm \leq \pi$.

When $\omega \rightarrow \infty$, the asymptotic value of integral (41) is formed by contributions within a small neighbourhood of the saddle point $\alpha = -\psi^\pm$. Let us take the standard approximation at this point

$$\Phi_{mn}(\tau_{mnp}, \psi^\pm + \alpha, \sigma) \approx \Phi_{mn}(\tau_{mnp}, 0, \sigma) \quad (42)$$

and for $|\sigma| \ll 1$ use the following relations

$$\begin{aligned} \tau_{mn}(\tau_{mnp}, \psi^\pm, \sigma) &\approx \tau_{mnp} - A \sin^2 \psi^\pm, \\ A &= \tau_{mnp} - \tau^* \approx \tau_{mnp} - \tau_m, \end{aligned} \quad (43)$$

$$\tau^* = \tau_{mn}(\tau_{mnp}, \psi^\pm, \sigma) \quad \text{with} \quad |\psi^\pm| = \pi/2,$$

$$|\psi^\pm| = \arcsin \sqrt{(\tau_{mnp} - \tau_{mn}) / (\tau_{mnp} - \tau_m)}. \quad (44)$$

Let us note that Eq. (43) may be derived by using a similar method as in the case $v_m = \text{constant}$ (Klem-Musatov, 1981a).

Let the contour L within a neighbourhood of the saddle point coincide with the steepest descent path

$$\text{Im}[i\omega\tau_{mn}(\tau_{mnp}, \psi^\pm + \alpha, \sigma)] = \text{Im}[i\omega\tau_{mn}(\tau_{mnp}, 0, \sigma)],$$

$$\text{Re}[i\omega\tau_{mn}(\tau_{mnp}, \psi^\pm + \alpha, \sigma)] < 0.$$

Using Eq. (43) gives the following equations for this part of the contour

$$\text{Re } \alpha + \psi^\pm = -\text{Im } \alpha \quad \text{for } \tau_{mnp} > \tau_m,$$

$$\text{Re } \alpha + \psi^\pm = \text{Im } \alpha \quad \text{for } \tau_{mnp} < \tau_m.$$

The integral exists if the contour approaches to the points $\text{Im } \alpha = \pm \infty$ within domains

$$\text{Re}[i\omega\tau_{mn}(\tau_{mnp}, \psi^\pm + \alpha, \sigma)] < 0.$$

Using Eq. (43) allows us to obtain this condition for $\text{Im } \alpha \rightarrow \infty$ in the following form

$$-\pi/2 < \text{Re } \alpha + \psi^\pm < 0 \quad \text{when } \tau_{mnp} > \tau_m,$$

$$0 < \text{Re } \alpha + \psi^\pm < \pi/2 \quad \text{when } \tau_{mnp} < \tau_m,$$

and for $\text{Im } \alpha \rightarrow -\infty$ in the form

$$0 < \text{Re } \alpha + \psi^\pm < \pi/2 \quad \text{when } \tau_{mnp} > \tau_m,$$

$$-\pi/2 < \text{Re } \alpha + \psi^\pm < 0 \quad \text{when } \tau_{mnp} < \tau_m.$$

The contours are shown in Fig. 5a for $\tau_{mnp} > \tau_m$ and in Fig. 5b for $\tau_{mnp} < \tau_m$ (in these figures, $q=2$ and term η must be replaced by ψ^\pm).

Then integral (41) may be written as

$$f^\pm = \Phi_{mn}(\tau_{mnp}, 0, \sigma) \Psi^\pm \exp(i\omega\tau_{mnp}),$$

$$\Psi^\pm = \frac{1}{4\pi i} \int_L [\cot[(\alpha - \psi^\pm)/2]] \exp[i\omega(\tau_m - \tau_{mnp}) \sin^2 \alpha] d\alpha. \quad (45)$$

To discuss the accuracy of this expression, we would repeat all that was said concerning Eq. (17). Equation (45) gives a satisfactory description within the so-called boundary layer where the amplitude of a tip wave changes rapidly. Within this domain the amplitude Φ_m may be considered as a function of the free point $\Phi_m(\tau_{mnp}, \psi^\pm, \sigma)$. The real accuracy of description is independent of the choice of the versions:

$$\Phi_m(\tau_{mnp}, 0, \sigma) \quad \text{or} \quad \Phi_m(\tau_{mnp}, \psi^\pm, \sigma).$$

Consideration of Eqs. (17), (34), (44), (45) and using identical mathematical transformations (for details, see Klem-Musatov, 1981a, b) allow us to write Eq. (33) in the form:

$$f_{mnp} = s_{mnp} \Phi_m H(\rho_{mnp}, \zeta_{mnp}) \exp(i\omega\tau_{mnp}), \quad (46)$$

$$H(\rho, \zeta) = W(\rho) \Psi(\rho, \zeta), \quad (47)$$

$$\Psi(\rho, \zeta) = \sin(2\zeta)/\pi \cdot \int_0^1 [x^2 - 2x \cos(2\zeta) + 1]^{-1} \cdot \exp[i\pi\rho^2(x + x^{-1} - 2)/8] dx,$$

$$\rho_{mnp} = \sqrt{2\omega(\tau_{mnp} - \tau_m)/\pi},$$

$$\zeta_{mnp} = \arcsin \sqrt{(\tau_{mnp} - \tau_{mn})/(\tau_{mnp} - \tau_m)}, \quad (48)$$

where H may be regarded as a tip diffraction coefficient. If $\tau_{mnp} < \tau_m$, we have $\rho_{mnp} = ix$, $x = \sqrt{2\omega(\tau_m - \tau_{mnp})/\pi}$, $H(ix, \zeta) = \overline{H(x, \zeta)}$, where \overline{H} denotes the complex conjugate of H . In these formulae we may use the analytical continuation of the amplitude Φ_m and the eikonals τ_m , τ_{mn} into the primary and secondary shadow zones by means of any type of parameterization of space.

Let us return to the example "sector" (Fig. 11). Now we can write the total wavefield in the form

$$f = f_1 + \sum_{n=1}^2 (f_{1n} + f_{1n1})$$

where f_{1n1} is a tip wave. Figures 12c and 13c show the tip waves computed by Eq. (46). It is possible to see in Fig. 13c the change of sign of their amplitudes at the secondary shadow boundaries 111 ($y=0.95$ km) and 121 ($y=1.66$ km). Figures 12d and 13d show the total wavefield formed by the interference of the reflected wave, two edge waves and two tip waves. The total field is regular everywhere.

If $0 \leq \rho < \infty$, $0 \leq \zeta \leq \pi/2$, we have the following approximate formulae

$$\Psi(\rho, \zeta) = \Psi(0, \zeta) - i\rho^2 \sin \zeta / 8 \cdot \ln(\pi\rho^2/8) \quad \text{when } \rho \rightarrow 0, \quad (49)$$

$$\Psi(\rho, \zeta) = (\sqrt{2\pi\rho})^{-1} (\zeta^{-1} - \cot \zeta) \exp(i5\pi/4) + W(\rho\zeta) + O(\rho^{-2}) \quad \text{when } \rho \rightarrow \infty, \quad (50)$$

$$\Psi(0, \zeta) = 1/2 - \zeta/\pi, \quad \Psi(\rho, 0) = 1/2, \quad \Psi(\rho, \pi/2) = 0 \quad (51)$$

where O is the symbol of asymptotic estimation. The point $\rho=0$ is the essential special point because the value of the function depends on the direction along which this point is approached. However, the total wavefield at this point is determined uniquely. Let us give the corresponding result.

Every single wavefield f_m has only two primary shadow boundaries within the small neighbourhood of the point $\rho=0$ (Fig. 15). Let us call their indices $n=a$ and $n=b$. Let γ_m be the dihedral angle between the tangent planes to the ma -th and mb -th primary shadow boundaries at a point of the line $\tau_{mnp} = \tau_{mn} = \tau_m$ with $n=a$ and $n=b$. This angle must be taken within (and not outside) the primary illuminated zone. Then at this point the following equality exists:

$$f_m + f_{ma} + f_{mb} + f_{map} + f_{mbp} = f_m \gamma_m / 2\pi. \quad (52)$$

Figure 16 shows the graphs of modulus and argument of the function (48).

Interrelation with known physical ideas

It was shown (Aizenberg, 1982) that Eq. (46) matches the classical theory of diffraction. If $v_m = \text{constant}$, the

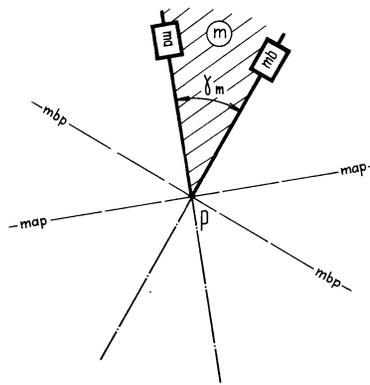


Fig. 15. The neighbourhood of a singular point of the primary illuminated zone (shaded region). Explanation in text

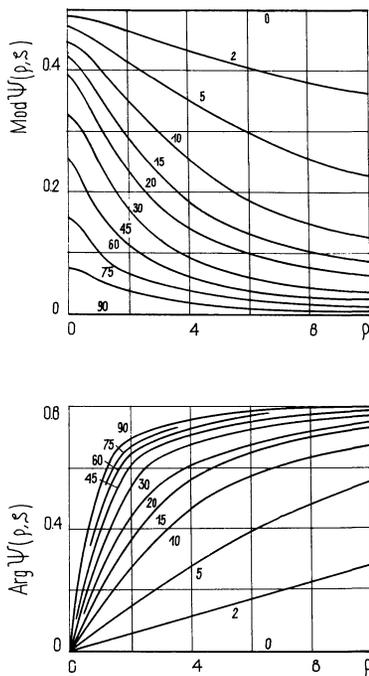


Fig. 16. Modulus and argument of function $\Psi(\rho, \zeta)$. Numbers at the curves give the values of ζ , expressed in degrees

tip wave can be found by asymptotic analysis of Kirchhoff's integral in the form of Eq. (46), where

$$H(\rho, \zeta) = \rho \cos \zeta / (2\pi) \int_{\rho \sin \zeta}^{\infty} (x^2 + \rho^2 \cos^2 \zeta)^{-1} \cdot \exp[i\pi(x^2 - \rho^2 \sin^2 \zeta)/2] dx. \quad (53)$$

It has been shown, both numerically and analytically, that this function is identical with the product in Eq. (47). It allows us to represent the superposition of Eqs. (17) and (46) in the form:

$$f_{mn} + f_{mnp} = f_m G(s_{mn} \sqrt{\pi/2} \rho \cos \zeta, s_{mnp} \sqrt{\pi/2} \rho \sin \zeta), \quad (54)$$

$$G(a, b) = a/(2\pi) \cdot \int_b^{\infty} (x^2 + a^2)^{-1} \exp[i(x^2 + a^2)] dx, \quad (55)$$

where $G(a, b)$ is the so-called generalized Fresnel integral (Clemmow and Senior, 1953) $\rho = \rho_{mnp}$, $\zeta = \zeta_{mnp}$.

Let us show with the simplest example that Eq. (46) complies with Fock's parabolic equation of *transverse two-dimensional diffusion* which describes diffusion of wave energy out of the secondary illuminated zone into the secondary shadow zone. Let the wave velocity be constant and wave f_m be plane ($\Phi_m = \text{constant}$) with its wave vector perpendicular to an edge containing a point of break. Let (R, θ, φ) be spherical coordinates, where R is the distance from the tip, θ is the angle between the tip ray and the ray $\theta=0$ which coincides with the intersection of the primary and secondary shadow boundaries, φ is the angle between the mnp -th secondary shadow boundary $\varphi=0$ and the plane which contains the given tip ray and the ray $\theta=0$.

By substitution of Eq. (46) into Helmholtz's equation $(\Delta + k_m^2)f_{mnp} = 0$, where $k_m = \omega/v_m$, neglecting all values of order less than k_m and using a linear approximation for trigonometric functions of small argument, we can obtain the equation of transverse diffusion

$$\begin{aligned} \frac{2ik_m}{R} \frac{\partial}{\partial R} (R\Phi_{mnp}) + \frac{1}{R^2} \frac{\partial}{\partial \theta} \frac{\partial}{\partial \theta} \Phi_{mnp} \\ + \frac{1}{R^2} \frac{\partial^2}{\partial \theta^2} \Phi_{mnp} + \frac{1}{R^2 \theta^2} \frac{\partial^2}{\partial \varphi^2} \Phi_{mnp} = 0. \end{aligned} \quad (56)$$

Using the following relations in the neighbourhood of the ray $\theta=0$

$$\begin{aligned} \rho_{mnp} &= \sqrt{2\omega(\tau_{mnp} - \tau_m)/\pi} \\ &= \sqrt{2k_m R(1 - \cos\theta)/\pi} \approx \theta \sqrt{k_m R/\pi}, \\ \zeta_{mnp} &= \arcsin \sqrt{(\tau_{mnp} - \tau_{mn})/(\tau_{mnp} - \tau_m)} = \arcsin \\ &\sqrt{[1 - (\cos^2\theta + \sin^2\theta \cos^2\varphi)^{1/2}]/(1 - \cos\theta)} \approx \varphi, \end{aligned} \quad (57)$$

we can represent Eq. (56) in the form

$$\begin{aligned} 2\pi i \rho^2 \Phi_{mnp} + (\pi i \rho^3 + \rho) \frac{\partial}{\partial \rho} \Phi_{mnp} + \rho^2 \frac{\partial^2}{\partial \rho^2} \Phi_{mnp} \\ + \frac{\partial^2}{\partial \zeta^2} \Phi_{mnp} = 0. \end{aligned} \quad (58)$$

The solution of this equation is $\Phi_{mnp} = H(\rho_{mnp}, \zeta_{mnp})$ where H is the integral (53). Thus, Eq. (46) describes a phenomenon of transverse diffusion in the form of the tip wave. Analysis shows the following mechanism of this phenomenon. The wave energy flows from the secondary illuminated zone through the secondary shadow boundary into the secondary shadow zone around the ray $\theta=0$. Hence, unlike the diffusion mechanism described earlier, the tip wave is formed by a type of three-dimensional eddy diffusion around the ray $\theta=0$.

Polarization

Now let us take the case where the amplitude of the wave (11) is the vector (25). Let us represent the tip wave (32) in the form:

$$f_{mnp} = \sum_{q=1}^3 \mathbf{j}_q f_{mnp}^{(q)}, \quad f_{mnp}^{(q)} = \varphi_{mnp}^{(q)} \exp(i\omega\tau_{mnp}) \quad (59)$$

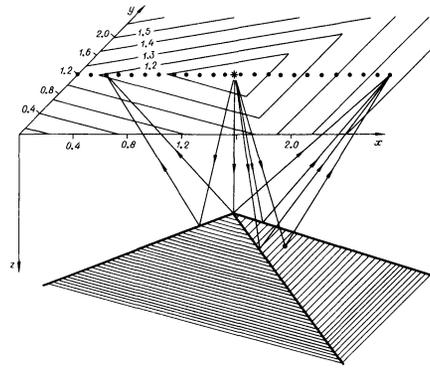


Fig. 17. Model of "pyramid". Reflecting interface is given by depth lines on the plane of observation

where $f_{mnp}^{(q)}$ are scalars. Let us use the same approach, Eqs. (33)–(46), for every function $\varphi_{mnp}^{(q)}(\tau_{mnp}, \psi^\pm, \sigma)$ which has been used for the scalar case. It allows us to determine three scalar functions

$$f_{mnp}^{(q)} = s_{mnp} \varphi_m^{(q)} H(\rho_{mnp}, \zeta_{mnp}) \exp(i\omega\tau_{mnp}) \quad (60)$$

with $q=1, 2, 3$.

Inserting Eq. (60) into Eq. (59) gives Eq. (46) again, where Φ_m is the vector (6).

This result may be interpreted in the same way as for the edge wave above. Let \mathbf{p}_{mnp} be a unit vector of polarization of the tip wave f_{mnp} . In accordance with the general theory, this vector must coincide with \mathbf{e}_{mnp} (for a longitudinal wave) or be perpendicular to \mathbf{e}_{mnp} (for a transverse wave). However, in Eq. (46) the vector \mathbf{p}_{mnp} coincides with \mathbf{p}_m . In fact, the real accuracy of description is independent of this discrepancy because the latter is of no importance within the boundary layer. The vector Φ_m may be considered as a function of a free point in space and continued analytically into the shadow zones.

The ray method including diffraction on edges and vertices

The present modification of the ray method results in the addition of edge, Eq. (17), and tip, Eq. (46), waves to the reflected/transmitted wavefield, Eq. (4),

$$f = \sum_m [f_m + \sum_n (f_{mn} + \sum_p f_{mnp})].$$

This approach can be used in any inhomogeneous media with piecewise-smooth edges of interfaces outside caustic zones. Here we shall show two simple examples of using this approach.

Let the interface be pyramid-shaped (Fig. 17). The total wavefield is given in the form

$$f = \sum_{m=1}^3 \left[f_m + \sum_{n=1}^2 (f_{mn} + f_{mn1}) \right]$$

where m is the index of a face of the interface and f_{mn1} is a tip wave. Figure 18 shows the wavefield scattered by this interface. Note, that all six tip waves have the

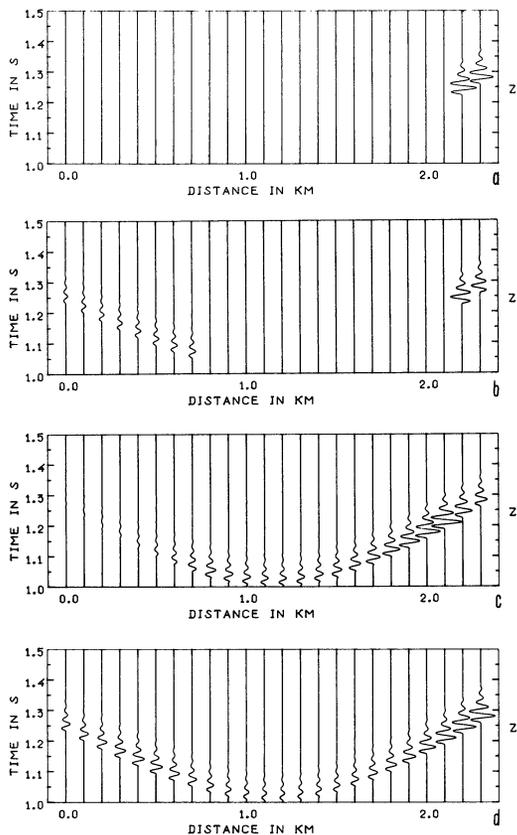


Fig. 18 a-d. Theoretical seismograms for model of "pyramid": **a** reflected wave (twice diminished), **b** edge waves (twice diminished), **c** tip waves (1.5 times enlarged), **d** total field

same eikonal and form a common diffracted wave scattered from the vertex of the interface.

Let us next consider the model "faults". It is known that, in seismic prospecting and deep seismic sounding, the observed waves often have a complex group (multiphase) character inspite of a rather simple source signal shape. The simplest example of the formation of multiphase groups is obtained from an examination of reflections from a boundary disturbed by a system of faults with small throw.

Let the interface be disturbed by two intersecting systems of faults, where each system contains four parallel faults of infinite extension (Fig. 19). Unlike the previous examples, the elements of the interface differ in their number of edges and the edges differ in their number of tips. Let us mark the elements with four edges by $m=1, 2, \dots, 9$, those with three edges by $m=10, 11, \dots, 21$, and those with two edges by $m=22, 23, \dots, 25$. Then the total wavefield may be written in the form

$$f = f^1 + f^2 + f^3,$$

$$f^1 = \sum_{m=1}^9 \left[f_m + \sum_{n=1}^4 \left(f_{mn} + \sum_{p=1}^2 f_{mnp} \right) \right],$$

$$f^2 = \sum_{m=10}^{21} \left[f_m + \sum_{n=1}^2 (f_{mn} + f_{mn1}) + f_{m3} + f_{m31} + f_{m32} \right],$$

$$f^3 = \sum_{m=22}^{25} \left[f_m + \sum_{n=1}^2 (f_{mn} + f_{mn1}) \right].$$

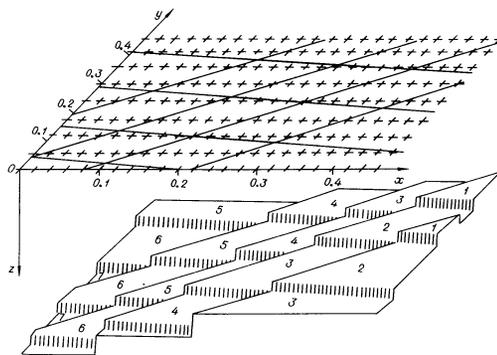


Fig. 19. Model of "faults". Projection of faults is given on the plane of observation. Further explanations in text

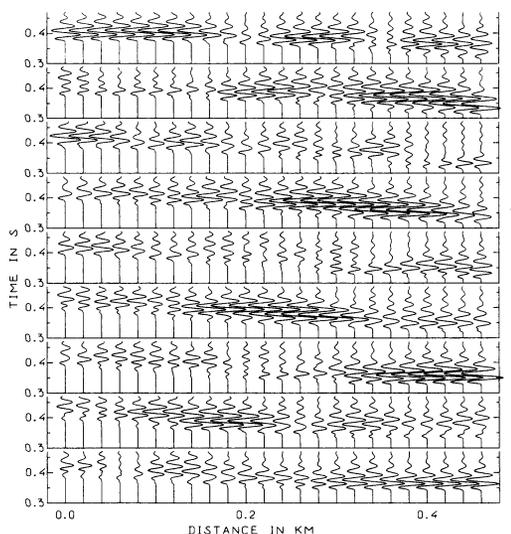


Fig. 20. Theoretical seismogram for model of "faults" ($h = 5 \lambda_p$)

Let the minimal distance of the interface from the plane of observation be $h = N \lambda_p$ where λ_p is the P wavelength. The depth of each block of interface is given by the formula $z = h + k \lambda_p / 4$ where k is given in Fig. 19. Figures 20 and 21 show seismograms of the wavefield scattered by the disturbed interface for $h = 5 \lambda_p$ and $h = 50 \lambda_p$, respectively.

These examples show that for relatively small depth ($h = 5 \lambda_p$) the lineups mainly represent the block structure of the interface with characteristic horizontal dimensions of the blocks of $2-3 \lambda_p$ (and, obviously, larger). Under these conditions, the character of the wave patterns is determined by that part of the field which is controlled by the laws of geometrical seismics (ray transport of energy). The diffraction components (mechanism of transverse diffusion) have a subordinate character, smoothing the characteristics of the field and complicating it by interference effects. With increasing depth of the interface, the role of diffraction components increases since the absolute dimensions of the zones of influence of diffusion mechanisms – the vicinity of the reflection point – increase. For relatively great depth ($h = 50 \lambda_p$), the diffusion mechanism plays the dominant role in the formation of fields from disturbed interfaces. Interference of diffraction components

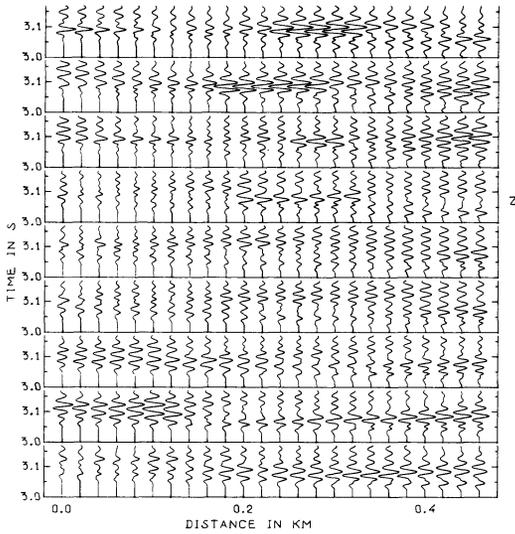


Fig. 21. Theoretical seismogram for model of “faults” ($h = 50 \lambda_p$)

generates multiphase trains, the character of which depends on the degree of disturbance of the interface.

The method of superposition of tip waves

Equations (4), (17) and (46) allow us to compute wavefields in 3-D inhomogeneous media with piecewise-smooth edges of interfaces outside caustic zones. There is an approach, based on the above formulae, which can be used within caustic zones as formed by the curvature of interfaces. Its idea is a generalization of the approach described in an earlier section.

According to the present approach, the wavefield scattered by a rectifiable interface can be replaced approximately by the field scattered by a piecewise-plane boundary, approximating it with a sufficiently large number of plane elements. But, unlike the method described earlier, it can be done by means of triangulation. In the first place, the interface must be divided into regular parts so that some regular surface curvilinear coordinates (ξ, η) can be introduced in each of them. Each regular part can be divided into a set of curvilinear tetragons by cutting out along both coordinate lines. Each tetragon can be divided into two triangular elements. Thus the element of the approximating boundary is the plane triangle. This approximation guarantees finite values of the term Φ_m in Eqs. (4) and (46) because these terms do not depend on the curvature of the initial interface. The wavefield scattered by the piecewise-plane boundary is a superposition of fields scattered by the plane elements composing it.

For a sufficiently small step of approximation the contributions from the elements, containing reflection/transmission points and points of edge diffraction, are small enough and may be neglected. Then the field scattered by the individual plane element will be a superposition of six tip waves diverging from its tips. The total field can be represented by superposition of just tip waves alone:

$$f = \sum_{m=1}^M \Delta f_m, \quad \Delta f_m = \sum_{n=1}^3 \sum_{p=1}^2 f_{mnp}, \quad m \neq m_0, \quad n \neq n_0 \quad (61)$$

where m is the index of an element, n – the index of an edge of the m -th element, p – the index of a tip of the mn -th edge, m_0 and n_0 are the indices of the elements containing the corresponding reflection/transmission points and the points of edge diffraction. Every term Δf_m of this sum is strictly bounded. That is why the sum in Eq. (61) is limited within caustic zones caused by the curvature of the initial interface.

To justify the above approach we take Eq. (61) to the limit of the initial interface, simultaneously letting the step of approximation tend to zero and increasing the number of elements. In this case

$$\lim_{\substack{\max(\Delta S_m) \rightarrow 0 \\ M \rightarrow \infty}} \sum_{m=1}^M \Delta S_m = \iint_S dS, \\ f = \iint_S F(\xi, \eta) \exp[i\omega\tau(\xi, \eta)] d\xi d\eta, \quad (62)$$

$$F(\xi, \eta) = \Phi_0(\xi, \eta) [-i\omega/(4\pi)] \left[\sqrt{(\tau_\xi - \tau_0)/(\tau - \tau_\xi)} + \sqrt{(\tau_\eta - \tau_0)/(\tau - \tau_\eta)} \right] \left| \frac{\partial \tau}{\partial \xi} \cdot \frac{\partial \tau}{\partial \eta} \right| (\tau - \tau_0)^{-1},$$

where $\Phi_0(\xi, \eta)$ and τ_0 are the amplitude and the eikonal of the analytical reflection/transmission for point (ξ, η) , τ_ξ and τ_η are the eikonals of the analytical edge diffraction. Here the analytical edge diffraction stands for the diffraction by the rectilinear edge which is tangent to the coordinate line $\xi = \text{constant}$ or $\eta = \text{constant}$ for point (ξ, η) .

Equation (62) can be regarded as a new asymptotic formulation of the concept of secondary sources. According to this formulation the value of the field at any point of the medium will be the superposition of secondary waves with amplitude $F(\xi, \eta) d\xi d\eta$ diverging from each point of the initial interface. Unlike the Fresnel and Kirchhoff descriptions, the amplitudes of the secondary waves in Eq. (62) are strictly bounded on the surface of integration.

The asymptotic analysis in the neighbourhood of the isolated stationary point of the first order (ξ_1, η_1) gives the wavefield (62) in the form of a regular wave

$$f = \Phi(\xi_1, \eta_1) \exp[i\omega\tau(\xi_1, \eta_1)], \\ \Phi(\xi_1, \eta_1) = F(\xi_1, \eta_1)/(i\omega h^{1/2}), \\ h = \frac{\partial^2 \tau(\xi_1, \eta_1)}{\partial \xi^2} \cdot \frac{\partial^2 \tau(\xi_1, \eta_1)}{\partial \eta^2} - \left[\frac{\partial^2 \tau(\xi_1, \eta_1)}{\partial \xi \partial \eta} \right]^2. \quad (63)$$

If the wave velocity is constant this formula coincides with the formula of the ray method. This justifies the present approach.

We illustrate this approach for the model “syncline” (Fig. 22). Its geometrical shape is given by the equation

$$z = 1 + 0.1 \exp[-32(x - 1.15)^2 - 128(y - 1.15)^2].$$

Figure 23 shows the wavefield scattered by this interface. It is seen that the typical “loop” structure of the wavefield occurs.

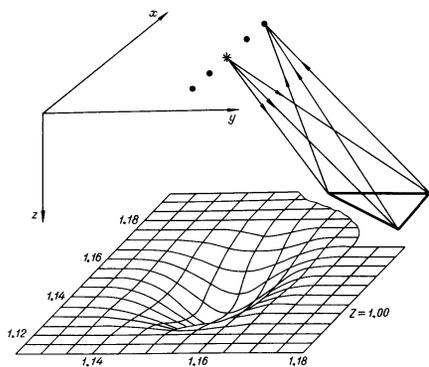


Fig. 22. Model of "syncline". Tip rays, scattering from an element of the interface, are shown. Further explanations in text

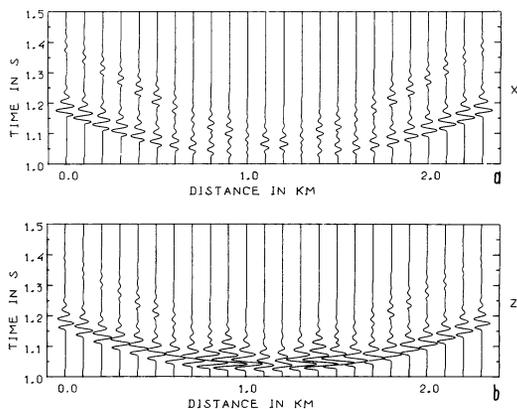


Fig. 23a and b. Theoretical seismograms for model of "syncline": **a** X-component (twice enlarged), **b** Z-component

Remarks

In conclusion, let us mention that there are many examples of mathematical modelling of wavefields in typical structures by the above method. There are theoretical seismograms for several types of pinch-out and low-amplitude faults (Klem-Musatov, 1980), interfaces of complex forms (Aizenberg and Klem-Musatov, 1980) and 3-D systems of intersecting faults (Klem-Musatov et al., 1982). The main principles of the generalization of the above approach for multiple diffraction are described in Klem-Musatov and Aizenberg (1984).

Acknowledgements. We wish to thank Prof. N.N. Puzyrev, who was the initiator of this study. Special thanks are due to our colleagues V.G. Chernyakov, who made the theoretical seismograms for model "pinch-out", and G.A. Klem-Musatova, who took a great part in writing computer programs and in seismic modelling. We are grateful to Professors A.S. Alekseyev, V.M. Babich, V. Červený and Dr. S.V. Goldin for useful discussions when this study was still in progress. We wish to thank Dr. P. Hubral, who invited us to submit the paper for publication.

Note added in proof. The secondary boundaries position in Fig. 11 is shown rough. Planes 111 and 121 must intersect profile II at the points $y=0.95$ km and $y=1.66$ km, respectively.

References

- Aizenberg, A.M.: Scattering of seismic waves by broken edge of a flat boundary. *Soviet Geol. and Geophys.* **23**, 74–82, 1982
- Aizenberg, A.M., Klem-Musatov, K.D.: Calculation of wave fields by the method of superposition of the edge waves. *Soviet Geol. and Geophys.* **21**, 79–94, 1980
- Babich, V.M., Alekseyev, A.S.: A ray method of computing wave front intensities. *Bull. Acad. Sci. USSR, Geophys. Ser.* **1**, 9–15, 1958
- Babich, V.M., Buldyrev, V.S.: Asymptotic methods in problems of diffraction of short waves (in Russian). Moscow: Nauka 1972
- Born, M., Wolf, E.: Principles of optics. Oxford: Pergamon Press 1968
- Červený, V.: Synthetic body wave seismograms for laterally varying layered structures by the Gaussian beam method. *Geophys. J.R. Astron. Soc.* **73**, 389–426, 1983
- Červený, V., Molotkov, I.A., Pšenčík, I.: Ray method in seismology. Praha: Universita Karlova 1977
- Claerbout, J.F.: Fundamentals of geophysical data processing. New York: McGraw-Hill 1976
- Clemmow, P.S., Senior, T.B.A.: A note on generalized Fresnel integral. *Proc. Camb. Phil. Soc.* **9**, 570–572, 1953
- Felsen, L.B., Marcuvitz, N.: Radiation and scattering of waves. New Jersey: Prentice-Hall, Englewood Cliffs 1973
- Fertig, J., Müller, G.: Approximate diffraction theory for transparent half-planes with application to seismic-wave diffraction at coal seams. *J. Geophys.* **46**, 349–367, 1979
- Fock, V.A.: Electromagnetic diffraction and propagation problems. New York: Pergamon Press 1965
- Hilterman, F.J.: Interpretative lessons from three-dimensional modelling. *Geophysics* **47**, 784–808, 1982
- Karal, F.C., Keller, J.B.: Elastic wave propagation in homogeneous and inhomogeneous media. *J. Acoust. Soc. Am.* **31**, 694–705, 1959
- Keller, J.B.: A geometrical theory of diffraction. *J. Opt. Soc. Am.* **52**, 116–130, 1962
- Kennett, B.L.N.: Reflection operator methods for elastic waves. I – Irregular interfaces and regions. *Wave Motion* 1984 (in press)
- Klem-Musatov, K.D.: The theory of edge waves and its applications in seismology (in Russian). Novosibirsk: Nauka 1980
- Klem-Musatov, K.D.: Scattering of waves of a point fracture of a diffracting edge. *Soviet Geol. and Geophys.* **22**, 118–127, 1981a
- Klem-Musatov, K.D.: Asymptotic formulae for the amplitude of wave scattered by a salient point on a diffracting edge. *Soviet Geol. and Geophys.* **22**, 108–114, 1981b
- Klem-Musatov, K.D., Aizenberg, A.M.: Ray method and the theory of edge waves. *Geophys. J.R. Astron. Soc.* **79**, 35–50, 1984
- Klem-Musatov, K.D., Aizenberg, A.M., Klem-Musatova, G.A.: An algorithm for mathematical modelling of three-dimensional diffraction fields. *Soviet Geol. and Geophys.* **23**, 116–121, 1982
- Popov, M.M.: A new method of computations of wave fields in high-frequency approximation. Report LOMI AN SSSR, E-I-81 Leningrad, 1981
- Trorey, A.W.: Diffraction for arbitrary source-receiver locations. *Geophysics* **42**, 1177–1182, 1977

Received August 28, 1984; Revised version February 28, 1985
Accepted March 20, 1985

The normal modes of a layered, incompressible Maxwell half-space

Detlef Wolf

Department of Physics, University of Toronto, Toronto, Ontario, Canada, M5S 1A7

Abstract. The theory describing the relaxation of an incompressible, layered Maxwell half-space is developed. The approach is based on the analytic solution of the associated elastic model and the subsequent application of the correspondence principle. The viscoelastic theory follows normal-mode theory, which allows the independent and exact determination of the relaxation-time and amplitude spectra for each mode of relaxation. The solution is tested by calculating the response of several models in the wavenumber and spatial domains. The examples are selected with regard to post-glacial adjustment in Fennoscandia and analyse effects caused by (a) varying lithospheric thickness, (b) adding an asthenosphere, (c) increasing lower-mantle viscosity, (d) permitting relaxation of the lower lithosphere or (e) introducing density contrasts at 400-km and 670-km depths.

Key words: Isostasy – Maxwell continuum – Normal modes

Introduction

The rheology of the solid Earth exhibits various forms of departure from perfect elasticity. Whereas, on a time-scale characteristic of seismic-wave propagation, anelastic effects are significant, on a much longer time-scale, the Earth's mantle is widely believed to support flow (creep). Recent experimental evidence of the steady-state creep properties of likely mantle materials suggest that the creep-rate limiting process is rate-dependent and the creep law therefore non-linear (e.g. Weertman and Weertman, 1975; Tullis, 1979). The inference of the creep properties of the Earth's mantle from such experiments is, however, beset with difficulties. One problem is that the chemical and mineralogical constitution of the mantle is poorly known. A severe limitation is also that laboratory creep experiments are necessarily carried out at creep rates which are orders of magnitude higher than the actual rates in the mantle. Large extrapolations from the experimental conditions are therefore necessary. In view of the uncertainties involved in this kind of reasoning it is therefore not clear whether a linear creep mechanism might not apply at the much lower creep rates characteristic of the Earth's mantle.

In the following we will pursue a pragmatic approach and will assume that the Earth's inelastic response is linear. More specifically, we will be concerned with the rheological model that is usually referred to as Maxwell continuum. This type of rheology has proved successful in interpreting the glacio-isostatic relaxation of the Earth's mantle (e.g. Cathles, 1975; Peltier and Andrews, 1976; Nakiboglu and Lambeck, 1982; Wu and Peltier, 1983) and the adjustments of the Earth's thermal lithosphere on a very long time-scale (e.g. Beaumont, 1978; Lambeck and Nakiboglu, 1980; Courtney, 1982). Clearly, our approach cannot prove that the material constituting the Earth's lithosphere or mantle does in fact respond linearly.

The theory describing the load-induced relaxation of a self-gravitating, compressible and pre-stressed Maxwell sphere has recently been summarized (Peltier, 1982). The gravitationally self-consistent model is recommended when analysing deformations associated with the Laurentide glaciation. In the investigation reported here we are, however, concerned with deformations of the Earth not exceeding the scale of the glacially induced depression in Fennoscandia. On this reduced scale, sphericity and self-gravitation are of subordinate importance (Wolf, 1984) and can therefore be neglected. Effects due to compressibility are analysed in Wolf (1985c). The results show that compressibility is significant only during the initial phases of relaxation. This special feature will therefore also be neglected. The significance of the pre-stress term in the equilibrium equations has been studied before (Wolf, 1985a, b) and needs no further discussion.

With these simplifications, the Earth model is reduced to an incompressible, pre-stressed Maxwell half-space. Special solutions for uniform or two-layer Maxwell models (e.g. Nakiboglu and Lambeck, 1982; Wolf, 1984, 1985b) are of some theoretical interest. For data interpretation the availability of a more versatile model is, however, of definite advantage. In the following we will therefore be concerned with the multi-layer Maxwell half-space model. The associated elastic model has recently been re-analysed by Ward (1984). His solution for the deformation also includes the modifications caused by an external gravity field. Previously, such effects, which, for incompressibility, virtually reduce to pre-stress advection, had usually been neglected in elastic half-space approximations (e.g. Kuo, 1969).

In the following section we will develop the theory governing the deformation of an incompressible, pre-stressed, layered elastic half-space. The solution for the associated Maxwell continuum is obtained by using the correspondence principle. The viscoelastic theory closely follows the normal-mode formulation developed by Peltier (1985). This method allows the independent and exact determination of the relaxation times and amplitudes of all normal modes characteristic of any specific model considered. The method is therefore distinctly superior to the approximate collocation method developed in Peltier (1976) and employed by Courtney (1982) in his study of the evolution of sedimentary basins.

After that, several numerical examples will be discussed. They are intended to illustrate the principal effects caused by (a) variations in lithospheric thickness, (b) the insertion of a low-viscosity channel (asthenosphere), (c) changes in lower-mantle viscosity, (d) the relaxation of the lower portion of the thermal lithosphere or (e) the presence of density discontinuities in the upper mantle. The relaxation of the individual models will be illustrated in various diagrams which show amplitude and relaxation-time spectra (wavenumber domain) or vertical surface deflections (spatial domain). This presentation constitutes a systematic and complete compilation of the response characteristics of the main Earth models under discussion, which has not been available before. The numerical values of the model parameters chosen are of relevance to the interpretation of glacio-isostatic adjustment in Fennoscandia.

Theory

We wish to derive the solution describing the deformation of a pre-stressed, multi-layer elastic half-space subject to an axisymmetric load. The differential equations governing this problem may be written in matrix form. Assuming incompressibility, we obtain the first-order system (Wolf, 1985c)

$$\begin{bmatrix} D & -k & -\frac{1}{\mu} & 0 \\ k & D & 0 & 0 \\ -4\mu k^2 & 0 & D & -k \\ 0 & 0 & k & D \end{bmatrix} \begin{bmatrix} \hat{u}_1 \\ \hat{w}_0 \\ \hat{\sigma}_{rz1} \\ \hat{\sigma}_{zz0} \end{bmatrix} = 0, \quad (1)$$

where $D=d/dz$. Symbols u , w , σ_{rz} and σ_{zz} denote the radial and vertical displacement components and the appropriate components of the total perturbation stress $\sigma_{ij} = \sigma_{ij}^{(e)} + \rho g w \delta_{ij}$, where $\sigma_{ij}^{(e)}$ is the *elastic* perturbation stress (Wolf, 1985a). Parameter μ is Lamé's second constant (shear modulus) and ρ denotes the density of the continuum. The external gravity field g is assumed to be directed in the positive z -direction. A circumflex denotes Hankel transformation of zeroth or first order, as indicated by the subscript, with k being the Hankel-transform variable or wavenumber. Equation (1) is formally equivalent to the first-order system appropriate to a non-gravitating elastic continuum, for which the general solution is well-known (e.g. Farrell, 1972; Lanczono, 1982, pp. 120–125). In terms of four

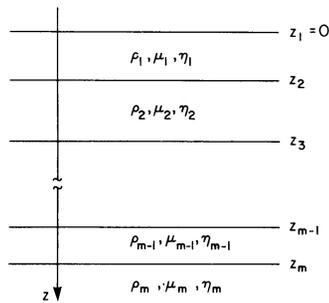


Fig. 1. Multi-layer Maxwell half-space

arbitrary constants, the solution is

$$\begin{bmatrix} \hat{u}_1 \\ \hat{w}_0 \\ \hat{\sigma}_{rz1} \\ \hat{\sigma}_{zz0} \end{bmatrix} = \begin{bmatrix} \pm 1 \\ 1 \\ -2\mu k \\ \pm 2\mu k \end{bmatrix} A_{1,2} \exp(\pm kz), \quad (2a)$$

$$\begin{bmatrix} \hat{u}_1 \\ \hat{w}_0 \\ \hat{\sigma}_{rz1} \\ \hat{\sigma}_{zz0} \end{bmatrix} = \begin{bmatrix} kz \pm 1 \\ \mp kz \\ \pm 2\mu k^2 z + 2\mu k \\ -2\mu k^2 z \end{bmatrix} B_{1,2} \exp(\pm kz). \quad (2b)$$

We first consider a homogeneous elastic half-space extending between $z_m < z < \infty$ (Fig. 1). Then A_1 and B_1 in Eq. (2) vanish. At $z = z_m$ the solution takes the form

$$\mathbf{Y}(z_m) = \mathbf{L}^{(m)} \mathbf{A}, \quad (3)$$

where column matrices

$$\mathbf{Y}(z) = [\hat{u}_1(z), \hat{w}_0(z), \hat{\sigma}_{rz1}(z), \hat{\sigma}_{zz0}(z)]^T, \quad (4)$$

$$\mathbf{A} = [C_1, C_2, 0, 0]^T \quad (5)$$

have been introduced. The symbols C_1 and C_2 designate new constants. Explicit expressions for the elements of the half-space propagator $\mathbf{L}^{(m)}$ are given in Appendix A, Eq. (26). In a similar fashion, the elements at the top $z = z_l$ of the l -th layer (Fig. 1) can be expressed in terms of the quantities at the base $z = z_{l+1}$. Then the general solution, Eq. (2), reduces to

$$\mathbf{Y}(z_l) = \mathbf{L}^{(l)} \mathbf{Y}(z_{l+1}). \quad (6)$$

The elements of the layer propagator $\mathbf{L}^{(l)}$ are given in Appendix A, Eq. (27).

Since *elastic* stress components are continuous across interfaces, it is convenient to modify the equations slightly. If we observe that the elastic stress $\sigma_{ij}^{(e)}$ is connected with the total stress σ_{ij} by $\sigma_{ij} = \sigma_{ij}^{(e)} + \rho g w \delta_{ij}$ (Wolf, 1985a), Eqs. (3) and (6) can be re-written in terms of the elastic stress components. With

$$\mathbf{Y}^{(e)}(z) = [\hat{u}_1(z), \hat{w}_0(z), \hat{\sigma}_{rz1}(z), \hat{\sigma}_{zz0}^{(e)}(z)]^T, \quad (7)$$

we obtain

$$\mathbf{Y}^{(e)}(z_m) = \mathbf{P}^{(m)} \mathbf{A} \quad (8)$$

for the half-space and

$$\mathbf{Y}^{(e)}(z_l) = \mathbf{P}^{(l)} \mathbf{Y}^{(e)}(z_{l+1}) \quad (9)$$

for the l -th layer. The relation between $\mathbf{L}^{(m)}$ and $\mathbf{P}^{(m)}$ is given in Eq. (28), that between $\mathbf{L}^{(l)}$ and $\mathbf{P}^{(l)}$ in Eq. (29), of Appendix A, respectively.

With the requirement that displacement components and elastic stress components be continuous across any interface, the solution for an arbitrary m -layer elastic continuum is now readily constructed from Eqs. (8) and (9). If we number the layers from the top to the base and observe that the m -th basal layer extends towards $z \rightarrow \infty$ (Fig. 1), the field quantities at the upper surface $z = z_l$ of the l -th layer may be represented as

$$\mathbf{Y}^{(e)}(z_l) = \mathbf{P}(z_l) \mathbf{A}, \quad (10)$$

where

$$\mathbf{P}(z_l) = \sum_{L=l}^m \mathbf{P}^{(L)}. \quad (11)$$

Equation (10) can be expressed as a combination of two linearly independent solutions. Remembering the definition of the column matrix \mathbf{A} , Eq. (5), and introducing

$$\mathbf{P}_1(z) = [P_{11}(z), P_{21}(z), P_{31}(z), P_{41}(z)]^T, \quad (12a)$$

$$\mathbf{P}_2(z) = [P_{12}(z), P_{22}(z), P_{32}(z), P_{42}(z)]^T, \quad (12b)$$

Eq. (10) becomes

$$\mathbf{Y}^{(e)}(z_l) = C_1 \mathbf{P}_1(z_l) + C_2 \mathbf{P}_2(z_l). \quad (13)$$

This constitutes the general solution for the field quantities at the top of the l -th layer.

The arbitrary constants C_1 and C_2 can be determined from the boundary conditions. As usual in geophysical applications, the stress components $\hat{\sigma}_{rz1}$ and $\hat{\sigma}_{zz0}$ are assumed to be known at the top $z = z_1 = 0$ of the upper layer. If the sub-matrix

$$\mathbf{M} = \begin{bmatrix} P_{31}(0) & P_{32}(0) \\ P_{41}(0) & P_{42}(0) \end{bmatrix} \quad (14)$$

is introduced and the definitions

$$\mathbf{B} = [\hat{\sigma}_{rz1}(0), \hat{\sigma}_{zz0}(0)]^T, \quad (15)$$

$$\mathbf{C} = [C_1, C_2]^T \quad (16)$$

are observed, the boundary conditions \mathbf{B} take the form

$$\mathbf{B} = \mathbf{M} \mathbf{C} \quad (17a)$$

or, after inversion,

$$\mathbf{C} = \mathbf{M}^{-1} \mathbf{B}. \quad (17b)$$

The inverse matrix \mathbf{M}^{-1} is given by

$$\mathbf{M}^{-1} = (\det \mathbf{M})^{-1} \mathbf{M}^A, \quad (18)$$

where \mathbf{M}^A is the adjugate matrix associated with \mathbf{M} . Equation (13) and the boundary conditions given by Eq. (17) completely determine the solution.

According to the correspondence principle (e.g. Cathles, 1975, pp. 25–29), Eqs. (13) and (17) can be interpreted as the Laplace-transformed quasi-static so-

lution appropriate to the associated Maxwell continuum subject to impulsive boundary conditions $\mathbf{B} \delta(t)$. In the time domain, the solution takes the form

$$\mathbf{Y}^{(ve)}(t, z_l) = \mathbf{Y}^{(e)}(z_l) \delta(t) + \sum_k \mathbf{Y}^{(v,k)}(z_l) s^{(k)} \exp(-s^{(k)} t). \quad (19)$$

The eigenvalues $s^{(k)}$ and associated eigenfunctions (normal modes) $\mathbf{Y}^{(v,k)}$ are complicated functions of the model parameters and boundary conditions. For simple Maxwell models they can be explicitly calculated (e.g. Wolf, 1984). For the multi-layer half-space considered here a more general approach is recommended. We employ the normal-mode method developed by Peltier (1985). The details of the calculations are outlined in Appendix B.

In the following we will be concerned with the vertical deflection $\hat{w}_0(t, z_l)$ subject to the boundary condition $\mathbf{B} = [0, -\hat{q}_0]^T$. Since the solution is linear in the load pressure \hat{q}_0 (Appendix B), we write

$$\hat{w}_0(t, z_l) = T^{(ve)}(t, z_l) \hat{q}_0, \quad (20)$$

where

$$T^{(ve)}(t, z_l) = T^{(e)}(z_l) \delta(t) + \sum_k T^{(v,k)}(z_l) s^{(k)} \exp(-s^{(k)} t) \quad (21)$$

is the viscoelastic transfer function for impulsive forcing. For a Heaviside unloading event $\hat{q}_0(k)[1 - H(t)]$ we obtain, upon convolution,

$$T^{(ve)}(t, z_l) = \begin{cases} T^{(e)} + \sum_k T^{(v,k)}, & t < 0 \\ \sum_k T^{(v,k)} \exp(-s^{(k)} t), & t > 0 \end{cases} \quad (22)$$

The Hankel transform of the gravity anomaly Δg at $z = 0$ associated with the deformation of a stack of m layers at $z > 0$ can be approximated by (e.g. Parker, 1972)

$$\Delta \hat{g}_0(t, 0) = G^{(ve)}(t) \hat{q}_0, \quad (23)$$

where

$$G^{(ve)}(t) = -2\pi\gamma \sum_{L=1}^m [(\rho_L - \rho_{L-1}) \exp(-kz_L) T^{(ve)}(t, z_L)] \quad (24)$$

and γ is the gravitational constant. For $L=1$, ρ_{L-1} denotes the density of the material (usually air or water) superimposed on the layered half-space.

Numerical examples and discussion

The relaxation of the Earth's surface in response to loads comparable in scale to the Fennoscandian ice-sheet has been widely assumed to be dominated by the viscosity of the upper mantle (e.g. Cathles, 1975, pp. 173–196). The response may, however, be modified by (a) the lithosphere, (b) the presence of a low-viscosity asthenosphere, (c) the viscosity stratification of the low-

Table 1. Parameters of Earth models employed

Layer	h (km)	ρ (kg m ⁻³)	μ (N m ⁻²)	η (Pa s)
Earth Model A.1				
1	100.0	3,380	0.67×10^{11}	∞
2	100.0	3,380	1.45×10^{11}	η_2
3	∞	3,380	1.45×10^{11}	1.0×10^{21}
Earth Model L.1				
1	h_1	3,380	0.67×10^{11}	∞
2	∞	3,380	1.45×10^{11}	1.0×10^{21}
Earth Model L.2				
1	75.1	3,380	0.67×10^{11}	∞
2	8.7	3,380	0.67×10^{11}	1.0×10^{25}
3	10.3	3,380	0.67×10^{11}	1.0×10^{23}
4	∞	3,380	1.45×10^{11}	1.0×10^{21}
Earth Model L.3				
1	51.8	3,380	0.67×10^{11}	∞
2	14.1	3,380	0.67×10^{11}	1.0×10^{25}
3	20.3	3,380	0.67×10^{11}	1.0×10^{23}
4	∞	3,380	1.45×10^{11}	1.0×10^{21}
Earth Model M.1				
1	100.0	3,380	0.67×10^{11}	∞
2	570.0	3,380	1.45×10^{11}	1.0×10^{21}
3	∞	3,380	1.45×10^{11}	η_3
Earth Model M.2				
1	100.0	3,380	0.67×10^{11}	∞
2	570.0	3,380	1.45×10^{11}	1.0×10^{21}
3	∞	3,770	1.45×10^{11}	1.0×10^{21}
Earth Model M.3				
1	100.0	3,380	0.67×10^{11}	∞
2	300.0	3,380	1.45×10^{11}	1.0×10^{21}
3	270.0	3,560	1.45×10^{11}	1.0×10^{21}
4	∞	3,950	1.45×10^{11}	1.0×10^{21}
Earth Model S				
1	100.0	3,380	0.67×10^{11}	∞
2	∞	3,380	1.45×10^{11}	1.0×10^{21}

er mantle, (d) relaxation near the base of the lithosphere or (e) density discontinuities in the upper mantle.

In this section we will be discussing the characteristic signatures produced by each of these special features successively. In order to have some reference, we proceed from a "standard" model, which is called Earth Model S. It is composed of a 100-km-thick elastic lithosphere overlying a uniformly viscous mantle with a dynamic viscosity of $\eta = 10^{21}$ Pa s. The model parameters are listed in Table 1.

Figure 2a shows the viscous transfer functions $T^{(v,k)}(0)$ associated with the two modes of Earth Model S as functions of angular order n . The latter quantity is formally defined by $n = ka$, with a the Earth's radius. The bimodal character of the viscous response of this model was discussed previously. It is characterized by a major mantle branch M0, for which the shear energy has a maximum in the interior of the mantle, and a subordinate lithospheric branch L0, for which the shear energy is concentrated immediately below the base of the lithosphere (Wu and Peltier, 1982; Wolf, 1984). One of the effects of the lithosphere

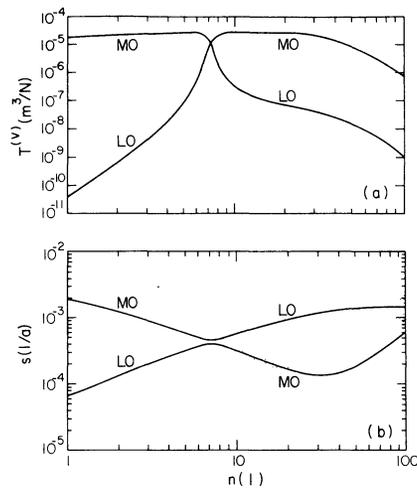


Fig. 2a, b. **a** Viscous transfer function $T^{(v,k)}$ and **b** inverse relaxation time $s^{(k)}$ as function of angular order n for Earth Model S; symbols denote relaxation mode

is that it accelerates the decay of short-wavelength deformations. This is evident from Fig. 2b, which shows that the relaxation time $1/s^{(k)}$ of the M0 mode is substantially shortened at $n > 30$.

Most interestingly, such a behaviour is displayed by a relaxation-time estimate based on a spectral decomposition of Pleistocene strandline data from Fennoscandia (McConnell, 1968). Walcott (1980) has, however, pointed out that the short-wavelength part of McConnell's spectrum could also reflect the effects of structural hinge zones on the tilt of the Pleistocene strandlines. Some caution as to the significance of McConnell's estimate is therefore indicated.

In Fig. 2b and the following relaxation-time diagrams the estimate for Fennoscandia is represented as a stippled band. McConnell (1968) based his decomposition on the assumption that the Fennoscandian uplift is governed by a single mode of relaxation. Although this is not strictly correct, his estimate can be compared with the theoretical relaxation-time spectrum of the fundamental mode M0, provided that the later dominates the theoretical response on the time-scale considered. This assumption holds for Earth Model S.

We first investigate the modifications introduced by varying lithospheric thickness (Table 1, Earth Model L.1). In Fig. 3a and b the thickness has been increased to 150 km, whereas in Fig. 3c and d the thickness is 200 km. Since the relative strength of the L0 mode remains insignificant, the relaxation is, as in Earth Model S, governed by the M0 mode. At higher wavenumbers the increase in lithospheric thickness causes a decrease in strength of the M0 mode and a shortening of its relaxation time. McConnell's (1968) relaxation-time estimate is best satisfied by a lithosphere of 100-km thickness. This implies that the wavenumber at which relaxation time reaches a maximum is used as the primary criterion for the goodness of the fit.

Figure 4 illustrates the modifications of the basic response of Earth Model S produced by inserting a 100-km-thick asthenosphere below the lithosphere (Table 1, Earth Model A.1). In Fig. 4a and b the asthenosphere has a viscosity of 5×10^{19} Pa s; in Fig. 4c and d

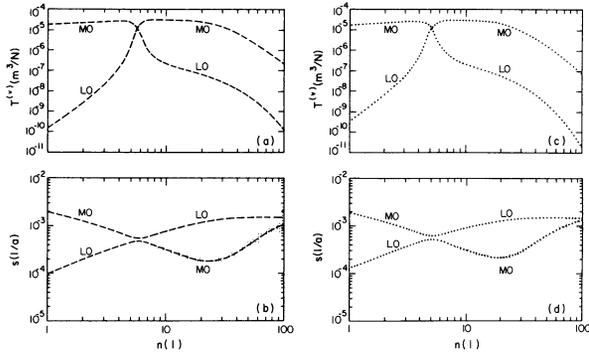


Fig. 3a–d. Same as Fig. 2 except for Earth Model L.1 with $h_1 = 150$ km (dashed) or $h_1 = 200$ km (dotted)

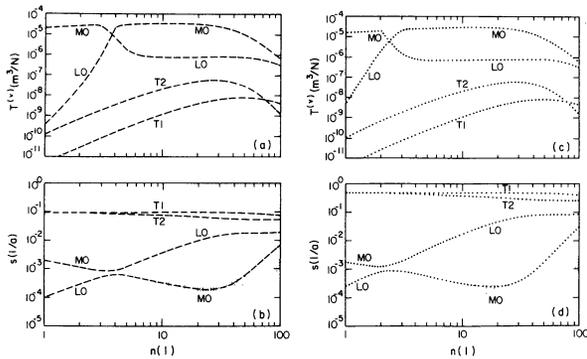


Fig. 4a–d. Same as Fig. 2 except for Earth Model A.1 with $\eta_2 = 5 \times 10^{19}$ Pa s (dashed) or $\eta_2 = 1 \times 10^{19}$ Pa s (dotted)

the viscosity is 1×10^{19} Pa s. The presence of an asthenosphere adds two modes of short relaxation time. Following Peltier (1976), we call such modes transition modes (“T modes”). Figure 4a and c shows that they are only poorly excited. The main modification introduced by the asthenosphere is therefore a shortening of the relaxation time of the fundamental mode M0 at $n > 30$ (see also McConnell, 1968). A comparison between Figs. 3 and 4 further suggests that there may exist a certain trade-off between lithospheric thickness and asthenospheric thickness so that the response is nearly unchanged, provided that the total thickness of lithosphere and asthenosphere remains approximately constant. To what extent this statement is correct will be discussed later, when the deformation in the spatial domain is discussed.

For completeness, the effects caused by increasing the viscosity of the lower mantle are also considered (Table 1, Earth Model M.1), although this model is physically equivalent to Earth Model A.1. In Fig. 5a and b the viscosity below a depth of 670 km is 2×10^{21} Pa s; in Fig. 5c and d it is 5×10^{21} Pa s. The main modification, compared with Earth Model S, are the longer relaxation times of the M0 mode at small wavenumbers (Fig. 5b and d). This effect is well-known (e.g. McConnell, 1965). The transition modes, which are not observed for Newtonian viscous continua, are again barely excited and can be neglected when calculating the response in the spatial domain.

The model of a perfectly elastic or *mechanical*

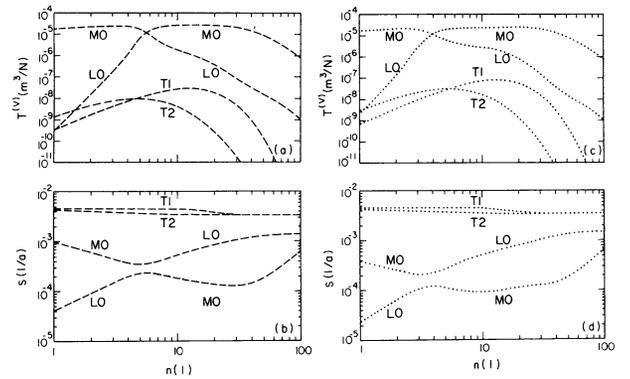


Fig. 5a–d. Same as Fig. 2 except for Earth Model M.1 with $\eta_3 = 2 \times 10^{21}$ Pa s (dashed) or $\eta_3 = 5 \times 10^{21}$ Pa s (dotted)

lithosphere employed so far is an idealization whose justification is governed by the time-scale of the external forcing. Whereas the model is clearly inadequate to sedimentary loads, it has generally been used when modeling glacio-isostatic adjustment.

In order to analyse the influence of relaxation near the base of the lithosphere, we develop a more realistic model. It is based on the fact that creep is a temperature-activated process. Then, from Appendix C, Eq. (47b), we have for the viscosity η of the material considered

$$\eta(T) = \eta_0 \exp[Q(1/T - 1/T_0)/R], \quad (25)$$

where $\eta_0 = \eta(T_0)$. In the present context Q denotes the activation energy of the lithospheric material, T is the absolute temperature in the lithosphere and R the gas constant. If a characteristic geotherm $T = T(z)$ is substituted, $\eta = \eta(z)$ is obtained. This method of estimating the viscosity-depth distribution in the lithosphere was previously employed by Courtney (1982), who studied the response of the thermal lithosphere in connection with the evolution of sedimentary basins.

Geologically, a large portion of Fennoscandia is a Precambrian shield. Therefore, the “old” continental geotherms discussed by Sclater et al. (1980) may serve as a guide-line when modeling subsurface temperatures in this region. For our purposes their geotherms may be approximated by a linear function which passes through the points $T = 0^\circ\text{C}$, $z = 0$ km and $T = 1,000^\circ\text{C}$, $z = 100$ km.

Recent laboratory estimates of the activation energy for materials believed to be typical of the Earth’s crust or mantle have, for example, been compiled by Tullis (1979). Most of the samples have activation energies which range between 200 and 500 kJ mol⁻¹. The lower values are usually associated with more silicic minerals, whereas the higher values are appropriate to olivine.

Figure 6 shows viscosity-depth distributions for different values of Q . If we define the *thermal* lithosphere as that part of the Earth’s outer shell where $\eta > 10^{21}$ Pa s, its thickness is 100 km in the models considered. The underlying mantle has a uniform viscosity of 10^{21} Pa s. If $Q \rightarrow \infty$ in the lithosphere, $\eta \rightarrow \infty$ and Earth Model S is recovered. For $Q = 500$ kJ mol⁻¹, the viscosity-depth distribution is less abrupt; if $Q = 200$

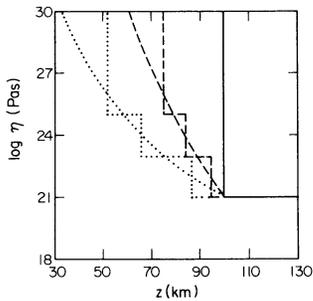


Fig. 6. Viscosity η as function of depth z for $Q \rightarrow \infty$ (solid), $Q = 500 \text{ kJ mol}^{-1}$ (dashed) or $Q = 200 \text{ kJ mol}^{-1}$ (dotted); exponential distributions are approximated by three uniform layers

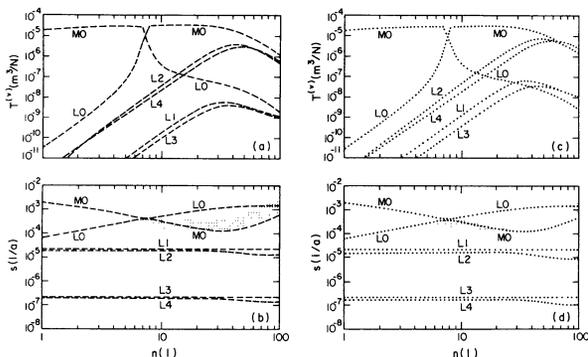


Fig. 7a-d. Same as Fig. 2 except for Earth Model L.2 (dashed) or Earth Model L.3 (dotted)

kJ mol^{-1} , the thickness of the region of intermediate viscosity near the base of the thermal lithosphere is further increased. Clearly, the value of Q is quite significant. Since silicic material is believed to be concentrated in the upper crust, $Q = 500 \text{ kJ mol}^{-1}$ might be more realistic for the lower lithosphere. For computational ease, the exponential distributions are approximated by three layers of constant viscosity (Fig. 6; Table 1, Earth Models L.2 and L.3). In the upper layer we have, as before, $\eta \rightarrow \infty$; in the middle layer $\eta = 10^{25} \text{ Pa s}$, whereas in the lower layer $\eta = 10^{23} \text{ Pa s}$.

The response of the two models is displayed in Fig. 7a and b (Earth Model L.2) and in Fig. 7c and d (Earth Model L.3). A comparison with Fig. 2 shows that the relaxation-time and amplitude spectra of the M0 and L0 modes remain virtually unchanged. Each layer, however, causes a pair of new modes. Since they are related to the rheological stratification of the lithosphere, they are of the L-type. We have chosen to number them in order of increasing relaxation time. If $n > 40$, the amplitudes of the L2 and L4 modes are comparable to the amplitude of the M0 mode, whereas their relaxation times are much longer. Due to the approximations involved in the previously mentioned collocation method, Courtney (1982) could not determine the strength of the weaker lithospheric modes. As shown here, their contribution is, however, insignificant.

The main modification of the relaxation pattern produced by internal, non-adiabatic density contrasts is that each such contrast is associated with a characteristic relaxation mode (Parsons, 1972; Peltier, 1976; Wu

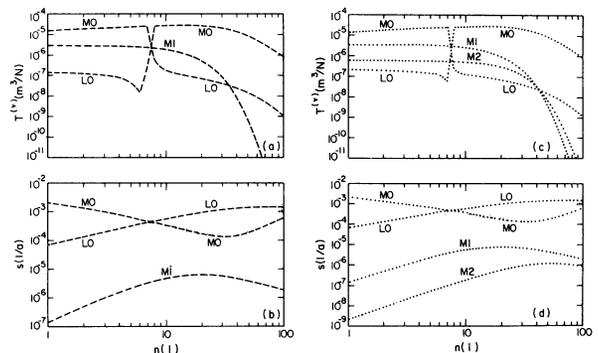


Fig. 8a-d. Same as Fig. 2 except for Earth Model M.2 (dashed) or Earth Model M.3 (dotted)

and Peltier, 1982). Since the density discontinuities in the upper mantle are about one order of magnitude smaller than the discontinuity at the Earth's surface, the buoyancy effects are also smaller. Thus, the modes associated with non-adiabatic density stratification in the interior are expected to decay much slower than the M0 mode.

In the following, density contrasts of 180 kg m^{-3} and 390 kg m^{-3} will be introduced at 400-km and 670-km depth, respectively (Table 1, Earth Models M.2 and M.3). This choice is consistent with the PREM model of the elastic structure of the Earth (Dziewonski and Anderson, 1981). The superimposed gradual increase in density with depth in the upper mantle is believed to be primarily due to adiabatic compression. It therefore does not interfere with our incompressible analysis.

In Earth Model M.2 only the larger internal density contrast at 670-km depth has been included (Fig. 8a and b). The M1 mode associated with this discontinuity has a relaxation time of the order of 1 Ma. For $n < 10$ it carries about 10% of the strength of the M0 mode. Deformations of shorter wavelength do not sample deeply enough to excite the M1 mode appreciably. In Earth Model M.3 the discontinuity at 400 km has been added (Fig. 8c and d). This causes a second internal mode M2 which, however, decays exceedingly slowly and carries even less energy than the M1 mode.

In the following, a systematic comparison of the response characteristics in the spatial domain of the different Earth models will be presented. Previously, only isolated cases were discussed in the literature. The examples discussed here apply to a square-edged disk load with a radius of $R = 600 \text{ km}$. This approximates the scale of the Fennoscandian ice-sheet during the period of stagnation between 10 and 12 ka B.P. (e.g. Cathles, 1975, p. 127). The thickness of the disk is $h_0 = 2 \text{ km}$. This corresponds to the average thickness of an ice-sheet of parabolic cross-section and 3-km axial thickness. The load density is $1,000 \text{ kg m}^{-3}$.

In calculating the deflection curves shown in Figs. 10 and 11, a Heaviside unloading event of the form $1 - H(t)$ has been assumed. The load is therefore assumed to have acted for an infinite period of time before it is instantaneously removed at $t = 0$ (Fig. 9a). This simulates the rapid disintegration of the Fennoscandian ice-sheet following the period of stagnation. The vertical surface deflection is obtained after taking the inverse Hankel transform (e.g. Wolf, 1985c).

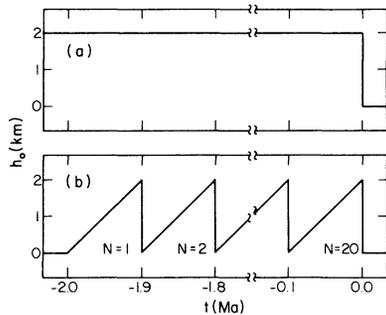


Fig. 9 a, b. **a** Heaviside unloading event applicable to Figs. 10 and 11 and **b** saw-tooth loading history applicable to Figs. 12-17

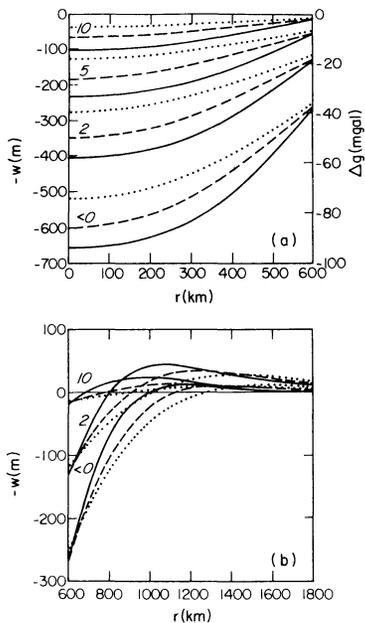


Fig. 10 a, b. Vertical surface displacement w as function of distance r from load axis for **a** central or **b** peripheral region and for several times (in units of ka) after load removal; deflection curves refer to Earth Model S (solid) or to Earth Model L.1 with $h_1 = 150$ km (dashed) or $h_1 = 200$ km (dotted); free-air gravity anomaly Δg is indicated for central region; results apply to Heaviside unloading event of Fig. 9a

In Fig. 10 the surface deflection for Earth Model S is compared with that for Earth Model L.1. The influence of lithospheric thickness on the magnitude of the deflection is substantial. This is a simple consequence of the filtering effect of the lithosphere, which, with increasing thickness, becomes less “transparent” for deformations of shorter wavelength (Fig. 3a and c). As the only density jump is at the Earth’s surface, the gravity anomaly can be determined directly from $\Delta g(t, 0) = 2\pi\gamma\rho_1 w(t, 0)$. Since $\rho_1 = 3,380 \text{ kg m}^{-3}$, a downward deflection of 70 m almost exactly corresponds to a free-air gravity anomaly of -10 mgal . The peak anomaly associated with the present ($t = 10 \text{ ka}$) degree of disequilibrium in Fennoscandia is probably around -15 mgal (Balling, 1980). From Fig. 10a it is thus obvious that, on the basis of the elementary model employed, lithospheric thicknesses in excess of 150 km are difficult to reconcile with the gravity data.

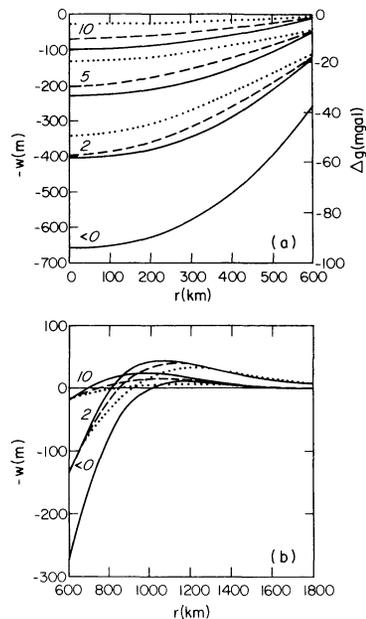


Fig. 11 a, b. Same as Fig. 10 except that Earth Model A.1 with $\eta_2 = 5 \times 10^{19} \text{ Pas}$ (dashed) or $\eta_2 = 1 \times 10^{19} \text{ Pas}$ (dotted) is compared with Earth Model S (solid)

The signature produced by an asthenosphere is illustrated in Fig. 11, which compares Earth Models S and A.1. As mentioned previously, the modification of the relaxation-time spectrum of the M0 mode produced by an asthenosphere is similar to the effect produced by an increase in the thickness of the lithosphere. Due to differences in the amplitude spectra, the response characteristics in the spatial domain are, nevertheless, distinct. As opposed to the modifications caused by increasing lithospheric thickness, there is, in particular, no effect of the asthenosphere on the initial deflection. This is because the equilibrium deflection at $t < 0$ must necessarily be independent of the viscosity stratification of the mantle and is only dependent on the thickness and the elastic structure of the lithosphere.

In the peripheral region the presence of an asthenosphere reduces the inward shift of the zero-crossing associated with the relaxation of Earth Model S considerably (Fig. 11b). The influence of the asthenosphere is, however, not strong enough to counteract the initially sympathetic uplift of this region for Earth Model S effectively (see also Cathles, 1975, pp. 184-191).

For a demonstration of the response characteristics produced by modifications (c), (d) and (e) (see beginning of this section) it is necessary to employ a higher approximation for the loading history. This is because the associated models support slowly decaying modes. Then the relaxation is no longer governed exclusively by the details of the deglaciation event but also markedly influenced by the long-term accumulation and ablation history of the ice-sheet.

Oxygen-isotope data from deep-sea sedimentary cores suggest that the recent ice age started approximately 2 Ma B.P. and consisted of individual glaciations of about 100-ka duration [see Imbrie and Imbrie (1979) for a summary]. A reasonable approximation to the complete sequence is the saw-tooth loading history shown in Fig. 9b. In the glaciation model we have

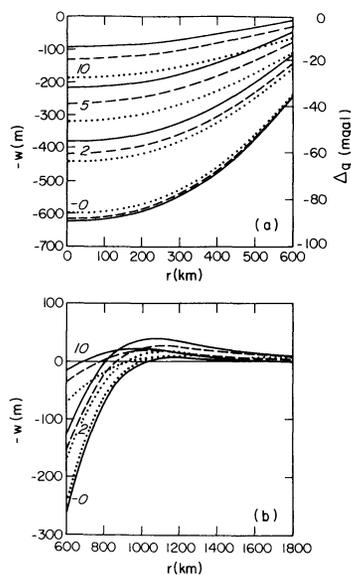


Fig. 12a, b. Vertical surface displacement w as function of distance r from load axis for **a** central or **b** peripheral region and for several times (in units of ka) after load removal; deflection curves refer to Earth Model S (solid) or to Earth Model M.1 with $\eta_3 = 2 \times 10^{21}$ Pa s (dashed) or $\eta_3 = 5 \times 10^{21}$ Pa s (dotted); free-air gravity anomaly Δg is indicated for central region; results apply to saw-tooth loading history of Fig. 9b

adopted exactly 20 cycles, where each cycle is assumed to have lasted for 100 ka. During each cycle the thickness of the straight-edged disk builds up linearly, whereas its radius is assumed to remain constant. This is a very good approximation to a more complicated model which incorporates effects due to variations of the radius (Wu and Peltier, 1983).

In order to facilitate comparisons with Figs. 10 and 11 the load radius has been kept at 600 km in the calculations underlying the following figures. It must, however, be pointed out that, at least at the time of the last glacial maximum about 18 ka B.P., the Fennoscandian ice-sheet extended south to North Germany. A somewhat larger radius might therefore be more appropriate.

Figure 12 shows the relaxation following the final cycle in the glaciation sequence of Fig. 9b for Earth Model M.1. As expected, the initial displacement is slightly reduced compared with Earth Model S. During relaxation this situation is, however, reversed, and Earth Model M.1 is characterized by considerably enhanced residual deformation. Since the significance of the lower mantle for post-glacial uplift in Fennoscandia has not always been appreciated in the past (e.g. Parsons, 1972), the demonstrated sensitivity of the response to the viscosity structure of the lower mantle should help further clarify this aspect. Figure 12b provides additional information and demonstrates that the inward movement of the zero-crossing of the displacement curve for Earth Model S is suppressed almost completely if the lower-mantle viscosity is increased by a factor of five. This characteristic feature was discussed previously with respect to relative-sea-level data from the North American east coast (Peltier, 1974).

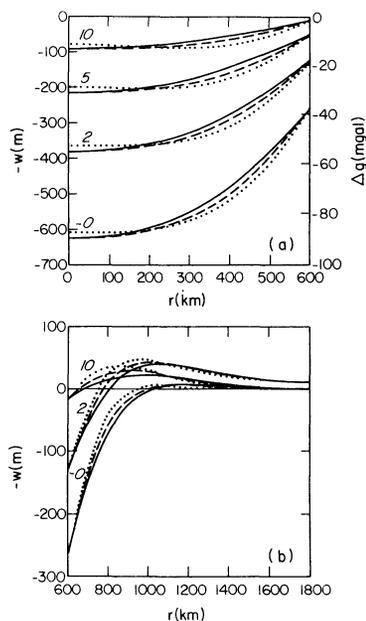


Fig. 13a, b. Same as Fig. 12 except that Earth Model L.2 (dashed) or Earth Model L.3 (dotted) is compared with Earth Model S (solid)

In Fig. 13 the modifications introduced by permitting the relaxation of the lower portions of the thermal lithosphere are shown. As expected, the relaxation after the final loading cycle resembles that associated with a perfectly elastic but thinner lithosphere. In our example the magnitude and tilt of the deflection in the marginal and peripheral regions of the load are very sensitive to the relaxation of the thermal lithosphere. This is consistent with the spectral-response characteristics (Fig. 7), according to which only shorter-wavelength deformations are markedly affected by the L1 and L2 modes. It is, however, important to realize that the details of the modifications of the response in the spatial domain depend strongly on the thickness of the thermal lithosphere and the lateral scale of the load adopted. Suitable combinations of both parameters may, therefore result in effects quite different from those described here.

In order to address the influence of the slowly decaying M1 and M2 modes associated with the 670-km and 400-km density discontinuities, respectively, we again employ the saw-tooth loading history (Fig. 9b). Figure 14a shows that the axial surface displacement associated with Earth Models M.2 and M.3 builds up gradually and, after about 10 cycles, has almost become stationary. The decay of the deflections at 400-km depth (Fig. 14b) and 670-km depth (Fig. 14c) with increasing number of load cycles can be understood from the fact that, physically, the M1 and M2 modes are buoyancy effects caused by disequilibrium at the interfaces. The internal modes will therefore “work” towards restoring equilibrium at the interfaces, which, in our case, corresponds to a plane interface.

The decay of the axial deflection after the final loading cycle is shown in Fig. 15. On the time-scale considered, the modifications introduced by the slowly decaying M1 and M2 modes just start to become visible in the surface deflection at $t = 10$ ka (Fig. 15a). At

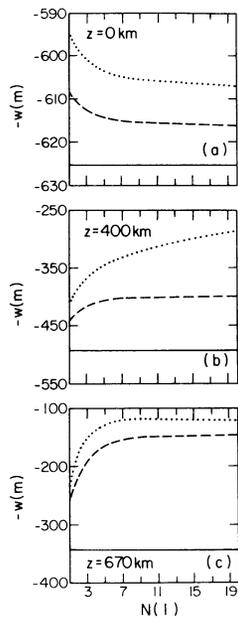


Fig. 14a-c. Axial displacement w as function of cycle number N (Fig. 9b) for Earth Model S (*solid*), Earth Model M.2 (*dashed*) or Earth Model M.3 (*dotted*); deflections refer to end of respective cycle but are represented as continuous lines; results apply to **a** $z = 0$ km, **b** $z = 400$ km or **c** $z = 670$ km

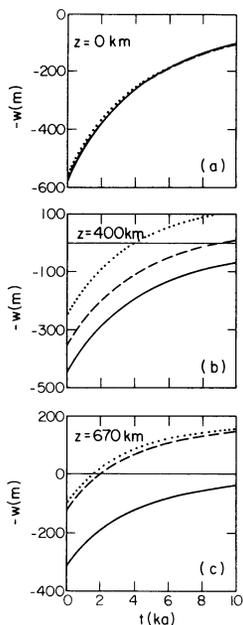


Fig. 15a-c. Axial displacement w as function of time t after load removal (Fig. 9b) for Earth Model S (*solid*), Earth Model M.2 (*dashed*) or Earth Model M.3 (*dotted*); results apply to **a** $z = 0$ km, **b** $z = 400$ km or **c** $z = 670$ km

400-km depth (Fig. 15b) and 670-km depth (Fig. 15c) the M0 mode results in upward (positive) vertical displacements after several thousand years. This in turn causes a positive contribution to the gravity anomaly. In the example discussed here its magnitude is, however, very small compared with the negative anomaly related to the first-order density discontinuity at the top surface (Fig. 16) and is therefore neglected in the following figure.

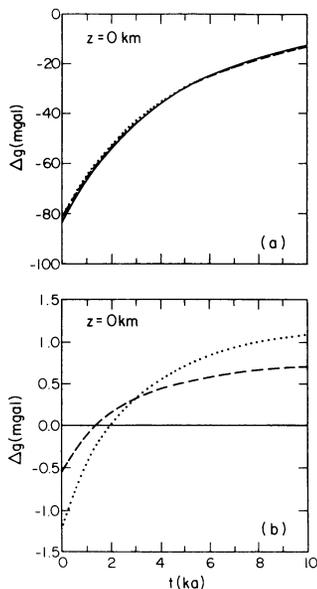


Fig. 16a, b. Axial free-air gravity anomaly Δg as function of time t after load removal (Fig. 9b) for Earth Model S (*solid*), Earth Model M.2 (*dashed*) or Earth Model M.3 (*dotted*); curves show gravity anomaly due to **a** surface density contrast or **b** density contrasts at 400-km and 670-km depths

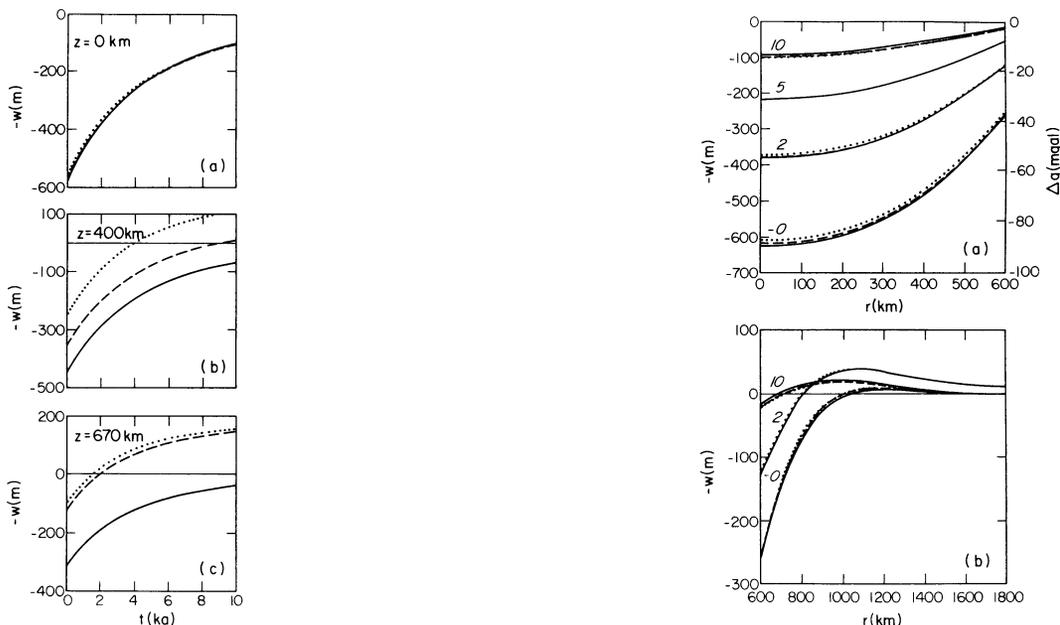


Fig. 17a, b. Same as Fig. 12 except that Earth Model M.2 (*dashed*) or Earth Model M.3 (*dotted*) is compared with Earth Model S (*solid*)

Figure 17 compares the final relaxation of the depression for Earth Models S, M.2 and M.3. The effect of the M1 and M2 modes is very small in our example. At larger times and for loads of larger diameter they will become more important, however. An appropriate example is the Laurentide glaciation, where the contribution of the M1 and M2 modes is essential for the explanation of the substantial free-air gravity anomaly correlated with this Pleistocene ice-sheet (Wu and Peltier, 1983).

Conclusion

Our discussion has demonstrated that, for loads comparable in dimension to the Fennoscandian ice-sheet, the influence of the lithosphere on the predicted surface deflection is pronounced. This is at variance with Cathles' analysis, who concluded "that the magnitude of the lithosphere's flexural rigidity is not sufficient to affect the central uplift of the larger Pleistocene loads (Fennoscandia or larger)" (Cathles, 1975, pp. 153–154). Whereas his statement is correct for the Laurentide glaciation (Wolf, 1984), post-glacial uplift of the central region in Fennoscandia may even be suitable for "measuring" lithospheric thickness.

Due to the reduced scale of the load, the Fennoscandian uplift should also be highly sensitive to the presence of an asthenosphere. As demonstrated, the modifications produced in the uplift pattern are distinct from those caused by increasing lithospheric thickness. The characteristic differences are, however, mainly confined to the initial period of isostatic recovery immediately after load removal. High-quality uplift data from this time interval are therefore required. Unfortunately, the initial period of recovery is also strongly affected by the largely unknown details of the deglaciation event (e.g. Wolf, 1985d).

The sensitivity of the response to the viscosity of the lower mantle is in accordance with results obtained by Nakiboglu and Lambeck (1982). Studying the relaxation of a viscoelastic channel underlain by a rigid half-space, they found that the response approaches the half-space limit if the channel thickness exceeds the load radius by about a factor of three.

The relaxation of the lower portions of the thermal lithosphere may turn out to be of some relevance for the interpretation of post-glacial-adjustment data. This may particularly apply if the thermal lithosphere is thicker than 100 km. As shown in the previous section, the axial deflection is substantially reduced by a thick lithosphere. Then lithospheric relaxation may also be efficient near the load axis. Since the relaxation of the higher L modes is exceedingly slow, relative-uplift observations after deglaciation are necessarily nearly unaffected by these modes. The free-air gravity anomaly, on the other hand, is a measure of the absolute deformation. It is therefore expected to be sensitive to lithospheric relaxation. This may be of some consequence for the interpretation of post-glacial-adjustment data from Fennoscandia, where the relation between relative-uplift and gravity observations has been controversial (e.g. Cathles, 1975, pp. 151–154).

Internal density contrasts have been shown to be of minor importance to the interpretation of post-glacial adjustment, provided that the load radius does not exceed 600 km. This result is expected to be slightly modified if somewhat larger load-scales are considered.

Appendix A

Propagator matrices for multi-layer elastic half-space

The non-zero elements of the half-space propagator $\mathbf{L}^{(m)}$ are

$$L_{11}^{(m)} = -1/(2\mu_m k), \quad (26a)$$

$$L_{12}^{(m)} = 1/(2\mu_m k), \quad (26b)$$

$$L_{21}^{(m)} = -1/(2\mu_m k), \quad (26c)$$

$$L_{31}^{(m)} = 1, \quad (26d)$$

$$L_{32}^{(m)} = -1, \quad (26e)$$

$$L_{41}^{(m)} = 1. \quad (26f)$$

Parameter μ_m denotes the shear modulus appropriate to the half-space.

The elements of the layer propagator $\mathbf{L}^{(l)}$ are

$$L_{11}^{(l)} = -k h_l \sinh(k h_l) + \cosh(k h_l), \quad (27a)$$

$$L_{12}^{(l)} = -k h_l \cosh(k h_l), \quad (27b)$$

$$L_{13}^{(l)} = -1/(2\mu_l k) [k h_l \cosh(k h_l) + \sinh(k h_l)], \quad (27c)$$

$$L_{14}^{(l)} = 1/(2\mu_l k) k h_l \sinh(k h_l), \quad (27d)$$

$$L_{21}^{(l)} = k h_l \cosh(k h_l), \quad (27e)$$

$$L_{22}^{(l)} = -k h_l \sinh(k h_l) + \cosh(k h_l), \quad (27f)$$

$$L_{23}^{(l)} = -1/(2\mu_l k) k h_l \sinh(k h_l), \quad (27g)$$

$$L_{24}^{(l)} = 1/(2\mu_l k) [k h_l \cosh(k h_l) - \sinh(k h_l)], \quad (27h)$$

$$L_{31}^{(l)} = -2\mu_l k [k h_l \cosh(k h_l) + \sinh(k h_l)], \quad (27i)$$

$$L_{32}^{(l)} = 2\mu_l k^2 h_l \sinh(k h_l), \quad (27j)$$

$$L_{33}^{(l)} = k h_l \sinh(k h_l) + \cosh(k h_l), \quad (27k)$$

$$L_{34}^{(l)} = -k h_l \cosh(k h_l), \quad (27l)$$

$$L_{41}^{(l)} = -2\mu_l k^2 h_l \sinh(k h_l), \quad (27m)$$

$$L_{42}^{(l)} = 2\mu_l k [k h_l \cosh(k h_l) - \sinh(k h_l)], \quad (27n)$$

$$L_{43}^{(l)} = k h_l \cosh(k h_l), \quad (27o)$$

$$L_{44}^{(l)} = -k h_l \sinh(k h_l) + \cosh(k h_l). \quad (27p)$$

Parameter μ_l denotes the shear modulus of the l -th layer and $h_l = z_{l+1} - z_l$ is its thickness.

The elements of $\mathbf{P}^{(m)}$ differing from the corresponding elements of $\mathbf{L}^{(m)}$ are

$$P_{41}^{(m)} = L_{41}^{(m)} - \rho_m g L_{21}^{(m)}, \quad (28a)$$

$$P_{42}^{(m)} = L_{42}^{(m)} - \rho_m g L_{22}^{(m)}, \quad (28b)$$

where ρ_m denotes the density of the homogeneous half-space.

The elements of $\mathbf{P}^{(l)}$ differing from the corresponding elements of $\mathbf{L}^{(l)}$ are

$$P_{12}^{(l)} = L_{12}^{(l)} + \rho_l g L_{14}^{(l)}, \quad (29a)$$

$$P_{22}^{(l)} = L_{22}^{(l)} + \rho_l g L_{24}^{(l)}, \quad (29b)$$

$$P_{32}^{(l)} = L_{32}^{(l)} + \rho_l g L_{34}^{(l)}, \quad (29c)$$

$$P_{41}^{(l)} = L_{41}^{(l)} - \rho_l g L_{21}^{(l)}, \quad (29d)$$

$$P_{42}^{(l)} = L_{42}^{(l)} - \rho_l g (L_{22}^{(l)} - L_{44}^{(l)}) - (\rho_l g)^2 L_{24}^{(l)}, \quad (29e)$$

$$P_{43}^{(l)} = L_{43}^{(l)} - \rho_l g L_{23}^{(l)}, \quad (29f)$$

$$P_{44}^{(l)} = L_{44}^{(l)} - \rho_l g L_{24}^{(l)}, \quad (29g)$$

where ρ_l denotes the density of the l -th layer.

Appendix B

Normal modes of layered Maxwell continuum

Following Peltier (1985) we write for Eqs. (13) and (17b), respectively,

$$\hat{\mathbf{Y}}^{(ve)}(s, z_l) = C_1(s) \mathbf{P}_1(s, z_l) + C_2(s) \mathbf{P}_2(s, z_l), \quad (30)$$

$$\mathbf{C}(s) = \mathbf{M}^{-1}(s) \mathbf{B}, \quad (31)$$

where the transformation $\mu \rightarrow \mu s/(s + \tau^{-1})$ (with $\tau = \eta/\mu$ the Maxwell time and s the Laplace-transform variable), and therefore $\mathbf{P} \rightarrow \mathbf{P}(s)$ and $\mathbf{M} \rightarrow \mathbf{M}(s)$, has been applied. Vector $\tilde{\mathbf{Y}}^{(ve)}$ is the Laplace transform of the "viscoelastic" column matrix, viz.

$$\tilde{\mathbf{Y}}^{(ve)}(s, z) = [\tilde{u}_1(s, z), \tilde{w}_0(s, z), \tilde{\sigma}_{rz1}(s, z), \tilde{\sigma}_{zz0}(s, z)]^T. \quad (32)$$

Since

$$C_i(s) = \sum_j [\det \mathbf{M}(s)]^{-1} [\mathbf{M}^A(s)]_{ij} B_j, \quad (33)$$

substitution for C_i in Eq. (30) yields

$$\tilde{\mathbf{Y}}^{(ve)}(s, z_l) = [\det \mathbf{M}(s)]^{-1} \mathbf{Q}(s, z_l), \quad (34)$$

where

$$\mathbf{Q}(s, z_l) = \sum_{ij} [\mathbf{M}^A(s)]_{ij} B_j \mathbf{P}_i(s, z_l). \quad (35)$$

From $\lim_{s \rightarrow \infty} \mu(s) = \mu$ it follows that

$$\mathbf{Y}^{(e)}(z_l) = \lim_{s \rightarrow \infty} \tilde{\mathbf{Y}}^{(ve)}(s, z_l). \quad (36)$$

With

$$\mathbf{Q}^{(v)}(s, z_l) = \det \mathbf{M}[\mathbf{Y}^{(ve)}(s, z_l) - \mathbf{Y}^{(e)}(z_l)], \quad (37)$$

Eq. (34) may then be written as

$$\tilde{\mathbf{Y}}^{(ve)}(s, z_l) = \mathbf{Y}^{(e)}(z_l) + [\det \mathbf{M}(s)]^{-1} \mathbf{Q}^{(v)}(s, z_l). \quad (38)$$

The time-domain solution associated with Eq. (38) is

$$\begin{aligned} \mathbf{Y}^{(ve)}(t, z_l) &= \mathbf{Y}^{(e)}(z_l) \delta(t) \\ &+ 1/(2\pi i) \int_B [\det \mathbf{M}(s)]^{-1} \cdot \mathbf{Q}^{(v)}(s, z_l) \exp(st) ds, \end{aligned} \quad (39)$$

where B denotes the Bromwich path. According to the residue theorem this, however, is equivalent to

$$\begin{aligned} \mathbf{Y}^{(ve)}(t, z_l) &= \mathbf{Y}^{(e)}(z_l) \delta(t) \\ &+ \sum \text{res} \{ [\det \mathbf{M}(s)]^{-1} \mathbf{Q}^{(v)}(s, z_l) \exp(st) \}. \end{aligned} \quad (40)$$

If $\mathbf{Q}^{(v)}(s, z_l) \exp(st)$ is regular and $\det \mathbf{M}(s)$ has simple zeros at $-s^{(k)}$, where $s^{(k)} > 0$, we therefore obtain

$$\begin{aligned} \mathbf{Y}^{(ve)}(t, z_l) &= \mathbf{Y}^{(e)}(z_l) \delta(t) \\ &+ \sum_k \frac{Q^{(v)}(-s^{(k)}, z_l) \exp(-s^{(k)}t)}{\left[\frac{d}{ds} \det \mathbf{M}(s) \right]_{s=-s^{(k)}}}. \end{aligned} \quad (41)$$

Defining

$$\begin{aligned} s^{(k)} \mathbf{Y}^{(v, k)}(z_l) &= \left\{ \left[\frac{d}{ds} \det \mathbf{M}(s) \right]_{s=-s^{(k)}} \right\}^{-1} \mathbf{Q}^{(v)}(-s^{(k)}, z_l), \end{aligned} \quad (42)$$

Eq. (41) may alternatively be written in the form

$$\begin{aligned} \mathbf{Y}^{(ve)}(t, z_l) &= \mathbf{Y}^{(e)}(z_l) \delta(t) \\ &+ \sum_k \mathbf{Y}^{(v, k)}(z_l) s^{(k)} \exp(-s^{(k)}t). \end{aligned} \quad (43)$$

This is the impulse response of the Maxwell half-space.

Appendix C

Viscosity stratification of lithosphere

The one-dimensional form of the stress-strain relation for linear creep can be written as (e.g. Weertman and Weertman, 1975)

$$\dot{\epsilon} = AD\sigma/(2\mu), \quad (44)$$

with σ the stress and $\dot{\epsilon}$ the strain rate. Parameter A is an empirical constant depending weakly on temperature. Parameter D is the diffusion coefficient of the material given by

$$D = D_\infty \exp[-Q/(RT)], \quad (45)$$

with Q the activation energy, R the gas constant and T the absolute temperature. Substituting for D in Eq. (45) yields

$$\dot{\epsilon} = AD_\infty \sigma \exp[-Q/(RT)]/(2\mu). \quad (46)$$

The effective viscosity is defined by $\eta = \sigma/(2\dot{\epsilon})$. Substituting for $\sigma/(2\dot{\epsilon})$ from Eq. (46) we obtain

$$\eta(T) = \eta_\infty \exp[Q/(RT)], \quad (47a)$$

where $\eta_\infty = \mu/(AD_\infty)$. If the viscosity $\eta_0 = \eta(T_0)$ is known, Eq. (47a) may alternatively be written as

$$\eta(T) = \eta_0 \exp[Q(1/T - 1/T_0)/R]. \quad (47b)$$

Acknowledgements. I would like to thank Richard Peltier, who provided a preprint which describes the normal-mode formalism and who called my attention to the possible significance of lithospheric relaxation. This research was financially supported by a Natural Sciences and Engineering Research Council of Canada Postgraduate Scholarship.

References

- Balling, N.: The land uplift in Fennoscandia, gravity field anomalies and isostasy. In: Earth rheology, isostasy, and eustasy, N.-A. Mörner, ed.: pp. 297–321. New York: Wiley 1980
- Beaumont, C.: The evolution of sedimentary basins on a viscoelastic lithosphere: theory and examples. *Geophys. J.R. Astron. Soc.* **55**, 471–497, 1978
- Cathles, L.M.: The viscosity of the Earth's mantle. Princeton: Princeton University Press 1975
- Courtney, R.C.: On the rheology of the oceanic and continental lithospheres. M.Sc. thesis, Dalhousie University 1982
- Dziewonski, A.M., Anderson, D.L.: Preliminary reference Earth model. *Phys. Earth Planet. Inter.* **25**, 297–356, 1981
- Farrell, W.E.: Deformation of the Earth by surface loads. *Rev. Geophys. Space Phys.* **10**, 761–797, 1972
- Imbrie, J., Imbrie, K.P.: Ice ages: solving the mystery. Hillside: Enslow 1979
- Kuo, J.T.: Static response of a multilayered medium under inclined surface loads. *J. Geophys. Res.* **74**, 3195–3207, 1969
- Lambeck, K., Nakiboglu, S.M.: Seamount loading and stress in the ocean lithosphere. *J. Geophys. Res.* **85**, 6403–6418, 1980
- Lanczano, P.: Deformations of an elastic Earth. New York: Academic Press 1982
- McConnell, R.K. jr.: Isostatic adjustment in a layered Earth. *J. Geophys. Res.* **70**, 5171–5188, 1965
- McConnell, R.K. jr.: Viscosity of the mantle from relaxation time spectra of isostatic adjustment. *J. Geophys. Res.* **73**, 7089–7105, 1968
- Nakiboglu, S.M., Lambeck, K.: A study of the Earth's re-

- sponse to surface loading with application to Lake Bonneville. *Geophys. J.R. Astron. Soc.* **70**, 577–620, 1982
- Parker, R.L.: The rapid calculation of potential anomalies. *Geophys. J.R. Astron. Soc.* **31**, 447–455, 1972
- Parsons, B.E.: Changes in the Earth's shape. Ph.D. thesis, Cambridge University 1972
- Peltier, W.R.: The impulse response of a Maxwell Earth. *Rev. Geophys. Space Phys.* **12**, 649–669, 1974
- Peltier, W.R.: Glacial-isostatic adjustment – II. The inverse problem. *Geophys. J.R. Astron. Soc.* **46**, 669–705, 1976
- Peltier, W.R.: Dynamics of the ice age Earth. *Adv. Geophys.* **24**, 1–146, 1982
- Peltier, W.R.: The LAGEOS constraint on deep mantle viscosity: results from a new normal mode method for the inversion of viscoelastic relaxation spectra. *J. Geophys. Res.* 1985 (in press)
- Peltier, W.R., Andrews, J.T.: Glacial-isostatic adjustment – I. The forward problem. *Geophys. J.R. Astron. Soc.* **46**, 605–646, 1976
- Sclater, J.G., Jaupart, C., Galson, D.: The heat flow through oceanic and continental crust and the heat loss of the Earth. *Rev. Geophys. Space Phys.* **18**, 269–311, 1980
- Tullis, J.A.: High temperature deformation of rocks and minerals. *Rev. Geophys. Space Phys.* **17**, 1137–1154, 1979
- Walcott, R.I.: Rheological models and observational data of glacio-isostatic rebound. In: Earth rheology, isostasy, and eustasy, N.-A. Mörner, ed.: pp. 3–10. New York: Wiley 1980
- Ward, S.N.: A note on lithospheric bending calculations. *Geophys. J.R. Astron. Soc.* **78**, 241–253, 1984
- Weertman, J., Weertman, J.R.: High temperature creep of rock and mantle viscosity. *Annu. Rev. Earth Planet. Sci.* **3**, 293–315, 1975
- Wolf, D.: The relaxation of spherical and flat Maxwell Earth models and effects due to the presence of the lithosphere. *J. Geophys.* **56**, 24–33, 1984
- Wolf, D.: Thick-plate flexure re-examined. *Geophys. J.R. Astron. Soc.* **80**, 265–273, 1985a
- Wolf, D.: On Boussinesq's problem for Maxwell continua subject to an external gravity field. *Geophys. J.R. Astron. Soc.* **80**, 275–279, 1985b
- Wolf, D.: The normal modes of a uniform, compressible Maxwell half-space. *J. Geophys.* **56**, 100–105, 1985c
- Wolf, D.: An improved estimate of lithospheric thickness based on a reinterpretation of tilt data from Pleistocene Lake Algonquin. *Can. J. Earth Sci.* **22**, 768–773, 1985d
- Wu, P., Peltier, W.R.: Viscous gravitational relaxation. *Geophys. J.R. Astron. Soc.* **70**, 435–485, 1982
- Wu, P., Peltier, W.R.: Glacial isostatic adjustment and the free-air gravity anomaly as a constraint on deep mantle viscosity. *Geophys. J.R. Astron. Soc.* **74**, 377–449, 1983

Received February 11, 1985; revised version April 23, 1985
Accepted April 25, 1985

A palaeomagnetic study of Turonian carbonates from the southeastern Münsterland area, NW Germany

M. Kerth

Geologisches Institut der Universität Bonn, Nußallee 8, D 5300 Bonn 1, W Germany

Abstract. A palaeomagnetic study of Turonian carbonates from the southeastern Münsterland area, NW Germany ($8^{\circ}15'E/51^{\circ}30'N$) has been carried out. The initial intensities of natural remanent magnetization of 450 collected specimens range from 5×10^{-5} to 3×10^{-4} A/m. The main carriers of magnetization are magnetites or titanomagnetites. Most of the samples show a strong VRM overprint probably caused by multidomain magnetite grains. Specimens from only one out of five sampling sites yield a stable component which is believed to represent the Late Cretaceous geomagnetic field direction. The mean direction of this remanent magnetization after demagnetization is $17^{\circ}/52^{\circ}$ (declination/inclination). The corresponding Upper Cretaceous palaeopole position is located at $68^{\circ}N/149^{\circ}E$ ($\alpha_{95} = 5.3^{\circ}$).

Key words: Palaeomagnetism – Carbonates – Turonian – Middle Europe – Viscous remanent magnetization

Introduction

Largely because of the very low intensity of magnetization in the Upper Cretaceous sediments of Middle Europe, only

Present address: Universität Essen, Fachbereich 9, Fach Geologie, Universitätsstr. 2, D 4300 Essen, W Germany

few palaeomagnetic pole determinations and no magnetostratigraphic data are currently available. Andreeva et al. (1965) carried out a palaeomagnetic study of rocks from the Bohemian Massif, CSSR, including samples of Upper Cretaceous age. However, the palaeopole position found in this study lies very close to the present geomagnetic pole and, therefore, may not be representative of the Late Cretaceous geomagnetic field. Heller and Channell (1979) sampled a number of sites in Upper Cretaceous rocks of the Münsterland area, NW Germany, and determined a reliable geomagnetic pole position.

The purpose of the present study of Turonian carbonates of the Münsterland basin, NW Germany, was to evaluate the suitability of these rocks for magnetostratigraphic work and to check the existence of a reversal within the Turonian as reported by Krumsiek (1982) from Turonian limestones of SW Morocco.

Geological setting

The Upper Cretaceous rocks of the Münsterland area were deposited in a basin which was formed in Albian and Cenomanian times by subsidence of the northern part of the Rhenish Massif (Arnold, 1964). During the Albian, the sea transgressed into the central and eastern parts of the basin. The basal sediments in the southwestern part are of Cenomanian and Turonian age. According to Arnold (1964) and



Fig. 1. Geological setting and location of sampling sites

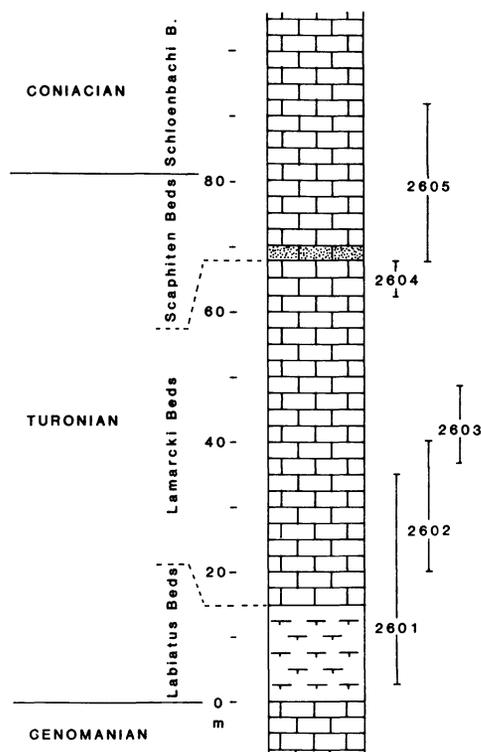


Fig. 2. Stratigraphy and stratigraphical extent of the sampling sites

Seibertz (1979a), the water depth during deposition of the Cretaceous sediments did not exceed 250 m.

A stratigraphically nearly complete section of Turonian carbonates in the southeastern Münsterland basin was sampled (Figs. 1 and 2). The total thickness of the Turonian sediments in this area is between 75 and 85 m.

The "Labiatus Beds" of Lower Turonian age comprise marls with nodular limestone layers. The "Lamarcki Beds" (Middle Turonian) are predominantly marly limestones. At the base of the "Scaphiten Beds" (Upper Turonian) a carbonaceous, glauconitic sandstone occurs, overlain by a sequence of marly limestones. These grade upwards into the "Schloenbachi Beds", which belong to the Coniacian (Seibertz, 1979b; Troeger, 1981).

The beds are tectonically undisturbed and dip northward at only 1° – 2° , so that a bedding correction of the palaeomagnetic data was unnecessary and a fold test could not be applied. At five sites (2601–2605), 70 independently oriented samples were drilled, each sample consisting of at least six specimens. The total number of specimens used in the present study is 450.

Natural remanent magnetization (NRM)

The NRM of the specimens was measured with a UGF-4 spinner magnetometer (Geofysika, Brno, CSSR) at the Department of Geology, University of Bonn. NRM intensities were between 5×10^{-5} and 3×10^{-4} A/m.

In the specimens from sites 2601–2603 and 2605 a viscous remanent magnetization (VRM) dominated and no stable magnetization could be determined. At these sites the main VRM component is aligned with the present magnetic field at the sampling site, but this is partly overprinted

by a component acquired in the laboratory field. Before measurement the cylindrical specimens were stored in an upright position with their $+z$ -axis downwards, but with randomly orientated x -axis. Since the inclination of the present magnetic field at Bonn is about 66° , the acquisition of a VRM in this field mainly resulted in a progressive increase in the component of magnetization along the z -axis. Due to different field orientations of the z -axis for each sample, the NRM vectors show a wide scatter of directions (Fig. 3). On an equal area projection the mean NRM vectors of all specimens of one sample lie on great circles passing through the direction of the present geomagnetic field at the sampling site and the z -axis for each sample (Fig. 4). Calculating the mean over all specimens of one sample averages out the component in the xy -plane. Furthermore, this component, due to the steep inclination angle of the laboratory field, is small in comparison to the component along the z -axis.

In marked contrast, the 44 specimens from site 2604 show a close grouping of the NRM vectors with a mean direction of $16^{\circ}/53^{\circ}$ (declination/inclination), $\alpha_{95} = 5.3^{\circ}$ (Fig. 5).

Stability tests

Storage test

After the first NRM measurement the specimens were again stored in the laboratory for a period of between 2 and 4 weeks, before a second measurement of the specimens selected for demagnetization was carried out. By this procedure acquisition of a VRM aligned parallel to the $+z$ -axis was identified, which continued in the specimens of sites 2601–2603 and 2605. After storage such a VRM could also be detected in specimens from site 2604, although the amount of this component was comparatively small.

Demagnetization experiments

Eighty selected specimens, at least one from each sample, were demagnetized by thermal or alternating field (AF) methods using a Schonstedt TSD-1 furnace and an AF demagnetizer developed in Bonn (Krumsiek, 1980, 1982). In some cases a combination of both methods was applied. Thus some specimens were thermally demagnetized up to 120° C to remove a component probably carried by goethite (Heller, 1978). Then demagnetization was continued using the AF method. Other specimens were thermally demagnetized up to 150° C, then AF demagnetized at 16 and 24 mT. After that, the specimens were further thermally demagnetized. This method can help to distinguish between hematite and magnetite as carriers magnetization (Krumsiek, 1982).

As a criterion for the stability of the remanence vectors during demagnetization, the change of angle between successive demagnetization steps was calculated. Stability was considered to be good if the total change between three successive remanence vectors was less than 10° .

Sites 2601–2603, 2605. During AF demagnetization, the specimens showed a rapid initial decrease of intensity. Some specimens appeared to acquire a magnetization at fields higher than 24 mT (Fig. 6a). Since a two-axis tumbler was used, this magnetization is, most likely, a rotational remanent magnetization (Stephenson, 1976).

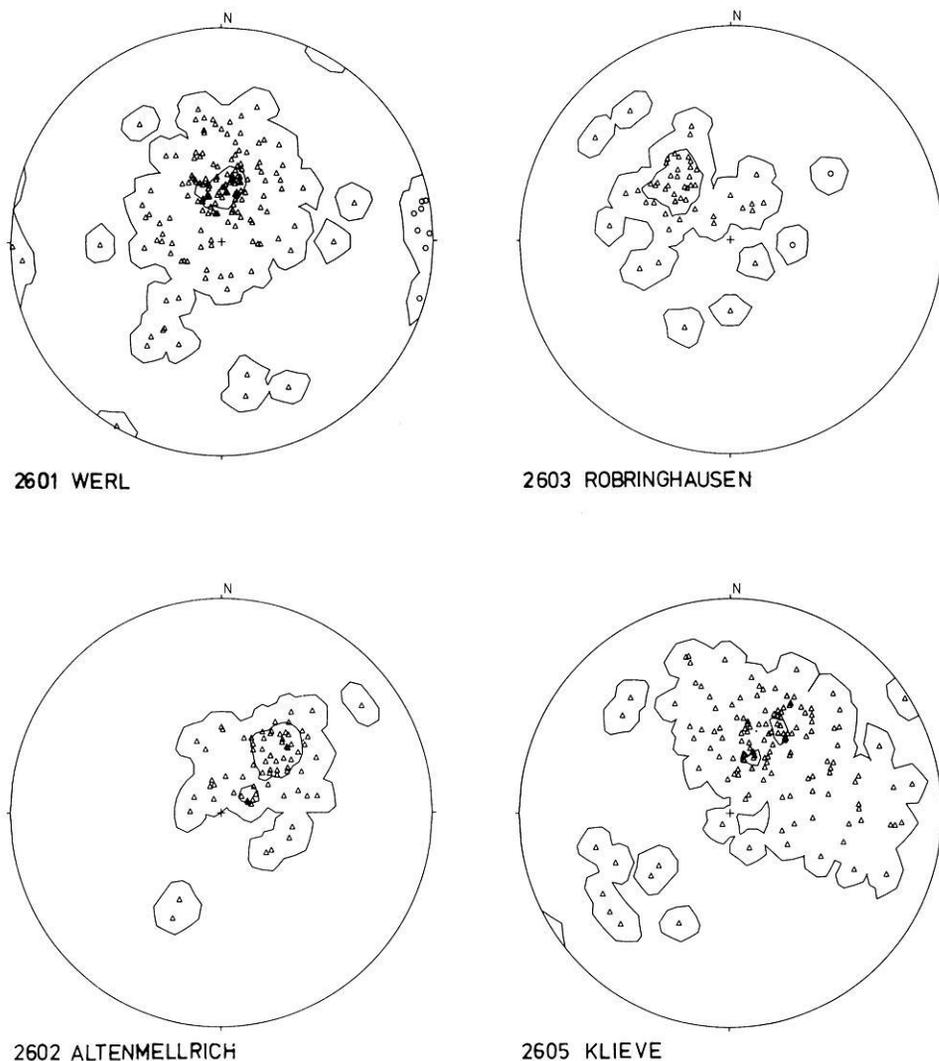
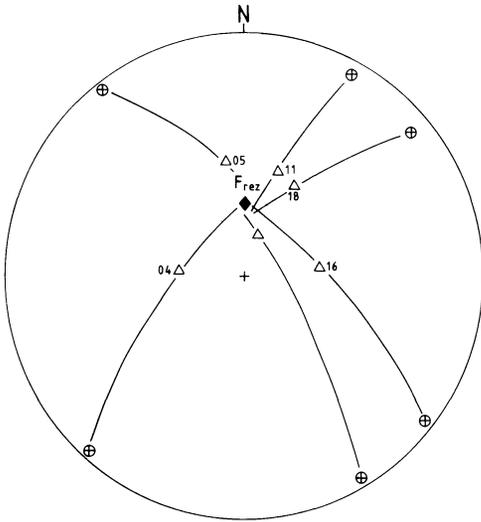


Fig. 3. NRM of all specimens from sites 2601–2603 and 2605. The vectors have been plotted with a Fortran program (Krumstiek and Siehl, 1976). The two lines are the 0% and 16% isolines of point density distribution. *Triangle*: Lower hemisphere, *circle*: Upper hemisphere. The vectors show a wide scatter which is due to a VRM overprint in the laboratory field. The overprint occurs mainly along the $+z$ -axis of the samples. Site 2602 has been sampled at a west-facing exposure, so that the $+z$ -axis of the specimens are pointing towards easterly directions. Site 2603 has been sampled at a mainly east-facing exposure. This results in a maximum of the NRM directional distribution to the east of north at site 2602 and to the west of north at site 2603. At sites 2601 and 2605 the sample orientation is variable resulting in a much larger scatter of NRM directions

During heating, the remanence intensity dropped at temperatures of 400°C . Above 250°C the specimens changed their colour from grey to red. Over 400°C a rapid increase of remanence intensity occurred (Fig. 6b), which is related to the formation of a new magnetic phase (see later) and its alignment in the very weak field of the furnace. Neither method revealed a stable remanence vector in the specimens of these sites and the change of angle between successive steps exceeded 10° for most of the specimens. Some specimens showed a decrease of the VRM component acquired in the laboratory field during demagnetization. This can clearly be seen by plotting the remanence vectors on an equal area projection (Fig. 6c). With increasing AF amplitudes or temperatures, the remanence vectors move towards the relatively more stable component aligned parallel to the present magnetic field at the sampling site. However, when the remanence vector after a particular demag-

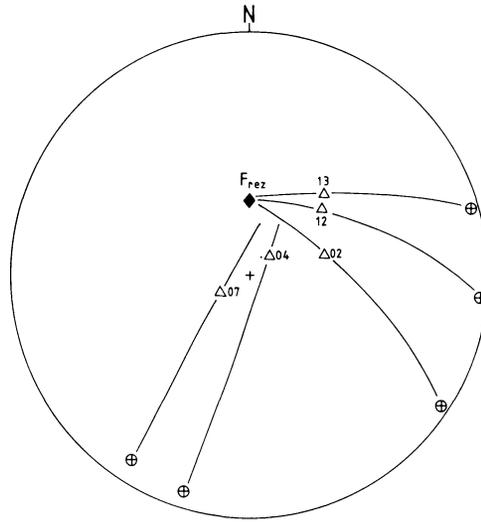
netization step was re-measured after 48 or 72 h, it moved back to a position nearer to the $+z$ -axis, thus showing again the acquisition of VRM along the $+z$ -axis during this time interval.

Site 2604. The behaviour of the remanence intensities during thermal demagnetization was similar to that of specimens from the other sites (Fig. 6d). However, the VRM component acquired during storage in the laboratory could be removed by heating up to 75°C . For all samples of site 2604 the $+z$ -axis field orientation is nearly horizontal, so that an acquisition of a component along the $+z$ -axis lowers the inclination angle of field-corrected data. In Fig. 6d it can clearly be seen that the inclination angle becomes steeper after the first thermal demagnetization step has been applied. Thus, this initial steepening of the inclination angle, which is typical for all demagnetized specimens

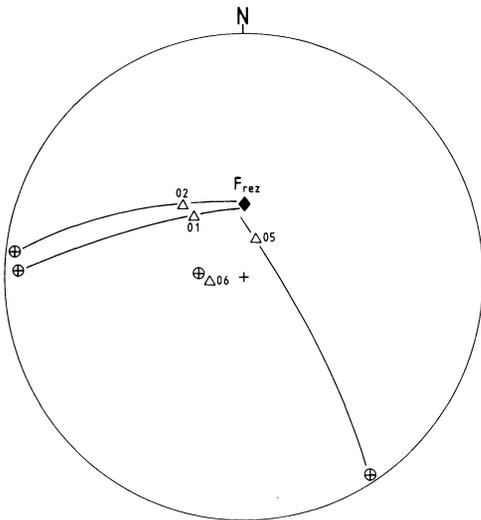


PROFIL 2601 WERL

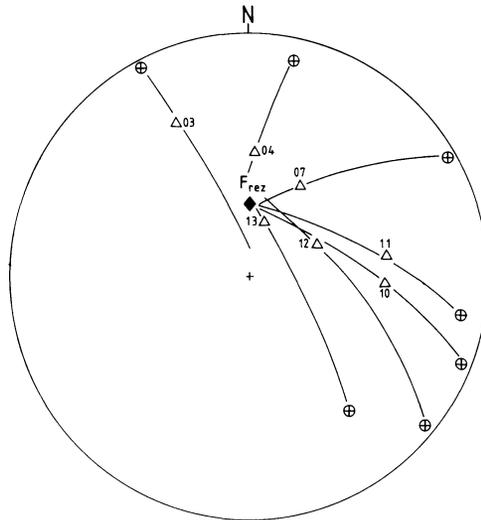
RM - direction Δ
z-axis \oplus



PROFIL 2602 ALTENMELLRICH

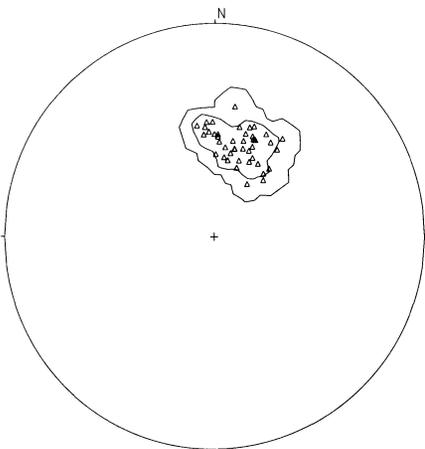


PROFIL 2603 ROBRINGHAUSEN



PROFIL 2605 KLIEVE

Fig. 4. Mean NRM vectors of all specimens of one sample for sites 2601–2603 and 2605, corresponding z-axes orientations and the present day field directions at the sampling sites. In each case, the three directions lie on great circles and therefore indicate a VRM acquired in the laboratory overprinting the VRM aligned with the present day field at the sampling site



2604 WESTEREIDEN KERNE

Fig. 5. NRM of all specimens from site 2604. Symbols: see comment on Fig. 3

of site 2604, is believed to represent the removal of the VRM acquired in the laboratory field.

Isothermal remanent magnetization (IRM)

Thirteen AF demagnetized or untreated and five thermally demagnetized specimens were subjected to IRM-acquisition experiments with a maximum field strength of 0.6 T. The thermally untreated specimens showed a strong initial increase of the remanence intensity in fields up to 0.1 T. Above 0.1 T the increase became very small and some specimens reached saturation magnetization in fields of about 0.2–0.3 T (Fig. 7a). Only specimen 260118 1 2, which has a reddish colour, showed a large component which did not saturate in fields up to 0.6 T.

After thermal demagnetization up to 500° C (50° C steps and 1 h heating-cooling cycle), the IRM intensities were 5–10 times higher than the values of the untreated speci-

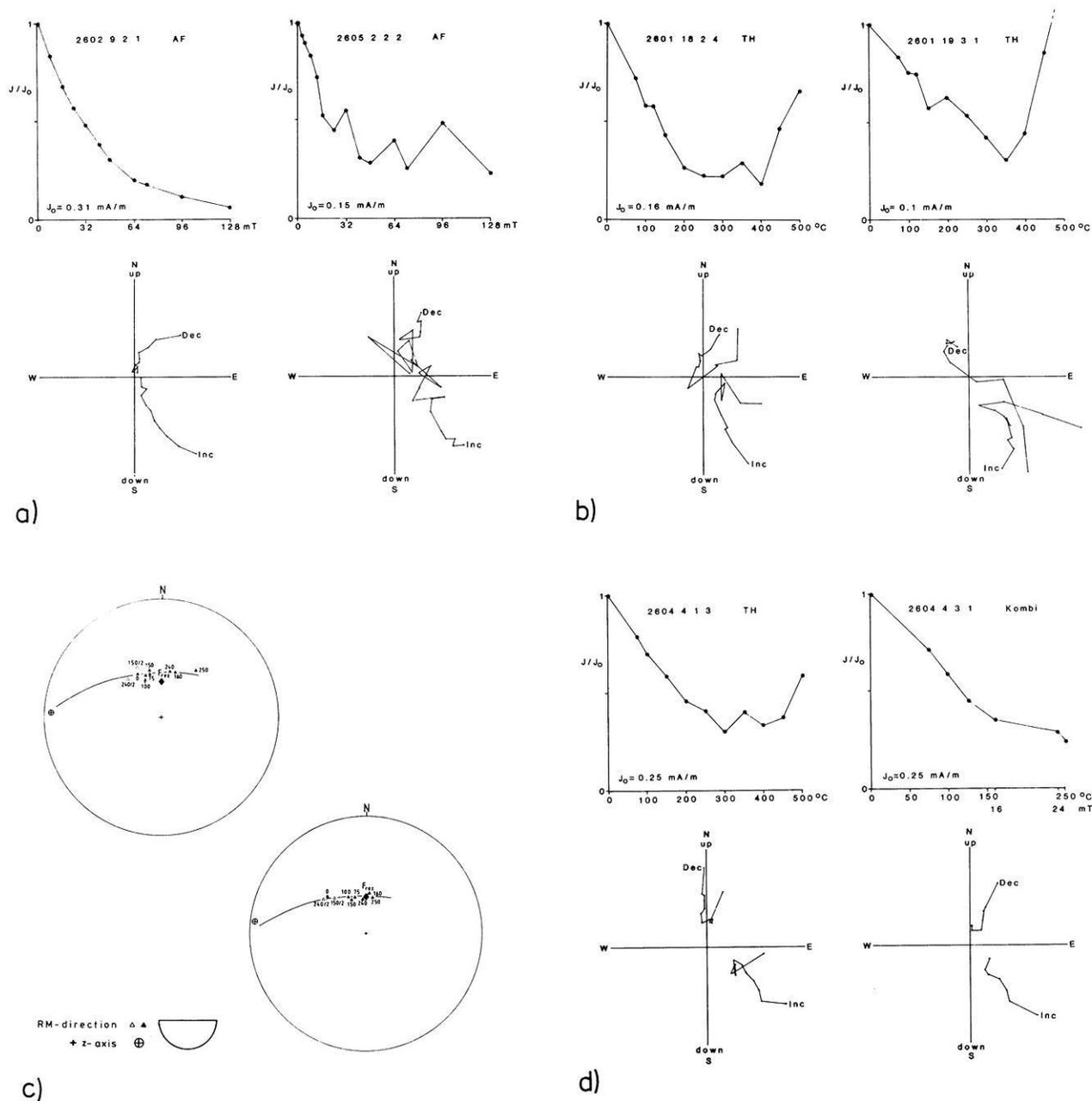


Fig. 6 a–d. AF demagnetization curves. Some of the samples show the acquisition of an artificially produced remanent magnetization. **b** Thermal demagnetization curves. The samples acquire a magnetization when heated above 400° C, which is related to the formation of new magnetic phases. **c** Remanence vectors of all demagnetization steps for two specimens in a stereographic projection. The (*open symbols*) vectors marked with 100/2, 150/2 etc. are measurements of the remanence vector at this particular demagnetization step repeated after 48 or 72 h storage in the laboratory field. **d** Demagnetization curves of specimens from site 2604

mens (Fig. 7b). Once again, the remanence strongly increased in fields up to 0.1 T, but none of the specimens reached saturation in the maximum applied field of 0.6 T.

Results

Magnetomineralogy

Both the AF and the thermal demagnetization experiments and the IRM acquisition experiments indicate magnetite or titanomagnetite as the main carrier of magnetization. In the AF demagnetization experiments no high coercivity

components have been recognized. According to the IRM acquisition curves, the remanent magnetization in some samples is entirely due to magnetite or titanomagnetite, while others contain small amounts of hematite or goethite [“type 1 limestone” and “type 4 limestone” of Heller and Channell (1979)].

The viscous behaviour of the specimens can be interpreted as the result of multidomain magnetites as the main carriers of magnetization. The movement of the remanence vectors during demagnetization away from a VRM component acquired in the laboratory field towards a component aligned parallel to the present day geomagnetic field at the

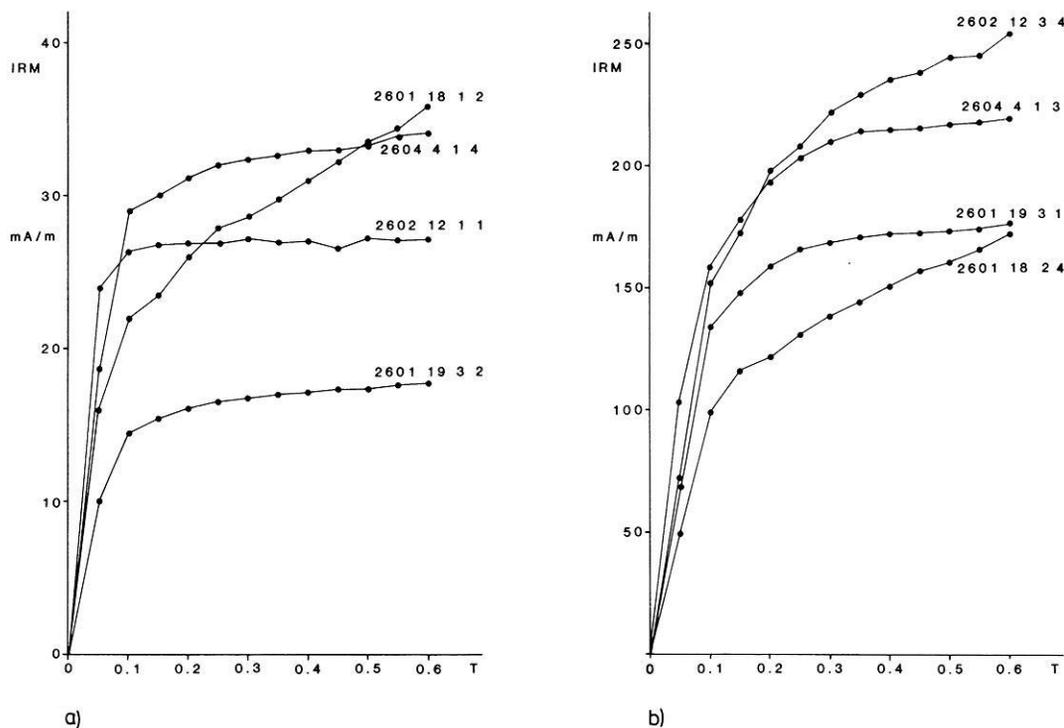


Fig. 7 a and b. IRM acquisition of a unheated limestone specimens, b specimens thermally demagnetized up to 500° C

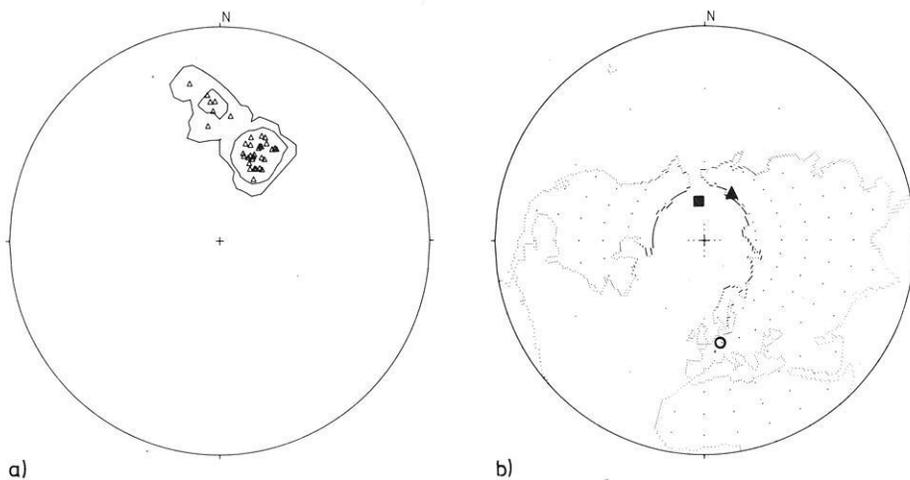


Fig. 8 a and b. Stereographic plot of all stable remanence directions from site 2604. Symbols: see comment on Fig. 3. b The sampling location (circle), the palaeopole position found in the present study (triangle) and the pole position from Heller and Channell (1979) (square)

sampling site indicates higher stability of the latter component. The higher stability can be explained by the acquisition of this VRM over a long period of time (e.g. throughout the Brunhes normal epoch).

The higher stability of remanence directions during demagnetization of specimens from site 2604 may be explained in terms of relatively smaller, pseudo-single domain magnetites being present.

The rapid increase of remanence intensity and the strong increase in IRM intensity after thermal treatment above 400° C can be explained by the alteration of pyrite, which is abundant in the Turonian carbonates, to magnetite (Kruczyk, 1977). The high coercivity component created

after thermal treatment might be caused by the desintegration of ironhydroxides formed during weathering of the carbonates at temperatures above 290° C (Hedley, 1968) and the formation of hematite pigment (Heller, 1978).

Palaeofield and palaeopole position

As discussed above, stable remanence directions up to 250° C were only found in the specimens from site 2604. No indication of a reversed field direction in the time interval represented by this site has been found. The stable remanence mean direction after demagnetization is N17°E/52° (declination/inclination). The corresponding palaeo-

magnetic pole position is located at 68°N/149°E, $\alpha_{95} = 5.3^\circ$ (from ten specimens out of the five samples of site 2604) (Fig. 8), which is in good agreement with the pole determined by Heller and Channell (1979).

Discussion

The suitability of the Turonian carbonates from the south-eastern Münsterland area for palaeomagnetic work is very limited. Only about 10% of the specimens yielded useful palaeomagnetic information. Therefore, the initial aim of the present study to check the occurrence of a reversal within the Turonian, recognized by Krumsiek (1982), was not achieved. However, all specimens which revealed a stable direction showed normal polarity. This is in agreement with the most recent Cretaceous timescale (Lowrie et al., 1980).

The palaeopole position found in the present study confirms the conclusions drawn by Heller and Channell (1979) such as the shortening of the Tethyan realm by the order of 1000 km since the Late Cretaceous and the rotation of the Iberian peninsula in its present position prior to Late Cretaceous times.

Kligfield and Channell (1981) described magnetic viscous behaviour of Helvetian limestones from the northern Alps similar to that found in the present study. They assume very small-sized pyrrhotites causing the strong viscosity of magnetization. However, their IRM curves have saturation values between 0.4 and 1.0 T, so that pyrrhothite as a carrier of magnetization in the limestones of the present studies appears to be excluded.

Mumme (1964) investigating Cenozoic basalts from Victoria, Australia found the instability of magnetization to related to an increasing titanium content of the titanomagnetites. Probably, the observed VRM in the Turonian limestones is not only related to grain size, but also to the composition of the magnetites.

Heller and Channell (1979) in their study of the Upper Cretaceous limestones of the Münsterland basin mention a large number of unstable specimens in the more southerly region of the Münsterland.

The viscous behaviour of the specimens in the present study might be related to carbonate facies factors or to weathering processes of the carbonates. A detailed study of the carbonate facies and its relation to magnetic properties of Cenomanian and Turonian rocks in the Münsterland basin is presently being carried out by Hambach (1985).

Acknowledgements. I am grateful to Prof. Dr. K. Krumsiek for the introduction into the techniques of palaeomagnetism and to

Dr. E.A. Hailwood for helpful suggestions and the revision of the translated manuscript.

References

- Andreeva, O., Bukha, V., Petrova, G.: Laboratory evaluation of magnetic stability of rocks in the Czech Massif. *Akad. Nauk. SSSR Izv., Earth Phys. Ser.*, 54–64, 1965
- Arnold, H.: Fazies und Mächtigkeit der Kreidestufen im Münsterländer Oberkreidegebiet. *Fortschr. Geol. Rheinld. Westf.* 7, 599–610, 1964
- Hambach, U.: Geologische und gesteinsmagnetische Untersuchungen im Cenoman und Turon des südlichen Münsterlandes. Unpubl. Diploma-thesis, University of Bonn, W Germany, 1985
- Hedley, I.G.: Chemical remanent magnetization of the FeOOH, Fe₂O₃ system. *Phys. Earth Planet. Int.* 1, 103–121, 1968
- Heller, F.: Rockmagnetic studies of Upper Jurassic limestones from South Germany. *J. Geophys.* 44, 525–543, 1978
- Heller, F., Channell, J.: Palaeomagnetism of Upper Cretaceous limestones from the Münster basin, Germany, *J. Geophys.* 46, 413–427, 1979
- Kligfield, R., Channell, J.: Widespread remagnetization of Helvetic limestones. *J. Geophys. Res.* B86, 1888–1900, 1981
- Kruczyk, J.: Examination of phase transitions of pyrite in the laboratory. *Publ. Inst. Geophys. Pol. Acad. Sci.* C3 III, 59–69, 1977
- Krumsiek, K.: Zur plattentektonischen Entwicklung des Indo-Iranischen Raumes. (Resultate paläomagnetischer Untersuchungen in Afghanistan). *Geotekt. Forsch.* 60, 1–223, 1980
- Krumsiek, K.: Cretaceous magnetic stratigraphy of Southwest Morocco. In: *Geology of the Northwest African Continental Margin*. Rad. U. Hinz, K., Sarntheim, M., Seibold, E., eds.: pp. 475–497. Berlin, Heidelberg, New York: Springer 1982
- Krumsiek, K., Siehl, A.: PALM – Ein Rechenprogramm für paläomagnetische Zwecke. *Z. dtsh. geol. Ges.* 127, 309–321, 1976
- Lowrie, W., Channell, J., Alvarez, W.: A review of magnetic stratigraphy investigations in Cretaceous pelagic carbonate rocks. *J. Geophys. Res.* 85, 3597–3605, 1980
- Mumme, W.G.: Stability of magnetization in Cainozoic basalts from Victoria, Australia. *Phil. Mag.* 8 Ser. 7, 1263–1278, 1964
- Seibertz, E.: Stratigraphisch-fazielle Entwicklung des Turons im südöstlichen Münsterland. *Newsl. Stratigr.* 8, 3–60, 1979a
- Seibertz, E.: Biostratigraphie im Turon des SE-Münsterlandes und Anpassung an die internationale Gliederung aufgrund von Vergleichen mit anderen Oberkreidegebieten. *Newsl. Stratigr.* 8, 111–123, 1979b
- Stephenson, R.W.: A study of rotational remanent magnetizations. *Geophys. J.R. Astron. Soc.* 47, 363–373, 1976
- Troeger, K.A.: Zu Problemen der Biostratigraphie der Inoceramen und der Untergliederung des Cenomans und Turons in Mittel- und Osteuropa. *Newsl. Stratigr.* 9, 139–159, 1981

Received October 29, 1984; Revised April 2, 1985

Accepted April 25, 1985

Finite parallel conductivity in the open magnetosphere

M.A. Volkov and Yu.P. Maltsev

Polar Geophysical Institute, Apatity, 184200, U.S.S.R.

Abstract. A pattern of convection is calculated for a magnetospheric model with open magnetic field lines going from the polar cap to the solar wind. A uniform anti-sunward flow at the open field lines is assumed as a source of convection. A layer with finite parallel conductivity is situated below the source. The ionospheric convection has a two-vortex structure, but the convective flow across the polar cap appears to be smaller than the source flow due to a rise of parallel electric fields in the field-aligned electric currents at the polar cap boundary. The return convective flow in the region of closed magnetic field lines is also smaller than the source flow. To conserve the flow continuity two narrow (with a width of about 10 km projected onto the ionosphere) intense return convective streams arise above the finite parallel conductivity layer at the boundary of open and closed magnetic field lines. The transverse potential difference in the streams is equal to the parallel potential difference in the field-aligned currents at the polar cap boundary. Another consequence of the finite parallel conductivity is a shift of centres of the ionospheric convection vortices into the polar cap.

Key words: Magnetosphere - Ionosphere - Electric field - Field-aligned current - Parallel conductivity - Convection

Introduction

Dungey (1961) supposed that reconnection of the Earth's and interplanetary magnetic fields in the framework of an open magnetosphere model gives rise to magnetospheric plasma convection. Intensification of the convection during periods of southward IMF, when the reconnection is most efficient, was confirmed experimentally (Heppner, 1972). One of the consequences of convection is field-aligned currents at the polar cap boundary. Iijima and Potemra (1976) named them the currents of Region 1. Usually these currents are theoretically studied under conditions of infinite conductivity parallel to the magnetic field. Meanwhile, parallel conductivity is often reduced so that a potential difference of up to 10 kV arises along high-latitude magnetic field lines (Mozer et al., 1980). Possible mechanisms of generation of electric fields E_{\parallel} were reviewed

by Block and Fälthammar (1976). The role of E_{\parallel} in large-scale convection and current was studied in a few papers. Lyons (1980) has assumed a potential distribution with $\text{div } \mathbf{E} \neq 0$ at some distance from the ionosphere and has obtained a broadening of the sheet of field-aligned currents under the action of E_{\parallel} . These calculations may be applied to Region 1 currents. Chiu et al. (1981) solved a similar problem but their task was to construct an auroral arc model. Maltsev (1985) has shown that E_{\parallel} in Region 2 currents produces two isolated convective vortices in the inner magnetosphere as well as narrow intense azimuthal convective streams along both sides of the sheet of Region 2 currents.

The aim of this paper is to investigate the effect of E_{\parallel} on convection near Region 1 of field-aligned currents. In this respect, our problem is similar to the one studied by Lyons (1980) but boundary conditions are different. Contrary to Lyons (1980) we assume that the sources of convection are localized in the area of open magnetic field lines. The parallel electric field will be taken into account by introducing a layer with finite parallel conductivity, the specific mechanism of which is of no importance to us.

Model and basic equations

A layer of thickness b with finite parallel conductivity σ_{\parallel} is situated not very far from the ionosphere. Above and below the layer the parallel conductivity is infinitely large. The area where open magnetic field lines cross the ionosphere is a circle of radius r_0 . Above the layer with finite parallel conductivity the following potential distribution is assumed:

$$\varphi_m = E_0 r \sin \lambda \quad \text{for } r < r_0, \quad (1)$$

where r and E_0 are the distance and the electric field, respectively, projected onto the ionosphere, $r=0$ at the pole, λ is longitude ($\lambda=0$ at midnight). The difference between the ionospheric φ_i and magnetospheric φ_m potentials is of the form

$$\varphi_i - \varphi_m = \frac{b}{\sigma_{\parallel}} j_{\parallel}, \quad (2)$$

where j_{\parallel} is the density of field-aligned currents. The ratio σ_{\parallel}/b is considered to be independent of the horizontal coordinates.

For the ionospheric potential we have the equation

$$\operatorname{div} \hat{\Sigma} \nabla \varphi_i = j_{\parallel}, \quad (3)$$

where $\hat{\Sigma}$ is the tensor of height-integrated ionospheric conductivity. For simplicity we assume that the ionosphere is a uniformly conducting plane; the magnetic field is perpendicular to the ionosphere. Hence Eq. (3) is rewritten as

$$\Sigma_P \Delta_{\perp} \varphi_i = j_{\parallel}, \quad (4)$$

where Σ_P is the Pedersen conductivity, Δ_{\perp} is the two-dimensional Laplace operator.

We are required to find the magnetospheric potential φ_m in the closed magnetic field line area ($r > r_0$) as well as the ionospheric potential φ_i . Equations (2) and (4) are insufficient for solving the problem. Just one more relationship between j_{\parallel} , φ_m and φ_i in the region $r > r_0$ is necessary. We shall consider two variants of the relationship. In the next section the simplest case when there are no current sources in the closed field line area is studied, hence $j_{\parallel} = 0$ for $r > r_0$. In the subsequent section, inertia currents are taken into account.

Calculation of the potential

Equations (2) and (4) yield

$$\Delta_{\perp} \varphi_i = k^2 (\varphi_i - \varphi_m), \quad (5)$$

where

$$k = \left(\frac{\sigma_{\parallel}}{\Sigma_P b} \right)^{1/2}. \quad (6)$$

Similar expressions were obtained by Chiu et al. (1981). In this section we neglect currents in the area of closed field lines. As a result, the equipotentiality of these lines takes place:

$$\varphi_m = \varphi_i \quad \text{for } r > r_0. \quad (7)$$

Substitution of φ_m from Eqs. (1) and (7) into Eq. (5), as well as the continuity of φ_i and $\partial \varphi_i / \partial r$, yield the ionospheric potential:

$$\begin{aligned} \varphi_i &= E_0 \left[r - \frac{2}{k} \frac{I_1(kr)}{I_0(kr_0)} \right] \sin \lambda \quad \text{for } r \leq r_0, \\ \varphi_i &= E_0 \frac{I_2(kr_0)}{I_0(kr_0)} \frac{r_0^2}{r} \sin \lambda \quad \text{for } r \geq r_0, \end{aligned} \quad (8)$$

where I is the modified Bessel function.

From Eqs. (1), (2), (7) and (8) we get the field-aligned current

$$\begin{aligned} j_{\parallel} &= -2 \Sigma_P E_0 k \frac{I_1(kr)}{I_0(kr_0)} \sin \lambda \quad \text{for } r < r_0, \\ j_{\parallel} &= 0 \quad \text{for } r > r_0. \end{aligned} \quad (9)$$

Substitution of Eqs. (9) and (2) into Eq. (6) yields the relation, valid for $kr_0 \gg 1$,

$$k \approx \frac{2E_0}{\varphi_{\parallel \max}}, \quad (10)$$

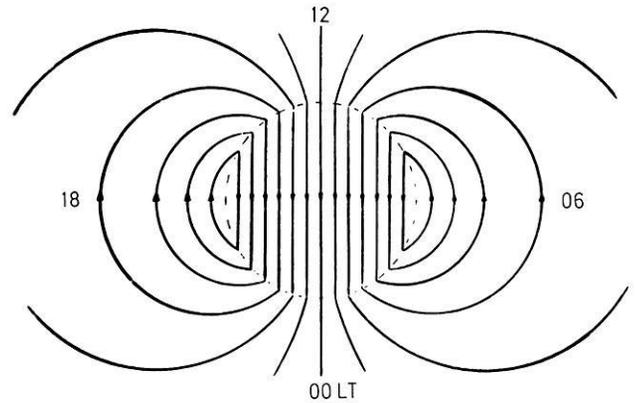


Fig. 1. The potential distribution in the ionosphere and magnetosphere in the case of infinitely large parallel conductivity. Arrows indicate the direction of the convective flow

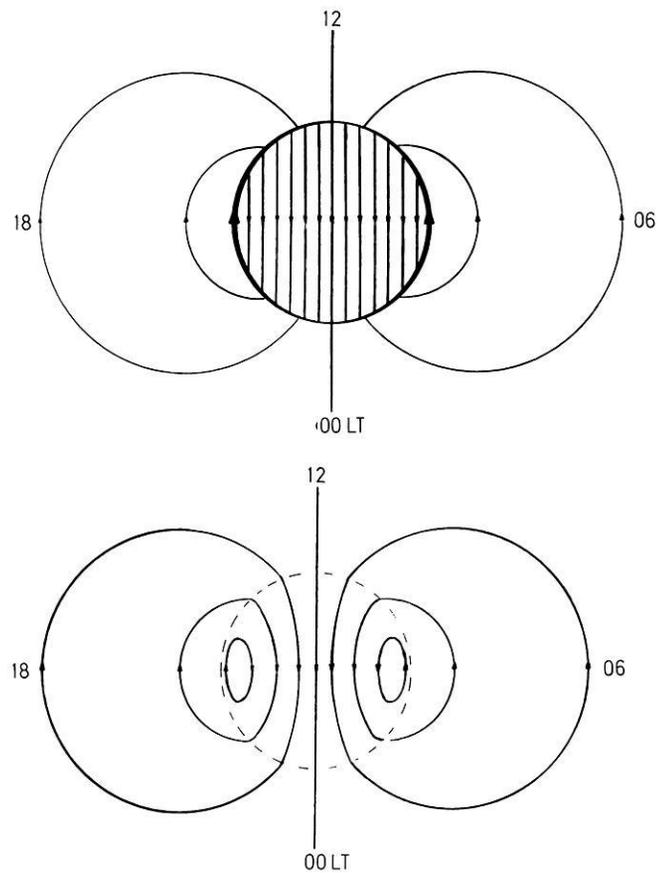


Fig. 2a, b. The potential distribution in the case of finite parallel conductivity: **a** the potential in the magnetosphere; **b** the potential in the ionosphere

where $\varphi_{\parallel \max}$ is the maximum magnitude of the parallel potential difference. Note that $\varphi_i - \varphi_m = \mp \varphi_{\parallel \max}$ at the points $r = r_0$, $\lambda = \pm \pi/2$, i.e. at the dawn and dusk boundaries of open magnetic field lines, respectively.

In Fig. 1 the potential distribution is shown for the case $\sigma_{\parallel} = \infty$. Potentials in the ionosphere and in the magnetosphere evidently coincide. The electric field is uniform inside the area of open magnetic field lines and resembles a field of a two-dimensional dipole outside this area.

In Fig. 2 the potential is shown for the case $kr_0 = 3$. It is interesting that the magnetospheric potential (Fig. 2a) has a discontinuity at the boundary of open and closed magnetic field lines. The magnitude of the discontinuity, taking Eq. (7) into account, is

$$\delta\varphi_m = \varphi_m|_{r=r_0-0} - \varphi_m|_{r=r_0+0} = (\varphi_m - \varphi_i)|_{r=r_0-0}, \quad (11)$$

i.e. it is equal to the parallel potential difference near the boundary of open field lines. A strong spike in the electric field $\delta E_{\perp} = \varphi_{\parallel}/\delta r$ occurs at the boundary. The characteristic scale δr is determined, in particular, by inertia of ions. Intense sunward convective streams are connected with the electric field spike.

The corresponding ionospheric potential is shown in Fig. 2b. The centres of the convective vortices appear to be shifted poleward. It is not difficult to show that the centres are situated at a distance (providing $kr_0 \gg 1$)

$$r \approx r_0 - \frac{\ln 2}{k}.$$

Influence of inertia currents

In the previous section, field-aligned currents in the area of closed magnetic field lines were neglected. As a result, an infinitely large spike of the magnetospheric electric field arose at the boundary of open and closed field lines. As seen from Fig. 2a, the convection lines cross the boundary. Plasma passing across the region of a strong electric field undergoes strong accelerations which give rise to inertia currents. The inertia currents, in their turn, produce additional small-scale field-aligned currents. Let us consider small-scale currents at the night side where the plasma drifts from the polar cap to the closed field line area.

The inertia current is determined by the well-known expression

$$\mathbf{j}_{\perp} = \frac{\rho c^2}{B^2} \frac{d\mathbf{E}}{dt}, \quad (12)$$

where ρ is the plasma density, B is the ambient magnetic field. The process is assumed to be stationary ($\partial/\partial t = 0$). The characteristic time of electric field variations in the coordinate system of drifting plasma is assumed to be large compared with the Alfvén resonance period, $T = 4l/V_A$, where l is the half-length of the magnetic field line and $V_A = B/\sqrt{4\pi\rho}$ is the Alfvén velocity. In this case we may neglect deformations of the magnetic field lines caused by the currents.

Integration of Eq. (12) along the magnetic field line from the finite parallel conductivity layer to the equatorial plane yields the total current transverse to a magnetic flux tube of unit ionospheric area

$$\mathbf{J}_{\perp} = \frac{c^2}{4\pi} (\mathbf{v}\nabla) \mathbf{E}_m \int_0^l \frac{dz}{V_A^2}. \quad (13)$$

The continuity condition gives the field-aligned currents at the ionosphere level to be

$$j_{\parallel} = \text{div } \mathbf{J}_{\perp} = -\frac{c^2}{4\pi} (\mathbf{v}\nabla) \Delta_{\perp} \varphi_m \int_0^l \frac{dz}{V_A^2}. \quad (14)$$

Combining Eqs. (2), (4) and (14), we get

$$\frac{c^2}{4\pi} \left(\int_0^l \frac{dz}{V_A^2} \right) (\mathbf{v}\nabla) \left(\frac{b}{\sigma_{\parallel}} \Delta_{\perp} j_{\parallel} - \frac{1}{\Sigma_p} j_{\parallel} \right) - j_{\parallel} = 0. \quad (15)$$

Let us suppose that the characteristic scale of changes of the current j_{\parallel} along r is much smaller than r . In this case, Eq. (15) is rewritten as

$$A \left(\frac{1}{k^2} \frac{d^3 j_{\parallel}}{dr^3} - \frac{dj_{\parallel}}{dr} \right) - j_{\parallel} = 0, \quad (16)$$

where

$$A = \frac{c^2}{4\pi} \frac{v_r}{\Sigma_p} \int_0^l \frac{dz}{V_A^2}. \quad (17)$$

Equation (16) will be solved under the following boundary conditions:

$$j_{\parallel}|_{r=r_0+0} = j_{\parallel}|_{r=r_0-0}, \quad (18)$$

$$j_{\parallel}|_{r \rightarrow \infty} = 0, \quad (19)$$

$$\int_{r_0}^{\infty} j_{\parallel} dr = 0. \quad (20)$$

Condition (18) is equivalent to the condition of absence of infinitely large spikes of the transverse electric field. Condition (20) means that the current in the magnetosphere does not flow from the open field line area into the closed field line area.

Let us assume $\Sigma_p = 1$ mho, $V_A = 10^3$ km/s, $T = 4l/V_A = 100$ s. For midnight, where $v_r \approx 0.5$ km, we get $A \approx 1$ km. For dawn and dusk meridians, where $v_r \approx 0$, we get $A \approx 0$. We shall study disturbances with horizontal dimensions much larger than 1 km, hence the second term in Eq. (16) may be neglected when compared with the third term. The solution of Eq. (16), under conditions (18)–(20), is

$$j_{\parallel}(r > r_0) = j_{\parallel}(r_0) \frac{2}{\sqrt{3}} \cdot e^{-(\alpha/2)(r-r_0)} \cos \left[\frac{\alpha\sqrt{3}}{2}(r-r_0) + \frac{\pi}{6} \right], \quad (21)$$

where

$$\alpha = \left(\frac{k^2}{A} \right)^{1/3}. \quad (22)$$

From Eq. (2), taking Eqs. (21) and (10) into account, we get the magnetospheric electric field (providing $kr_0 \gg 1$)

$$E_{mr} = E_{0r} \quad \text{for } r < r_0, \\ E_{mr} = -E_{0r} \left\{ \frac{r_0^2}{r^2} + \left(\frac{4\varphi_{\parallel \max}}{E_0 A} \right)^{1/3} \cdot \frac{2}{\sqrt{3}} e^{-(\alpha/2)(r-r_0)} \cos \left[\frac{\alpha\sqrt{3}}{2}(r-r_0) + \frac{\pi}{3} \right] \right\} \quad (23)$$

for $r > r_0$,

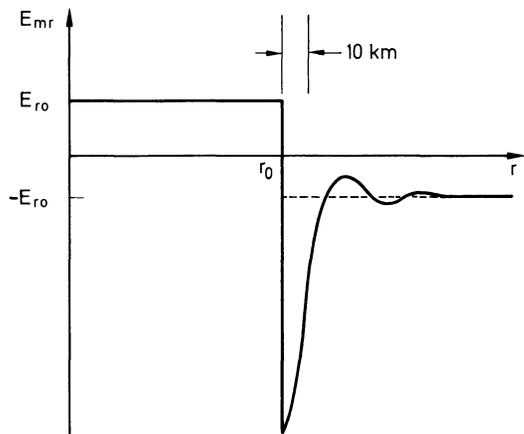


Fig. 3. The magnetospheric electric field distribution near the boundary of open and closed magnetic field lines

where $E_{0r} = -E_0 \sin \lambda$ is the radial component of the electric field in the polar cap. In Fig. 3 the $E_{mr}(r)$ dependence is shown. The following parameters are chosen: $\varphi_{\parallel \max} = 2$ kV, $E_0 = 30$ V/km, $A = 1$ km. In this case we have $k \approx 3 \times 10^{-2} \text{ km}^{-1}$, $\alpha \approx 10^{-1} \text{ km}^{-1}$. Quickly damping oscillations of the electric field arise near the boundary. The characteristic scale of oscillations is $\alpha^{-1} \approx 10$ km. The first spike is about six times the large-scale field. It should be noted that there are no small-scale spikes of the field in the ionosphere.

Oscillations shown in Fig. 3 are the result of superposition of Alfvén waves carried by the convection and undergoing repeated reflections from the conjugate hemispheres. The waves are damped due to ohmic losses in the ionosphere and in the finite parallel conductivity layer.

Discussion

In Fig. 4 the potential distribution in the model by Lyons (1980) (Fig. 4a) is compared with that in the present model (Fig. 4b). The solid line is the magnetospheric potential, the dashed line is the ionospheric one. The potential in Fig. 4a is continuous. The magnetospheric potential in Fig. 4b has a jump at the boundary of open and closed lines. The potential jump must be registered by satellites as a strong spike of the transverse electric field.

Spikes of the transverse electric field of the magnitude of several hundred millivolts per metre with a width of ~ 10 km are often observed at altitudes of 2,000–8,000 km (Mozer et al., 1980). Their generation mechanism is not clear because fields of such magnitude are observed neither in the ionosphere nor in the solar wind. A possible cause of the spikes is an inhomogeneity of field-aligned current flowing through the layer with finite parallel conductivity. In the open magnetosphere model the field-aligned current is, most likely, sharply inhomogeneous and, consequently, is capable of generating spikes of E_{\perp} .

Chmyrev et al. (1983) reported observations of two electric field spikes of magnitude ~ 250 mV/m with a width of 1–10 km at altitudes of 800–900 km. The electric spikes were accompanied by magnetic ones with an

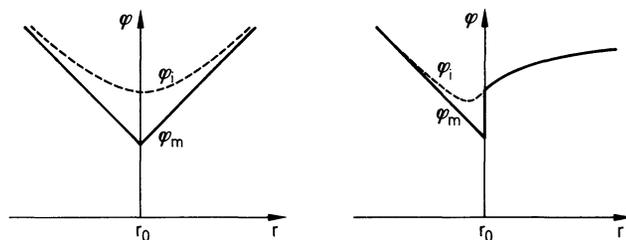


Fig. 4a, b. Distribution of the magnetospheric (solid lines) and ionospheric (dashed lines) potential in different models: **a** the model of Lyons (1980); **b** our model

amplitude of ~ 300 nT. Three analogous spikes of magnitudes of 200–500 nT and with a width of 2–8 km were observed at the same altitudes according to Volokitin et al. (1984). In all cases the magnetic disturbances were transverse to the ambient magnetic field, hence they were caused by field-aligned currents. It is not difficult to show that the field-aligned currents, Eq. (14), give the magnetic disturbance

$$B_{\lambda} = -\frac{4\pi \Sigma_p A}{c} \frac{\partial E_{mr}}{\partial r}. \quad (24)$$

Assuming $\Sigma_p = 1$ mho, $A = 1$ km, $\alpha = 0.3 \text{ km}^{-1}$, $E_{mr} = 200$ mV/m we get magnetic oscillations with amplitude $B_{\lambda} \approx 100$ nT. It should be mentioned that the magnetic spikes in the given model must be observed both above and below the finite parallel conductivity layer, whereas the electric spikes only above it.

Conclusions

A decrease of the parallel conductivity in an open magnetosphere model gives rise to the following consequences:

- 1) The centres of convective vortices are not located at the boundary of the open field line area. They are shifted into this area.
- 2) Narrow intense convective sunward streams arise above the finite parallel conductivity layer at the dawn and dusk boundaries of the closed field line area. The characteristic width of the streams is about 10 km. The electric field in the streams is several hundred millivolts per metre. Small-scale (~ 10 km) variations of field-aligned currents and of transverse magnetic disturbances with magnitude of ~ 100 nT are connected with the streams.

References

- Block, L.P., Fälthammer, C.-G.: Mechanisms that may support magnetic-field-aligned electric fields in the magnetosphere. *Ann. Géophys.* **32**, 161–174, 1976
- Chiu, Y.T., Newman, A.L., Cornwall, J.M.: On the structures and mapping of auroral electrostatic potentials. *J. Geophys. Res.* **86**, 10029–10037, 1981
- Chmyrev, V.M., Oraevsky, V.N., Isaev, N.V., Vilichenko, S.V., Stanev, G., Teodosiev, D.: Thin structure of intensive small-scale electric and magnetic fields in high-latitude ionosphere observed by “Intercosmos-Bulgaria 1300” satellite. Preprint IZMIRAN N° 25 (436), Moscow 1983

- Dungey, J.W.: Interplanetary magnetic field and the auroral zones. *Phys. Rev. Lett.* **6**, 47-48, 1961
- Heppner, J.P.: Electric field variations during substorms: OGO-6 measurements. *Planet. Space Sci.* **20**, 1475-1498, 1972
- Iijima, T., Potemra, T.A.: The amplitude distribution of field-aligned currents at northern high latitudes observed by Triad. *J. Geophys. Res.* **81**, 2165-2174, 1976
- Lyons, L.R.: Generation of large-scale regions of auroral currents, electric potentials, and precipitation by the convection electric field. *J. Geophys. Res.* **85**, 17-24, 1980
- Maltsev, Yu.P.: Effect of finite parallel conductivity on the magnetospheric convection. *Planet. Space Sci.* **33**, 493-498, 1985
- Mozer, F.S., Cattell, C.A., Hudson, M.K., Lysak, R.L., Temerin, M., Torbert, R.B.: Satellite measurements and theories of low altitude auroral particle acceleration. *Space Sci. Rev.* **27**, 155-213 1980
- Volokitin, A.S., Krasnoselskikh, V.V., Mishin, E.V., Tyurmina, L.O., Sharova, V.A., Shkolnikova, S.I.: On small-scale structure of intensive field-aligned currents in high latitudes. *Kosmicheskie issledovaniya* **22**, 749-755, 1984

Received December 3, 1984; Revised February 21, 1985
Accepted March 28, 1985

Evaluation of the effectiveness of theoretical model calculation in determining the plasmopause structure

K. Kurita* and M. Hayakawa

Research Institute of Atmospherics, Nagoya University, Toyokawa, Aichi, 442, Japan

Abstract. The relative position of the VLF/ELF emission region with respect to the plasmopause is of essential importance in studying their generation and propagation mechanism. On occasions when whistler data are not available, providing extensively the experimental determination of the plasmopause, we are obliged to rely on the theoretical model calculation or, alternatively, on the empirical formulas.

The present paper deals with the evaluation of the effectiveness of the use of a theoretical model calculation in estimating the plasmopause location with reference to its comparison with in-situ electron density measurements and empirical formulas, during a specific geomagnetic storm. It is concluded that the temporal evolution with the present theoretical calculation, under a more acceptable convection electric field model, would yield a sufficiently reliable value for the plasmopause configuration rather than the empirical formulas. It can be used in the study of wave-particle interactions when whistler data are not available and also in the study of the erosion of the plasmasphere itself.

Key words: Plasmopause – Plasmasphere – Magnetosphere – Convection – VLF/ELF emissions

Introduction

The plasmopause is known to play an important role in wave-particle interactions or in the generation of VLF/ELF and ULF emissions (Kaiser et al., 1977). When one studies the generation and propagation mechanism of VLF/ELF emissions based on ground-based measurements, the relative position of the emission occurrence region with respect to the plasmopause is of great importance (Kaiser and Bulough, 1975; Gendrin, 1975; Foster et al., 1976; Hayakawa et al., 1977; Hayakawa et al., 1981) and we normally deduce the plasmopause location by using whistler data simultaneously observed (Corcuff, 1975). However, on some occasions when the whistler data or in-situ measurement for the plasmopause are not available, we are obliged to make use of empirical formulas obtained at specific local times as a function of magnetic activity (Rycroft and Thomas, 1970; Carpenter and Park, 1973) or alternatively to theoret-

ically estimate the plasmopause structure based on the temporal K_p variation during the period preceding the observation.

The present paper is concerned with the calculation of the theoretical evolution of the plasmopause structure such as the erosion of the plasmasphere for a specific geomagnetic storm of 16–19th December, 1971, under a more realistic model for the convection electric field than the uniform field model previously adopted by Grebowsky (1970) and Chen and Wolf (1972). Then this temporal evolution of the theoretical plasmopause configurations is compared with the in-situ density measurement aboard the S³-A satellite (Maynard and Cauffman, 1973) and with empirical formulas, in order to evaluate the effectiveness of the use of the present theoretical model calculation in approximating the actual plasmopause location.

Theoretical calculation of the plasmopause structure in a time-dependent convection electric field

The most important quantity in determining the plasmopause configuration and particle trajectory is the model of convection electric field. A semi-empirical model for the convection electric field has been proposed by Volland (1973) and Stern (1974, 1975), who assumed that the electrostatic field could be described by a scalar potential ϕ_E of the following form.

$$\phi_E = A R^\gamma \sin \phi - 91.5 R_e/R, (\text{kV}) \quad (1)$$

$$E = -\nabla \phi_E \quad (2)$$

where ϕ is the local time measured from the midnight sector, R the radial distance from the Earth and R_e the Earth's radius. The first and second terms in Eq. (1) represent the convection and corotating electric field, respectively. the parameter γ in the convection electric field indicates the degree of screening of the field from the inner magnetosphere. $\gamma=1$ means a uniform electric field, which was adopted in the previous plasmopause calculation by Grebowsky (1970) and Chen and Wolf (1972). $\gamma=2$ is adopted in the present paper as the optimum value, based on the theoretical evidence that the convection electric field is partially shielded from the inner magnetosphere (Volland, 1973; Jaggi and Wolf, 1973; Southwood, 1977) and also on the comparison with a variety of observational results on particle injections (Ejiri et al., 1978; Kivelson et al., 1978; Kaye and Kivelson, 1979). The amplitude factor A

* Present address: Kawasaki Heavy Industries, Kakamihara, Gifu, 504, Japan

Offprint requests to: M. Hayakawa

for the convection field in Eq. (1) should be expressed as a function of K_p index and we have utilized the expression by Maynard and Chen (1975), whose relationship of A with K_p index is given in the following, based on the observational data aboard S³-A satellite.

$$A = \frac{0.045}{(1 - 0.159 K_p + 0.0093 K_p^2)^3} \quad (\text{kV}/R_e^2). \quad (3)$$

The magnetic field is assumed to be that of a centred dipole model. The trajectories of energetic particles in these field configurations can be traced by a method similar to that developed by Ejiri (1978), on which the calculation of plasmopause structure is essentially based.

Much more complete theoretical models have in fact been constructed (e.g. Blanc, 1983; Spiro et al., 1981). These more elaborate physical models exhibit some features that are not present in the simple Volland-Stern model, but do seem to be present in the observations. Of course, they are too complicated to be useful for the purpose of this paper; namely, an establishment of a simple, easy-to-use method of estimating the plasmopause position for a magnetospheric event. The Volland-Stern model remains the best semi-empirical formula for estimating the electric field in the inner magnetosphere.

The method of calculating the temporal evolution of the plasmopause is based on the original idea by Grebowsky (1970) and Chen and Wolf (1972), which traces the trajectory of cold electrons (energetic electrons with zero energy) backward in time from the universal time of interest. Since the flux tubes are depleted of plasma when they open to interplanetary space and are effectively filled with plasma when they are closed, the density in a closed flux tube at a specific time will depend roughly on the total time the flux tube has been closed and on the dayside of the Earth where solar ionization produces the plasma. By computing this "dayside closure time" at all (L , LT) coordinates, we can deduce the plasmopause structure in the (L , LT) space, on the criterion that the dayside closure time is more than 5 days and the field lines corresponding to L greater than 10 are open. The previous authors (Grebowsky, 1970; Chen and Wolf, 1972; Grebowsky et al., 1974) adopted the spatially invariant ($\gamma = 1$) convection electric field, whose magnitude is varied in step with K_p . While, in the present paper, we choose the more acceptable model for the convection electric field, as discussed previously.

Empirical formulas for the plasmopause location

Two main formulas for the estimation of the plasmopause location have been proposed by Rycroft and Thomas (1970)

and Carpenter and Park (1973). Rycroft and Thomas (1970) derived the following empirical formula for the plasmopause location at a specific L.T. sector (L.T. = 0 h, midnight), based on the measurement of the midlatitude trough of the electron density on spacecrafts.

$$L_p = 5.64 - (1.09 \pm 0.22) \sqrt{K_p}, \quad (4)$$

where K_p is the value at the observation time. So this formula is seen to include no information of the past history of the geomagnetic activity.

On the other hand, using the statistics of the plasmopause estimation by means of ground-based knee whistlers, Carpenter and Park (1973) deduced the following empirical formula of the plasmopause at dawn (L.T. = 6 h).

$$L_p = 5.7 - 0.47 K_{pm}, \quad (5)$$

where K_{pm} is the maximum K_p value during 12 h preceding the observation time, and their formula includes, in part, a hysteresis effect as is clearly involved in the theoretical plasmopause.

Higel and Lei (1984) have recently made a quite thorough study of the plasmopause characteristics based on the statistical investigation using the GEOS density measurements and, furthermore, they have discussed the previous empirical formulas of the plasmopause locations, as discussed here. Readers are advised to consult their paper for a detailed review.

Theoretical plasmopause structure during the geomagnetic storm (16–19th December 1971) and its comparison with in-situ density measurements and the empirical formulas

Figure 1 illustrates the temporal evolution of the geomagnetic activity measured by K_p index during the storm of 16–19th December, 1971 and the interval preceding it. Figure 2(a)–(j) shows the successive variations of the theoretically calculated plasmopause configuration with an interval of 6 h. At 18 h on the 16th (times are all in U.T.) the plasmopause is found to be nearly circular with its L value around 4.0 (Fig. 2a). The shape begins to change noticeably after 6 h (i.e. 0 h on the 17th in Fig. 2b) such that the plasmasphere exhibits a bulge in the afternoon sector. As the bulge rotates from the afternoon to dusk sector (Fig. 2d), the bulge becomes a thin plasma-tail. At 18 h on the 17th and 0 h the 18th (Fig. 2e and f) when we have the maximum K_p value of 7, the plasma-tail appears at the earlier L.T. sector of 14–15 h and wraps round the main body of the plasmasphere (as in Fig. 2g–j) and corotates with the Earth.

For the geomagnetic storm we are dealing with, the

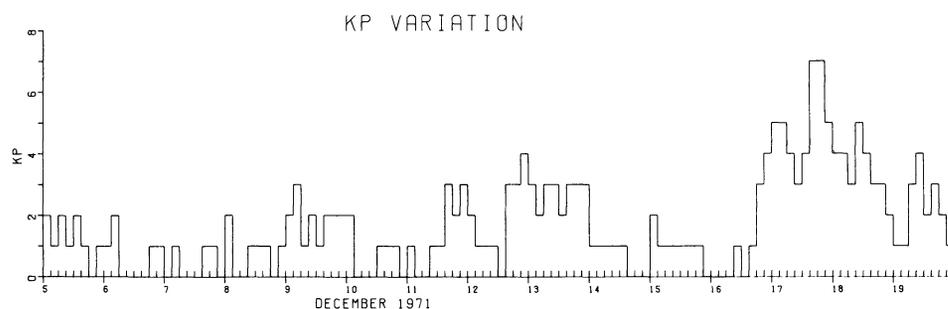


Fig. 1. Temporal evolution of K_p index. There are geomagnetic storms during the period from 16th to 19th December, 1971

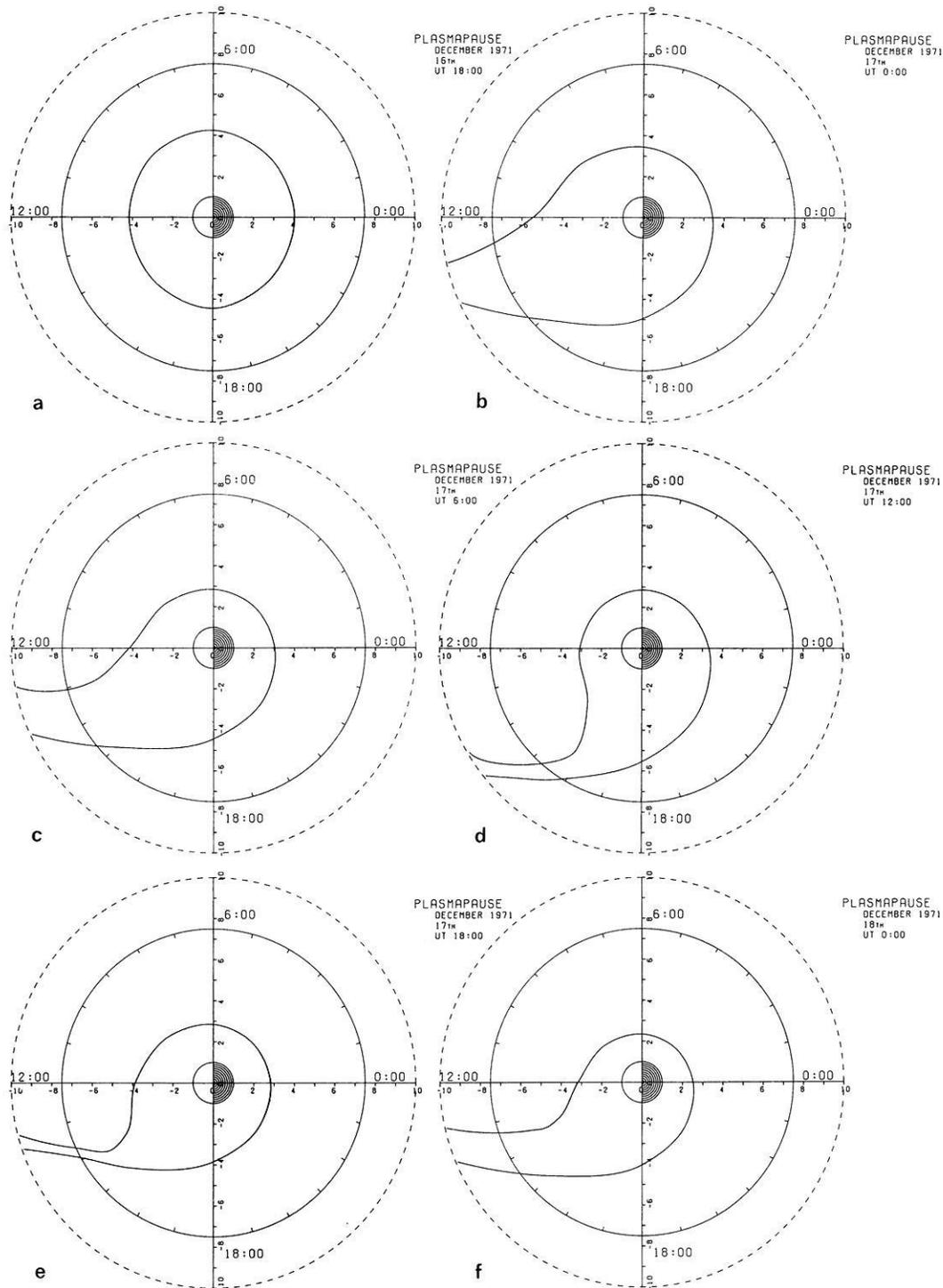


Fig. 2a-j. Temporal evolution of theoretically calculated plasmopause structure with an interval of 6 hours from 18 h on 16th (U.T.) (a) to 0 h on 19th (j)

data on the equatorial electron density are available from the S³-A satellite (Maynard and Cauffman, 1973). The double floating probe measurement designed to measure dc electric fields is also used as a crude plasmopause detector. Figure 3 presents the comparison of the plasmopause location determined by the S³-A satellite (in full lines) and by the theoretical model calculation (in chain lines) at the two local times, L.T.=18 h, dusk and L.T.=0 h, midnight. One can find two different values of the plasmopause at some

times on the full lines, which may be the consequence of the two-step structure as discussed by Chappell (1972) or the existence of the detached plasma. While the similar two values on the chain line result from the plasma-tail as found in Fig. 2. The following features have emerged from the comparison.

A comparison of the plasmopause at the dusk sector indicates that the *L* value of the theoretical plasmopause is, on some occasions, larger than the experimental one

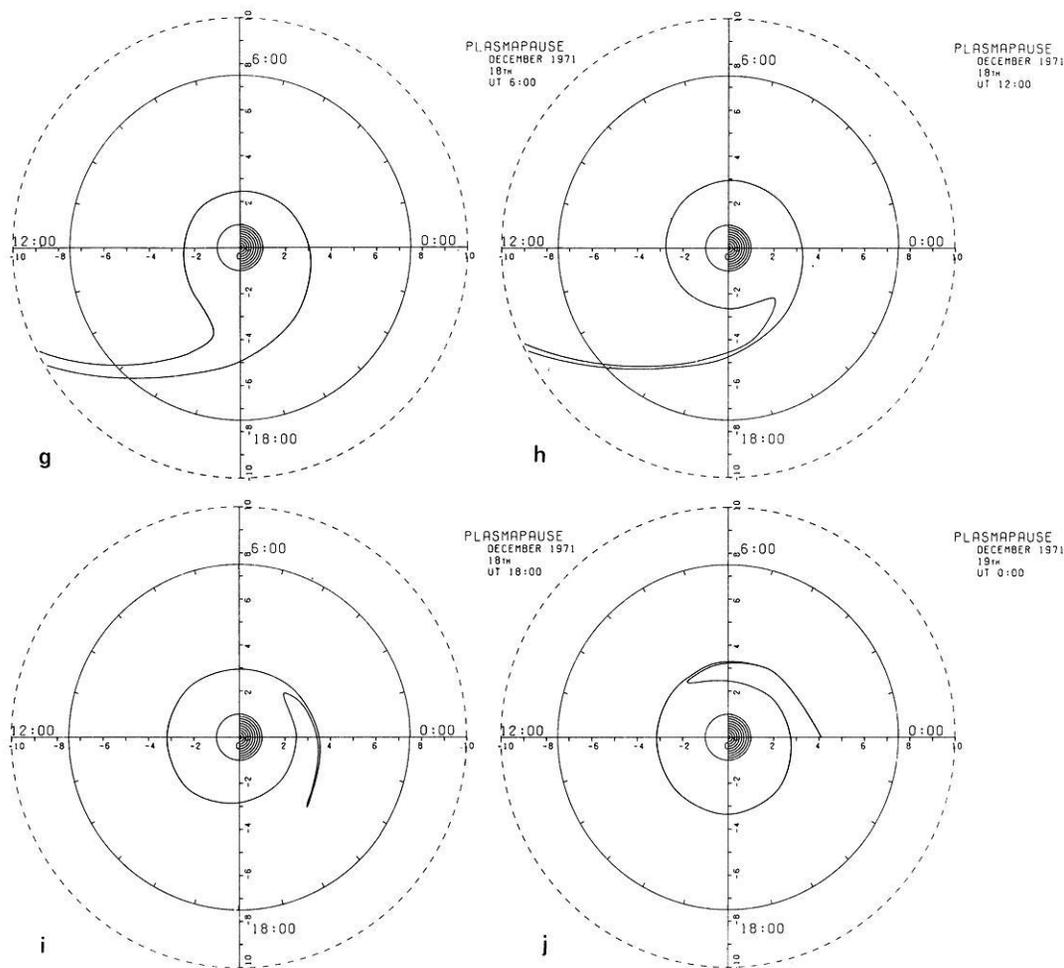


Fig. 2g-j.

and we have the reverse situation on other occasions. However, considering that dusk is the sector exhibiting the most complicated behaviours, we can say that the theoretical model calculation and the in-situ result are, as a general tendency, in good agreement. The discrepancy is not larger than $1.0 R_e$ and it is normally about $0.5 R_e$, except at the orbit numbers of 103 and 104.

A comparison of the temporal variation of the plasma-pause locations in the midnight sector indicates that both of the two plasma-pause locations show an excellently parallel variation. This very nature is very important, suggesting that the theoretical model calculation will yield a very reliable measure of the plasma-pause location in the midnight sector. The experimental value is always larger than the theoretical one by no more than $1.0 R_e$ and normally about $0.5 R_e$. Hence, in this example, the experimental plasma-pause at midnight can be estimated by adding $\sim 0.5 R_e$ to the theoretical one.

The theoretical model calculation is now compared with the empirical formulas. Figure 4 illustrates the comparison with the formula, Eq. (4), by Rycroft and Thomas (1970) in the midnight sector (L.T.=0 h) indicated by two full lines. The chain line refers to the theoretical model calculation. The empirical formula, itself, has an uncertainty range as given by Eq. (4). At the two times; 18 h on the 18th and 0 h on the 19th, the theoretical plasma-pause of the

main body of the plasmasphere is found to the outside the uncertainty range; however, at other times, the theoretical plasma-pause is found to be nearly within that range. An excellent parallel nature is recognizable for midnight between the theory and in-situ result in Fig. 3. So, if we find a highly parallel variation between the empirical and observational plasma-pause, then we would expect, consequently, a parallel relationship between the theoretical and empirical plasma-pause in Fig. 4. However, the parallel nature becomes much less obvious, since we take, in Fig. 4, the smaller L values for the last two times corresponding to the plasma-pause of the main body of the plasmasphere. This means that the use of the present theoretical model calculation would provide a more reliable plasma-pause location at the midnight sector than the empirical formula.

Next, the dawn-side plasma-pause is compared (Fig. 5) between the theoretical estimation (in chain line) and the empirical formula, Eq. (5), by Carpenter and Park (1973) (in full line). Very good agreement is obtained in the period of the main and early recovery phase of the storm (i.e. 18 h on 17th to 6 h on the 18th). Otherwise, their empirical value seems to be $\sim 0.5 R_e$ larger than the theoretical value. The general impression is that the parallel nature of the theoretical plasma-pause with the empirical one seems to be improved in Fig. 5, which might be associated with the inclusion of a hysteresis effect in their formula.

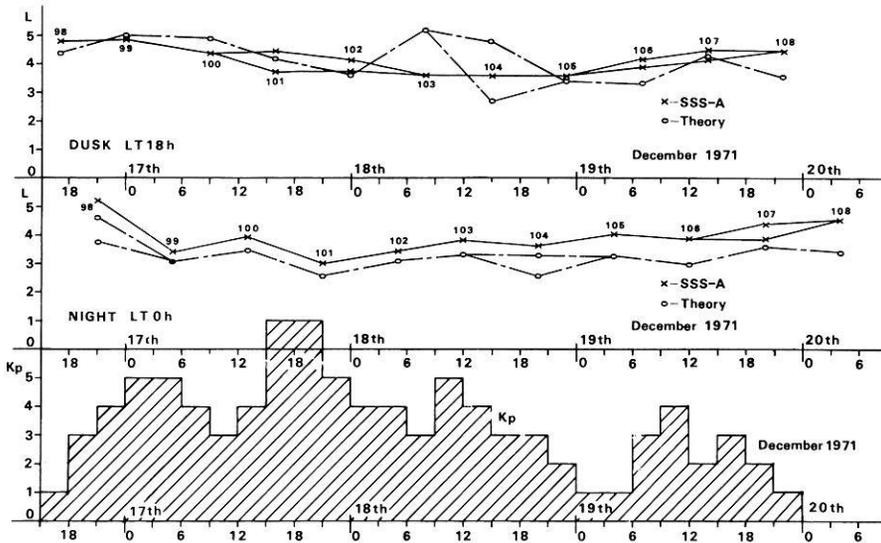


Fig. 3. Comparison of the theoretically calculated plasmopause with the in-situ electron density measurement by S³-A satellite (after Maynard and Cauffman, 1973) at two different local times. The full line linked by crosses refers to the satellite measurement, and the chain line linked by open circles indicates the theoretical plasmopause

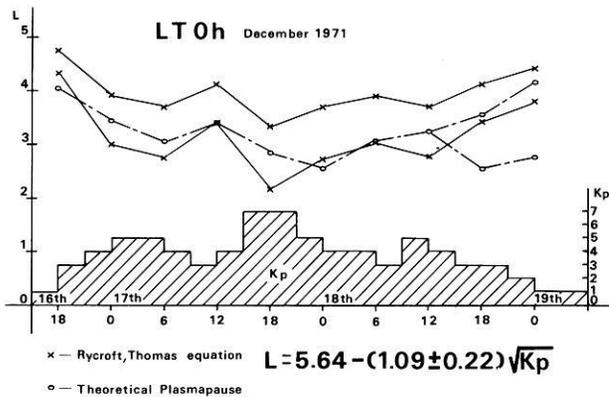


Fig. 4. Comparison of the theoretically calculated plasmopause with the empirical formula at the midnight sector. The chain line connected by open circles corresponds to the theoretical plasmopause and the two full lines connected by crosses refer to the boundaries of the empirical formula by Rycroft and Thomas (1970)

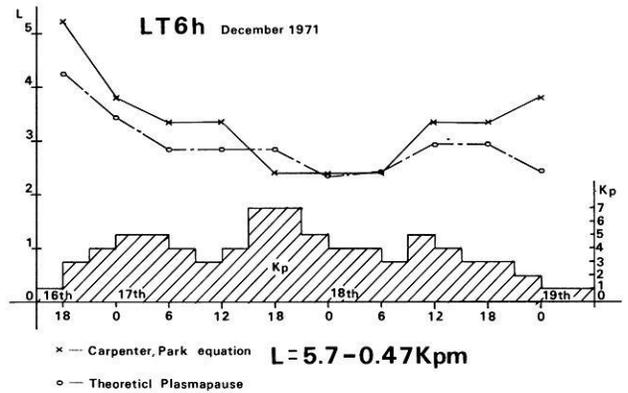


Fig. 5. Comparison of the theoretical plasmopause with the empirical formula at the dawn sector. The chain line has the same meaning as in Fig. 4, while the full line connected by crosses indicate the Carpenter and Park (1973) 's formula with taking into account a hysteresis effect

Concluding remark

The plasmopause configuration can be theoretically calculated under the more realistic and acceptable convection electric field model than the previous spatially invariant model. This approach is useful in describing the erosion of the plasmasphere during a magnetospheric event. The present study is based on a comparison of the theoretical plasmopause with the in-situ density measurement for one specific geomagnetic storm only. However, the study seems to show that the theoretical calculation, taking into account the past history of K_p index, provides a considerably reliable plasmopause location and hence it can be utilized in specifying the plasmopause location in VLF/ELF emission studies even when the simultaneous whistler data are not available.

References

- Blanc, M.: Magnetospheric convection effects at mid-latitudes. 3. Theoretical derivation of the disturbance convection pattern in the plasmasphere. *J. Geophys. Res.* **88**, 235–251, 1983
 Carpenter, D.L., Park, C.G.: On what ionospheric workers should know about the plasmopause-plasmasphere? *Rev. Geophys. Space Phys.* **11**, 133–154, 1973

- Chappell, C.R.: Recent satellite measurements of the morphology and dynamics of the plasmasphere. *Rev. Geophys. Space Phys.* **10**, 951–979, 1972
 Chen, A.J., Wolf, R.A.: Effects on the plasmasphere of a time-varying convection electric field. *Planet. Space Sci.* **20**, 483–509, 1972
 Corcuff, Y.: Probing the plasmopause by whistlers. *Ann Geophys.* **31**, 53–68, 1975
 Ejiri, M.: Trajectory traces of charged particles in the magnetosphere. *J. Geophys. Res.* **83**, 4798–4810, 1978
 Ejiri, M., Hoffman, R.A., Smith, P.H.: The convection electric field model based on Explorer 45 observations. *J. Geophys. Res.* **83**, 4811–4815, 1978
 Foster, J.C., Rosenberg, T.J., Lanzerotti, L.J.: Magnetospheric conditions at the time of enhanced wave-particle interactions near the plasmopause. *J. Geophys. Res.* **81**, 2175–2182, 1976
 Gendrin, R.: Is the plasmopause a preferential region for proton precipitation? *Ann. Géophys.* **31**, 127–136, 1975
 Grebowsky, J.M.: Model study of the plasmopause motion. *J. Geophys. Res.* **75**, 4329–4333, 1970
 Grebowsky, J.M., Tulunay, Y.K., Chen, A.J.: Temporal variations in the dawn and dusk midlatitude trough and plasmopause position. *Planet. Space Sci.* **22**, 1089–1099, 1974
 Hayakawa, M., Bullough, K., Kaiser, T.R.: Properties of storm-time magnetospheric VLF emissions as deduced from the Ariel

- 3 satellite and ground-based observations. *Planet. Space Sci.* **25**, 353–368, 1977
- Hayakawa, M., Tanaka, Y., Iwai, A., Ohtsu, J., Storey, L.R.O., Béghin, C., Jorgensen, T.S.: Simultaneous spaced direction finding measurement of medium latitude VLF/ELF emissions. *Planet. Space Sci.* **29**, 505–520, 1981
- Higel, B., Lei, W.: Electron density and plasmopause characteristics at $6.6 R_E$: A statistical study of the GEOS 2 relaxation sounder data. *J. Geophys. Res.* **89**, 1583–1601, 1984
- Jaggi, R.K., Wolf, R.A.: Self-consistent calculation of the motion of a sheet of ions in the magnetosphere. *J. Geophys. Res.* **78**, 2852–2866, 1973
- Kaiser, T.R., Bullough, K.: VLF and ELF emissions. *Ann. Géophys.* **31**, 137–141, 1975
- Kaiser, T.R., Orr, D., Smith, A.J.: Very low frequency electromagnetic phenomena: Whistlers and micropulsations. *Phil. Trans. Roy. Soc. London* **B279**, 225–235, 1977
- Kaye, S.M., Kivelson, M.G.: Time dependent convection electric fields and plasma injection. *J. Geophys. Res.* **84**, 4183–4188, 1979
- Kivelson, M.G., Kaye, S.M., Southwood, D.J.: The physics of plasma injection events. UCLA Inst. Geophys. Planetary Phys., Publication No. 1879. Prepared for presentation at the AGU Chapman conference on Magnetospheric Substorms, 1978
- Maynard, N.C., Cauffman, D.P.: Double floating probe measurements on S³-A. *J. Geophys. Res.* **78**, 4745–4750, 1973
- Maynard, N.C., Chen, A.J.: Isolated cold plasma regions: Observations and their relation to possible production mechanisms. *J. Geophys. Res.* **80**, 1009–1013, 1975
- Rycroft, M.J., Thomas, J.O.: The magnetospheric plasmopause and the electron density trough at the Alouette 1 orbit. *Planet. Space Sci.* **18**, 65–80, 1970
- Southwood, D.J.: The role of hot plasma in magnetospheric convection. *J. Geophys. Res.* **82**, 5512–5520, 1977
- Spiro, R.W., Harel, M., Wolf, R.A., Reiff, P.H.: Quantitative simulation of a magnetospheric substorm 3. Plasmaspheric electric fields and evolution of the plasmopause. *J. Geophys. Res.* **86**, 2261–2272, 1981
- Stern, D.P.: Models of the Earth's electric field. NASA/GSFC Document 602-74-160, May, 1974
- Stern, D.P.: The motion of a proton in the equatorial magnetosphere. *J. Geophys. Res.* **80**, 595–599, 1975
- Volland, A.: A semiempirical model of large-scale magnetospheric electric fields. *J. Geophys. Res.* **78**, 171–180, 1973

Received July 26, 1984; Revised version Febr. 27, 1985

Accepted April 10, 1985

First results and preliminary interpretation of deep-reflection seismic recordings along profile DEKORP 2-South

DEKORP Research Group

Contributors: **R.K. Bortfeld, J. Gowin, M. Stiller** (Institut für Geophysik, Arnold-Sommerfeld-Str. 1, 3392 Clausthal-Zellerfeld); **B. Baier** (Institut für Geophysik, Feldbergstr. 47, 6000 Frankfurt/M.); **H.J. Behr, T. Heinrichs** (Institut für Geologie und Dynamik der Lithosphäre, Goldschmidtstr. 3, 3400 Göttingen); **H.J. Dürbaum, A. Hahn, C. Reichert, J. Schmoll** (Bundesanstalt für Geowissenschaften und Rohstoffe/Niedersächsisches Landesamt für Bodenforschung, Stilleweg 2, 3000 Hannover 51); **G. Dohr** (Preussag AG, Arndtstr. 1, 3000 Hannover 1); **R. Meissner, R. Bittner, B. Milkereit** (Institut für Geophysik, Olshausenstr. 40–60, 2300 Kiel 1); **H. Gebrande** (Institut für Allgemeine und Angewandte Geophysik, Theresienstr. 41/IV, 8000 München 2)

Abstract. In 1984 the DEKORP reflection seismic crustal studies have been started on a large scale in the Federal Republic of Germany as a contribution to the International Lithosphere Project, the European Geotraverse and the German Deep Drilling Program. The first 250 km long profile DEKORP 2-South has been observed between the Nördlinger Ries and the Taunus crossing the boundaries between the Variscan belt units Moldanubian, Saxothuringian and Rhenohercynian. This paper gives an overview of the measurements and data processing. From the steep-angle observations on the main profile a seismic section was prepared which is dominated by diffraction events. These events are analyzed with respect to velocities and structural information. Also wide-angle observations on the main profile and on a parallel profile have been carried out. A geological/geophysical interpretation of the near-vertical reflection data has been attempted and is presented in this paper: there are strong indications of large-scale horizontal tectonics affecting the upper and lower crust down to the Moho depth of 26–29 km.

Key words: Deep reflection seismic profiling – DEKORP – Mid-European Variscides – Crustal structure – Diffraction clusters – Ductile shear tectonics – Intracrustal thrust system – Mid German Crystalline High

1. Introduction and concept of DEKORP

1.1 Overview

DEKORP = *Deutsches kontinentales reflexionsseismisches Programm* (German continental reflection seismic program) was initiated and planned by a group of earth scientists from the Federal Republic of Germany in 1982. It is tightly connected to the German contribution to the International Lithosphere Program, especially to the German Deep Drilling Program (KTB = *Kontinentales Tiefbohrprogramm*) (Althaus et al., 1984) and the *European Geotra-*

verse (EGT) (Giese, 1983). DEKORP has been conceived as a network of vertical reflection seismic traverses crossing the major geotectonic units in roughly strike and dip direction and thus contributing to the German part of the EGT study area.

The main objectives of DEKORP are:

- A detailed investigation of the Earth's crust in Germany by seismic reflection measurements along seismic profiles of great length.
- A combination of the reflection measurements with other geophysical and geological studies in order to reveal the tectonic structure and development of the Variscan orogenesis.

In the northern part of Germany, in the Upper Rhine Valley and in the Alpine Molasse Basin many seismic reflection lines have been observed in the last 40 years mostly in prospecting for hydrocarbons, some of them collecting additional valuable information down to more than 30 km (Dohr, 1957a; Dohr, 1957b; Bally, 1983). In the Variscan area, i.e. the German Mittelgebirge, abundant seismic refraction profiles exist (Giese et al., 1976). Also some short special crustal reflection profiles have been observed in this area in the last 20 years (Angenheister and Pohl, 1969; Dohr and Meissner, 1975; Bartelsen et al., 1982; Meissner et al., 1982, 1983). The planned profile network of DEKORP with its long lines should provide a regional overview of the crustal structure, define the various tectonic boundaries and combine the information from the north with those in the southern part. The DEKORP network is shown in Fig. 1 together with some of the previously observed shorter reflection profiles and some general geological information. The DEKORP profiles are arranged about parallel and perpendicular to the strike direction of the Variscan mountains. They have a typical length of several hundreds of kilometers. A close cooperation between industry and research institutes is considered an essential requirement.

The following gives a rough idea of the structure of DEKORP's organization: All important decisions are taken by DEKORP's steering committee consisting of

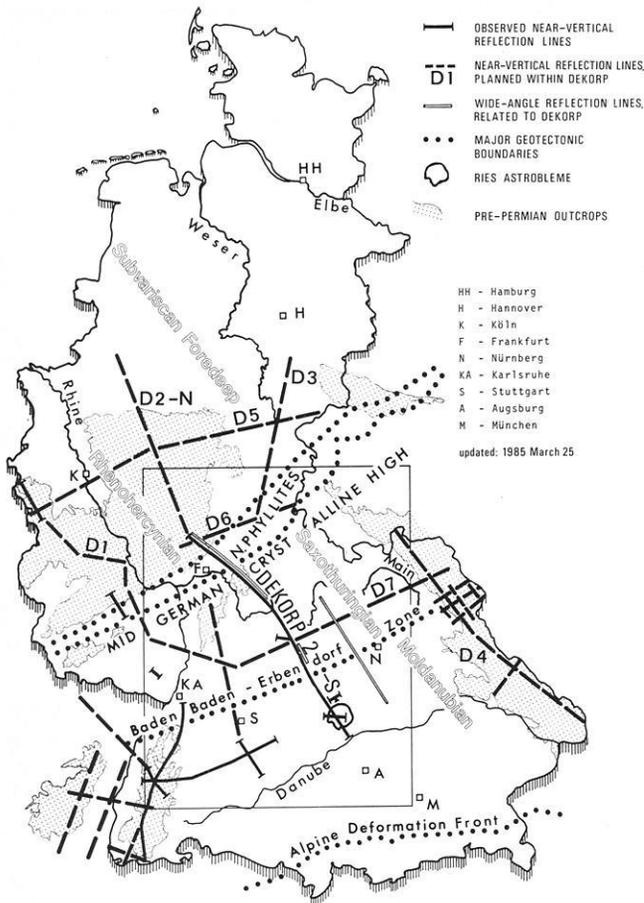


Fig. 1. Location map of observed and planned near-vertical reflection lines in the German Variscides

members from universities, industry and geological surveys. The execution is controlled by a project management group. The field work is carried out by a contractor and various groups of scientists from universities and geological surveys. Most of the data processing is done at the DEKORP Processing Center (DPC) at the Technische Universität, Clausthal. The detailed planning of the profiles and the interpretation of the results is performed by regional working groups set up by the steering committee. Also various scientists of research institutes are working on the data in order to improve methods of processing and analysis. Each year the results are presented at a DEKORP workshop. The entire project is under the auspices of the Geological Survey of Lower Saxony (NLFb), Hannover, and is financed by the Ministry of Research and Technology (BMFT), Bonn.

1.2 Profile DEKORP 2-S and the accompanying seismic activities

The first DEKORP line with a total length of 250 km was observed between April 4 and May 17, 1984. It represents the southern part of the central cross-strike profile called DEKORP 2-S(outh).

From the Danube in the SSE to the Taunus mountains in the NNW, it crosses the Ries astrobleme, the boundary between Moldanubian and Saxothuringian, the Spessart

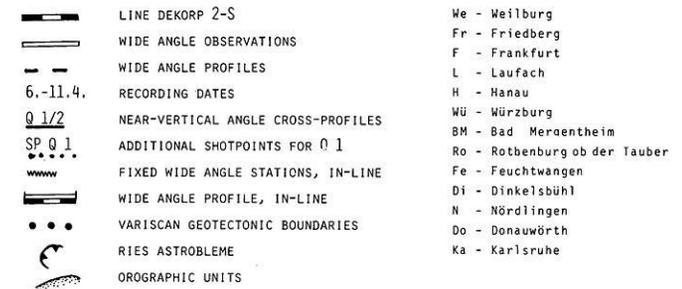
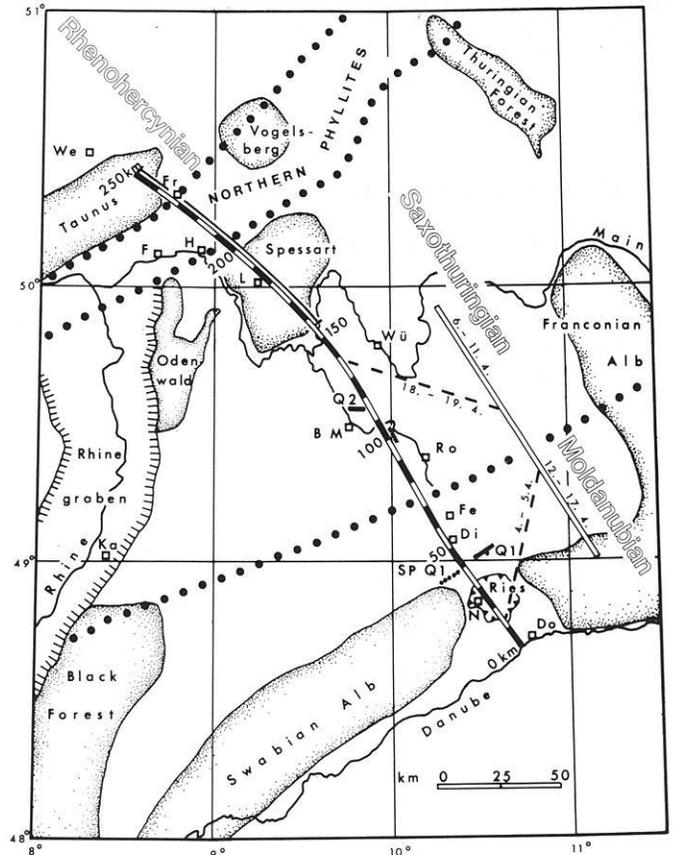


Fig. 2. Location map of the DEKORP 2-S main profile and related seismic measurements

mountains, the Hessian trough and the boundary between Saxothuringian and Rhenohercynian. Figure 2 shows the location map of DEKORP 2-S and the sites of the additional observations to be described later. With exception of a 5-day-experiment with VIBROSEIS in the northern Spessart, explosives in shallow boreholes were used throughout the operation as sources of seismic energy. Charges of 30 kg were fired at depths of 30 m. Average spacing between shotpoints was 320 m. Larger charges up to 90 kg were used for every 4th shot in the southern part of the line and for every 10th shot in the northern part in order to provide sufficient energy for the wide-angle observations.

On the receiving side, 24 geophones per trace in a linear configuration with a total length of 80 m parallel to the profile were used. 200 geophone groups with a spacing of 80 m resulted in a 16 km long geophone spread corresponding to a 1:2 ratio with regard to the relation between spread length and investigation depth, thus providing a basis for the calculation of crustal velocities and a theoretical 25

fold coverage in the subsurface. Data were digitized, multiplexed and telemetrically transmitted by a 16 km long cable to the recording truck. The field work was performed by Prakla-Seismos GmbH, Hannover; the recording unit was a 200 channel Sercel SN 348. For the communication between the up to 100 members of the field crew special transmitter and relay stations had to be installed along the profile.

The additional seismic measurements to be described in detail in chapter 4 consisted of

- Near-vertical observations along perpendicular and oblique reflection spreads in the vicinity of the main profile by means of reflection equipment of university institutes and other research institutions (Sect. 4.1; Fig. 2, profile Q₁ and Q₂).
- In-line wide-angle observations along the main profile by means of conventional reflection equipment (Sect. 4.2).
- In-line and off-line wide-angle observations with portable refraction stations. In the southern part of the DEKORP line a specially stacked wide-angle profile parallel to the main profile has been constructed and, moreover, some in-line observations with fixed stations have been carried out (Sect. 4.3). In the northern part of the profile semi-continuous in-line observations were performed (Sect. 4.4). 80 observers mainly from universities took part in these accompanying observations.

A comparative test using explosives and Vibroseis simultaneously along the same part of the main profile was performed in the crystalline part of the Spessart. Although the use of explosives provided a higher signal to noise ratio for a given coverage than the Vibroseis method, the quality of the Vibroseis record section was considered sufficient to warrant a successful use of the Vibroseis source in future deep reflection surveys, especially in view of a significant improvement by means of higher coverage. On the basis of this experience the Vibroseis technique was chosen for the deep reflection presite survey of the KTB location "Black Forest" in autumn 1984. Also many parameters (e.g. sweep characteristics and number of vertical stacks) tested in the Spessart could be used immediately in the Black Forest and yielded good results.

1.3 Geological situation and objectives of profile DEKORP 2-S

It was the immediate goal of our studies to investigate the crustal structure of the Saxothuringian Zone (ST) and of its transition into the adjacent Moldanubian (MN) and Rhenohercynian Zones (RH) of the Variscan Belt. Today's definition of these zones on the basis of their distinct tectonothermal evolution originally goes back to Kossmat (1927, see also Behr et al., 1984).

The RH is characterized by its continuous sedimentary cover of a mildly folded northwest-facing Paleozoic sequence that is anchimetamorphic or just reaches a very low grade of metamorphism.

The ST exhibits a structural pattern of high mobility. The zone contains remnants of an Upper Proterozoic to Lower Carboniferous cover of varying vergence and generally of very low to medium grade metamorphism together with largely allochthonous polymetamorphic inliers, some of them including granulites and high pressure eclogites.

Most of the MN is made up of polymetamorphic high grade gneisses pervaded by a multitude of late to post-

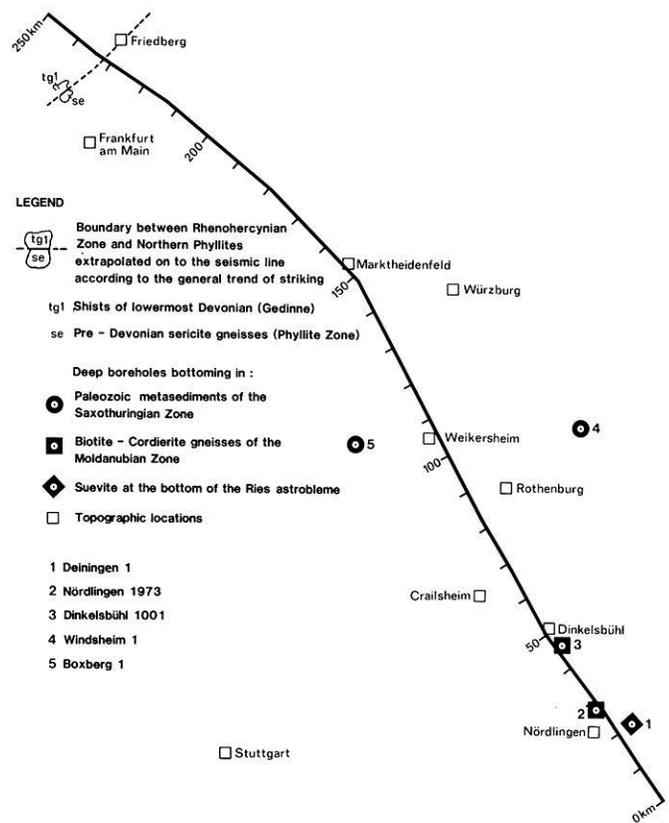


Fig. 3. Location map of important drillings confining the position of the suture between Moldanubian and Saxothuringian zone

kinematic granitoids. Its relic suprastructure consists of usually slightly deformed paleozoic sediments.

During the last decade the concepts of nappes and horizontal tectonics for the RH-, ST- and MN-Zones of the Variscan belt have been revived (Ziegler, 1978; Martin and Eder, 1983; Behr et al., 1984). Certainly, also near-vertical normal faulting is observed along the profile, and for quite a long time it has been attempted to explain the Variscan tectonic processes as mainly vertical tectonics with no considerable crustal shortening. It is one of the main subjects of this seismic experiment to obtain as much evidence as possible about the crustal structure in this area in order to support or modify the current geological concepts of the tectonics of the Variscan belt.

DEKORP 2-S traverses the Franconian Platform (Carlé, 1955; Ziegler, 1982) from south of the Ries astrobleme toward northwest ending beyond the northeast branch of the Rhine Graben within the Taunus mountains. Along the seismic line the Variscan is covered by up to 1.6 km of Permian and younger sediments except for the Ries, the Northern Spessart and the Taunus mountains. Within the Ries astrobleme the impact excavated crystalline basement slivers of the Moldanubian Zone. In the Northern Spessart late Variscan uplift and Tertiary rejuvenation exposed a segment of the ST that represents part of the Mid German Crystalline High (MGCH). Similar processes led to the exposure of the southern margin of the RH in the Taunus mountains at the northern end of the profile.

Additional information from deep wells penetrating the Mesozoic platform cover constrains the approximate position of the MN-ST boundary to between kms 45 and 85 of the profile (see Fig. 3), whereas the RH-ST boundary

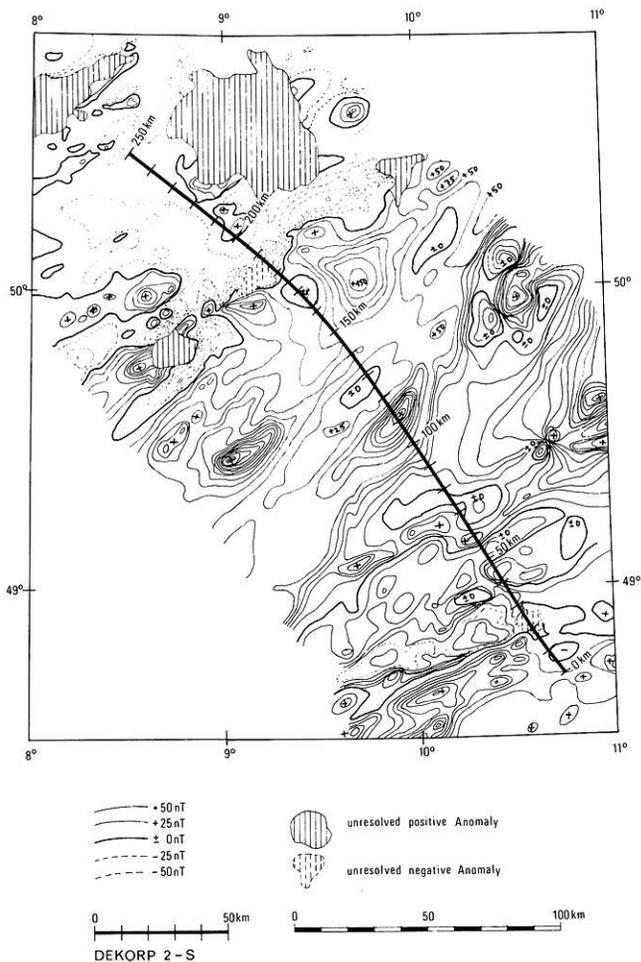


Fig. 4. Location of DEKORP 2-S with adjacent anomalies of the total magnetic intensity. Note that between km 50–100 isolines with an interval of 5 nT are added. From: *Aeromagnetische Karte der Bundesrepublik Deutschland 1:1.000.000* (Eberle, 1973) and *Karte der Anomalien der Totalintensität des erdmagnetischen Feldes in der Bundesrepublik Deutschland 1:500.000* (Bundesanstalt für Geowissenschaften und Rohstoffe, 1976)

is extrapolated from nearby outcrops to occur close to km 233 beneath Lower Permian of the Hessian trough and Tertiary sediments of the NE branch of the Rhine Graben.

The following is intended as a brief outline of the geological evolution of the study area.

ST sediments exposed, among others, in the Bohemian Massif and Thuringian Forest indicate an Early Paleozoic basin with a shallow carbonate platform to the south. Ordovician and younger bimodal volcanism suggests an early development of this ST basin on rifted crust, the southern fringe of which was affected by the Ligerian Convergence at the turn of the Silurian-Devonian.

A restricted deep water basin persisted to the north which itself was bounded in the north by the MGCH, emerging probably from the mid-Devonian onward. The final destruction of the ST basin occurred during the Suetetic convergence probably from 330 my onward following the inferred closure of the basin of the Northern Phyllites. According to current concepts (Behr et al., 1984) to be tested by DEKORP, the convergence is an expression of the northward migration of an orogenic front lasting from Ordovician to Carboniferous times and being driven by

a southward subduction of continental lithosphere (Ziegler, 1978).

During late Carboniferous and the Early Permian the post-Variscan collapse led to the formation of intramontane basins whose depocenter alignments were largely controlled by the Variscan structural trend notwithstanding an as yet unproven contribution of late-Variscan wrench faulting to their development (e.g. Arthaud and Matte, 1975). DEKORP 2-S crosses two of these NE trending basins, namely the Oos-Stockheim trough (Carlé and Wurm 1971), extending between kms 100 and 170, and the Hessian trough (Falke, 1971) between kms 200 and 235.

Also, the thickness of lower Permian strata preserved below marine Zechstein as deduced from reflection seismics and deep wells (Trusheim, 1964) indicates that the amplitudes of intra-Rotliegend down-faulting and warping did not exceed 900 m in the Oos-Stockheim trough or 600 m in the Hessian trough.

The Mesozoic cover of the Franconian platform was deformed along preexisting discontinuities as the whole Platform is pervaded by faults of Hercynian and Rhenish trends.

Particularly toward the NE of the platform the Hercynian faults exhibit strongly compressional geometries. Some of them were active during the Permian with throws partly exceeding 1,000 m. They were reactivated from the Lias onward and persisted probably to the Late Tertiary.

By contrast, faults with Rhenish trends are concentrated in the western segment of the platform. They exhibit mainly extensional geometries or intermittently left-lateral slip. They were active mainly from Eocene to Holocene and probably accompany the formation of the Central European Rift System in its Bresse-Rhine-Leine-Segment.

This intra-platform deformation largely attributable to the Alpine cycle affected the area of the DEKORP 2-S line only by a slight warping with amplitudes up to about 100 m (e.g. Carlé, 1955). One of the latest events to affect the area under investigation was the Ries impact about 15 my ago. It led to the excavation of a shallow crater about 25 km in diameter. Earlier refraction seismic observations demonstrated brecciation of the Moldanubian basement down to at least 4 km below the crater floor (Angenheister and Pohl, 1969, 1976).

The position of the profile referring to magnetic and Bouguer gravity anomalies is shown in Figs 4 and 5, respectively.

With regard to the MN/ST boundary a straight chain of small magnetic anomalies exhibiting their maxima in the south and their minima in the north of the boundary is observed, crossing DEKORP 2-S at about km 75. In Fig. 5 this zone coincides with the southern margin of a major positive Bouguer anomaly, the Neckar-Tauber anomaly.

In the interval between km 100 and km 140 two positive magnetic anomalies are crossed which belong to magnetic lineaments striking N 45° E. These lineaments can be followed from the Main river to the Vosges mountains, at least. The southern one crosses the Haslach area in the Black Forest. The tops of the corresponding bodies beneath the profile DEKORP 2-S are not deeper than 5 km. In the neighbourhood of the profile both magnetic lineaments are running parallel with positive Bouguer gravity anomalies. Thus, it can be concluded: along the MN/ST boundary there are relatively small magnetized bodies situated on the MN side. In the area along the profile from its southern

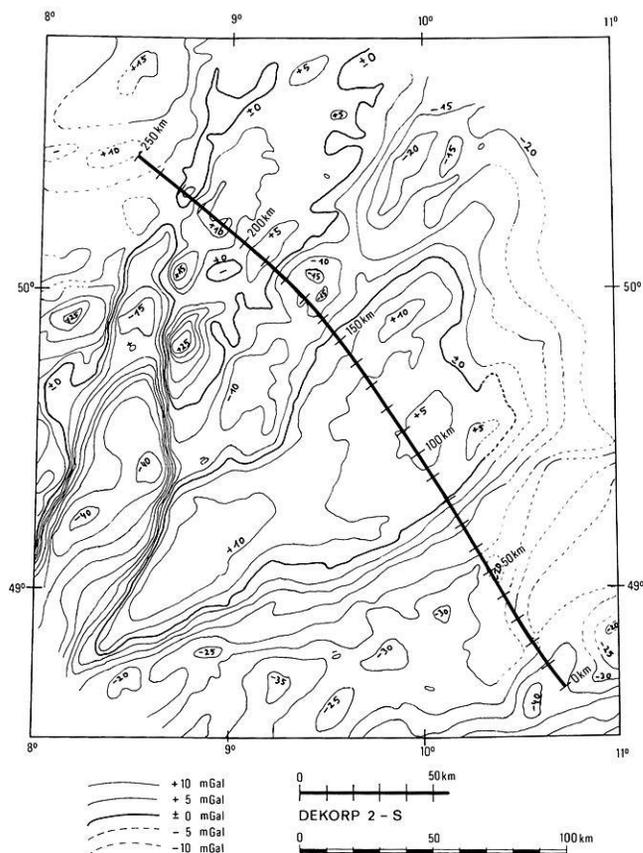


Fig. 5. Location of DEKORP 2-S with adjacent Bouguer anomalies of gravity. From: Schwerekarte von Westdeutschland (Gerke, 1957)

boundary to km 150 the ST shows a higher average density and a stronger average magnetization than the MN north of the Nördlinger Ries.

Between km 155 and km 185 the profile crosses a band of negative Bouguer anomalies striking about 45° N. It runs along the outcrop of the Middle and Lower Bunter from the Rhön mountains to the Odenwald. In the area crossed by the profile the gravity minimum line coincides with a magnetic minimum. At km 178 the profile enters a zone of strong negative magnetic anomalies. They are arranged in an irregular band some 15 km wide with a general strike direction of 50°–60° N. It seems to indicate a large magnetized mass being located south of this band and having its northwestern edge along the zero-line of the anomalies with a steeply dipping boundary face; the top should be outcropping or subcropping. The Bouguer gravity map shows in this interval (km 178–193) only a weak increase towards a maximum at km 210.

At the western border of the Hessian trough at km 225–230 a weak gravity minimum is situated. The magnetic isolines beyond km 193 do not show any remarkable anomalies.

One of the objectives of DEKORP is to better understand the causes of the gravity and magnetic anomalies, too.

1.4 Organization and techniques of the field survey

The near-vertical reflection measurements along the profile DEKORP 2-S were performed using explosives. The decision to do so was based on two reasons:

Table 1. Field parameters for near-vertical reflection measurements

Equipment	SERCEL SN 348, telemetric, 200 traces
Sampling rate	4 m s
Record length	20 s
Low cut filter	8 Hz, 18 db/oct
High cut filter	62.5 Hz, 72 db/oct
Preamplification	2 ⁴ = 24 db
Number of geophone groups	200
Spacing of geophone groups	80 m
Geophone pattern	24 fold, in-line
Pattern length	80 m
Shotpoint spacing	a) 320 m (near vertical angle) b) 1,280 m (wide angle)
Depth of charge	30 m
Size of charge	a) 30 kg, partly reduced to 5 kg (near vertical angle) b) 90 kg, partly reduced to 60 kg (wide angle, every 4th or 10th shot, resp.)
Spread	off-end-shooting
Spread length	15.92 km
Offset	
1. in-line	a) 40 m (near vertical angle) b) 57 km (wide angle)
2. off-line	max. ± 500 m
Coverage	a) 25 fold (near vertical angle) b) 6 fold (wide angle)

– No experience had been collected so far in Germany with the Vibroseis method for deep crustal studies outside the large sediment basins.

– The observations in the wide-angle range by university groups made the use of explosive sources necessary because Vibroseis cannot be recorded properly by the equipments of the institutes. Thus, the additional observations could be well integrated into the continuous profiling work of the contractor along the main profile.

Table 1 displays the most important field parameters. The maximum shot-receiver distance of 16 km was chosen in order to get reliable information about the crustal velocities. The in-line 80 m long geophone pattern should have been approximately twice as long because analysis of noise waves has yielded wave lengths of up to 300 m; by a geophone group setup of 160 m and by twofold vertical stacking during the processing procedure wave lengths up to 320 m could have been extinguished. On the other hand, the chosen setup provided a simple and economical way for laying out and collecting geophones within just one trace interval.

Preparations for the measurements started two months ahead with geodetic surveying and permitting. The measurements began in the south on April 4, 1984, rolling along to the north, and ended on May 17, 1984 after 32 working days and a break between April 19 and May 2. For economical reasons a progress of 10 km per day was considered an optimum. The contractor's crew was increased to 103 persons for the continuous profiling plus 28 persons for the Vibroseis experiment and 6 persons for the wide-angle shots. Up to 16 drilling rigs were in operation simultaneously.

The prerequisite for the intended progress was a careful planning and a tight organizational schedule, especially in view of the rough terrain in the Spessart and in the Taunus mountains. Serious drilling problems arose by the presence

of massive limestones of the Muschelkalk formation and of hard gneisses and quartzites in the crystalline part of the Spessart. These difficulties were overcome by the in-hole hammering technique.

An important condition for a smooth operation was a reliable radio-communication between all participating groups, sometimes operating at distances of up to 80 km in rough terrain and tamber forests. Up to 6 moving relay stations were in operation simultaneously.

In spite of all these problems and poor weather conditions with ice and snow at the beginning of the measurements, the progress of 10 km per day turned out to be achievable on the average. On 26 working days 729 shots were fired along the 250.8 km long profile. 11% of the planned shotpoints had to be discarded for security reasons, reducing the mean coverage from 25 to 22. On the average, 28 shots were fired on a 10-hour working day with an average line progress of 9.62 km per day.

2. Data processing of the main profile

A first preliminary treatment of line DEKORP 2-S has been carried out at the DEKORP-Processing-Center (DPC) in the Geophysical Institute of the Technische Universität, Clausthal.

For data processing a seismic computing system – type: Phoenix I – was at disposal, hardware consisting of a Raytheon minicomputer – type: RDS 500 (64 K × 16 bit core memory) – with floating point and array processor. Mass storage was provided by 4 Wangco tape units (75 ips, 1,600/800 bpi) and by a CDC disk unit (80 Mbyte). The appropriate software comprises a multitude of programs for analysis, processing and display of seismic data so that all standard procedures customary in exploration seismics can be applied.

In the following each step of the processing and its importance especially with regard to deep seismic data processing will be explained. Also the first steps of more refined processing procedures on profile DEKORP 2-S – not yet completed – are described.

2.1 Common shotpoint gathers

Scaling. In a first step the samples had to be rearranged from a time-multiplexed into a trace-sequential format (demultiplexing), yielding common shotpoint gathers. The Sercel equipment compensates automatically for the large dynamic range of about 120 dB of the movement of the ground by instantaneous floating point recording. This procedure has been retrieved in order to get true reflection amplitudes (amplitude recovery). Figure 6a shows the result. Due to the extremely high differences between amplitudes at the beginning and at the end of the seismogram practically only the first arrivals and the strong surface waves are recognizable in Fig. 6a. Then, the amplitude values have been multiplied by the factor $k \cdot T \cdot e^{aT}$ (where $k = \text{const.}$, $T = \text{traveltime}$ and $a = \text{constant of absorption}$) in order to compensate for the systematic amplitude decrease due to spherical divergence ($\sim 1/T$) and absorption ($\sim e^{-aT}$) and to obtain a balanced seismogram. The value of a has been assessed empirically for processing purposes only and has no direct physical meaning.

In order to find suitable values for a and k by least

squares fitting numerous analyses of energy in single records along the profile have been carried out. The best result is obtained when at 5 s two way traveltime (TWT) the exponential function is replaced by a constant value because from thereon a constant noise level is prevailing. Typical values of a cover the range from 1.1 s^{-1} to 1.8 s^{-1} . An exception is represented by the Nördlinger Ries data where a reduction to 0.8 s^{-1} is observed.

In order to correct for influences of coupling between ground and geophone or shot, respectively, a constant scaling factor for each trace has been applied to get the same root mean square energy for all traces (trace equalization). Figure 6b shows an amplitude-corrected shot with a uniform energy behaviour.

Data quality. Part of the recorded seismograms exhibits a high noise level by technical reasons or by field conditions. A first quality control has been provided by monitor playbacks in the field.

According to the type and strength of the noise one of the following methods has been chosen for noise discrimination: frequency filtering, partial zeroing, change of polarity or complete elimination of traces.

In order to discriminate first arrivals and other undesired waves initial muting of traces extends from about 200 ms (near the shotpoint) to 6–9 s (at the end of the spread).

On the other hand, the low frequency high energy surface waves (Fig. 6b) often coincide with reflections and cannot be muted. Even by low-cut filtering, only a weakening can be achieved because the frequency spectra of the surface waves (about 5–20 Hz) and of the reflections (about 10–40 Hz) overlap. Figure 7 exhibits that this holds true especially for near-shotpoint traces and in the first few seconds. Since surface waves were sufficiently weakened by stacking, time consuming procedures for their elimination, e.g. wavenumber filtering, were not applied in the first preliminary data processing.

In Fig. 8 a completely processed common shotpoint gather is shown with the muting curve used. It contains a few distinct reflections and many strong diffractions, above all. All common shotpoint gathers have been plotted at the DPC. Data quality – except in the above mentioned cases – is good. Along the major part of the profile reflections are scarce in the upper crust and a distinct beginning of strong energy at about 5 s TWT is observed. A better resolution of the weak reflections in the early traveltime range makes special processing necessary.

Common-midpoint-sorting. Sorting of common shotpoint gather data into common midpoint gather data (CMP) was performed on the basis of geodetic data made available on magnetic tapes. The intended coverage on profile DEKORP 2-S was 25 fold. Therefore, every 4th receiver point was used as shotpoint location for the given spread. Shots cancelled by security reasons and the elimination of noisy traces yielded a real coverage of 20–22 fold after editing and sorting. Extreme values of coverage along the profile were 9 and 31, respectively. The product of 200 traces per record, 5000 samples per trace and more than 20 fold coverage in connection with the small capacity of disk storage at the present computing equipment made CMP gathering the most time consuming step of the processing.

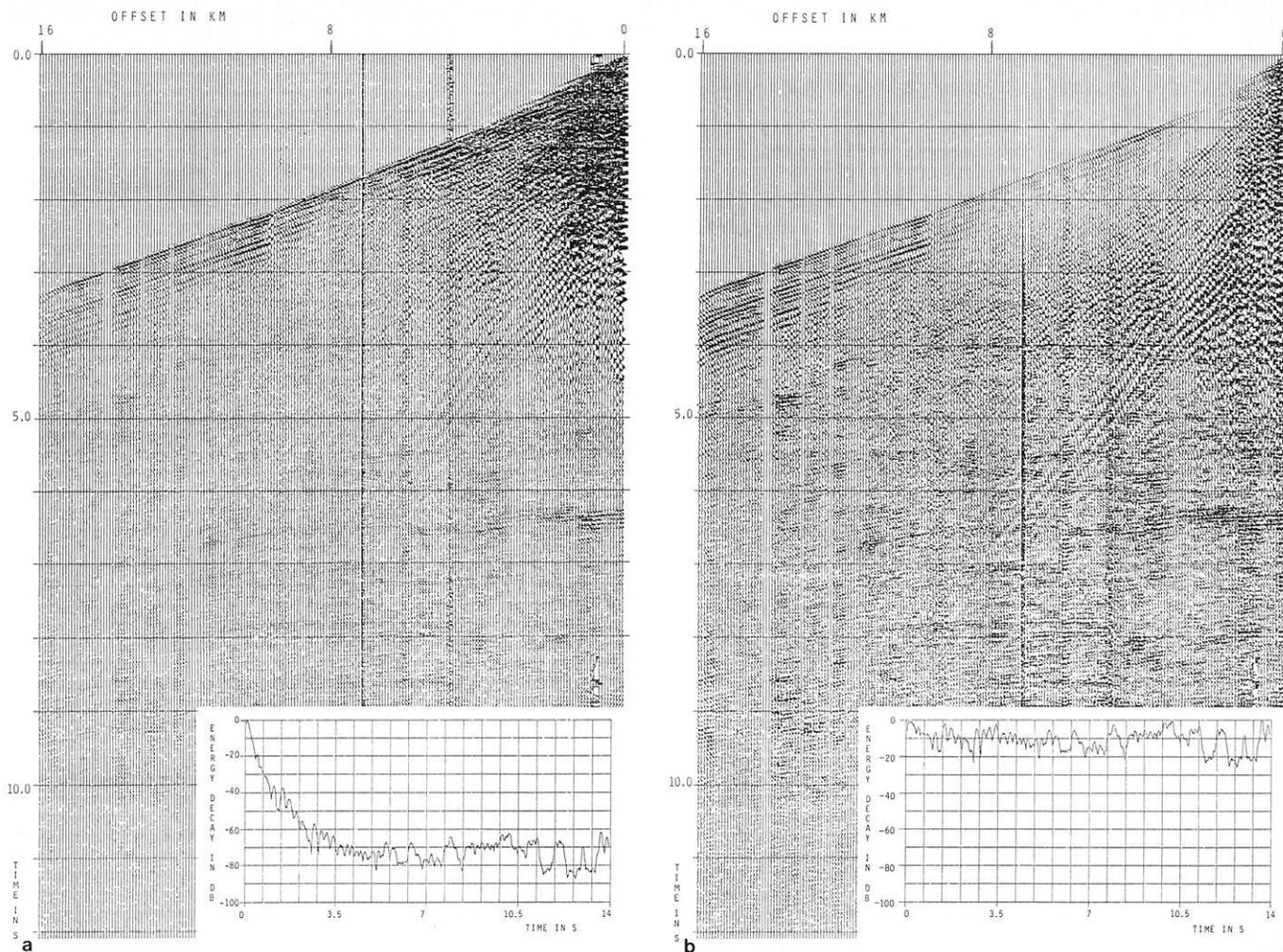


Fig. 6. **a** Common shotpoint gather before correction by analytical gain function, i.e. true amplitudes are shown. In the *lower part* of the figure the time dependent energy behaviour is shown; **b** Common shotpoint gather after correction by analytical gain function. In the *lower part* of the figure the time dependent energy behaviour is shown

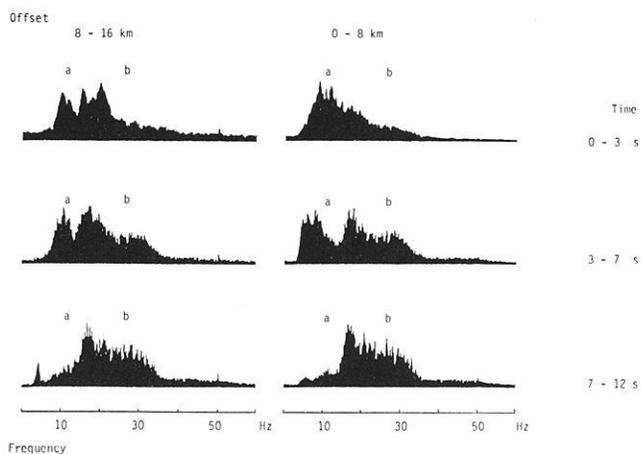


Fig. 7. Amplitude spectra of a common shotpoint gather averaged in time and space for different offset ranges: (a) surface waves, (b) signal

2.2 Common midpoint gathers

Static corrections. Static corrections have been calculated from uphole times and from topographic elevations by the

field crew. In order to obtain the correction velocities and to interpolate between shotpoints the first arrivals up to distances of 2 km or more have been used. The datum level was chosen to be 400 m above seal level. The removal of the uppermost layers with low velocities has been performed in most cases down to the first high velocity layer. Hence, low velocity layers below the limestones of the Malm or Muschelkalk in the southern part of the profile have not been taken into account. In the Nördlinger Ries the "Seeton" (limnic) sediments have been corrected for, only, but not the Suevite which is characterized by medium velocities.

With regard to future processing of data, especially more subtle velocity analyses, it is planned to evaluate all first arrivals up to the maximum offset of 16 km and to compensate for lateral variations of velocity down to some km depth by seismic stripping. In the region of the Nördlinger Ries appropriate velocity models are already available.

Deconvolution. The optimal parameters for deconvolution have been estimated by calculation of autocorrelation functions and amplitude spectra. The tests resulted in a prediction length of 4 ms (spike deconvolution), and by noise

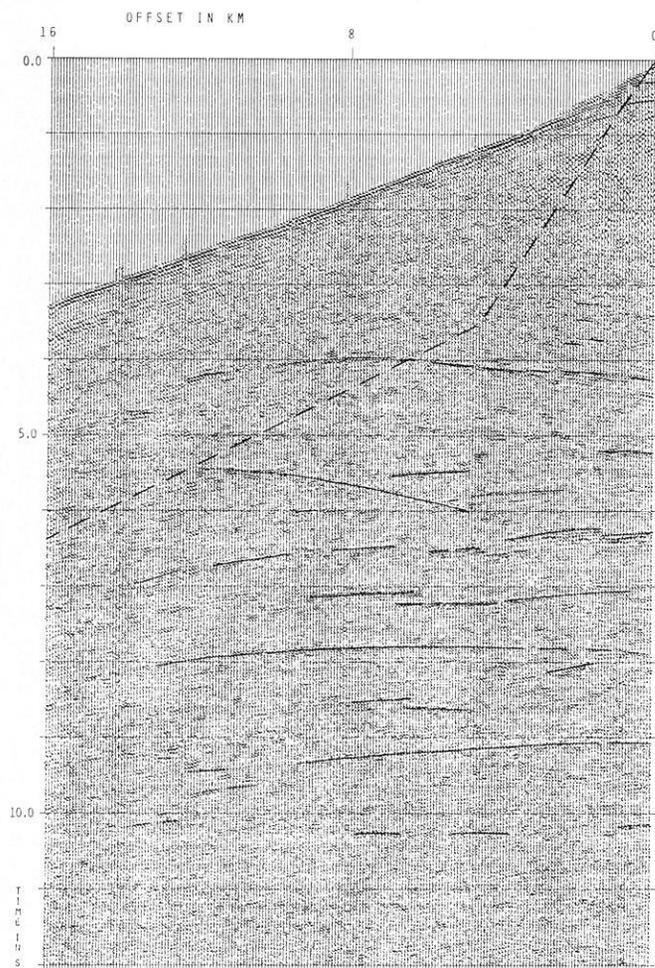


Fig. 8. Common shotpoint gather of Fig. 6b after frequency filtering and scaling (AGC) with interpretation and with muting curve (dashed)

mixing of 0.2% steady state inverse filters with operator lengths of 120–180 ms were obtained. Figure 9a shows one trace of a CMP gather before deconvolution with the corresponding autocorrelation functions and amplitude spectra summed over all CMP-traces in three different time windows. Figure 9b shows how the long seismic signals have been contracted after deconvolution. The side lobes of the autocorrelation functions corresponding to the amplitude maxima and minima of the signal have disappeared. The rest is a narrow central maximum indicating a statistical series of reflecting signals. The amplitude spectra are nearly whitened corresponding to the appearance of spike-type signals in the time traces. By subsequent bandpass filtering the pulses have been transformed into (symmetrical) zero-phase signals.

Dynamic corrections and CMP stacking. In order to estimate stacking velocities stacking tests have been carried out every 3 km along the profile with groups of 11 CMP's applying 48 constant velocities. An important problem was posed by the appearance of numerous diffractions (see Fig. 8). The optimal stacking velocities for diffractions are higher than those for subhorizontal reflections. Because dif-

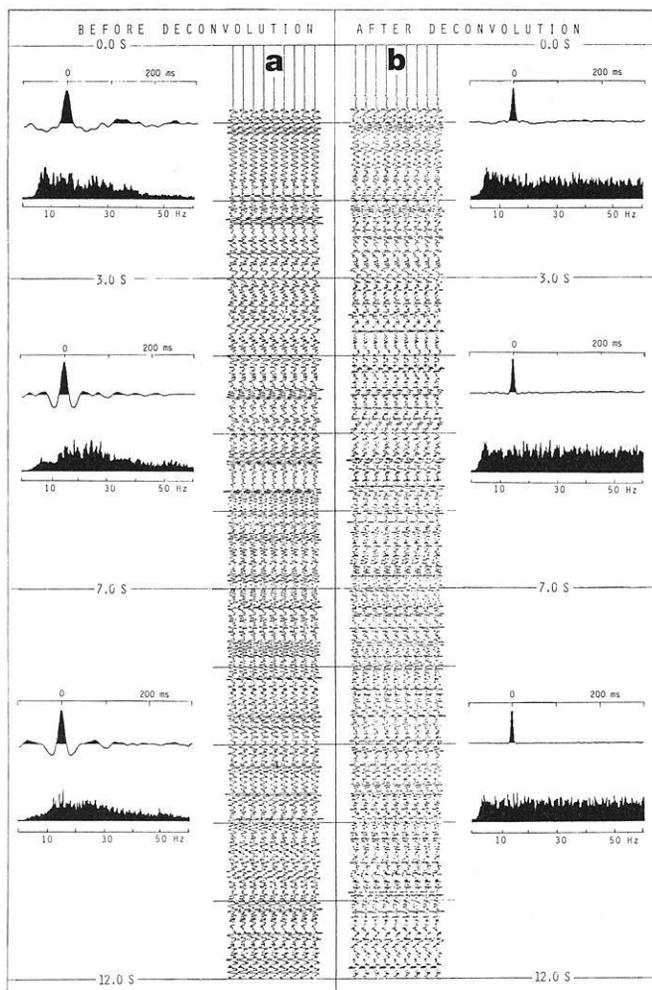


Fig. 9. **a** One trace of a CMP (repeated 9 times for better readability) with the corresponding autocorrelation functions and amplitude spectra summed over all CMP-traces in three different time windows before deconvolution. **b** The same as in **a**, but after deconvolution

fractions prevail by number and by energy the recovery of the weaker reflected events in the stacking tests of 11 CMP's often was difficult. Frequently two or three stacking maxima (see Fig. 10) are observed, when diffractions appear beside reflections. In addition problems arise by different dips of reflecting elements. In order to improve the classification of the seismic signals stacking tests with a greater number of CMP's are being performed. This work has not been finished, yet.

The deviation of reflection traveltime curves from ideal hyperbolas is very slight in our case despite the maximal offset of 16 km. It has been shown by calculations that this deviation amounts to only 12 ms at 16 km offset for reflections from 10 km depth, and for 30 km depth the result is 3 ms. These values correspond to less than half a wavelength in the observed frequency range. The reason is that the vertical gradient of the velocity is very small which implies that the effects of refraction cannot be very strong. Therefore no difficulties can arise with regard to the stacking procedure. On the other hand, problems are encountered by the length of the spread in connection with lateral variations of the geological conditions close to the surface.

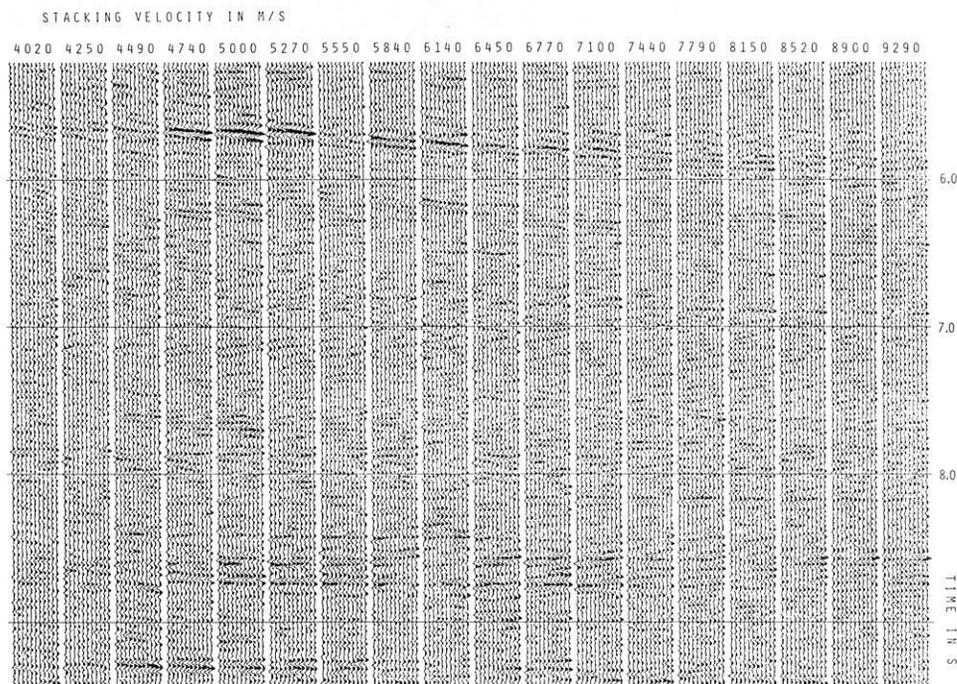


Fig. 10. Part of a velocity analysis (constant velocity stack). Note the multifold alignments for different velocities. Correction velocities are given on top of the figure

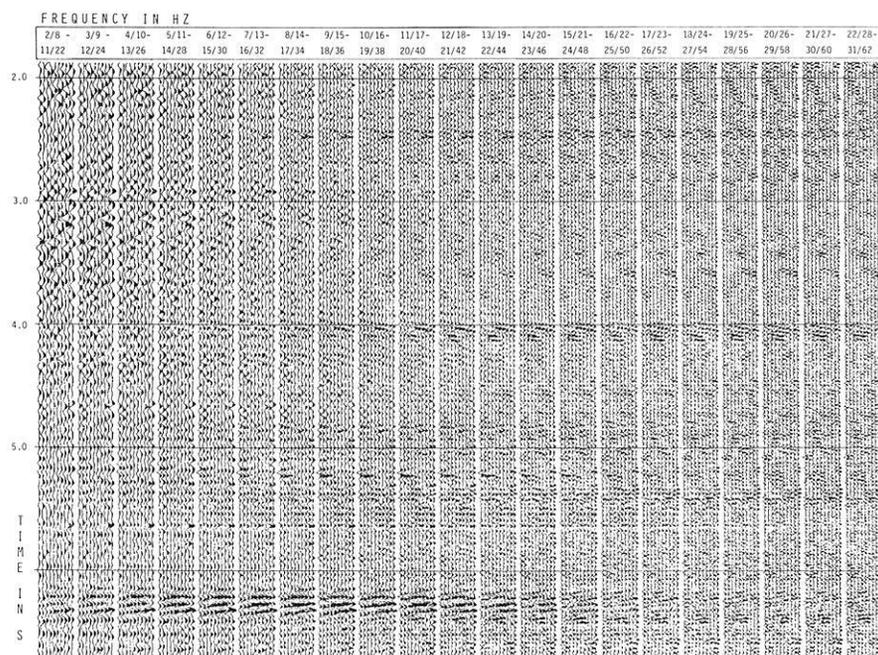


Fig. 11. Part of a filter test (stepwise shifting of a narrow bandpass of constant width over the entire frequency range). Filter characteristics are given on top of the figure

2.3 Stacked section

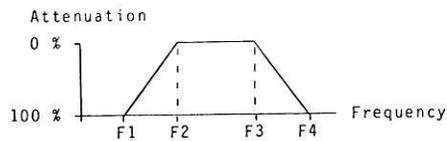
Frequency filtering. After scaling (AGC) and stacking different bandpass filters have been tested by application on groups of 10 CMP's which are representative for a larger part of the profile. Two ways have been used in order to get the optimum filter parameters:

- Stepwise shifting of a narrow bandpass of constant width over the entire frequency range yielding the main frequencies of the reflections.
- variation of the low cut-off frequency, fixing the high cut-off frequency, and vice versa yielding the low- and high-cut boundaries.

Figure 11 shows part of a filter test after the first method. The main frequencies of the signals are in the ranges 12–30 Hz for shorter traveltimes and 6–24 Hz for longer ones. Therefore, it was necessary to choose the shape of the trapezoidal bandpass filters carefully. Figure 12 shows the average filter parameters used. A stacked and filtered time section is shown in Fig. 20; such sections represent an essential basis for the interpretation.

Migration. As a last step a migration of the entire profile was performed. Along km 24–72 of the profile the Kirchhoff-, finite-difference-(FD-) and frequency-wavenumber-(FK-) migration methods have been tested. The FD-migra-

Schematic transfer function



Time interval in s	Characteristic filter points			
	F1	F2	F3	F4 in Hz
0 - 2.2	8 / 15	-	20 / 40	
2 - 5	6 / 12	-	19 / 38	
4.5 - 11	4 / 9	-	18 / 36	
10 - 20	3 / 8	-	16 / 32	

Fig. 12. Sketch of bandpass characteristics and table of bandpass limits used (averages)

Table 2. Velocity vs. Time function used for migration (v_{rms} = RMS velocity; v_{int} = interval velocity)

TWT (s)	v_{rms} (m/s)	v_{int} (m/s)
0.0		3,500
0.5	3,500	4,720
2.5	4,500	6,040
6.5	5,500	7,240
10.3	6,200	8,240
14.0	6,800	

tion provided the best results with regard to the signal/noise ratio, but it required, unfortunately, the longest processing time.

The interval velocities derived from stacking velocities varied strongly and showed partly too high values. For this reason they could not be used as migration velocities. Therefore an approximate velocity function has been derived from smoothed and reduced stacking velocities and from other informations about the crustal structure (Table 2). This function has been used for the entire profile. The resulting section is shown in Fig. 23 and is compared with another migration by the same method shown in Fig. 22.

Because of the restricted computer capacity the whole body of data has been resampled with a time increment of 8 ms. No considerable loss of information occurred because of the prevailing low frequencies. Still, the profile had to be migrated in 6 parts. Due to sufficiently large overlapping the boundaries of the single parts match well. A restriction to the first 15 s TWT was economical because below the Moho reflections little useful information is present.

The result is satisfactory for the entire section in spite of the laterally constant velocity function. Above all, the focussing of the diffractions yields a significantly clearer picture for interpretation, e.g. in the Spessart section (Fig. 13).

2.4 Special processing

The standard processing described in the previous sections could not remove all observed and already mentioned diffi-

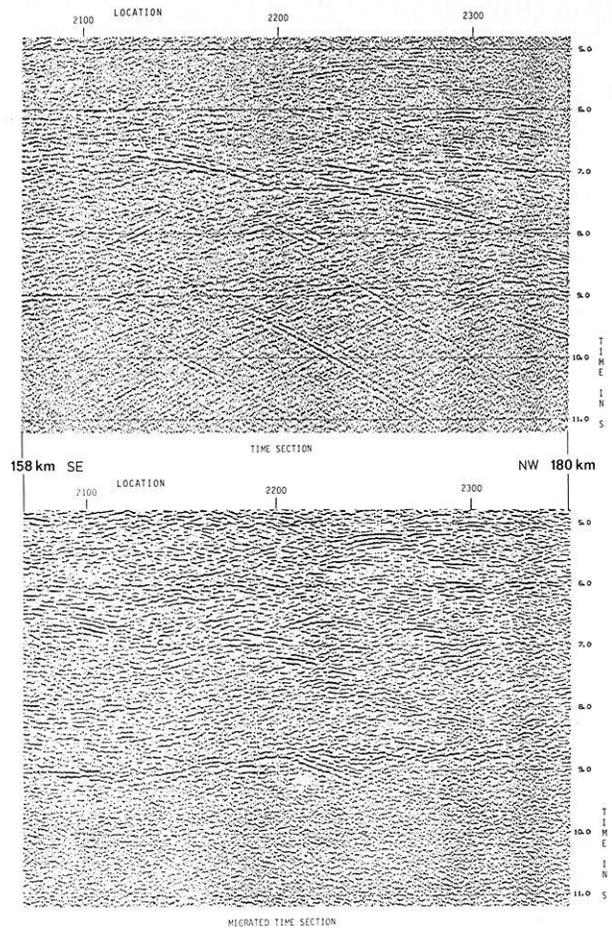


Fig. 13. Part of the DEKORP 2-S stacked section from the Spessart area (km 158–180) before migration (*top*) and after migration (*bottom*)

culties. In the following special processing of selected parts of the profile is presented using three particular examples.

Frequency-wavenumber filtering. The appearance of strong surface waves in many DEKORP seismograms obscured weak reflections in near-surface regions. The overlapping of the frequency bands of signal and noise makes the application of conventional bandpass filtering ineffective in many cases. Frequency-wavenumber filtering is an efficient but also very time-consuming tool for the elimination of noise waves. After 2-dimensional Fourier transformation of the seismograms from the travelttime-distance ($t-x$) domain into the frequency-wavenumber ($f-k$) domain the signal amplitudes are obtained in dependence on frequency and apparent wavenumber.

Figure 14 shows the $f-k$ transform of the common shot-point gather of Fig. 6b. In the $f-k$ spectrum waves with higher velocity are displayed as straight lines with greater slope than those with lower velocity. Therefore, refractions ($v=5,000$ m/s) can be clearly distinguished from surface waves ($v=1,600$ m/s). Both wave types appear only on the left side of the figure because of the single-sided spread configuration. Steep angle reflections show up as a point cluster around $k=0.0$ m^{-1} and $f=18.0$ Hz. Due to the relatively large geophone group interval (80 m) spatial aliasing appears, for $v=1,600$ m/s even from 10 Hz on.

If the $f-k$ spectral parts above the dashed filter limits

in Fig. 14 are rejected (i.e. all waves with $v < 3,600$ m/s are suppressed) a complete elimination of surface waves in the $x-t$ domain is achieved (see Fig. 15).

It is not reasonable to restrict the bandpass in the $f-k$ domain further, because the reflection hyperbolas could be affected which exhibit also steeper slopes at greater offsets. Therefore, the refraction arrivals have to be removed by appropriate muting in the common shotpoint gather.

Nevertheless, the $f-k$ filtering of the first 40 km of the DEKORP profile (Nördlinger Ries) did not yield the expected enhancement in the early traveltime range, probably due to the absence of reflections. Considering the fact that the surface waves are already weakened by stacking and considering the uneconomical relation between effort and result it did not seem to be reasonable to process the entire profile in this way.

Nördlinger Ries. From earlier investigations using the reflection seismic method in the Nördlinger Ries very clear reflections from the boundary between limnic "Seeton" sediments and Suevite at about 250–400 ms are known (Angenheister and Pohl, 1969). A comparable result from the DEKORP measurements cannot be expected because due to the spread configuration not even a single fold coverage is obtained in this time range and because the low velocity limnic sediments have been removed already in the correction process as mentioned earlier in Sect. 2.2. The following steps were made in order to reveal the base of the limnic sediments: (1) Careful editing of bad traces, (2) application of $f-k$ filtering in order to eliminate the surface waves especially in the Ries, (3) selection of an optimal interpolated muting curve in order to retain maximal coverage and (4) replacement of the "Seeton" corrections. Thus, a clear enhancement was achieved: The base of the limnic sediments is more clearly recognized now (see Fig. 16).

In summary it must be stated that it is not possible to resolve simultaneously the near-surface geological units as well as the deeper subsurface especially if distinct lateral velocity inhomogeneities occur as observed in the Ries. The resolution of structures in the lower crust makes necessary deeper reaching seismic stripping or undershooting of the inhomogeneous near-surface areas by means of a different spread configuration.

True amplitude stacking. The automatic gain control (AGC) applied before and after stacking implied the loss of true reflection amplitudes in favour of a balanced seismogram picture. In order to get an assessment of the real temporal behaviour of energy a so-called true amplitude stack (TA) was performed for the profile range from 77–97 km. In this process the amplitude correction involves only an analytical gain function (correction for spherical divergence and absorption) and a trace equalization (correction for different shot charges and coupling between geophone and ground) but no AGC. Before the derivation of the parameters of the $k \cdot T \cdot e^{aT}$ -function and their application on the seismic traces low-cut filters had to be applied in order to suppress strong surface waves.

Further, new muting curves were chosen very carefully in combination with revised velocities. The result of this true amplitude processing is shown in Fig. 17a. For the sake of comparison the same part of the profile has been stacked with predictive deconvolution (16 ms prediction length) and with AGC, too, see Fig. 17b. Whereas conven-

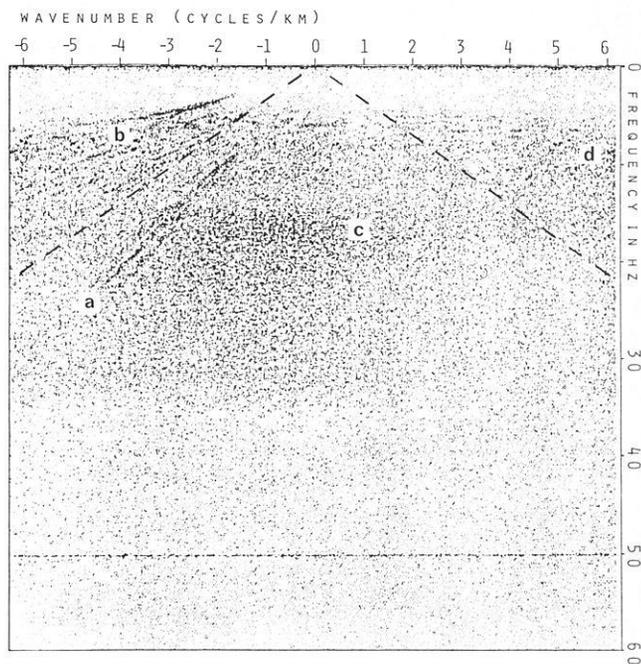


Fig. 14. Frequency-wavenumber ($f-k$) transform of the common shotpoint gather shown in Fig. 6b: **a** first arrivals, **b** surface waves, **c** signal, **d** aliasing of **b**; dashed lines indicate $f-k$ filter limits

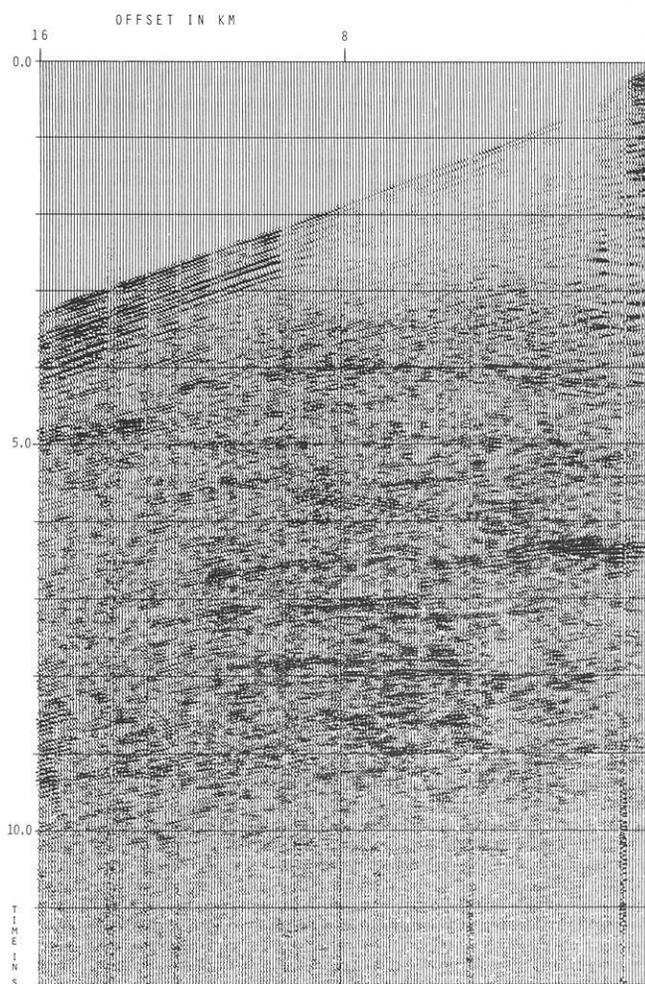


Fig. 15. Common shotpoint gather of Fig. 6b after frequency-wavenumber filtering

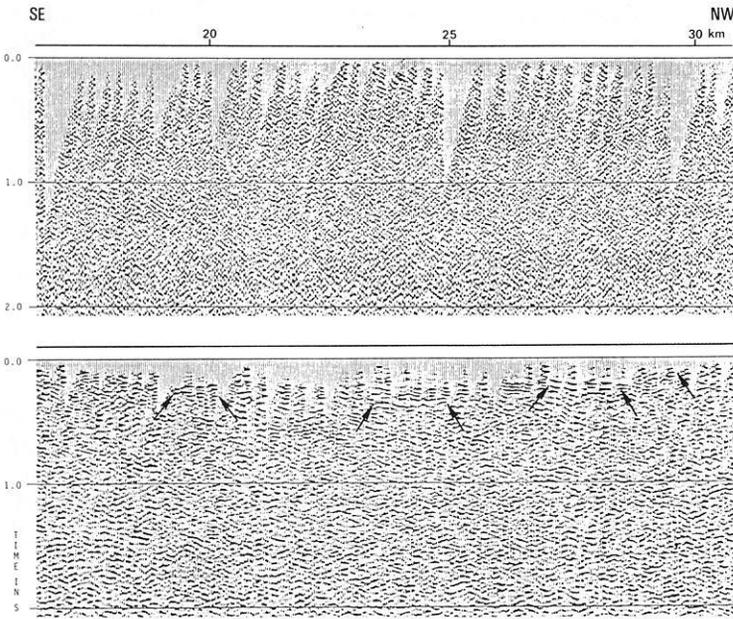


Fig. 16. Part of DEKORP 2-S (Nördlinger Ries). *Top*: first stack. *Bottom*: reprocessed stack with replacement of the limnic ("Seeton") sediments ("Seeton" base is indicated by arrows)

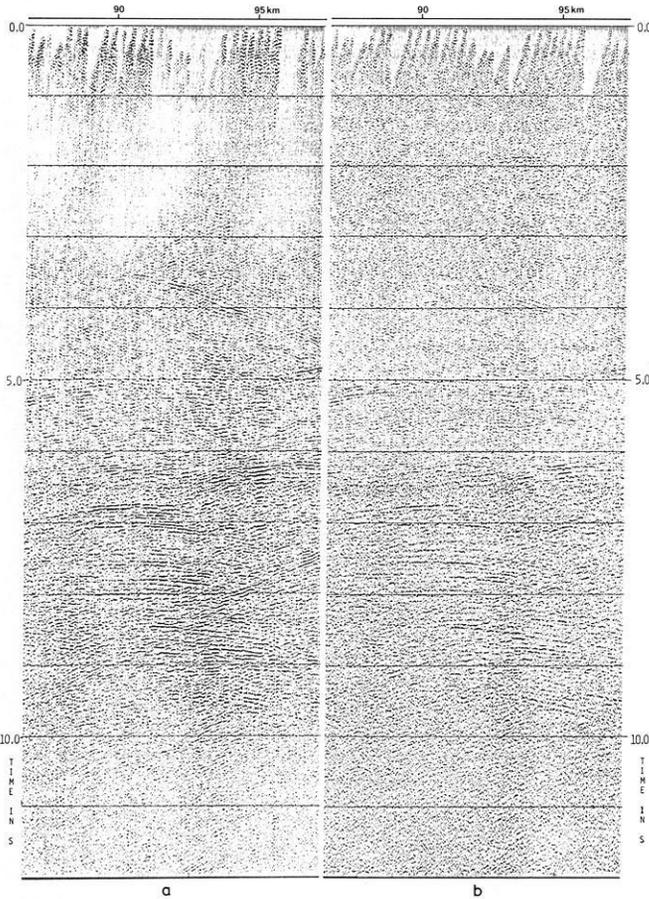


Fig. 17 a, b. Comparison of a true amplitude stack (a) with a conventionally processed stack (b)

tional processing emphasizes all events more or less equally the TA processing shows a high dynamic range. Besides the predominant diffractions only a few reflections with high contrast of impedance are observed. Many weak events can

hardly be recognized. The frequency content of the reflections reaches up to 50 Hz.

2.5 The velocity problem

In order to obtain reliable velocity information from the deeper crust a maximum shot-receiver distance of 16 km had been chosen. This configuration yields a maximum normal moveout of about 0.35 s for a reflecting interface at 30 km depth which might be a solid base for deriving crustal velocities from stacking velocity analyses. However, a serious problem arises from the dominance of strong diffraction patterns and the weakness and discontinuity of reflections. For this reason the preliminary stacking velocities might often be too high. Special efforts had to be made in order to solve the velocity problem.

In a restricted part of the DEKORP 2-S profile velocity functions can be derived from in-line wide-angle observations. The procedure and the results are given in Sect. 4.2.

In addition, diffraction patterns in the lower crust have been analyzed with respect to their velocity information. In a first attempt, a section of 32 km length north of the Ries astrobleme was selected, where one of the strongest diffraction clusters has been observed. The various apices of the diffraction curves in 83 consecutive shotpoint gathers were determined with the help of maximum convexity master curves. The mean effective velocities \bar{v} (from the surface to the apex) have been determined neglecting refraction effects. Master curves for discrete center-values of 5.0, 5.5, 6.0 and 6.4 km/s were used for the analysis. A better resolution than ± 0.25 km/s was not feasible by eye-inspection. Possible off-line locations of the diffraction sources have not been taken into account in this evaluation. The spatial position of some of these diffractors was investigated by a different approach (Sect. 4.1).

Figure 18 shows a diagram of the selected section where the position of the specific apices and the related values of the mean effective velocity \bar{v} are plotted in a TWT versus distance diagram. The individual \bar{v} -values are denoted by different symbols. Reflection elements are indicated by up-

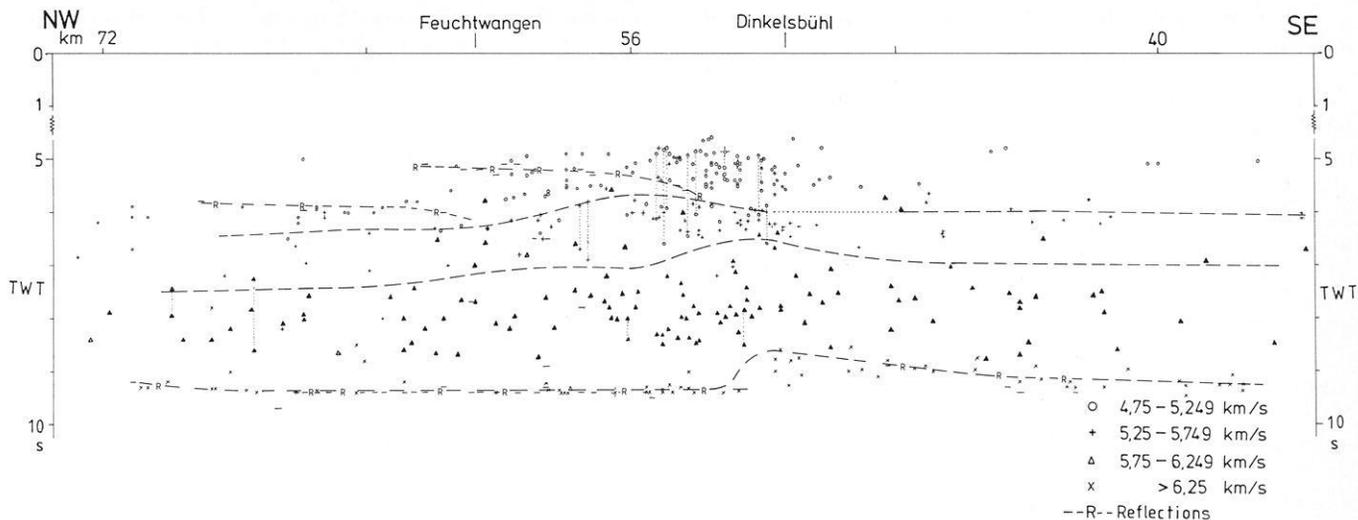


Fig. 18. Time distance diagram of apex positions and mean effective velocities picked from 83 shotpoint gathers. For further explanation see text

percent “R” and narrowly dashed lines. The dotted vertical lines indicate a corresponding arrangement of apices which could not be resolved in more detail. The clustering of apices with \bar{v} -values of 5.0 ± 0.25 km/s in the center of the figure (near Dinkelsbühl) is remarkable (see the unmigrated section in Fig. 21). In a next step areas of apices with the same \bar{v} -value were separated from each other by widely dashed lines. In the time range from about 5–6 s TWT surprisingly low \bar{v} -values appear (4.75–5.25 km/s) except for very few singular apices related to higher velocities (5.25–5.75 km/s). One possible explanation of these low mean effective velocities is a lateral offset of the diffractors; another one is the presence of rocks with rather low velocities.

In deeper regions, especially in the Moho range, relatively high \bar{v} -values are obtained. These surprising results might be caused by edge-diffractors: crossing them by profiles under oblique angles generates diffraction hyperbolas yielding higher apparent velocities than the true average velocity ($v_{\text{apparent}} = v_{\text{true}} / \sin \phi$, where ϕ is the angle between profile direction and the edge-diffractor).

A more detailed processing of the profile concerning, among others, the velocity problem is under way at the DPC. Also tests using other processes than the conventional stacking method are being performed. In addition, a tomography analysis of the large first-arrival data set has been started.

3. Presentation of first seismic results and preliminary geological interpretation

3.1 Description of seismic results

Figure 19 gives an example of a typical common shotpoint gather observed along DEKORP 2-S, with 200 traces at a location near Dinkelsbühl (shotpoint 641.5). The first 12 seconds two-way traveltimes (TWT) (out of 20 s recorded) are shown:

- In the upper 6 s TWT reflections are very poor or absent.
- There are several good reflections between 6 and 10 s TWT.

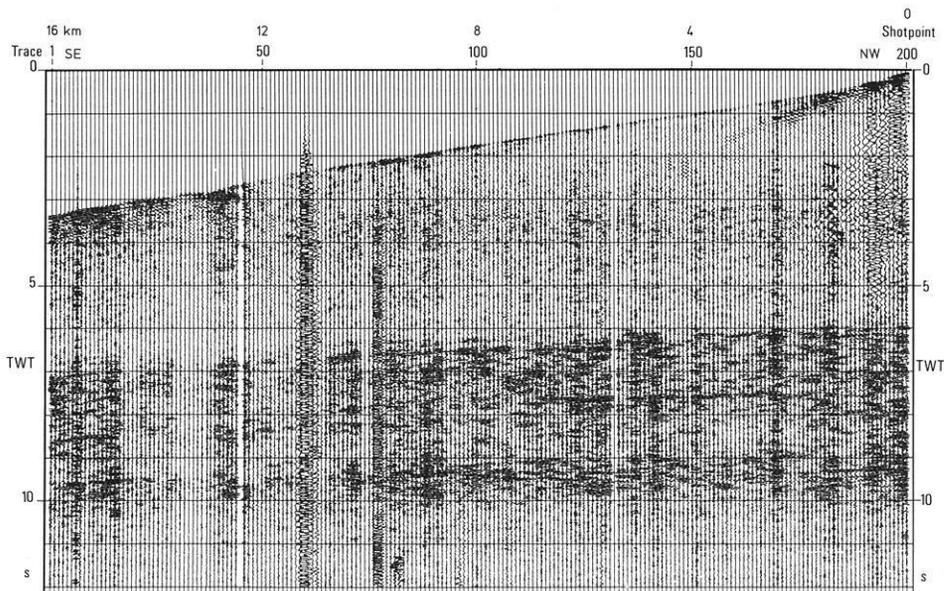


Fig. 19. Example of a DEKORP 2-S seismicogram (common shotpoint gather from SP 641.5), unfiltered. Scaling function is $K \cdot T^n \cdot e^{aT}$ with $K=1.0$, $n=1.0$, $a=1.1 \text{ s}^{-1}$ for $0.0 \leq T \leq 3.5 \text{ s}$ and $K=164.5$, $n=0.0$, $a=0.0 \text{ s}^{-1}$ for $3.5 \leq T \leq 12.0 \text{ s}$ (see Sect. 2.1)

– Below 10 s TWT the seismic traces are again void of reflections.

The common shotpoint gather of Fig. 19 is typical in so far as generally – with a few exceptions – the early parts of the seismograms contain little information. The TWT at which a considerable increase of reflection density and quality is observed, changes along the profile coming close to only 2 s in the Spessart area. Nearly on all seismograms a strong reflection-energy band is observed between 9 and 10 s TWT which represents the last pronounced event. This event comes from the zone of the Mohorovičić discontinuity (Sect. 3.2).

The first arrivals are clearly defined along the 16 km-spread. They have been picked for all shotpoint gathers, and the information is being analyzed in order to obtain the velocity distribution in the uppermost crust.

Figure 20 shows the CMP-stacked record section of DEKORP 2-S unmigrated and uninterpreted. Figure 21 presents the same stack with some interpretation. The dominant tectonic units, a stratigraphic division and the position of major topographic units and localities are marked on the top of the figures.

The general picture of the section is dominated by numerous strong diffractions, rarely seen with such a density and strength in crustal sections. The diffractions concentrate between 5 and 10 s TWT, i.e. in the time range of the highly reflective lower crust. In the northern part of the Franconian Platform and below the Spessart, however, they ascend up to 2 s TWT. Some of these diffractions cluster strongly near the area of Dinkelsbühl (km 40–70) and below the Spessart between the Main river and the town of Laufach (km 150–190). A weaker concentration is observed between the Tauber river and Würzburg (km 110–140), and an oblique SE dipping alignment of diffractions occurs between Rothenburg and the Tauber river (km 65–110).

Picking of continuous seismic reflections is not possible over longer distances. Even the Moho band appears in the form of rather short reflecting elements. At certain locations there are pronounced dipping events, not only in the reflective lower crust but also in the upper crust, sometimes approaching the surface, such as in the Spessart and Tauber area and near Dinkelsbühl.

Dips to the southeast dominate. Some alignments of diffraction apices occur which seem to dip towards the southeast although vertically arranged clusters, such as that below Dinkelsbühl, are also observed. Most of the reflections in the DEKORP 2-S section are subhorizontal. With their clustering in the lower crust they fit well into the general picture of reflectivity from records in the Variscan belt, showing a highly reflective lower crust embedded between a poorly reflective upper crust and upper mantle (Meissner et al., 1984).

Several locations along the profile show poor quality reflections. In the Ries area such a decrease of quality may be related to lateral variations in velocity and wave-scatter-

ing caused by the brecciated basement. The reflectivity of the lower crust below the Ries seems to be similar to that in its vicinity, as indicated by the short E-W reflection profile through the Ries crater described by Angenheister and Pohl (1969, 1976).

BEB¹ provided the first migration of the entire profile (Fig. 22). Another one was processed at the DEKORP Processing Center at Clausthal later on (see Fig. 23 and Sect. 2.3) where higher velocities in the time range 0–6 s TWT have been used. The result differs in so far as some reflection events have become more pronounced so that a clearer overall picture has been obtained in Fig. 23. However, this effect might be due to slightly different parameters (e.g. scaling function or filter limits) in the processing sequence which was essentially the same for both migrations. The interpretation is based on both results.

The migrated sections contain still some curved events which could partly correspond to curved reflectors or to residual diffraction patterns. The latter ones could have been generated by edge diffractors. Some “overmigrated” events (i.e. too high migration velocity for this particular diffraction) can be produced by point diffractors at positions not vertical below the profile. The maximum spatial extent of all point diffractors has been estimated by the Fresnel zone formula to a diameter of 4.9 km at a depth of 30 km for a frequency of 15 Hz and a velocity of 6 km/s. At a depth of 15 km the diameter decreases to 3.5 km.

South of Rothenburg remarkable southeast dipping reflecting elements at the Moho level are observed which seem to originate from the sub-horizontal Moho zone. Pronounced, relatively shallow, southeast dipping events are seen below the Spessart. Generally speaking the majority of reflection elements is accumulated in the region of 50–200 km of the section with the densest concentration in the regions of the diffraction clusters of the unmigrated section. From the Spessart mountains in the north to the Ries in the South the top of the crustal reflection zone descends from 1.5 s to 5 s.

Based on the migrated sections a line-drawing, containing the most reliable reflections, is presented in Fig. 24. The positions of the apices of the most prominent diffractions taken from the unmigrated section have been added. The southeast dipping reflectors below the Spessart are clearly defined. Their dip decreases at about 5 to 6 s TWT, but there are again strongly dipping events in the lower crust below Rothenburg. In the upper crust a zone of increased reflectivity is seen northwest of Würzburg.

3.2 Depth and structure of the Moho

Reflections have been observed along DEKORP-profile 2-S down to a TWT between 8.5 and 9.5 s. Below the corresponding depth only very rare seismic events have been

¹ BEB Brigitta and Elwerath Betriebsführungsgesellschaft mbH, Hannover

Fig. 20. Uninterpreted and unmigrated stack of DEKORP 2-S. Vertical exaggeration approximately 1.5:1. Mean coverage about 22 fold. Processing parameters: predictive deconvolution with 188 ms operator length and 4 windows from 3,000–12,000 ms, AGC window: 1,000 ms. Plot parameters: 2 fold vertical stack, 12 Hz (24 dB/oct) – 45 Hz (42 dB/oct) filter, AGC window: 400 ms. Note that diffraction hyperbolas are the dominating events

Fig. 21. Unmigrated stack of DEKORP 2-S with some interpretation. Parameters as explained in caption of Fig. 20

NW

SE

RHENOHERC.

SAXOTHURINGIAN

MOLDANUBIAN

Devonian | Tertiary | Crystalline Basement | Bunter | Muschelkalk | Keuper | Impact Breccia | Malm

TAUNUS | WETTERAU | SPESART | FRANCONIAN PLATFOM | RIES | SWAB. ALB

Friedberg

Hanau

Laufach

Main

Würzburg

Tauber

Rothenburg

Dinkelsbühl

Nördlingen

Donauwörth

0

0

5

5

10

10

s
(TWT)

s
(TWT)

15

15

20

20

km 250

200

150

100

50

0 km

5

5

10

10

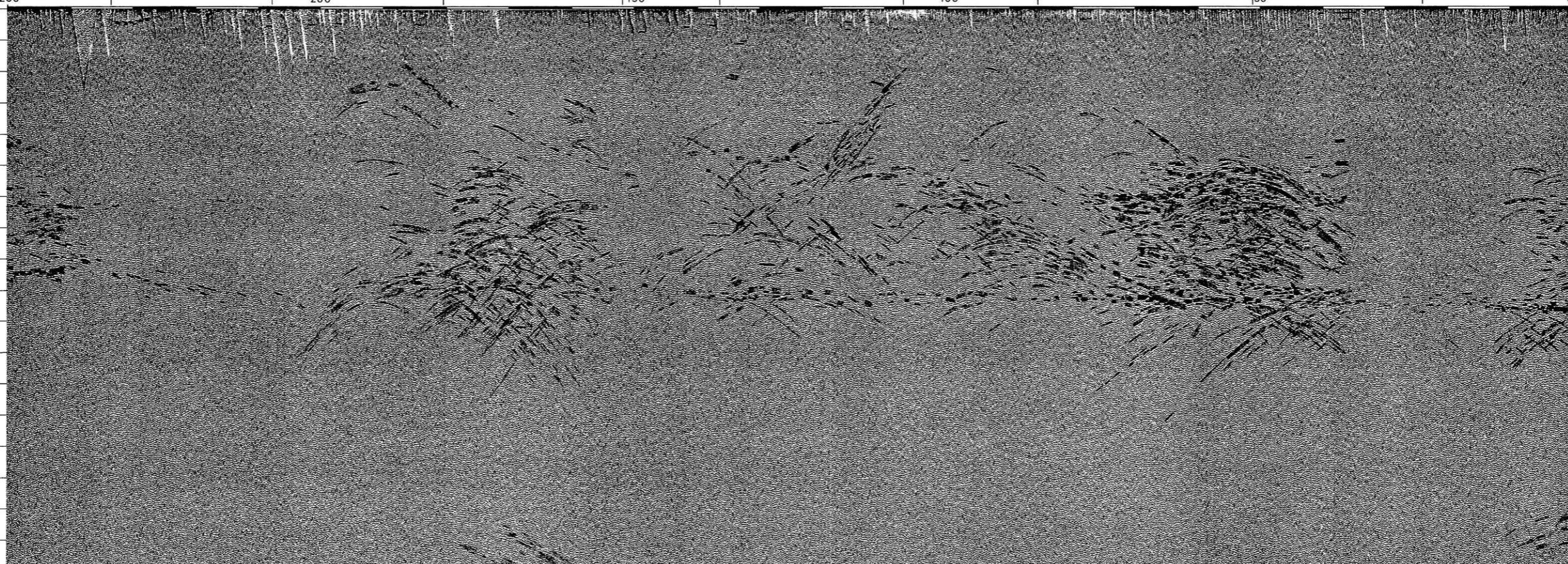
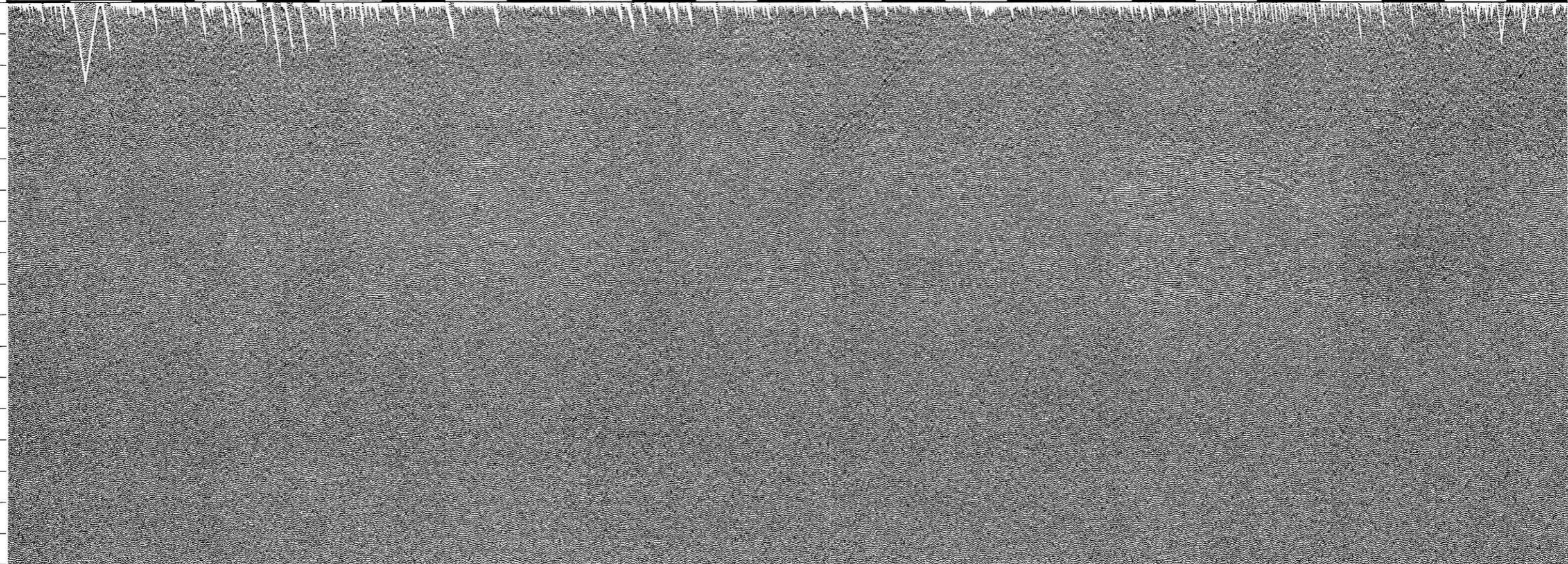
s
(TWT)

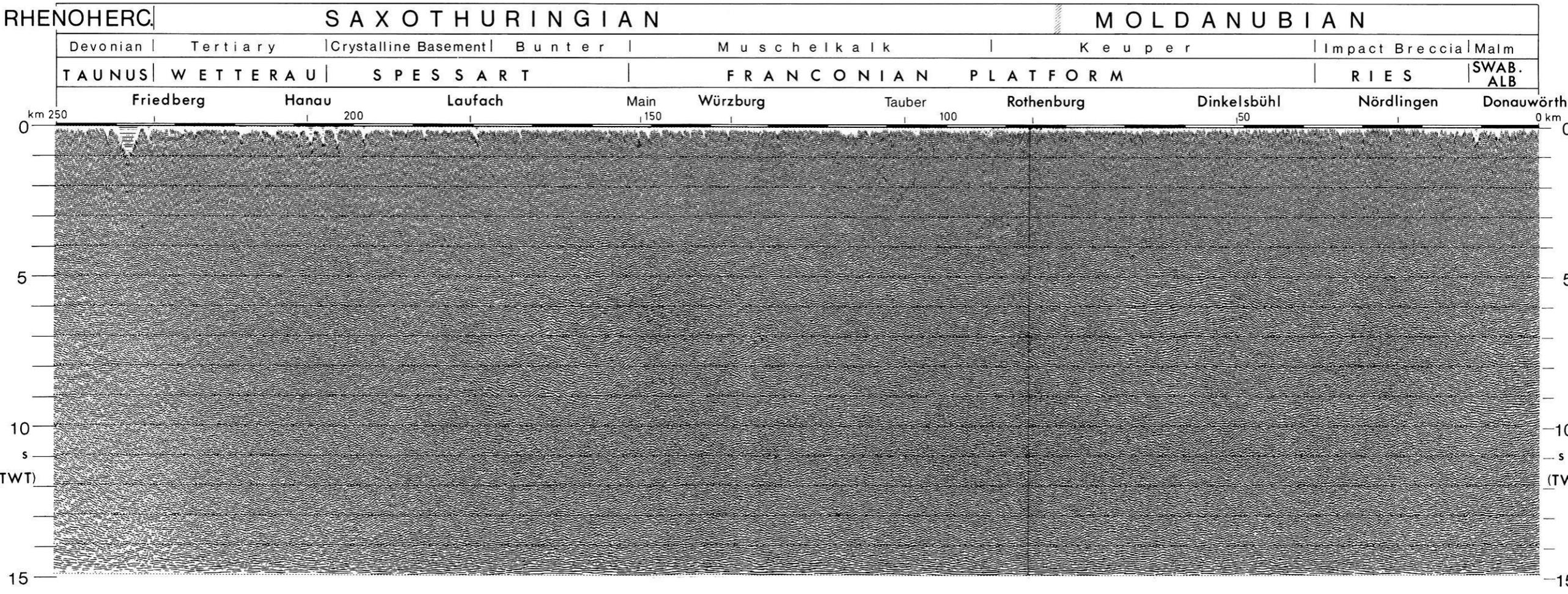
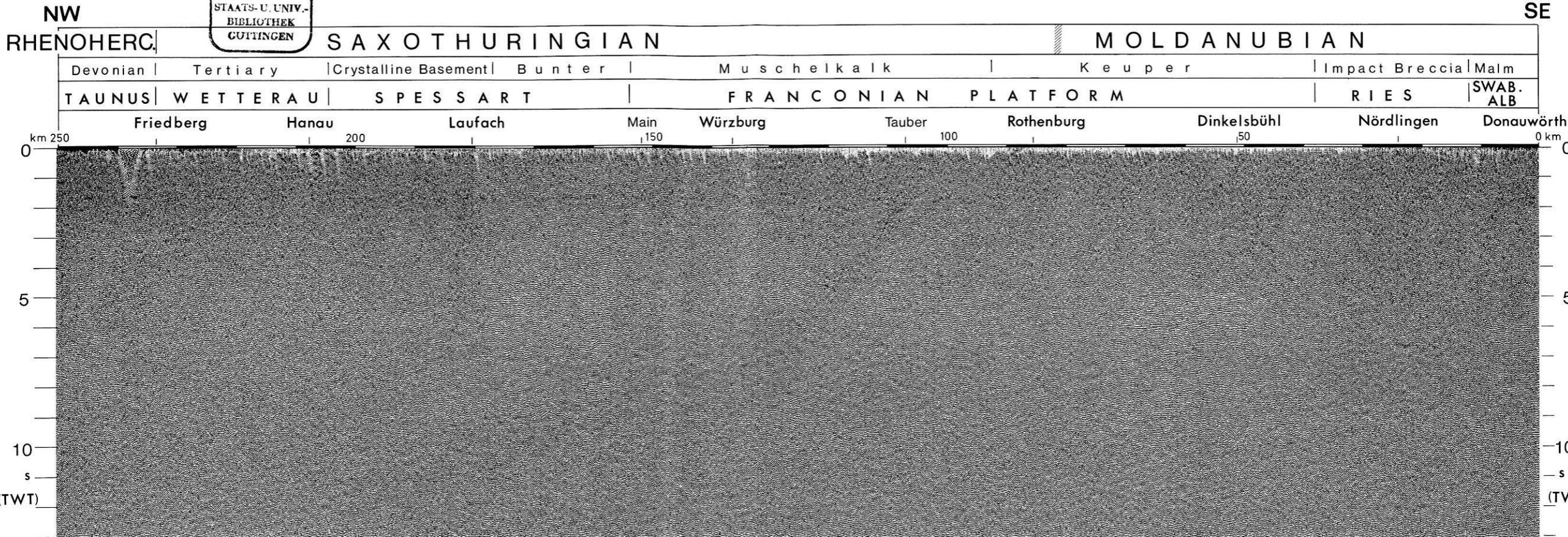
s
(TWT)

15

15

21





recorded. The reflections between 8.5 and 9.5 s TWT are not continuous, but a band of reflections can be defined which can be followed all along the profile with some interruptions. Converting to depth with an approximate average velocity of 6,100 m/s (see Sect. 4.3 and Bartelsen et al., 1982; Meissner et al., 1982), this band of reflections corresponds to a discontinuity zone at 26–29 km depth. Regarding the information about the depth of the Moho-discontinuity from the most recent publications (Mostaanpour, 1984; Gajewski and Prodehl, 1985), depths of 26 (near the Taunus) up to more than 28 km in the area of Nördlingen are given for locations closest to the profile.

It is intended to carry out much more detailed velocity studies from the reflection data. Only with more reliable velocity data at hand, one should discuss the question whether there is a discrepancy between the Moho as determined from reflection and refraction data, respectively, and whether this is due to anisotropy or other reasons. At this stage, we state that the somewhat discontinuous band of reflections at 8.5–9.5 s TWT at the base of the reflective crust can be considered as the Moho zone, and that there is no prominent anisotropy effect.

On the DEKORP 2-S profile the Moho below the Ries astrobleme seems to exhibit a weakly anticlinal structure. The updoming effect would still be reinforced if the correct velocities obtained from the more detailed analysis would be lower than those used so far. Such effects have not been observed in the earlier E-W profile through the western rim of the Ries (Angenheister and Pohl, 1969). Thus, no safe conclusions can be drawn at this stage.

The Moho band reveals some tectonic structure in the Dinkelsbühl area; this is shown in more detail in Sect. 4.2 where the diffraction data of that area are discussed. At the NW end of the DEKORP 2-S profile the Moho band is faulted with the upthrown side below the Taunus. The overall geotectonic situation can be solved only after the extension of profile 2 to the north.

3.3 Attempt of tectonic interpretation

It is evident that there is more information from reflected energy in the central part of the profile, approximately between Dinkelsbühl (km 50) and Hanau (km 210), than at both ends. As this part is the general area of the Saxothuringian (ST) zone, this might be the seismic expression of the different types of crustal structure in the Rhenohercynian (RH), ST and Moldanubian (MN) zones of the Variscan belt as described in Sect. 1.3. The higher reflectivity of the central section might have its cause in the remnants of highly active horizontal tectonics with crustal shortening in the ST zone. The upper crust in the Dinkelsbühl area is nearly void of any reflections, thus the suture zone between ST and MN might be dipping SE, and there is no contradiction between the MN-gneisses drilled and the seis-

mic character of the lower crust in the same area (see Fig. 3). With the clustering of reflections in the lower crust the DEKORP 2-S section fits well in the general picture of reflectivity in the Variscan system showing a highly reflective lower crust embedded between the poorly reflective upper crust and upper mantle (Meissner et al., 1984).

In the northernmost part of the ST zone, a narrow belt of lowgrade rocks is exposed in the southern Taunus. This "Northern Phyllite Zone" is a zone of intense tectonic imbrication and polyphase deformation (Weber and Behr, 1983; Weber 1984). Since phyllitic rocks occur as tectonic shavings at the base of greywacke nappes in the RH, the Phyllite Zone is regarded as the root zone of these nappes. This implies a distance of transport of at least 60 km. At the junction of RH and ST zones (Behr, 1978; Weber and Behr, 1983) the Mid-German Crystalline High (MGCH) overrides the Phyllite zone, and both are thrust northward over the lower-grade rocks of the RH zone proper.

At the southern border of the ST a similar geological situation is encountered, where at an analogous position of the tectonic framework in the western Bohemian Massif supracrustal nappe piles were generated and transported to the northwest (e.g. the Münchberg nappes). The predominance of higher-grade metamorphic rocks in the overriding units is an expression of the same principle of metamorphic and stratigraphic inversion, which is also inferred for the intracrustal thrust-sheets that contain granulites, eclogites, garnet-peridotites and serpentinites. These phenomena could probably also be explained if one regards the overriding units at the margins as small-size crustal elements (microcontinents or exotic terranes, see, e.g. Ziegler, 1982) which during the process of collision underwent a high degree of tectonic-metamorphic reactivation, so that their margins are transitional between an intracontinental suture and a nappe-thrust.

Principally in all Mid-European pre-Permian crystalline regions the following steps of development can be observed:

- A medium-pressure/high-temperature metamorphic event (granulites, eclogites with integrated tectonic slices of high-pressure rocks).

- A subsequent mylonitization, shearing and intracrustal thrust tectonics with ductile and brittle deformation affecting the entire basement. This type of deformation is penetrative and can be observed on a regional scale. The corresponding structures in the ST are known since the beginning of this century and were studied and described in connection with the Variscan nappe tectonics. Recently, similar structures were recognized in the Black Forest and at the western border of the Bohemian Massif. Interpretation of their structural evolution is supported by numerous micro-tectonic studies (Behr, 1978, 1983). They comprise thrust sheets and large-scale antiforms of granulites with lateral extensions exceeding 50 km which later pierced through the overlying rocks (e.g. "Granulitgebirge"). They also include huge "horses" (phacoids), thrust slices, refolded simplex-

Fig. 22. Migrated stack of DEKORP 2-S by use of wave equation migration (type: finite difference, 52 ms interval). Parameters as explained in caption of Fig. 20. Migrated and plotted by BEB Brigitta und Elwerath Betriebsführungsgesellschaft mbH, Hannover

Fig. 23. Migrated stack of DEKORP 2-S processed by DEKORP Processing Center, Clausthal, and plotted by BEB Brigitta Elwerath Betriebsführungsgesellschaft mbH, Hannover. Processing parameters: before FD-migration bandpass filtering as given in Fig. 12 and additional 30–45 Hz antialiasing high-cut filter; after FD-migration random scaling with 1,000 ms window and 500 ms shift; bandpass filtering with 5–10 Hz low-cut and 20–35 Hz high-cut. Plot parameters as explained in caption of Fig. 20

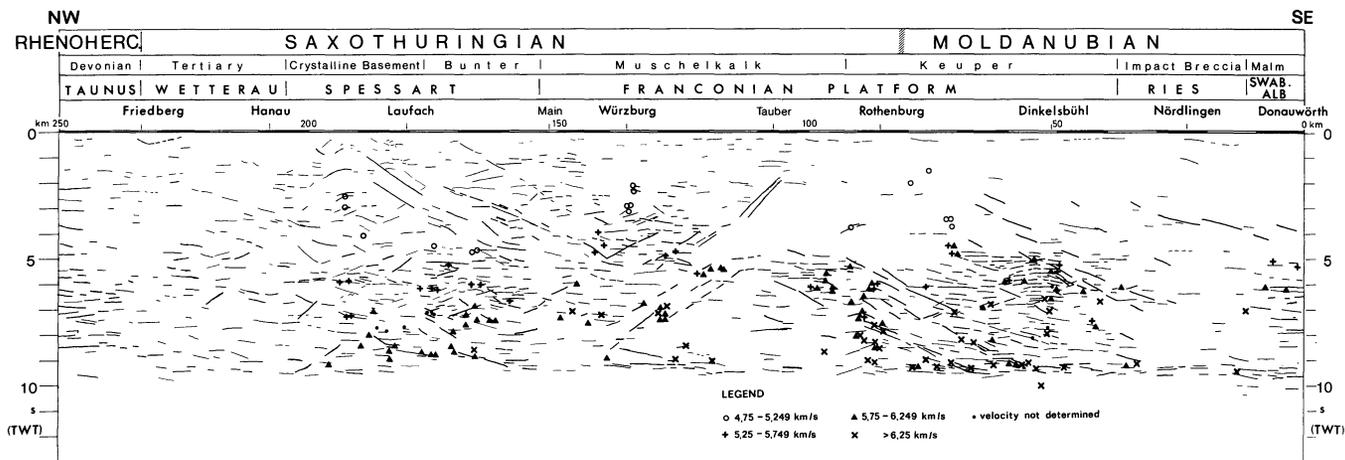


Fig. 24. Line drawing of good and strong reflections picked from the migrated sections in Figs. 22 and 23. The symbols indicate the position of apices of diffraction hyperbolas and estimated mean effective velocities obtained from analysis of the stacked section in Fig. 20 with the aid of maximum convexity master curves. Vertical exaggeration 1.5:1

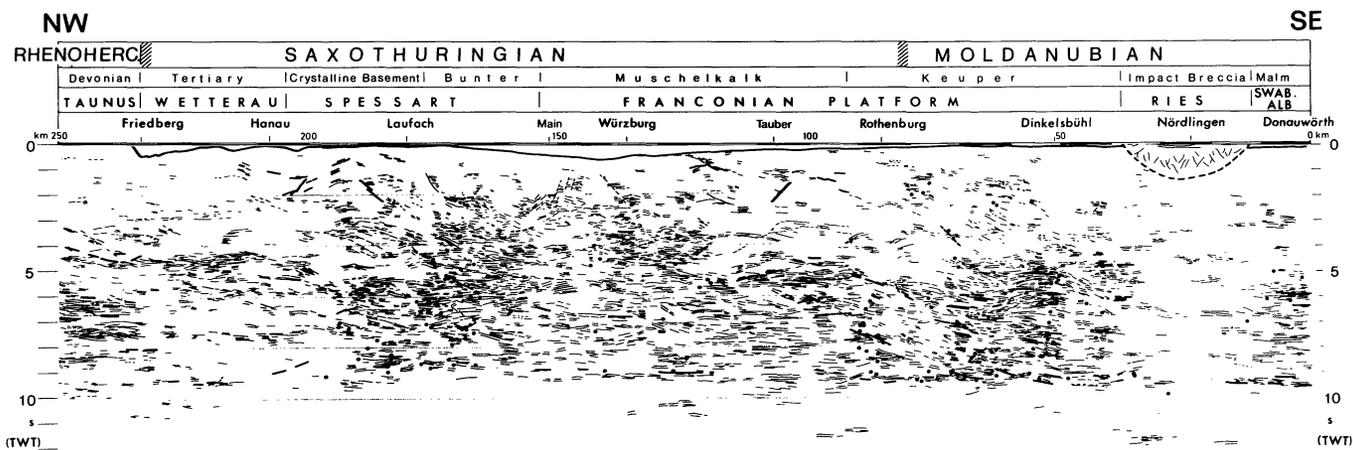


Fig. 25. Line drawing of profile DEKORP 2-S derived from the migrated section which contains all recognizable reflections including also questionable ones. Thick lines indicate high quality reflections; points are apparent diffraction apices as determined from the unmigrated stack with the aid of master curves. The continuous line below datum represents the base of the Permian derived from seismic and drilling results. The dashed line below the Ries outlines the brecciated basement. Vertical exaggeration 1.5:1

and duplex-structures, blastomylonitic belts and zones of cataclases at various scales down to the microscopic.

– Finally, there follows a high-temperature metamorphic event producing migmatite gneisses and granitoides or partially overprinting most earlier deformation structures by recrystallization or blastesis.

The following preliminary interpretation is based on all the geophysical and geological information presented above. In order to provide more detail than in Fig. 24 it has been tried as an intermediate step in Fig. 25 to include also a large number of short and weak reflection elements; many of them might be questionable. The result of the geological interpretation is shown in Fig. 26.

At the northern border of the ST a large accretionary wedge-type interface between ST and RH is expected (Weber and Behr, 1983). First indications of high-pressure (10–12 kbar) relics in the Northern Phyllites (Masonne and Schreyer, 1983) make probable a situation like the one shown in our interpretation. The assumed depth extension of RH and Northern Phyllite Zone metasediments definitely

needs further testing by detailed velocity analyses. We further assume that out of the suture zone a decollement developed toward north that was initially tied to the base of the weakly metamorphic RH. Out of this horizon numerous listric shears formed that correspond to reverse faults at the surface.

The RH zone seems to be distinguished by its originally at least 6–8 km thick metasedimentary cover overlying a rheologically distinctly different basement. This, we suppose, renders it likely that the displacement was largely transferred by a single thrust zone at the base of the supra-structure. It would imply less structural perturbation in the RH basement and thus more persistent subhorizontal reflections could be expected north of the RH/ST suture.

At the border ST/MN the situation is different. Here, two polymetamorphic crustal segments collided. In contrast to the northern border, the shear zone dips steeper and causes stronger rotations and more intensive interstacking, starting in the basement and continuing into the supra-structure. Therefore, particularly in the southern part of the

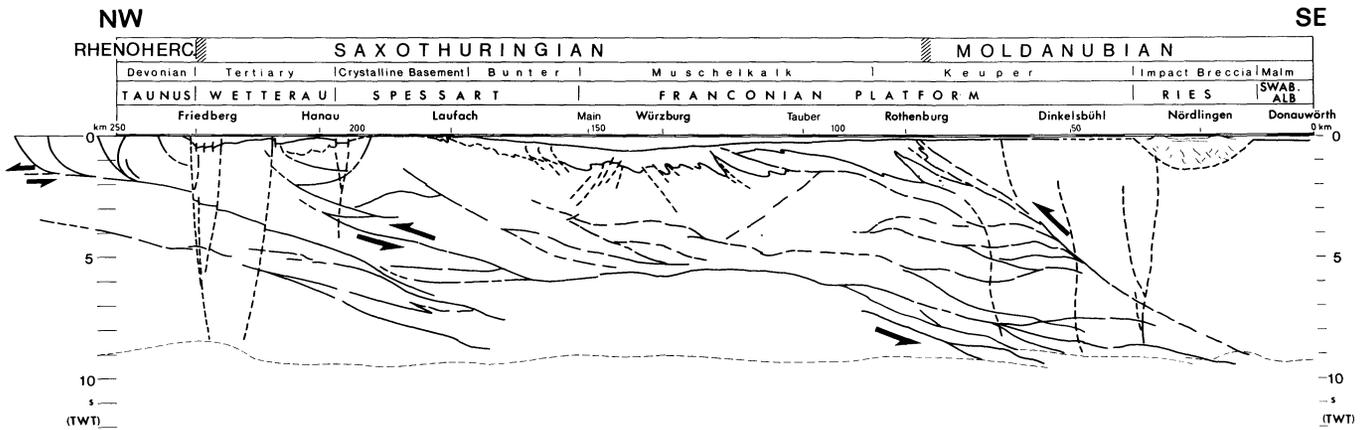


Fig. 26. Preliminary structural model along profile DEKORP 2-S based on a thrust concept. The first line below the datum shows the base of the Permian. Dashed lines indicate hypothetical downward extensions of post-Permian normal faults. For further explanation see text. Vertical exaggeration 1.5:1

border zone isolated fragments from the lower crust and upper mantle were transferred even into the top of the upper crust. This is also supported by the southeast dipping strong reflections reaching down to the Moho in the Rothenburg-Dinkelsbühl area. An equivalent situation is known at this border zone in the Bohemian Massif (Behr et al., 1984).

The assumption of a thrust-body below the Tauber is only vaguely supported by the seismic data. There are only some short reflections. But there are strong magnetic and gravity highs in the Tauber area (Figs. 4 and 5 in Sect. 1.3). This hypothetical structure may be regarded as an analogue to the "Münchberger Gneissmasse" situated in an equivalent tectonic position farther east.

With regard to the basement depth of the ST sedimentary cover, no information can be retrieved from the seismic section. It was assumed on the basis of general knowledge about the structural style and approximate thickness of the ST suprastructure.

The optimal estimation of exact interval velocities in combination with gravity model calculations could help to decide whether the crust was doubled along the south dipping structures on a large scale. However, there exists certainly a large system of thrusts between the northern and southern border of the ST. This is indicated by reflection structures as well as by the arrangements of diffraction apices along SE dipping lines, e.g. below Rothenburg and the Spessart (Fig. 24 and 25). The middle portion of the shear zone seems to parallel the top of the lower crust. At its northern extremity it should accommodate considerable displacement. For instance, the crystalline Spessart should have been uplifted along those shears from the level of 4–5 kbar (Matthes, 1954).

Regarding the southern part of the shear system there should be a direct connection to the Moho. This might also be indicated by the diffraction and reflection elements in single seismograms recorded in the area north of the Ries astrobleme (see Fig. 18 and Sect. 2.5). The seismic picture below Dinkelsbühl and the Spessart may be explained by the deformation processes that have been active at the ramps of this thrust system.

Ductile shears next to brittle deformation may develop due to:

- differences in viscosity due to strong composition contrasts, e.g. ultramafic next to felsic,

- dry next to H_2O -rich compositions,
- inhomogeneous migration of fluids as well as fluid controlled heat transfer,
- melting due to pressure decrease during uplift,
- mineral reactions entailing volume changes and dehydration.

Crustal domains showing strong inhomogeneities with respect to the above parameters will, therefore, tend to develop into an aggregate of sheared bodies of limited extension (e.g. huge horses). Map patterns from the neighbouring basement outcrops of the western Bohemian Massif and the Black Forest reflect such complex situations. The exposed basement appears as a complex puzzle of lithologically diverse tectonically laminated units displaying arcuate patterns of planar and linear fabric elements with mylonites and cataclasites intervening where attitudes of layering and foliation change abruptly.

It has been tried also to derive a fracture pattern which attempts to explain Variscan basin formation and deformation purely by normal faulting (Fig. 27). Though we believe that the ST basin was initiated by crustal stretching and rifting, we also know from surface geology that subsequent deformation led to considerable crustal shortening. Some of the preexisting, probably listric, extensional faults may have been reactivated with opposite shear sense during the latter event. However, the hypothetic fracture pattern as shown in Fig. 27 can by no means accommodate the shortening necessary to bring for example eclogites into the suprastructure as exposed in the Münchberger Massif which represents an analogous tectonic position. In domains of high pore pressures and concomitant low effective shear stresses extensional faults and veins may form locally even in shear zones where ductile rheology changes into a brittle one. In this way diffraction points and edges (Fig. 24) may also be explained.

The late- and post-Variscan development led to the subsidence of Permian and Tertiary basins through extensional faulting. This may explain slight crustal thinning and Moho steps. The apparent arrangement in a near-vertical pattern and in the strike direction of the Variscides and the concentration of diffracting elements in the lower crust below Dinkelsbühl (Fig. 21 and 24) may point to magmatic intrusions along such vertical extensional features, too. We consider this a possible alternative cause for the Dinkelsbühl and

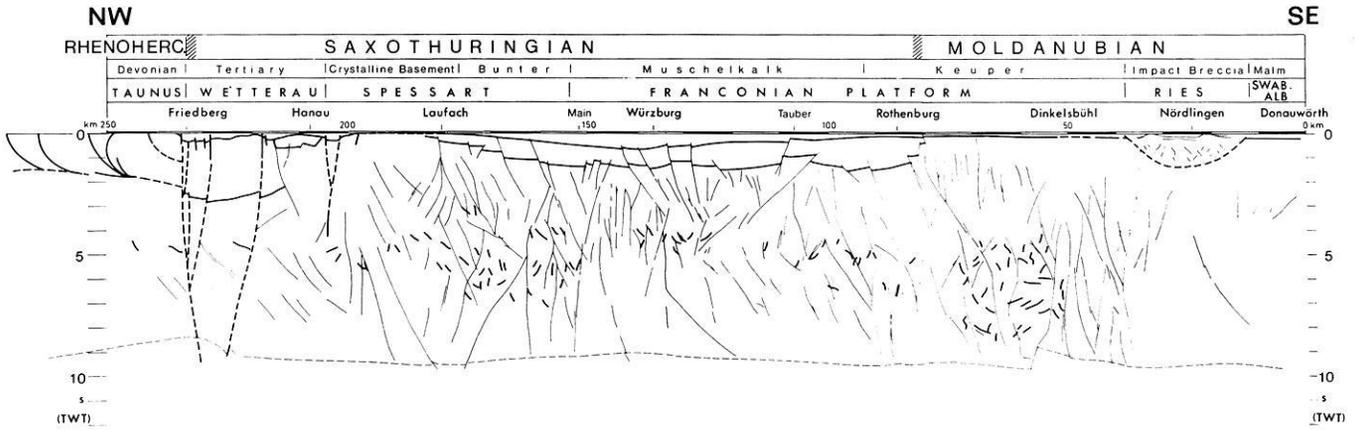


Fig. 27. Alternative structural model based on extensional tectonics concept. For further explanation see text. Vertical exaggeration: 1.5:1

Spessart diffraction clusters, though Permian and younger basins are missing from the above areas. The fault pattern north of the Spessart results from an analysis of surface geologic data (Wenz, 1936; Golwer, 1968) and of a shallow reflector interpreted as the top of the pre-Permian basement.

4. Accompanying seismic activities and their first results

4.1 Near-vertical observations of crews from universities and other research institutions

The objectives of the additional reflection observations on lines Q_1 and Q_2 (Fig. 2) were:

- Studying the lateral continuation of crustal reflection data in order to obtain a quasi 3-D picture in a certain area, especially the construction of parallel and perpendicular profiles and the determination of true strike and dip angles.
- Collecting data from undershooting the Ries astrobleme.

The additional reflection observations were carried out in the area north of the Ries astrobleme (line Q_1) by means of five individual digital reflection recording units² with a total capacity of 120 channels. Using the shots along the main profile and five additional shots west of it, a three-dimensional control of the subsurface was obtained in an area of about 200 km² size east of the main profile (see Fig. 28). The major part of the additional geophone groups, i.e. 108 traces, were placed perpendicular to the main profile, the other 12 groups parallel to the main line (Fig. 2). Because of the different directions between shots and geophones an areal pattern of 24 geophones with a rather constant angular characteristic was used for all 120 traces. This pattern remained fixed throughout the operation. A grid with squares of 40 × 40 m² was applied for the construction of “common midpoints” (CMP) in the area, while in Fig. 28 only the coverage by 320 × 320 m² squares is shown. Thus, a six-fold coverage was obtained on average. Amplitudes were normalized to those of the first arrivals, and no further automatic gain control (AGC) was necessary, an effect of the large offsets. A band-pass filter from 12–45 Hz was used, the dominant signal frequency being around 20 Hz.

² Operated by teams from Clausthal, Hamburg, Hannover and Kiel

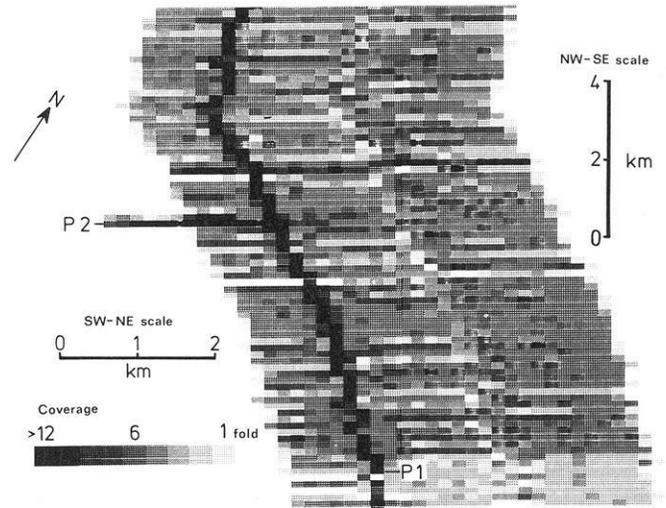


Fig. 28. Degree of coverage by 320 × 320 m² CMP squares for the additional near-angle observations of the university teams. Vertical to horizontal scale ratio is 2:1. Black squares indicating the highest degree of coverage display the position of the parallel profile P1 (see Figs 30–32)

The construction of a parallel profile P1. The 12 parallel geophone groups and the shotpoints along the main profile were used as a basis for the construction of the first parallel profile P1, parallel with DEKORP 2-S (see also Fig. 29). It starts in the center of the Ries astrobleme, at about km 20 of the main profile, and ends at km 72 in the north. Two consecutive traces were interpolated and plotted at the common midpoints along the line. The maximum offset-difference for stacking was 3 km. Stacking velocities from refraction data in the vicinity and from reflection data along the Urach profiles (Bartelsen et al., 1982; Meissner et al., 1982) were used. Special corrections were applied for the area of the Ries astrobleme in order to take account of the near-surface low-velocity-layers.

The stacked seismogram section of P1 is shown in Fig. 30, while Fig. 31 shows a line drawing of the major reflections (solid lines) and diffractions (broken lines). The density of reflections and diffractions is similar to those along the main profile, and most of these events can be correlated. The reflections and the diffractions concentrate between 5.5 and 9.5 s TWT. Figure 32 shows an “energy

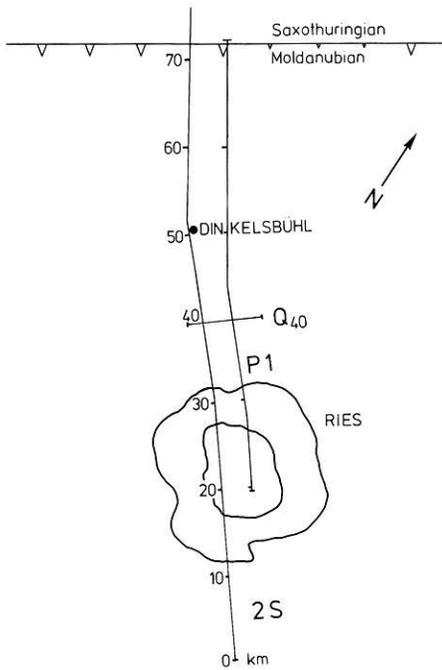


Fig. 29. Location map of the parallel profile P1 and perpendicular profile Q40

section" of P1. For this section an average of 8 consecutive traces equivalent to 320 m was formed and a moving time window of 200 ms was applied. Squared amplitudes have been normalized with regard to the maximum amplitude per trace. This energy section is dominated by the energy of diffractions which can be followed across the whole P1-section, being especially strong in the northern sector. Also in this presentation Conrad- and Moho-bands are clearly defined, sometimes also a sub-Conrad, even below the Ries astrobleme in the south. But also individual correlations of reflections from the Moho- and Conrad-bands (Fig. 31) as well as of diffractions can be made all along P1.

The perpendicular profile Q40. Profile Q40 is in the strike direction of the Variscides and is one of six perpendicular profiles constructed so far. It crosses the main profile at km 40 (as seen in Fig. 29) and makes use of 5 additional shotpoints west of the main profile and the 108 geophone groups east of it (see Q1 in Fig. 2). The 5 additional shots and 3 shots from the main profile around km 40 were used for stacking. The width of the grid was increased to $80 \times 80 \text{ m}^2$.

Figure 33a presents the profile Q40 without interpretation. Its quality is slightly better than that along P1, especially reflections come out clearer than the diffraction events. Only some of these diffractions can be tied to and correlated with those along the main profile, profile P1, and the additional perpendicular profiles (not shown here). See also the line drawing section, shown in Fig. 33b. Again, reflection bands from Conrad and Moho mark the beginning and the termination of the highly reflective lower crust. Also the "energy section" (Fig. 33c) shows these two reflection bands.

On the positioning of diffraction sources. Special attention was given to the Dinkelsbühl diffraction pattern along the DEKORP 2-S profile which can also be detected along P1 (Figs. 30–32). In a first attempt, a correlation of the main

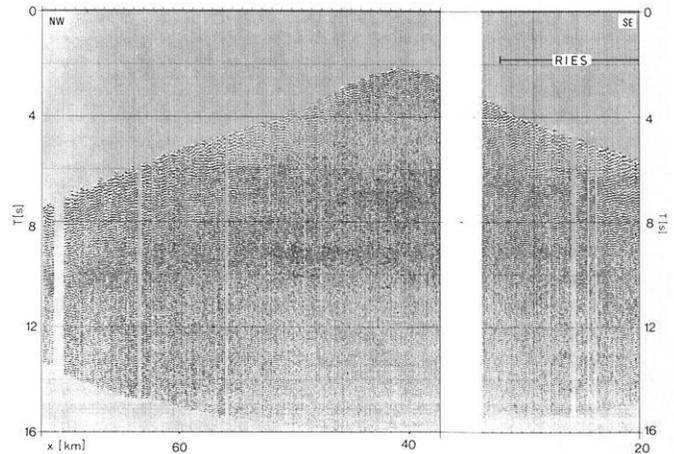


Fig. 30. Uninterpreted and unmigrated stack along the parallel profile P1. Horizontal exaggeration approximately 1.5:1. The gap is caused by technical reasons. For further explanation see text

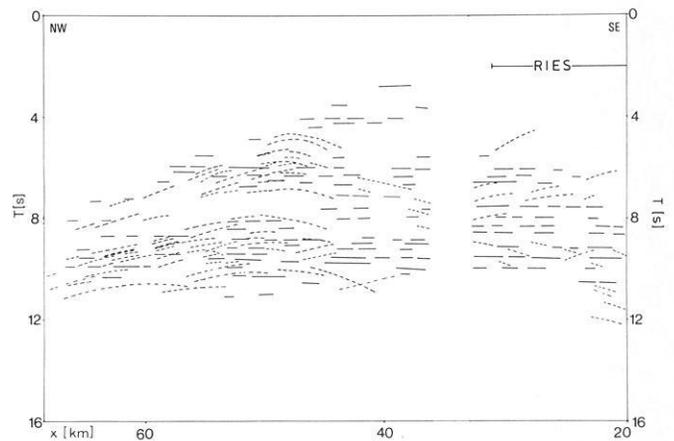


Fig. 31. Line drawing of the parallel profile P1. Horizontal exaggeration approximately 1.5:1. For further explanation see text

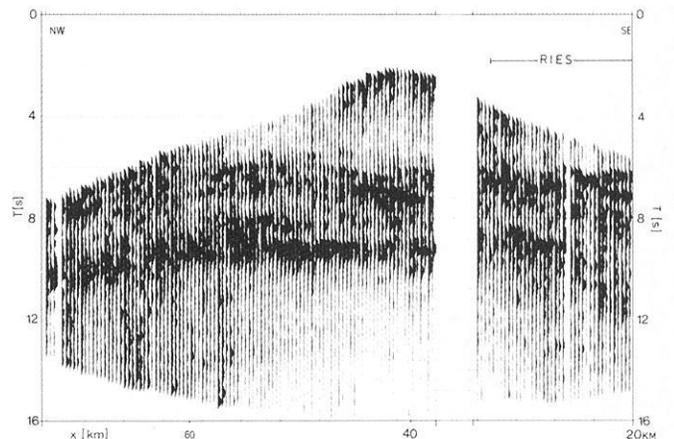


Fig. 32. Energy section along the parallel profile P1. Horizontal exaggeration approximately 1.5:1. For further explanation see text

events from DEKORP 2-S to Q40 (at km 40) and again to P1 was tried, but only bands from prominent diffractions or reflections like those of the Conrad and Moho, not single phases, could be successfully correlated around the three lines. Some of the Dinkelsbühl diffractions could be followed onto Q40, and here their dip could be determined. Some of them appear as horizontal events, indicating a

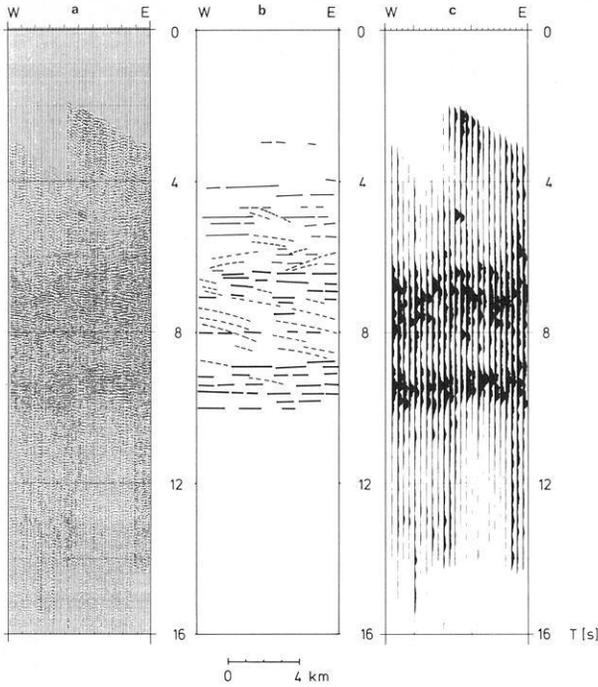


Fig. 33a–c. The perpendicular profile Q40: uninterpreted version (a), line drawing (b) and energy section (c). Horizontal exaggeration approximately 1.5:1

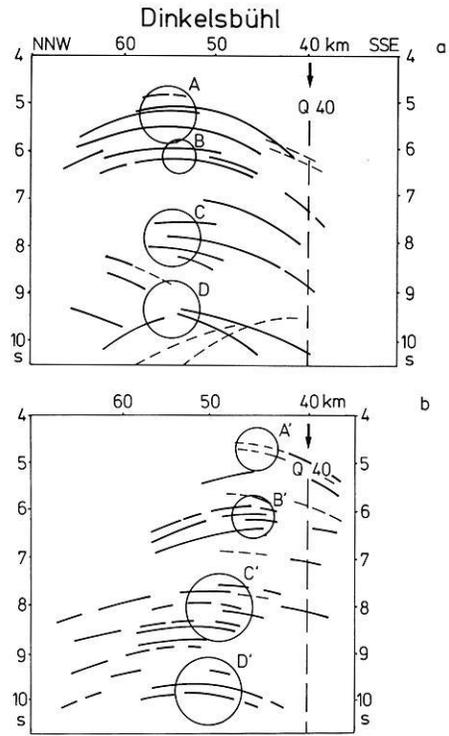


Fig. 35a, b. Significant diffraction curves with clusters indicated along DEKORP 2-S (a) and along P1 (b)

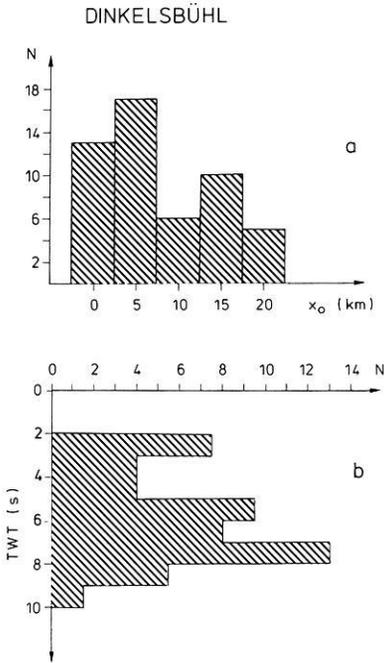


Fig. 34a, b. Distribution of lateral offsets of diffraction apices (a) and frequency-time distribution of corrected apices (b) (x_0 = lateral offset; N = number of apices)

line diffractor in the Variscan strike direction parallel to Q40. The early diffractions come from sources on both sides of the main profile; the later ones show westerly components, i.e. their apices lie east of the main profile.

For a better positioning of the apices, several diffraction charts (= curves of maximum convexity) were constructed for offsets of 0, 5, 10, 15, and 20 km from the profiles.

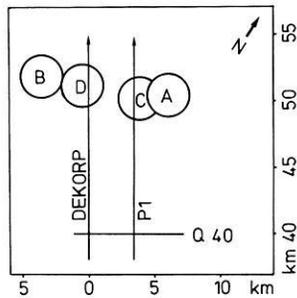


Fig. 36. Location map of diffraction clusters (A) to (D) and seismic lines

For the Dinkelsbühl diffraction cluster along DEKORP 2-S the best fit of the curvature of the events allowed to determine the offset of the individual diffractions. The E-W ambiguity could be resolved by means of the profile Q40, as mentioned above.

Figure 34 shows the distribution of the Dinkelsbühl diffraction apices as a function of lateral offsets and of TWT. Apices in Fig. 34b are corrected for lateral offsets. The peaks of the distribution resemble those of the Conrad and Moho bands. Figure 35 presents a line drawing of the most powerful and correlatable diffractions along the DEKORP 2-S and the P1 profiles and some clusters A to D. Although no exact phase correlation could be performed, the correlation along “bands” and the similarity of diffraction clusters on the DEKORP and the P1 line definitely permit a positioning of the main clusters. Their location is shown in Fig. 36. The various aspects of the interpretation of the diffraction sources are discussed at the end of Sect. 3.3.

4.2 In-line wide-angle observations by means of the reflection spreads on the main profile

27 additional shots with charges of 90 kg were fired in-line along the main profile (km 105–148). They had a distance of 65 km from the center of the 16 km long reflection spread and were recorded by the 200 trace contractor's equipment. This test was carried out in order to obtain additional structural and velocity information by comparing the wide angle with the near vertical reflection data using the same advanced technology of recording. The average distance between these shots was 1.5 km providing a five-fold coverage of the CMP elements.

The processing of the data presented here was performed by aid of a VAX 11/780 computer at the Federal Institute for Geosciences and Natural Resources (BGR) in Hannover. Two different methods for velocity determination were chosen. Both methods suffer from the large offset with unknown refraction effects and lateral irregularities in the velocities.

The first method makes use of the assumption of true hyperbolas as traveltimes branches extrapolated to zero offset times. The trial stacks resulted in a broad scatter of calculated velocity values, and only an average and two extreme velocity-time functions can be given (see Fig. 37, lines a, b and c).

In the second method trial stacks were related to an offset of 65 km. Straight lines and hyperbolas were used for the stacking of the data, and a uniform reduction of these stacks to the zero offset t_0 -time was performed as indicated by line d in Fig. 37.

As a next step four trial zero-offset stacks according to the four velocity functions a, b, c and d in Fig. 37 were made. Surprisingly the best results were obtained by using curve c, i.e. the high velocity extreme of the first method. The value of the used stacking velocity at 9.5 s TWT (i.e.

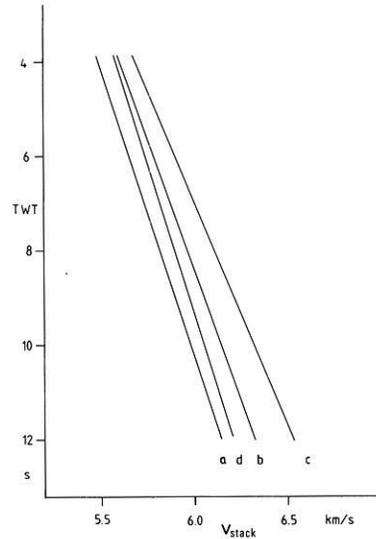


Fig. 37. Stacking velocity functions obtained from the evaluation of the in-line wide-angle observations. For further explanation see text

the Moho range) is 6.25 km/s which seems to coincide roughly with the values given in Sect. 4.3. The resulting stacked section using velocity function "c" is shown in Fig. 38. It has a 2.5:1 ratio of the horizontal to the vertical scale. Many features of the near-vertical sections (see Figs. 20–23) can be observed in the wide-angle section, too: the beginning of the highly reflective lower crust at 6 s TWT (especially in the southern part) and its termination at the Moho at about 9.5 s TWT. In the central part of the section a group of southeast dipping reflections at 7.5–8.5 s TWT is observed similar to those in the near-vertical section which shows diffractions in this region (Fig. 21). The zones of poor reflection quality are similar in both sections. Even

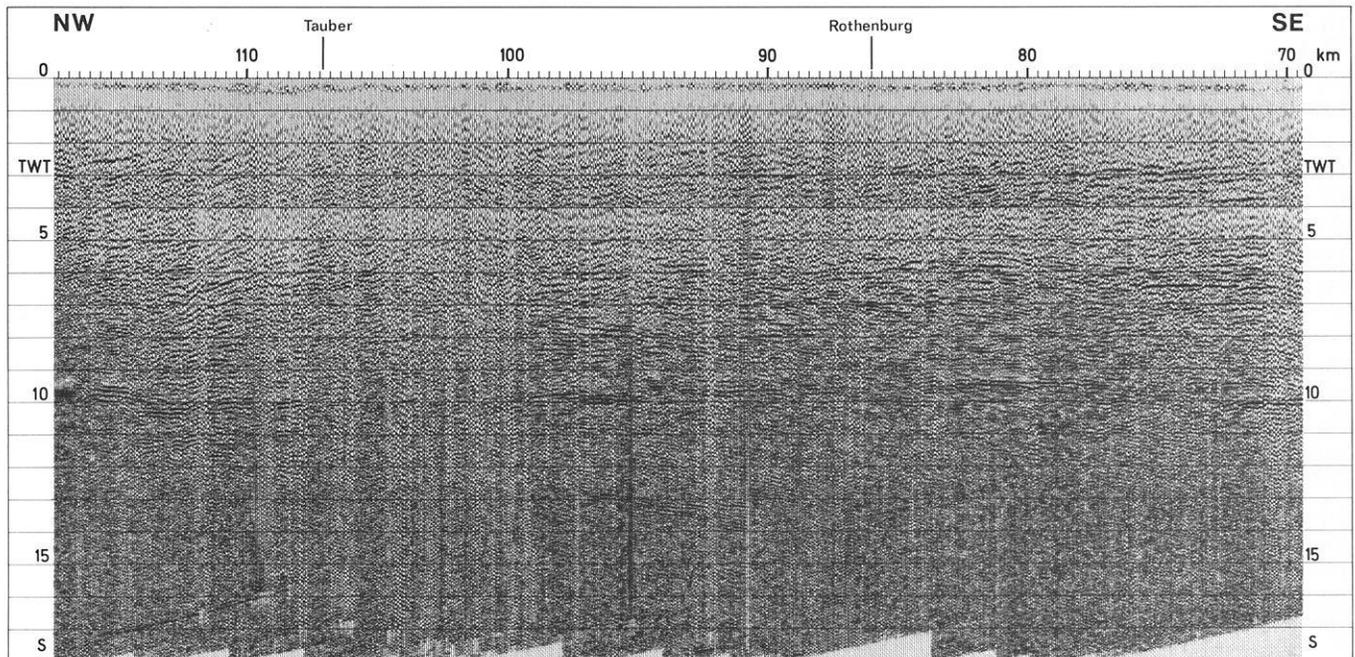


Fig. 38. Stack of the in-line wide-angle data by use of velocity function c (see Fig. 37). 6–25 Hz bandpass filter. Mean coverage about 5 fold. Horizontal exaggeration approximately 2.5:1. For further explanation see text

the continuity of reflections with a maximum phase correlation up to 15 km is comparable. Generally, the Moho comes out clearer in wide angle than in near vertical stacks, most probably an effect of a velocity gradient zone in the range of the crust/mantle transition.

4.3 In-line and off-line wide-angle studies of universities in the southern sector

The DEKORP shots in the southern sector were used additionally for refraction and wide-angle reflection observations with arrays of 10 and 24 MARS stations (Berckhemer, 1970), respectively³. The 10 station array remained stationary near km 100 of the main line (see Fig. 2) and thus a multiply covered in-line refraction profile was observed with shot-receiver distances up to 100 km; a total of 3020 3-component seismograms was recorded, which could only partly be processed so far.

A specially designed wide-angle experiment was carried out by the 24 station array along a parallel line with lateral offset of 50 km from the main profile; details are shown in Fig. 39. Between April 6 and 11, while shots moved from the center of the Ries astrobleme towards the presumed boundary between Moldanubian and Saxothuringian, the mobile 24 km long MARS array moved simultaneously from the center of the wide-angle line southeast of Neustadt to the northwest keeping the mean shot-receiver distance fixed at the critical Moho-distance of about 70 km. Starting April 12 the array was deployed at the southeastern end of the wide-angle line and followed the shots with the same mean distance, both units proceeding to the northwest. At the end of the experiment the stations occupied again their starting positions.

The main goal of the experiment was the investigation of the spatial structure of the crust along a traverse as extension of the structures observed on the main profile. The velocity and depth information derivable from wide-angle observations can be attached in a first approximation to the shot-receiver midpoints. They line up midway between the shot and station profiles, i.e. 25 km northeast of the main profile. For example, recording of a shot northwest of Dinkelsbühl by a 24 km long station array northwest of Neustadt yields information beneath a 12 km long midpoint profile "CMP" as indicated in Fig. 39. The same midpoints are again covered, when the shotpoint moved to Weikersheim and the station array is located east of Ansbach; ray paths in the latter case are more or less perpendicular to the ones in the first case. This offers several advantages in velocity and depth determination, when both sets of observations are jointly evaluated. However, these advantages can be exploited only after all data have been digitized and processed, and this has not yet been accomplished. At the beginning and at the end of the experiment the 24 MARS stations observed two additional refraction profiles obliquely connecting the main near-vertical profile with the wide-angle line as shown in Fig. 2.

Due to organizational and technical reasons it was necessary to record not only the 90 kg shots, as originally planned, but also all the smaller ones. In the southern part many of the small shots gave excellent results, but with shotpoints proceeding to the northwest seismic efficiency

³ Operated by teams from Berlin, Bochum, Clausthal, Karlsruhe, Kiel, Munich and Münster. The radio-communication was provided by the NLFb

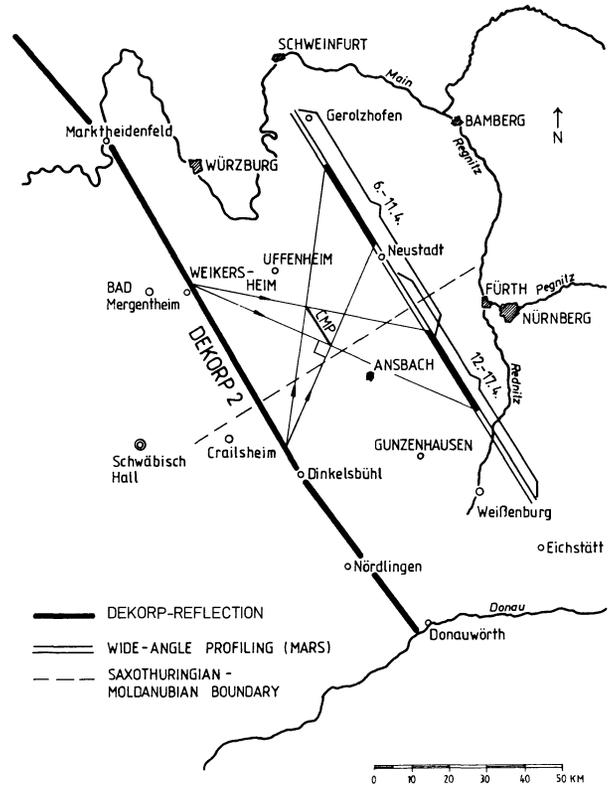


Fig. 39. Location map of the off-line wide-angle observations. The common midpoints are covered by perpendicular intersecting families of rays

became poorer. In the Muschelkalk area, where most shots had to be fired in partly dry limestone, even many 90 kg shots failed to overcome the mean ambient noise level of about 1.5×10^{-5} cm/s at wide angle distances. On the other hand, Moho reflection amplitudes up to 10^{-3} cm/s (at frequencies between 10 and 15 Hz) were obtained for many low-charge shots fired in the wet sediments filling the Ries astrobleme.

The following discussion is restricted to the external wide-angle line and is based on data from the first half of the experiment. About 50% of the 3,400 3-channel seismic records of this part of the experiment have been digitized so far. The rest was inspected but rejected from further processing due to insufficient signal to noise ratio, although within a series of less efficient shots some records of rather small signal to noise ratio have been retained for the sake of a uniform coverage. After digitizing and demultiplexing further processing was performed on a PDP 11/40 multipurpose computer of the Institut für Allgemeine und Angewandte Geophysik at the University of Munich.

Figure 40 shows a detailed location map of the stations and shots processed so far. Stations with odd numbers were equipped with 3-component seismometers and stations with even numbers had 3 vertical seismometers with a spacing of 300–400 m. On the basis of raw seismogram sections for every station poor seismograms were sorted out and the remaining 1892 vertical traces were passed through a high-cut 25 Hz-filter for further noise reduction. Dominant frequencies of most wide-angle seismograms are about 10–15 Hz.

In addition to the shotpoints and recording sites Fig. 40 shows the locations of the shot-geophone midpoints of the

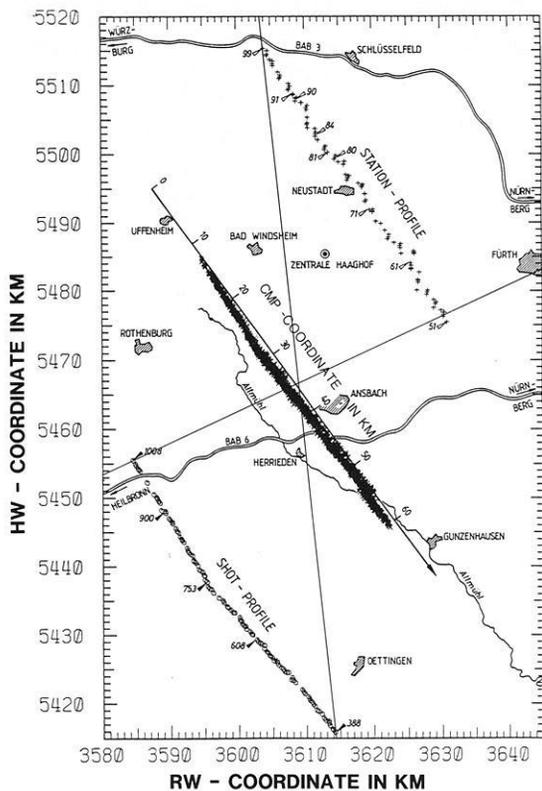


Fig. 40. Detailed location map of shotpoints, recording sites and common midpoints for the wide-angle data processed so far

selected vertical seismograms. Due to the multifold coverage the midpoints form an elongated black cloud. They can be described by a linear common-midpoint (CMP) coordinate, the origin of which has been fixed arbitrarily.

The large amount of data allows the construction of common-station (CST) and common-shotpoint (CSP) as well as common-offset (COF) and common-midpoint (CMP) seismogram sections. Some typical examples are shown in Figs. 41–48.

The most outstanding result is the strong crustal heterogeneity, which is reflected by the seismogram sections of adjacent shots and stations. For instance, station 84 (Fig. 41) shows clear P_g onsets and a pronounced and continuous Moho reflection (P_{MP}) with a reasonable apparent velocity of 7.8 km/s, but almost no correlatable signals in between. Station 92 (Fig. 42), however, located only 8 km apart gives a completely different picture with a disintegrated P_g phase, a complicated multiple Moho phase with extremely large apparent velocities and very strong intracrustal reflections of undulatory shape. Comparison of the common shotpoint sections 608 (Fig. 43) and 753 (Fig. 44) reveals similarly large differences within short distances. When more examples are studied, a trend of first decreasing and then increasing complexity is discernible with shots or stations moving from the southeast to the northwest. Short reflection segments between P_g and P_{MP} exhibiting very large apparent velocities on the shot sections but only moderate apparent velocities on the reversed station sections are frequently observed. This indicates the existence of generally southeast dipping reflecting elements in the middle and lower crust, in qualitative agreement with the near-vertical reflection results on the main profile. A detailed

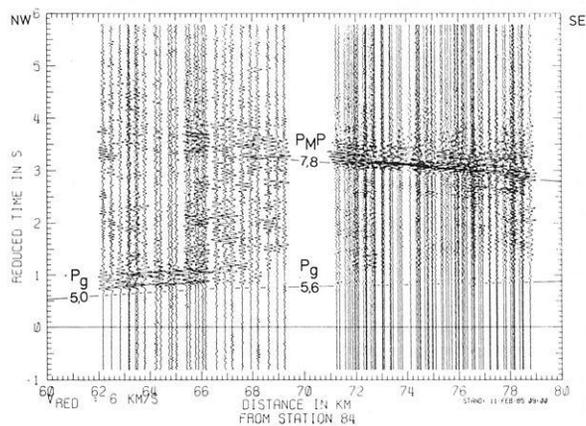


Fig. 41. Wide-angle record section for station 84 and moving shot-points. The Moho shows up as prominent reflector; the crust looks almost transparent. Maximum amplitudes in all traces are the same; this applies also to Figs. 42–48

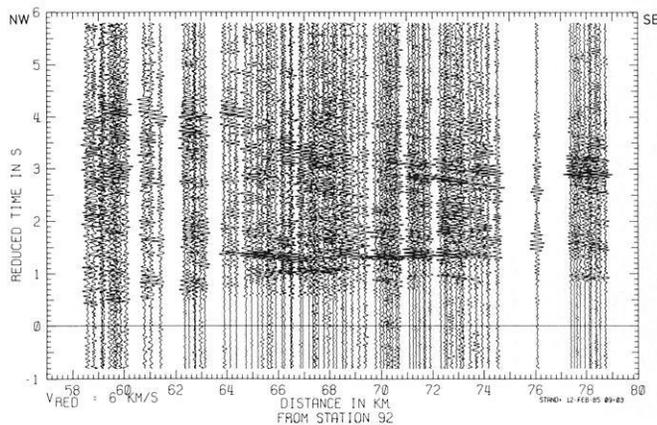


Fig. 42. Wide-angle record section for station 92, only 8 km apart from station 84 but exhibiting quite different crustal properties

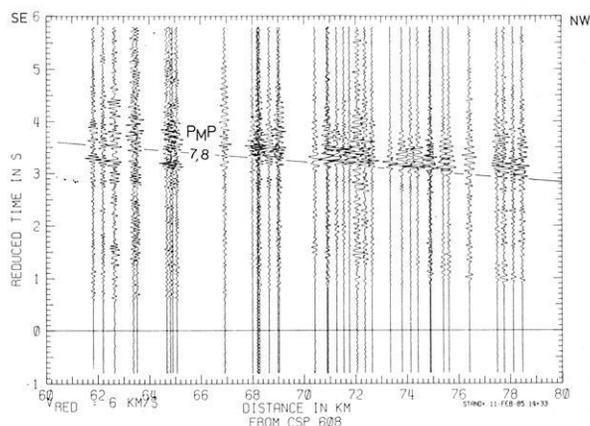


Fig. 43. Example of a common shotpoint (CSP) record section (SP 608) which is partly inverted in comparison to Fig. 41

evaluation of these phenomena will require 3-dimensional modelling and extensive ray-tracing calculations and is beyond the scope of this first report.

A more direct approach can be based on common offset and common midpoint sections, as shown in Figs. 45 and

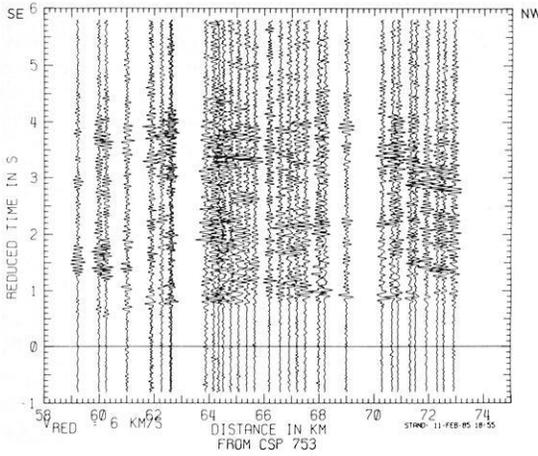


Fig. 44. Wide-angle record section for shotpoint 753, 10 km apart from shotpoint 608

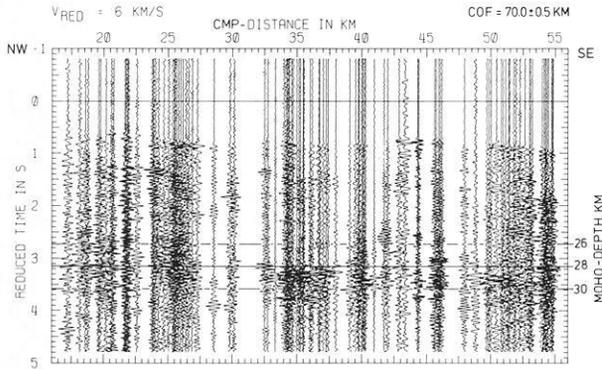


Fig. 45. Example of a common-offset wide-angle section. Offset is 70 km. Seismograms are plotted downward at the CMP-coordinate defined in Fig. 40

46. COF wide-angle sections yield a picture of crustal structure comparable to zero-offset reflection data and can be interpreted similarly. Figure 45 gives an example for an offset of 70 km. (Seismograms with shot-receiver distances within 70 ± 0.5 km have been corrected to the nominal offset according to an average apparent velocity of 7.0 km/s). For easier comparison with the near-vertical reflection data the reduced time has been plotted downward in Fig. 45; the horizontal axis is the CMP coordinate defined in Fig. 40. The Moho can be recognized at reduced times between 3.0 (CMP 30) and 3.4 s (CMP 50) and shows pronounced variations in depth, sharpness and continuity with in short distances. (Such variations would hardly be detectable with classical methods of deep seismic sounding involving few shotpoints and station spacings of several kilometers.) Some southeastward dipping structures within the crust are also discernible in Fig. 45. By tuning the COF distance some of them can be imaged even more clearly.

In order to convert COF time sections into depth sections at least effective average velocities are required. They can be obtained by the analysis of CMP wide-angle sections, an example of which is shown in Fig. 46. This section contains all vertical seismograms with CMP coordinates of 35 ± 0.5 km (see Fig. 40). CMP sections focus on much

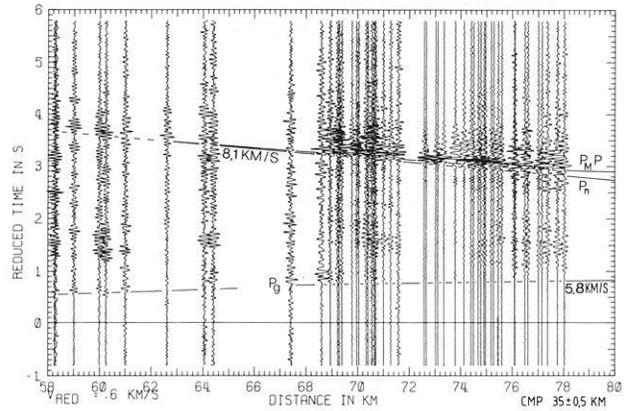


Fig. 46. Example of a common-midpoint wide-angle record section suited for velocity and depth determinations (CMP 35). This section is typical for the central part of the CMP-profile; other parts show more complexity. Distance corresponds to azimuthal directions N26°E to N5°

smaller subsurface elements than CSP or CST sections; they are therefore less affected by lateral heterogeneities and allow some reliable velocity and depth determinations.

Since the seismic profiles have been observed perpendicular to the strike of the main geological structures and geophysical anomalies, the effective average velocity \bar{v} down to a reflector and the vertical depth z_{CMP} beneath the CMP can be calculated according to the simplified formulae:

$$\bar{v} = \cos \alpha \sqrt{\frac{\Delta v_{\text{CMP}}}{T}} \approx \sqrt{\frac{\Delta v_{\text{CMP}}}{T}}$$

and

$$z_{\text{CMP}} = \frac{\Delta}{2} \sqrt{\frac{T v_{\text{CMP}}}{\Delta} - 1 - \left(\frac{D}{\Delta} \tan \alpha\right)^2} \approx \frac{\Delta}{2} \sqrt{\frac{T v_{\text{CMP}}}{\Delta} - 1}$$

where Δ , T and v_{CMP} are distance, traveltime and apparent velocity of a wide-angle reflection in a CMP gather, α is the unknown in-line slope of the reflector and D is the normal offset between the shot and station profiles, i.e. 50 km in our case. By neglecting the weak influence of dip one obtains upper bounds for \bar{v} and z_{CMP} . (This remains true, even if additional vertical velocity variations exist.)

Applying these formulae to the data of Fig. 46 gives an average crustal velocity $\bar{v} = 6.0$ km/s and a Moho depth of 28 km at CMP 35. If this average crustal velocity is assumed to be constant for the whole CMP line, the Moho traveltimes of Fig. 45 can be converted into Moho depths according to the depth scale indicated on the right of this figure.

Velocity determinations for other CMP gathers more to the north and to the south gave slightly higher values (6.1–6.2 km/s), but seem to be less accurate because the corresponding CMP sections are less clear than Fig. 46. In view of Fig. 45 this may not be surprising.

Attempts to resolve separate average velocities for the upper and lower crust have failed so far, because sufficiently continuous intracrustal wide-angle reflections could not yet be identified. Their shape is additionally disturbed by so far uncorrected sedimentary layers of variable thickness and velocity. It seems necessary to first correct these influences

by using all available refraction information before more details can be extracted from the CMP velocity analysis.

4.4 In-line wide-angle studies of universities in the northern sector

In-line refraction and wide-angle observations were continued between May 7 and 17 in the northern sector⁴. In addition, six automatic magnetic longtime recording stations (MLR) were in operation in the southern Spessart and in the northern Taunus. The area covered by the wide-angle observations is indicated in Fig. 2. The southernmost station was near the Main river. Nine of the southern stations along the eastern flank of the Spessart mountains were equipped with MARK 3-component geophones, the other stations used FS 60 instruments with 2 vertical and 1 horizontal component operating. The distance between stations was 540 m providing a spread length of 11 km.

During the first 8 days of operation the stations remained fixed while the shotpoints moved from the northern end of the spread towards northwest up to the critical distance of the Moho reflection. In the last two days of operation the refraction spread was moved together with the shotpoints, i.e. about 10 km per day.

At each station about 250 shots were recorded by FM technique. More than 6,000 seismograms or about 9,000 traces are available, and the digitization is not finished yet. In general, the data quality is good; even the signals from the small shots of 10 kg are clearly recognizable in the critical distance range. Only after the shotpoint had crossed the Hessian trough and moved into the Devonian of the Rhenish Massif the signals from the smaller shots could not be observed any more. The frequency range of the signals is between 5 and 30 Hz with a maximum between 7 and 20 Hz.

Two examples of record sections are presented in Figs. 47 and 48, one is related to an MLR station located in the Spessart and the other to a station placed in the Taunus mountains. Only about $\frac{1}{3}$ of all shots in between the two stations is shown. No clear intracrustal reflectors but a strong $P_M P$ – even in the subcritical range – can be observed at the Spessart station in the SE (Fig. 47) whereas the Taunus station (Fig. 48) in the NW shows two strong intracrustal reflectors and a clear beginning of the $P_M P$ wave from the Moho.

Certainly the velocity-depth structure in the northwestern part of the observed profile section, represented by the Taunus station, is quite different from that of the southwestern part. From the near-vertical data, shown in Figs. 20–23, it becomes clear that major lateral heterogeneities are present, especially below the Spessart mountains, which should show up predominantly in the Spessart profile. The migration of reflection points may then result in a smearing and smoothing of the events as is seen in this seismogram section. The Taunus profile, on the other hand, represents more subhorizontal layering, although the Hessian trough provides also lateral inhomogeneities and a large travelt ime delay, especially for the shorter distances.

The interpretation of the complete data set can be expected to result in a more detailed picture of crustal velocities in the northwestern part of DEKORP 2-S, which dis-

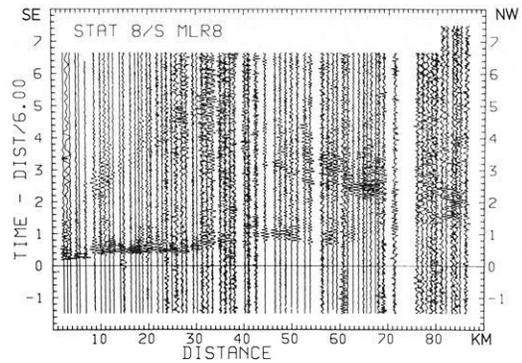


Fig. 47. Reduced record section of a magnetic longtime recording station (MLR) in the Spessart mountains. 4–30 Hz bandpass filter. NW is to the right!

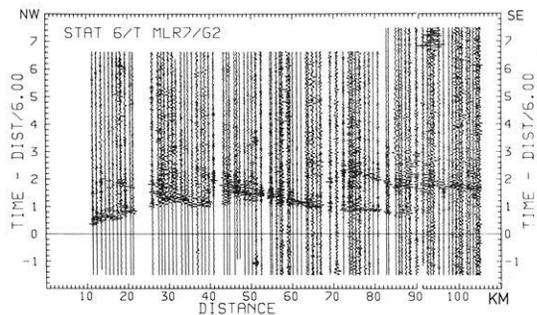


Fig. 48. Reduced record section of a MLR station in the Taunus mountains. Shots between this station and the Spessart station (see Fig. 47) provide a reversed observation scheme. SE is to the right!

plays one of the most interesting tectonic developments along the whole profile.

5. Conclusions

The DEKORP 2-S profile is the first long range near-vertical reflection profile for crustal studies in Central Europe. Running perpendicular to the Variscan strike it crosses at least two significant geotectonical boundaries and provides many new data. The use of explosives allowed important near-vertical as well as wide-angle experiments yielding valuable additional data sets. The establishment of an independent DEKORP processing center proved to be a major advantage for the handling and interpretation of data.

A major achievement with regard to the tectonic structure along the profile is the confirmation of generally south-east dipping thrust systems. In particular, the boundary between the Saxothuringian and Rhenohercynian zones is marked clearly by strongly dipping reflectors, supposed to represent thrust zones in the upper crust. There are several indications that also the lower crust was involved in the thrust tectonics although this evidence might not be as strong as that of the BIRPS' SWAT profiles or the first Ecos line.

The appearance of several strong diffraction clusters, concentrating in the lower crust, was the greatest surprise of the survey. Although the position of prominent diffractions could be assessed in the case of the Dinkelsbühl clusters by means of a quasi 3-D observation, only speculations

⁴ By teams from Bochum, Frankfurt and Munich

can be provided with regard to their origin and nature. The predominantly vertical arrangement of large sets of these diffractions suggests an origin by extensional tectonics, possibly related to the post-Variscan extensional or wrench faulting (Ziegler, 1978; Arthaud and Matte, 1975). This could explain their arrangement along near-vertical planes in the strike direction of the Variscides. On the other hand, the clustering of the diffractions could also be explained by the idea of ramps and huge horses of the large thrust system.

Compared to previously observed crustal reflection profiles in the Variscides (Bartelsen et al., 1982; Meissner et al., 1982, 1983, 1984) the reflectivity pattern is nearly identical. An upper crust, void of reflections except for those of fault zones, is underlain by a dense lamellae-like, highly reflective lower crust. The beginning of this zone at about 4.5–5 s TWT may be related to the Conrad level, their abrupt termination between 9 and 10 s TWT is generally attributed to the Moho. Only very few sporadic reflections are observed from the Upper Mantle. Compared to the Urach profiles a tendency for less correlatable and shorter reflection segments with a dominance of diffraction patterns seems to emerge for the profiles perpendicular to the strike direction, like the DEKORP 2-S line.

One of the major questions not yet answered is the velocity problem. The dominance of diffractions and the shortness and weakness of correlatable reflections made the calculation of reliable stacking velocities extremely difficult. Additional attempts to calculate velocities from the curvature of diffractions, from the various wide-angle experiments and from other processes than the conventional stacking method are well under way and reveal the strong lateral velocity inhomogeneities along and around the main profile. The mapping of exact velocity-depth functions and their application to the tectonic problems of the Variscides is the major challenge for the on-going interpretation.

Such problems are for instance: What does the deep structure of the Northern Phyllite Zone look like? Does it extend beneath the Mid-German Crystalline High? What is the depth extent of the Saxothuringian cover, and what are its structural relations to the basement?

Several groups of scientists in the Federal Republic of Germany are continuing to work on the DEKORP 2-South data, its better processing, analysis and interpretation. Additional ideas will come from the results on other DEKORP-/KTB-profiles surveying the same geological units in analogous tectonic positions. The hope is justified that questions which had to be left open at this time will be answered soon due to these combined geoscientific efforts in the frame of the DEKORP project.

Acknowledgements. The funding of the DEKORP-Project by the Bundesministerium für Forschung und Technologie, Bonn, is gratefully acknowledged. The administrative services are provided by the Niedersächsisches Landesamt für Bodenforschung, Hannover. The excellent performance of the Prakla-Seismos GmbH field crew under the command of Mr. Ceranski has to be emphasized, also the smooth cooperation between all the various groups in the field. Mobil Oil AG, Celle, and Seismograph Service LTD., London, provided the hard- and software, respectively, used at the DEKORP-Processing-Center at Clausthal. Preussag AG and BEB Gewerkschaften Brigitta und Elwerath Betriebsführungsgesellschaft mbH, both in Hannover, supported the work with their modern plotting facilities. The authors thank also the large number of students participating in the field and processing work.

References

- Althaus, E., Behr, H.J., Eder, F.W., Goerlich, F., Maronde, D., Ziegler, W.: Kontinentales Tiefbohrprogramm ("KTB") (Continental Deep Drilling Program) of the Federal Republic of Germany. *Terra cognita*, **4**, 389–397, 1984
- Angenheister, G., Pohl, J.: Die seismischen Messungen im Ries 1948–1969. *Geol. Bavar.*, **61**, 304–326, 1969
- Angenheister, G., Pohl, J.: Results of seismic investigations in the Ries crater area (Southern Germany). In: Giese, P., Prodehl, C., Stein, A., eds.: *Explosion seismology in Central Europe, data and results*. Springer, Berlin Heidelberg New York, 1976
- Arthaud, F., Matte, Ph.: Le décrochement tardi-hercynienne du sud-ouest de l'Europe – Géométrie et essai de reconstitution des conditions de la déformation. *Tectonophysics*, **25**, 139–171, 1975
- Bally, A.W.: Seismic expression of structural styles – A picture and work atlas; Vol. I The layered earth. AAPG Studies in Geology # 15, Tulsa/Oklahoma, 1983
- Bartelsen, H., Lueschen, E., Krey, Th., Meissner, R., Schmoll, J., Walter, Ch.: The Combined Seismic Reflection-Refraction Investigation of the Urach Geothermal Anomaly. In: Haenel, R. (ed.): *The Urach Geothermal Project (Swabian Alb – Germany)*. E. Schweizerbart'sche Verlagsbuchhandlung (Nägele u. Obermiller), Stuttgart, 1982
- Behr, H.-J.: Subfluenzprozesse im Grundgebirgsstockwerk Mitteleuropas. *Z. Dtsch. geol. Ges.*, **129**, 291–326, 1978
- Behr, H.-J.: Intracrustal and subcrustal thrust-tectonics at the northern margin of the Bohemian Massif. In: Martin, H., Eder, F.W., eds.: *Intracontinental fold belts*. Springer, Berlin Heidelberg New York Tokyo, 1983
- Behr, H.-J., Engel, W., Franke, W., Giese, P., Weger, K.: The Variscan belt in Central Europe: main structures, geodynamic implications, open questions. In: Zwart, H.J., Behr, H.-J., Oliver, J.E., eds.: *Appalachian and Hercynian fold belts*. *Tectonophysics*, **109**, 15–50, 1984
- Berckhemer, H.: Mars 66: Eine Magnetbandapparatur für seismische Tiefensondierung. *Z. Geophys.*, **36**, 501–518, 1970
- Carlé, W.: Bau und Entwicklung der süddeutschen Großscholle. *Beih. Geol. Jahrb.*, **16**, 272 pp., 1955
- Carlé, W., Wurm, F.: Die wissenschaftlichen Ergebnisse der Tiefbohrung Allmersbach/Weinberg, Landkreis Backnang, Baden-Württemberg. *Jahrb. Geol. Landesamt Baden-Württemberg*, **13**, 171–221, 1971
- Dohr, G.: Zur reflexionsseismischen Erfassung sehr tiefer Unstetigkeitsflächen. *Erdöl Kohle*, **10**, **5**, 278–281, 1957a
- Dohr, G.: Ein Beitrag der Reflexionsseismik zur Erforschung des tieferen Untergrundes. *Geol Rundsch.*, **46**, 17–26, 1957b
- Dohr, G., Meißner, R.: Deep Crustal Reflections in Europe. *Geophysics*, **40**, 25–39, 1975
- Eberle, D.: *Aeromagnetische Karte der Bundesrepublik Deutschland 1:1.000.000*. Hannover: Bundesanstalt für Bodenforschung, 1973
- Falke: Zur Paläographie des kontinentalen Perm in Süddeutschland. *Abh. Hess. L.-A. Bodenforsch.*, **60**, 223–234, Wiesbaden: 1971
- Gajewski, D., Prodehl, C.: Crustal structure beneath the Swabian Jura, SW Germany, from seismic refraction investigations. *J. Geophys.*, **56**, 69–80, 1985
- Gerke, K.: Die Karte der Bouguer-Isanomalen 1:1.000.000 von Westdeutschland. Frankfurt/Main: Institut für Angewandte Geodäsie, 1957
- Giese, P.: Die Europäische Geotraverse (EGT). *Geowissenschaften: DFG, Dt. Forschungsgemeinschaft, Verlag Chemie, Weinheim*, 1983
- Giese, P., Prodehl, C., Stein, A., eds.: *Explosion seismology in Central Europe, data and results*. Springer, Berlin Heidelberg New York, 1976
- Golwer, A.: Paläogeographie des Hanauer Beckens im Oligozän und Miozän. *Notizbl. Hess. Landesamtes Bodenforsch. Wiesbaden*, **96**, 157–184, 1968

- Karte der Anomalien der Totalintensität des erdmagnetischen Feldes in der Bundesrepublik Deutschland 1:500.000. Hannover: Bundesanstalt für Geowissenschaften und Rohstoffe, 1976
- Kossmat, F.: Gliederung des variszischen Gebirgsbaues. Abh. Sächs. Geol. Landesamt, **1**, 39 pp, 1927
- Martin, H., Eder, F.W., eds.: Intracontinental fold belts. Springer, Berlin Heidelberg New York Tokyo, 1983
- Masonne, H.J., Schreyer, W.: A new experimental phengite barometer and its application to a Variscan subduction zone at the southern margin of the Rhenohercynicum. *Terra cognita*, **3**, 187, 1983
- Matthes, S.: Die Paragneise im mittleren kristallinen Vorspessart und ihre Metamorphose. Abh. Hess. Landesamt Bodenforsch., **8**, 1–86, 1954
- Meissner, R., Bartelsen, H., Murawski, H.: Thin-skinned tectonics in the Northern Rhenish Massif – Germany. *Nature*, **290**, 399–401, 1981
- Meissner, R., Bartelsen, H., Krey, T., Schmoll, J.: Detecting velocity anomalies in the region of the Urach geothermal anomaly by means of new seismic field arrangements. In: Cermak, V., Haenel, R., eds.: *Geothermics and geothermal energy*. 285–292, Stuttgart, E. Schweizerbart'sche Verlagsbuchhandlung, 1982
- Meissner, R., Springer, M., Murawski, H., Bartelsen, H., Flueh, E.R., Duerschner, H.: Combined seismic reflection-refraction investigations in the Rhenish Shield and their relation to recent tectonic movements. In: Fuchs, K., Murawski, H., eds.: *Plateau uplift*. 276–287. Springer, Berlin Heidelberg New York, 1983
- Meissner, R., Springer, M., Flueh, E.: Tectonics of the Variscides in North-Western Germany based on seismic reflection measurements. In: Hutton, D.H.W., Sanderson, D.J., eds.: *Variscan Tectonics of the North Atlantic region*. Blackwell, 1984
- Mostaanpour, M.M.: Einheitliche Auswertung krustenseismischer Daten in Westeuropa – Darstellung von Krustenparametern und Laufzeitanomalien. *Berliner Geowiss. Abh.*, Reihe B, Heft 10, Dietrich Reimer, Berlin, 1984
- Trusheim, F.: Über den Untergrund Frankens – Ergebnisse von Tiefbohrungen in Franken und Nachbargebieten 1953–1960. *Geol. Bavar.*, **54**, 1964
- Weber, K.: Variscan events – Early palaeozoic continental rift metamorphism and late palaeozoic crustal shortening. In: Hutton, D.H.W., Sanderson, D.J., eds.: *Variscan tectonics of the North Atlantic region*. *Geol. Soc., Spec. Publ.*, London: 1984
- Weber, K., Behr, H.-J.: Geodynamic Interpretation of the Mid-European Variscides. In: Martin, H., Eder, F.W., eds.: *Intracontinental fold belts*. Springer, Berlin Heidelberg New York Tokyo, 1983
- Wenz, W.: Erläuterungen zur geologischen Karte von Hessen – Blatt Rodheim Nr. 5718. 61 pp, Darmstadt: 1936
- Ziegler, P.A.: North West Europe Tectonics and Basin Development. *Geol. Mijnbouw*, **57**, 589–626, 1978
- Ziegler, P.A.: *Geological Atlas of Western and Central Europe*. Elsevier, Amsterdam, 1982

Received June 20, 1985; revised version August 1, 1985

Accepted August 12, 1985

Inverse dynamic problems and seismic methods for determination of the structure of a medium

A.S. Alekseev and G.F. Zherniak

Computing Center, USSR Academy of Sciences, Siberian Division, Prospekt Akademika M.A. Lavrentieva, 6, 630090, Novosibirsk, USSR

Abstract. In this work we discuss the problems of applying dynamic inversion methods in seismic exploration. The linearized inversion methods, rapidly developed in recent years, are still of very limited practical use. This is associated particularly with the presence, in real data, of powerful regular-noise background which is neglected in linearized solutions. At the same time, the usual treatment by the CDP-migration does not provide the required quality in images of seismic sections in complicated situations, particularly in the case of a 3-D strongly inhomogeneous structure of a medium.

The proposed method of imaging consists in reducing the inverse dynamic problem in its classical statement to a problem of reconstructing the stable functionals of the velocity structure of a medium, i.e. the dynamic images of its inhomogeneities. The relative redundancy of multifold seismic observation systems is used for dynamic filtration of regular noise. If a medium is complex-structured, the physical parameters of the equation of the useful part of the field are corrected in the process of treatment. The imaging transformations are supplemented here with specific continuation of the field in the inhomogeneous medium. The practical treatment is interactive. In connection with 3-D inverse dynamic imaging it is sufficient to observe the field, using an areal system of parallel profiles of multifold coverage.

Key words: Seismic image inversion – Double wave migration – Weighted spatial-frequency stacking

Introduction

At the present time, common-depth-point (CDP) and wave migration methods are most widely used for processing and interpretation of wave fields. However, recently in seismic exploration problems have arisen that extend beyond the limits of these seismic data processing methods. These are problems of seismic stratigraphy, problems of searching for geological structures in complicated spatial-configuration regions, etc. Kinematic simplification of schemes defining useful waves,

representation of an inhomogeneous medium by average-velocity models and other assumptions used in the CDP-migration methods may, more often than not, substantially distort results of processing.

The most complete and exact information on physical characteristics of inhomogeneous media can be provided by solving inverse dynamic problems in general formulation. Here we mean problems of defining variable coefficients (physical parameters of a medium) in differential equations according to wavefields described by these equations. Theory and methods of numerical solution of such problems developed up till now, and results obtained, allow us to firmly restore the structure of vertically inhomogeneous media (Alekseev, 1967; Alekseev and Dobrinsky, 1975). An important issue which remains to be considered is the development of methods for multidimensional inverse dynamic problems and the applicability of corresponding transformations in real data processing. A serious handicap, besides the mathematical difficulties, remains in the rigid requirements of acceptable models of 2-D and 3-D media to details of observations and accuracy of dynamic wavefield recordings.

The theory of solving inverse dynamic problems admits the presence of irregular noise in the initial data. It is assumed that minimization of discrepancies between an observed and calculated field (using an iterative method), in the case of a unique solution of the inverse problem, is a sufficient condition to find an acceptable solution (Aki and Richards, 1980). However, if the equations used are inadequate for real fields the solutions thus obtained can intrinsically differ from real ones, even if there is negligible irregular noise. This is because part of the recorded waves turns into the category of regular noise, not taken into account in the solution, which is the cause of incorrectness of the inversion. An example of incorrectness is the use of exact algorithms of acoustic inversion for processing an elastic wavefield. Obviously, minimization of discrepancies between values of the simulated acoustic field and the recorded elastic wavefield is not sufficient for the acoustic solutions to converge to some approximate values of the parameters of an elastic medium. Realistic seismic fields can be appreciably affected by imperfect elasticity, porosity, anisotropy and other physical properties of rocks. Disregarding the wave phenomena involved can also be the cause of incorrectness of exact

solutions of inverse problems for the Lamé equations describing the seismic fields.

Approaches to the solution of dynamic problems in seismic exploration

We will consider schematically how a problem of wave-noise filtration in seismic exploration is solved from the point of view of inverse problems.

Let the observed field $U=Lv$ be described as the sum

$$U=U_0+U_1=L_0v_0+L_1v_1,$$

where L, L_0 are, respectively, operators of a full (realistic) and particular (simulated) solution of the forward dynamic problem; v and v_0 are physical parameters of the realistic and model medium, respectively; $U_1=L_1v_1$ is the part of the observed field considered to be noise. We assume that an inverse operator L_0^{-1} of the simulated problem is constructed which is stable with respect to irregular noise and such that $L_0^{-1}U_0=v_0$. Applying it to the field $U=U_0+U_1$ will yield a distorted solution if, in the domain of definition of v_0 , the values $\|L_0^{-1}U_1\|=\|L_0^{-1}L_1v_1\|$ (i.e. the result of its influence on the field of regular noise) are comparable to or exceed the values of $\|v_0\|$ in some norm. Assume that there exist filtering operators M reducing the recorded field U to $MU=U_m+\varepsilon_m$ so that $U_m\approx L_mv_0$, where L_m is the operator of some special solution of the forward problem (in the general case, L_m can be different from L_0). To filter out the regular noise-field and find v_0 , it is necessary to find operators M and L_m^{-1} such that $L_m^{-1}MU\approx v_0+L_m^{-1}\varepsilon_m$ and, in the chosen norm, $\|v_0\|\gg\|L_m^{-1}\varepsilon_m\|$. Typical of this statement is that under special filtration of the observed field the inverse problem can be posed for an equation that is more easily handled than the initial (model) one (Alekseev, 1967). When part of the regular noise-field has characteristics which differ only slightly from those of the useful field, full filtration of noise may prove impracticable. In this case transformation of M and L_m^{-1} will lead to reconstruction of some functionals $c(v)$ of the realistic physical parameters of a medium. Though they do not imply an exact solution to the inverse problem the information about the medium carried by them may be of great importance in seismic prospecting.

An example of an extremely simple implementation of this approach can be CDP-migration processing. When seismograms are processed by the CDP method one usually classifies multiple, surface and, sometimes, converted waves as noise. CDP transforms, based on simple kinematic schemes, are used to process a full dynamic wavefield, i.e. the method is not purely kinematic. Filtration of the regular noise-field by stacking of seismograms along CDP time-distance curves is made possible due to redundancy of multifold observation systems (standard seismic exploration systems). For example, the 2D CDP section $U_m(l, t)=M_{\text{CDP}}U(x_0, x, t)$ is formed from the field $U(x_0, x, t)$ (recorded on a profile line), dependent on three variables (here M_{CDP} is a CDP-stacking operator). The redundancy is used to stack the signal component of the field and to suppress the part of irregular wave

noise whose kinematics is different from that of useful waves. However, the kinematic simplification of the CDP scheme often gives rise to a dynamic distortion of useful reflections on CDP sections. Besides, the size of the common reflection element is fairly small only for an insignificant dip of the reflecting boundaries. With the increase of the dip of the boundaries, dynamic distortion intensifies and the locality property of CDP sections is lost.

In further treatment of seismic data, the CDP sections are assumed to be a filtered field of primary reflected waves (to be recorded at points of wave excitation) against remaining noise background. In migrating the CDP sections they are taken to be the result of radiation of secondary wave sources distributed along reflecting boundaries. The inverse problem of migration usually reduces to determining values of the field at the moment $t=0$ when these secondary radiations are "turned on". This allows one to obtain an image of the inhomogeneities of the medium, i.e. to image it. As shown by Alekseev and Tsybul'chik (1978), there is a certain connection between problems of imaging the structure of a medium and inverse problems of wave theory. So far an obvious correspondence between them has been found only in the case of simple models of a medium. Nevertheless, from our experience with using the CDP-migration methods it may be presumed that constructive approaches being employed in seismic exploration can be helpful in developing approximate methods for inverse dynamic problems.

By means of CDP-migration methods the inverse problem to determine elastic parameters of an inhomogeneous medium reduces, in fact, to simple problems of multidimensional filtration (stacking) of the field and to the subsequent reconstruction of special functionals of its structure (seismic sections). Methods of calculation of "velocity spectra" make it possible to determine also the mean (effective) velocity of wave propagation in a layered medium. Moreover, correlation between peculiarities of the pattern of stacking sections and the geology of sedimentary rocks is observed. Therefore, the formulation and solution of a more general problem, an inverse dynamic problem of reconstruction of multidimensional functional-images of complexly structured media, on the basis of actual wavefields of redundant observation systems is of theoretical and practical interest.

The imaging dynamic inversion

We will describe the propagation of waves in a spatially inhomogeneous medium by the equation

$$\Delta U - \frac{1}{v^2(\mathbf{R})} U_{tt} = -\delta(\mathbf{R} - \mathbf{R}_0) f(t); \quad (1)$$

$$\mathbf{R} = (x, y, z), \quad \mathbf{R}_0 = (x_0, y_0, z_0),$$

where $v(\mathbf{R})$ is the wave propagation velocity, $f(t)$ is the form of the wave radiated by the wave source [$f(t)\equiv 0$ for $t<0$] which may be at any present point \mathbf{R}_0 . Let the velocity $v(\mathbf{R})$ take on a constant value v_0 outside some inhomogeneous restricted domain D . Having per-

formed a Fourier transform over t , we give the resulting Helmholtz equation in the equivalent integral form:

$$U(\mathbf{R}, \mathbf{R}_0, k) = \Phi(kv_0) \frac{e^{-ik|\mathbf{R}-\mathbf{R}_0|}}{4\pi|\mathbf{R}-\mathbf{R}_0|} - \frac{k^2}{4\pi} \int_D \tilde{v}(\mathbf{R}_1) U(\mathbf{R}_1, \mathbf{R}_0, k) \frac{e^{-ik|\mathbf{R}-\mathbf{R}_1|}}{|\mathbf{R}-\mathbf{R}_1|} d\mathbf{R}_1. \quad (2)$$

Here $\Phi(\omega) = F[f(t)]$ is the spectrum of the wave, $k = \omega/v_0$ is the wavenumber, $\tilde{v}(\mathbf{R}) = 1 - v_0^2/v^2(\mathbf{R})$ is the function of relative velocity variations in a medium, $d\mathbf{R}_1 = dx_1 dy_1 dz_1$.

Let there be on the plane $z=0$ a recorded field, excited on $z_0 = \text{const.}$ by independent sources. The inhomogeneous domain D lies below the planes of wave excitation and recording, i.e. $\tilde{v}(\mathbf{R})=0$ for $z<0$, $z<z_0$. We will assume that the signal $f(t)$ in the sources and the value of the constant velocity v_0 near the observation system are known. The frequency spectrum of the recorded field

$$U(\mathbf{r}, \mathbf{r}_0, k) = U(\mathbf{R}, \mathbf{R}_0, k)|_{z_0 = \text{const.}, z=0} \\ [\mathbf{r} = (x, y), \mathbf{r}_0 = (x_0, y_0)]$$

is determined according to Eq. (2), which is looked upon as a non-classical integral equation concerning the velocity function $\tilde{v}(\mathbf{R})$. The unknown field in the medium, $U(\mathbf{R}_1, \mathbf{R}_0, k)$, which enters into the expression under the integral in Eq. (2), depends on the medium's velocity structure $v(\mathbf{R})$, which in turn must be determined from the field on the surface $z=0$. Hence, it is necessary to specify an assumption about the form of the field in the inhomogeneous medium. We will assume that in some part D_1 of the medium below sources and receivers it has a composition that is sufficiently similar to the downward field of sources in an effectively homogeneous volume, i.e.

$$U(\mathbf{R}_1, \mathbf{R}_0, k) \approx \Phi(kv_0) \frac{e^{-ik|\mathbf{R}_1-\mathbf{R}_0|}}{4\pi|\mathbf{R}_1-\mathbf{R}_0|} + U_1(\mathbf{R}_1, \mathbf{R}_0, k). \quad (3)$$

Within the framework of the problem of reconstruction of functional-images of the medium's structure, stated below, the condition of "sufficient similarity" is expressible by a Rayleigh criterion. According to this criterion, the system which forms the image is assumed to be fairly good if phase distortions of the field introduced by it are not greater than $\pi/4$. In our case, this means that the phase distortion of the downward field in D_1 due to inhomogeneity of the upper part of the medium must be less than $\pi/4$. From here there follow certain limitations on the degree of inhomogeneity and sizes of the domain D_1 (with respect to the wavelengths of the primary field). The composition of the velocity structure of the medium remains arbitrary enough to include both scattering elements and reflecting boundaries of any configuration.

In a series of papers (e.g. Alekseev, 1967; Zapreev, 1977; Raz, 1981; Clayton and Stolt, 1981) a low-frequency approximation (the condition of weak wave scattering inside a medium, called the Born-Rayleigh approximation) is used to solve inverse problems reduc-

ing to equations similar to Eq. (2). According to the weak scattering condition, the field of secondary waves in a medium [$U_1(\mathbf{R}_1, \mathbf{R}_0, k)$ in Eq. (3)] must be negligibly small compared to the downward source field

$$\mathcal{E}_0(\mathbf{R}_1, \mathbf{R}_0, k) \\ = \Phi(kv_0) \exp(-ik|\mathbf{R}_1-\mathbf{R}_0|)/(4\pi|\mathbf{R}_1-\mathbf{R}_0|).$$

This requirement is fulfilled only in the case of low frequencies or insignificant variations of velocity all over the medium, i.e. if $\omega \rightarrow 0$ or $\tilde{v}(\mathbf{P}) \ll 1$, $\mathbf{R} \in D$ and if the dimensions of the whole inhomogeneity domain D are fairly small. Hence, the assumption about weak scattering substantially limits the form of the model of a medium allowed to solve respective inverse problems. Besides, as was shown by Zapreev (1977), if the conditions of the Born approximation are fulfilled and the noise background is present, we can stably determine not the function $\tilde{v}(\mathbf{R})$ itself, but only its low-frequency (smooth) part. Here we will not require that the secondary field U_1 be small in the medium as compared to the downward field of sources. This implies that relative velocity variations in the inhomogeneous domain D on the whole are not assumed to be small. The influence of the field U_1 in the proposed nonlinearized statement of the inverse problem will be looked upon as an effect of strong regular noise. On account of this, the solution of the problem of determining the structure of a medium must include a procedure for suppressing the wave noise in the observed field. The possibility of directional filtration of a regular noise-field is determined by the redundancy of the observation system considered (the recorded 5- D field and the 3- D structure of the medium). Relative redundant information in the data allows us to include, in the solution algorithms, procedures of dynamic stacking of useful waves (those described by adopted wave equations) and suppressing the remaining noise.

As shown by Zapreev and Cheverda (1981), part of the secondary field U_1 in a medium may be similar to the primary downward field. That is, the field U_1 in Eq. (3) may have a common (generally, significant) component of the form $\tilde{c}(\mathbf{R}_1) \mathcal{E}_0(\mathbf{R}_1, \mathbf{R}_0, k)$ which depends on unknown wave properties $\tilde{c}(\mathbf{R})$ of the inhomogeneous medium in the vicinity of any point $\mathbf{R} \in D_1$. Thus, as a result of the inversion, in the general case, we will determine not the velocity function $\tilde{v}(\mathbf{R})$ itself, but a functional of the form $c(\mathbf{R}) = \tilde{v}(\mathbf{R})[1 + \tilde{c}(\mathbf{R})]$ which depends on local properties of the medium. Because the values of the functional show precisely local variations of the velocity structure of the medium, it is defined as an image which reflects inhomogeneities of the medium. The physical sense of the function of secondary scattering $\tilde{c}(\mathbf{R})$ [hence the functional $c(\mathbf{R})$] can be defined more precisely by assuming a law of re-radiation of the field in a medium, i.e. the physical model of the strong field $U_1(\mathbf{R}_1, \mathbf{R}_0, k)$. To do this, we may suppose that all elements in D_1 are rather smooth and sloping reflecting boundaries or they scatter the field according to a well-known law. However, such a definition considerably narrows the possibilities of using the corresponding inversion algorithms in practice.

After we substitute Eq. (3) into the integral of

Eq. (2), considering all the adopted requirements and assumptions, we obtain the integral equation with respect to the functional $c(\mathbf{R})$

$$U(\mathbf{r}, \mathbf{r}_0, k) = -\frac{k^2 \Phi(k v_0)}{(4\pi)^2} \int_{D_1} c(\mathbf{R}_1) \frac{e^{-ik|\mathbf{R}_1 - \mathbf{R}_0|}}{|\mathbf{R}_1 - \mathbf{R}_0|} \cdot \frac{e^{-ik|\mathbf{r} - \mathbf{R}_1|}}{|\mathbf{r} - \mathbf{R}_1|} d\mathbf{R}_1 + \tilde{U}(\mathbf{r}, \mathbf{r}_0, k). \quad (4)$$

Here \tilde{U} is the total noise-field due to both the secondary interactions in D_1 and the scattering (reflection) of waves in the other part of D . This implies that we include in the noise-field anything which cannot be described to a required accuracy (according to the Rayleigh criterion) by the integral term in Eq. (4).

Solution of the integral equation, Eq. (4), is developed in the range of spatial frequencies. Having performed a Fourier transform of the field $U(\mathbf{r}, \mathbf{r}_0, k)$ over $\bar{\mathbf{r}}, \bar{\mathbf{r}}_0$ of the shots and receivers we arrive at

$$U(\boldsymbol{\kappa}, \boldsymbol{\kappa}_0, k) = F_{\mathbf{r}, \mathbf{r}_0} [U(\mathbf{r}, \mathbf{r}_0, k)] = \frac{\Phi(k v_0)}{4} \int c(\boldsymbol{\kappa} + \boldsymbol{\kappa}_0, z_1) \cdot \frac{e^{-ik[(z_1 - z_0)\sqrt{1 - \kappa_0^2/k^2} + z_1\sqrt{1 - \kappa^2/k^2}]}{\sqrt{1 - \kappa_0^2/k^2} \sqrt{1 - \kappa^2/k^2}} dz_1 + F'_{\mathbf{r}, \mathbf{r}_0} [\tilde{U}(\mathbf{r}, \mathbf{r}_0, k)], \quad (5)$$

where $\boldsymbol{\kappa} = (\kappa_x, \kappa_y)$, $\boldsymbol{\kappa}_0 = (\kappa_{x_0}, \kappa_{y_0})$ are spatial-frequency variables. In deriving Eq. (5) we make use of the Weyl formula and consider the requirement that $c(\mathbf{R}) = 0$ if $z < 0, z_0$; $[c(\mathbf{R}) = 0, \text{ if } \tilde{v}(\mathbf{R}) = 0]$. Introducing the focusing transformation of the observed field for the plane $z > 0, z_0$ gives us

$$W_0(\boldsymbol{\kappa}, \boldsymbol{\kappa}_0, z) = \int U(\boldsymbol{\kappa}, \boldsymbol{\kappa}_0, k) \Psi(\boldsymbol{\kappa}, \boldsymbol{\kappa}_0, k, z) dk, \quad (6)$$

in which the function Ψ for the total areal observation has the form

$$\Psi = 2 \frac{\sqrt{1 - \frac{\kappa_0^2}{k^2}} + \sqrt{1 - \frac{\kappa^2}{k^2}}}{\pi \Phi(k v_0)} \Pi\left(\frac{\boldsymbol{\kappa}_0}{k}\right) \Pi\left(\frac{\boldsymbol{\kappa}}{k}\right) \cdot e^{+ik\left[(z - z_0)\sqrt{1 - \frac{\kappa_0^2}{k^2}} + z\sqrt{1 - \frac{\kappa^2}{k^2}}\right]}, \quad (7)$$

$$\Pi(\mathbf{X}) = 1, \quad |\mathbf{X}| \leq 1; \quad \Pi(\mathbf{X}) = 0, \quad |\mathbf{X}| > 1.$$

The presence of the factors $\Pi(\boldsymbol{\kappa}_0/k)$ and $\Pi(\boldsymbol{\kappa}/k)$ indicates that solution of Eq. (4) is in the range of steadily recorded homogeneous plane waves of the full spatial field spectrum. It appears that such truncation of spatial frequencies may not lead to "smearing" of the solution $c(\mathbf{R})$. Integration over k in Eq. (6) is performed in the range of steady recording of temporal frequencies.

Let us turn to spatial-frequency variables $\mathbf{k}_r = \boldsymbol{\kappa} + \boldsymbol{\kappa}_0$, $\boldsymbol{\gamma} = \boldsymbol{\kappa} - \boldsymbol{\kappa}_0$ and Fourier-transform the field $W_0(\mathbf{k}_r, \boldsymbol{\gamma}, z)$ over the coordinate z . As a result of this

transform in the full spectral domain $(\mathbf{v}, \boldsymbol{\gamma})$ [$\mathbf{v} = (\mathbf{k}_r, k_z) = (k_x, k_y, k_z)$] we obtain (Zherniak, 1982):

$$W_0(\mathbf{v}, \boldsymbol{\gamma}) = F_z [W_0(\mathbf{k}_r, \boldsymbol{\gamma}, z)] = [c(\mathbf{v}) + \xi(\mathbf{v}, \boldsymbol{\gamma})] \Pi\left(\frac{\mathbf{k}_r \cdot \boldsymbol{\gamma}}{k_z^2}\right). \quad (8)$$

Here $c(\mathbf{v}) = F_{\mathbf{R}} [c(\mathbf{R})]$ is the 3-D spectrum of the functional-image to be determined, $\xi(\mathbf{v}, \boldsymbol{\gamma})$ is the spectrum of the transformed noise-field in Eq. (4). Equation (8) shows that the useful part $c(\mathbf{v})$ of the field transformed in this manner does not depend on the variable $\boldsymbol{\gamma}$, while the noise-field $\xi(\mathbf{v}, \boldsymbol{\gamma})$ from $W_0(\mathbf{v}, \boldsymbol{\gamma})$ remains dependent on it. In the absence of the noise background, $c(\mathbf{v})$ can be defined by $W_0(\mathbf{v}, \boldsymbol{\gamma})$ at any fixed value of $|\boldsymbol{\gamma}|$, i.e. $\boldsymbol{\gamma}$ is a free parameter in Eq. (8). This appears to be the consequence of redundancy of the observation system of multifold coverage considered.

One can take advantage of the fact that the spectrum of the solution, $c(\mathbf{v})$, we are seeking is independent of $\boldsymbol{\gamma}$ to filter out the regular noise-field $\xi(\mathbf{v}, \boldsymbol{\gamma})$ entering into $W_0(\mathbf{v}, \boldsymbol{\gamma})$. Suppression of the noise-field can be performed, for instance, by integrating (stacking) $W_0(\mathbf{v}, \boldsymbol{\gamma})$ over $\boldsymbol{\gamma}$ with the weighting function $\varphi(\mathbf{v}, \boldsymbol{\gamma})$ satisfying the requirement

$$\int \Pi\left(\frac{\mathbf{k}_r \cdot \boldsymbol{\gamma}}{k_z^2}\right) \varphi(\mathbf{v}, \boldsymbol{\gamma}) d\boldsymbol{\gamma} = 1, \quad (9)$$

which implies that suppression of the noise from $W_0(\mathbf{v}, \boldsymbol{\gamma})$ must not lead to a distortion of the reconstructed image $c(\mathbf{R})$. Equation (9) appears to have a variety of solutions $\varphi(\mathbf{v}, \boldsymbol{\gamma})$. For instance, it can be satisfied by functions

$$\varphi = (\mathbf{k}_r) / 2k_z^2, \quad (\mathbf{k}_r) / [\pi \sqrt{k_z^2 + (\mathbf{k}_r \cdot \boldsymbol{\gamma})^2}],$$

etc. It allows one to select, among stacking functions φ , those which maximize the signal/noise ratio on reconstructed images. Thus, when selecting φ one can establish the condition of minimization of relative energy of the remaining regular noise, i.e.

$$E_{\text{rel}} \left[\int \Pi\left(\frac{\mathbf{k}_r \cdot \boldsymbol{\gamma}}{k_z^2}\right) \xi(\mathbf{v}, \boldsymbol{\gamma}) \varphi(\mathbf{v}, \boldsymbol{\gamma}) d\boldsymbol{\gamma} \right] = \min. \quad (10)$$

In real data processing there is, as a rule, a priori information about the composition of regular and irregular noise. Using this information and special techniques for the analysis of the wavefield and reconstructed images, one can select the most effective stacking function φ .

The general form of transformation of the space-time spectrum $U(\mathbf{k}_r, \boldsymbol{\gamma}, k)$ of the observed field to the image $W(\mathbf{R})$ of the inhomogeneous medium in variables $\mathbf{k}_r, \boldsymbol{\gamma}, k, z$ is as follows:

$$W(\mathbf{R}) = F_{\mathbf{k}_r}^{-1} \left[\iint U(\mathbf{k}_r, \boldsymbol{\gamma}, k) \Psi(\mathbf{k}_r, \boldsymbol{\gamma}, k, z) \cdot \varphi(\mathbf{k}_r, \boldsymbol{\gamma}, k) dk d\boldsymbol{\gamma} \right]. \quad (11)$$

The value of the function φ in variables $\mathbf{k}_r, \boldsymbol{\gamma}, k$ can be derived from its values in the frequency range $\mathbf{v}, \boldsymbol{\gamma}$, if we substitute k_z by

$$k \left[\sqrt{1 - (\mathbf{k}_r + \boldsymbol{\gamma})^2 / 4k^2} + \sqrt{1 - (\mathbf{k}_r - \boldsymbol{\gamma})^2 / 4k^2} \right].$$

Having interpreted the results of the imaging transformations, it is possible to properly define the structure of the domain D_1 of the inhomogeneous medium where expansion (3) is appropriate. Usually this is the upper part of the medium, lying below the observation system, where the downward source field is described fairly accurately (in phase) by the function $\mathcal{E}_0(\mathbf{R}, \mathbf{R}_0, k)$. The image of the other parts of the medium is, in this case, suppressed and geometrically distorted, as the field re-radiated outside D_1 at this stage of solution refers to the noise background. By means of the visual analysis of the image obtained it is possible, to establish approximately where the lower boundary of the area D_1 is located. This continuation corresponds to a transfer of the entire system of observations below D_1 , which makes it possible to correct parameters of the field equation for useful waves, reflected and scattered by deeper inhomogeneities of the medium, to solve a similar imaging problem for the next subregion D_2 , and so on.

If the velocity structure of the medium varies strongly in a horizontal direction, images of its domains, which then form the total picture, are reconstructed in the same order.

A limited range of temporal frequencies

We have obtained the solution of the above integral equation, Eq. (4), assuming that $U(\mathbf{r}, \mathbf{r}_0, k)$ is steadily recorded for all $\omega = kv_0$. This assumption disagrees with real conditions in, practically, a finite range of temporal frequencies. The higher the maximum frequency, the smaller the size of any defined region D_n of the inhomogeneous medium must be, according to the Rayleigh criterion used by us. On the other hand, with the decrease of the frequencies in the field being processed, the resolving capability and the details of the images become worse. Varying the range of frequencies, involved in the processing, will make it possible to firstly reconstruct crude, smoothed images of large domains of a medium and then, adding high frequencies, to define more exactly and in more detail the structure of its separate parts.

We take the limitation on the range of frequencies $|\omega| < \Omega$ into account by introducing the characteristic function $\Pi(\omega/\Omega)$ into the imaging transformation. The field in this case is written as follows:

$$U(\mathbf{r}, \mathbf{r}_0, k, K) = U(\mathbf{r}, \mathbf{r}_0, k) \Pi(k/K), \quad K = \Omega/v_0. \quad (12)$$

Applying transformations (5)–(8) to the field $U(\mathbf{r}, \mathbf{r}_0, k, K)$ gives us

$$W_0(\mathbf{v}, \gamma, K) = [c(\mathbf{v}) + \xi(\mathbf{v}, \gamma)] \Pi\left(\frac{\mathbf{k}_r \cdot \gamma}{k_z^2}\right) \Pi\left(\frac{k_{1,2}}{2K}\right), \quad (13)$$

where

$$k_{1,2} = \pm \sqrt{k_z^2 + k_r^2 + \gamma^2 + (\mathbf{k}_r \cdot \gamma)^2/k_z^2}.$$

This relation differs from Eq. (8) obtained earlier only by the additional term $\Pi(k_{1,2}/2K)$. Just as in the case of an infinite temporal frequency band, we are free here only to choose stacking functions $\varphi(\mathbf{v}, \gamma, K)$ whose form depends now on the limiting frequency $\Omega = Kv_0$. The

solution is constructed in a similar way when the spectrum of the field being processed is defined in the limited frequency band $\Omega_{\min} < |\omega| < \Omega_{\max}$. The field W_0 obtained has a form of the difference

$$W_0(\mathbf{v}, \gamma, K_{\max}) - W_0(\mathbf{v}, \gamma, K_{\min}).$$

Analysis of Eq. (13) shows that, in the finite range of temporal frequencies $|\omega| < \Omega$, non-zero values of the spatial spectrum $c(\mathbf{v})$ of the image are determined only inside the sphere $|\mathbf{v}| \leq 2K$. This allows us to find limiting estimates of the resolving capability of the image $W(\mathbf{R}, K)$. According to the Rayleigh criterion, the limiting resolving capability is approximately half the width of the central maximum of the function of the system's response to the field of an elementary radiator of waves. In this case the limiting resolution, uniform over x, y, z , is equal to $0.35\lambda_{\min}$, where $\lambda_{\min} = 2\pi v_0/\Omega$ is the minimum wavelength in the spectrum of the processed field. It is possible to improve the resolution over one of the coordinates x, y, z a little by reducing it over other coordinates in truncation of the respective spatial frequencies.

Other systems of observation

When we solve 3-D seismic problems particular difficulties of technological character arise in the use of the total areal system of observations of the 5-D field $U(\mathbf{r}, \mathbf{r}_0, t)$. The algorithm described for the 3-D imaging inversion can be modified for use in less redundant areal systems of observations of the field over four space-time coordinates. An example of such a system is a set of parallel profiles of multifold coverage. Under the same assumptions, the temporal spectrum of the field, recorded only on straight lines along which its sources shift, will be described by Eq. (4), if $y_0 = y, z_0 = 0$. The Fourier transform of the field over the coordinates x_0, x, y of sources, receivers and profiles has the form

$$\begin{aligned} U(\kappa_x, \kappa_{x_0}, k_y, k) &= F_{x, x_0, y} [U(x, x_0, y, k)] \\ &= \frac{k^2 \Phi(kv_0)}{4\sqrt{2\pi}} \int \frac{c(\kappa_x + \kappa_{x_0}, k_y, z_1)}{\sqrt{z_1}} \\ &\quad \cdot \frac{e^{-i \operatorname{sig}(k)[z_1 \sqrt{(\alpha + \beta)^2 - k_y^2} - \pi/4]}}{[\alpha \beta \sqrt{(\alpha + \beta)^2 - k_y^2}]^{1/2}} dz_1 \\ &\quad + F_{x, x_0, y} [\tilde{U}(x, x_0, y, k)], \end{aligned} \quad (14)$$

where $\alpha = \sqrt{k^2 - \kappa_x^2}$, $\beta = \sqrt{k^2 - \kappa_{x_0}^2}$. This relation was obtained using the stationary phase method, the errors of this approach being negligibly small if $kz_1 \gg 1$.

Applying Eqs. (6) and (8) to $U(\kappa_x, \kappa_{x_0}, k_y, k)$ and using the function

$$\begin{aligned} \psi &= \frac{4 \Pi(\kappa_x/k) \Pi(\kappa_{x_0}/k) \Pi[k_y/(\alpha + \beta)]}{\sqrt{2\pi} |k| \Phi(kv_0)} \\ &\quad \cdot \frac{(\alpha + \beta)^2 e^{+i \operatorname{sig}(k)[z \sqrt{(\alpha + \beta)^2 - k_y^2} - \pi/4]}}{[\alpha \beta \sqrt{(\alpha + \beta)^2 - k_y^2}]^{1/2}}, \end{aligned} \quad (15)$$

we obtain the full spatial spectrum of the field $W_0(k_x, k_y, \gamma, z)$:

$$W_0(\mathbf{v}, \gamma) = [c_1(\mathbf{v}) + \xi(\mathbf{v}, \gamma)] \Pi \left(\frac{\gamma k_x}{k_y^2 + k_z^2} \right). \quad (16)$$

Here $c_1(\mathbf{v}) = F_{\mathbf{R}}[c(\mathbf{R})/\sqrt{z}]$, $k_x = \kappa_x + \kappa_{x_0}$, $\gamma = \kappa_x - \kappa_{x_0}$, $\mathbf{v} = (k_x, k_y, k_z)$. The effect of band-limiting the temporal frequencies here is the same as before.

Thus, the solution of this areal system turns out to be similar to that considered above. This example shows that it is possible to modify the transformations described for the case of various redundant systems. In particular, if the medium's structure does not change in the y -direction (perpendicular to the profile), it is sufficient to have data, recorded only along one profile $y = 0$, in order to determine the 2-D imaging inversion. One must only set $k_y = 0$ in Eqs. (14) and (15) and select respective stacking functions.

Continuation of wavefields

Within the framework of a dynamic imaging problem the inverse continuation of the field is a means to continue the entire observation system outside the reconstructed domain of an inhomogeneous medium. This implies that one must compute the values of such a field whose "sources" and "receivers" would be at some surface below the reconstructed part D_1 of the medium and the homogeneous domain would be above the surface. In this case it is sufficient for the field to consist only of homogeneous plane waves, i.e. to have no peculiarities in its secondary "sources". Thus, in continuation of the observed field it is necessary that the influence on it of the known inhomogeneities of the upper part of the cross-section be compensated for as accurately as possible.

We can use the procedures of continuation of the field in a hypothetical inhomogeneous medium in the finite-difference migration algorithms based on various approximate solutions of the wave equation (Claerbout, 1976; Kosloff and Baysal, 1983). However, in the case of such field continuation, the attenuation of wave amplitudes due to their travel through inhomogeneities is not compensated for. The amplitude distortions increase due to the inverse travel of waves through the inhomogeneities.

Zherniak (1983) proposed an algorithm for a recurrent local application of Kirchhoff's integral formula to continue the field in a spatially inhomogeneous medium. The structure of a medium is approximated by a set of thin phase variable-velocity plates. In the case of consistent continuation of the field through the plates we consider only phase variations of homogeneous plane waves from its spatial spectrum. The approximations and the form of the algorithm are in agreement with the approximations and the form of the imaging inversion. The local phase continuation of the field allows us to take into consideration wave refractions in an inhomogeneous medium and partially compensate for an undesirable attenuation of their amplitudes. This, in fact, appears to be a specific method of regularization of the inverse problem in the field reconstruction.

Let us represent the formula of phase continuation of the field $U(\mathbf{r}, z, \omega)$ from the upper to the lower boundary of an inhomogeneous thin layer ($z, z + \Delta z$) in the space-frequency domain as follows:

$$U(\boldsymbol{\kappa}, z + \Delta z, \omega) = \sum_m \{ [U(\boldsymbol{\kappa}, z, \omega) \otimes S_m(\boldsymbol{\kappa}, z)] e^{i\Delta z \omega \sqrt{\frac{1}{v_m^2} - \frac{\kappa^2}{\omega^2}}} \}, \quad (17)$$

where $U(\boldsymbol{\kappa}, z, \omega) = F_{\mathbf{r}}[U(\mathbf{r}, z, \omega)]$, S_m is the spatial spectrum of the characteristic function of the m -th plate of the layer components ($z, z + \Delta z$), v_m is the velocity in it. The distortions resulting from the approximate formula, Eq. (17), are insignificant if $\Delta z \left| \frac{\omega}{v_m} - \frac{\omega}{v_n} \right| \ll 1$ for

any m and n . Under continuous vertical and horizontal variations of the velocity in a medium, Eq. (17) becomes:

$$U(\boldsymbol{\kappa}, z + \Delta z, \omega) = \int_{z=\text{const}} U(\mathbf{r}, z, \omega) e^{i \left[\Delta z \omega \sqrt{\frac{1}{v^2(\mathbf{r}, z)} - \frac{\kappa^2}{\omega^2}} - \boldsymbol{\kappa} \cdot \mathbf{r} \right]} dx dy. \quad (18)$$

Such downward continuation of the field, accomplished simultaneously over the coordinates of both the receivers ($\mathbf{r} \rightarrow \boldsymbol{\kappa}$) and the sources ($\mathbf{r}_0 \rightarrow \boldsymbol{\kappa}_0$), corresponds approximately to transferring the entire system of observation below the inhomogeneous domain of the medium with partial removal of the effect of upper inhomogeneities. Combined with the imaging transformations, this allows us to organize an interaction process, a means for a continuous and more precise definition of the structure of a highly inhomogeneous medium.

To continue the field in the inhomogeneous medium, we must know at least the mean (interval) values of the velocity in it. These values might be obtained from a well log or from velocity analysis methods, analogous to the "velocity spectrum" algorithm in the CDP method. It appears that maximum "brightness" of the image elements reconstructed by means of imaging transformations can be obtained if the velocity at which they are reconstructed coincides with the actual interval velocity of the domains of the medium. So we can firmly obtain the mean velocity values in any domain by some selection method.

Results of numerical experiments

Possibilities of the algorithms of the inverse dynamic imaging were tested by means of numerical modelling (Alekseev and Zherniak, 1983). A model of a 2-D medium and observation system is shown in Fig. 1. Arrangement of the 64 receivers here was the same for any of the 64 positions of the source of cylindrical waves. Forward dynamic problems were solved by the accurate method of boundary integral equations (Voronin, 1978). Two of the 64 common receiver trace gathers, where direct events were excluded, are shown in Fig. 2. The computed wavefield $U(x, x_0, t)$ was decomposed into Fourier harmonics with respect to the discrete variables $t, x_0, x \rightarrow \omega, \kappa_0, \kappa$.

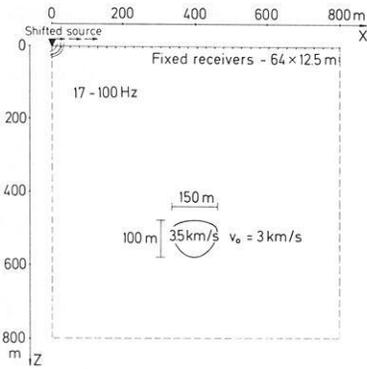


Fig. 1. Model of the medium and the observation system

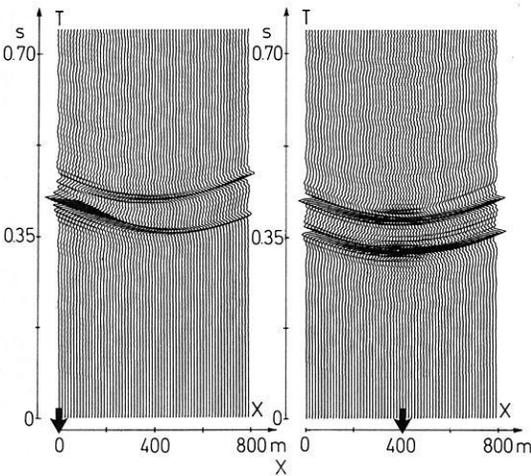


Fig. 2. Two of the 64 common receiver trace gathers for the model in Fig. 1

The inverse imaging problem was solved in an interactive regime. In the first approximation, the medium was assumed to be homogeneous. The velocity in it was taken equal to that near the sources. The reconstruction of the image $W(x, z)$ was carried out by means of the 2-D variant (for cylindrical waves) of Eq. (11). The focusing function, Eq. (7), with $\kappa = \kappa_x = (k_x + \gamma)/2$, $\kappa_0 = \kappa_{x_0} = (k_x - \gamma)/2$, $z_0 = 0$ was used. $\varphi \sim 1 + 5 \sin^2 [\pi(k_x/2k)^2]$ was chosen to be a rough approximation of the weighting function of γ -stacking. Stacking over 64 values of $k = \omega/v_0$ was performed in the frequency range $17 < |\omega/2\pi| < 100$ Hz. Only the upper part of the reconstructed image corresponds to the real position of the object (indicated by dots in the lower part of Fig. 3).

After that, the information obtained about the object configuration and the velocity in it was used. The medium was imaged together with double field continuation under the chosen upper part of the curvilinear boundary of the object. For the downward continuation of the field we used Eq. (17) with $m=0, 1$; $v_0=3$, $v_1=3.5$ km/s. At every step Δz through the depth of the object the computation of $U(\kappa, \kappa_0, z, \omega)$ was carried out by means of inverse and direct Fourier transforms over $\kappa \rightarrow x$, $\kappa_0 \rightarrow x_0$. We substituted the continued field $U(\kappa, \kappa_0, z, \omega)$ in Eq. (11) in place of the field observed on the line $z=0$, and took $z=z_0=0$ in the Ψ -function of Eq. (7). The resulting depth cross-section is shown in

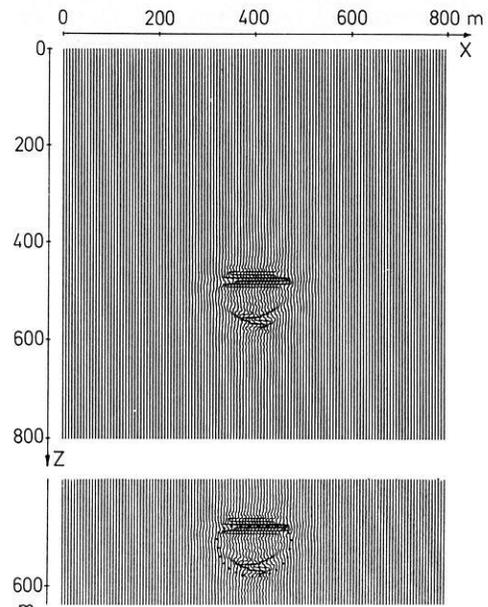


Fig. 3. Image of the object in the first approximation

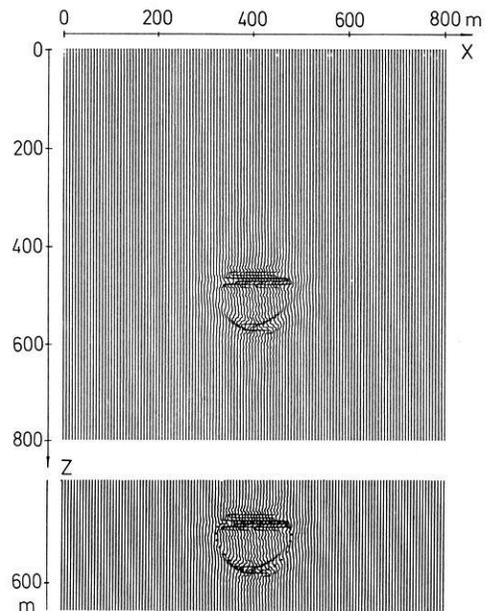


Fig. 4. The result of imaging the structure of the medium in the second approximation

Fig. 4. The image of the step parts of the closed boundary was cancelled out because the rays reflected from them do not strike the fixed arrangement of the receivers.

This undesirable effect can be removed if we use conventional offset-spread systems shifting over long distances. An example of the image reconstruction in this case is shown in Fig. 5. An object here is a sloping boundary with two small scattering inclusions above it. The structure of the medium here does not vary across the profile, $v_0=3$ km/s, $v_1 \ll v_0$, the size of the scattering elements is $\sim \frac{1}{4} \lambda_{\min}$. The source radiates spherical waves in the frequency range 8–64 Hz. With the step

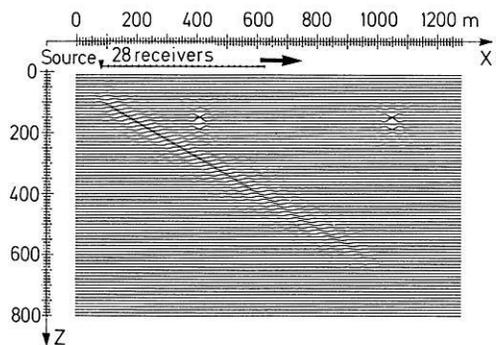


Fig. 5. Image of the 2-D structure of the medium reconstructed on the data of a shifting offset spread

$\Delta x_0 = \Delta x = 20$ m, the source together with 28 receivers shifts along the profile over a distance of 1,200 m. The reconstruction of the image was carried out directly by Eq. (11). The focusing function, Eq. (15), with $k_y = 0$ was used with the same approximation for φ . Similar experiments were also performed for simple 3-D models. The wavefield in this case was observed on a multifold system of independent profiles.

Conclusion

The imaging transforms, resembling to some extent the spectral form of the CDP-migration transformations (Zherniak, 1984) can be looked upon as their dynamic generalization. In comparison to the CDP-migration, with their help one can solve a wider range of seismic problems used when both steeply dipping reflecting boundaries and scattering elements are present in a medium. The choice of the optimum stacking functions guarantees a high signal/noise ratio in connection with the reconstruction of two- and three-dimensional images in seismic sections. Taking the dynamics of reflected and diffracted waves in the observed field more accurately into consideration allows us to obtain a higher resolution in images and to improve the mapping of the physical properties of an inhomogeneous medium. The efficiency of the imaging transformations was tested in a series of numerical experiments.

The solutions of the problem of reconstruction of functional-images obtained can be looked upon as a particular case of quite a general approach to the solution of dynamic inversion problems, using data of overdetermined observation systems. One may choose other equations of the useful part of the field to formulate problems of reconstruction of other stable functionals of the structure of the medium. Similar problems may be posed, for instance, for the equations of elasticity theory. In this case, in order to develop the solution one has to find such transformations of the redundant field as a result of which the useful part of the transformed field (the part corresponding to a for-

mula of the signal component of the field) would be independent of a spatial or a spatial-frequency coordinate. All this makes it possible to find procedures of specific filtration of irregular noise, which do not distort the dynamics of useful signals, from which the structure of the inhomogeneous medium will be determined.

References

- Aki, K., Richards, P.G.: Quantitative seismology. Theory and Methods. San Francisco 1980
- Alekseev, A.S.: Inverse dynamic seismic problems. In: Some methods and algorithms for interpretation of geophysical data (in Russian), pp. 9–84. Moscow: Nauka 1967
- Alekseev, A.S., Dobrinsky, V.I.: Some problems of practical application of inverse dynamic seismic problems. In: Mathematical problems in geophysics (in Russian), Vol. 6, pt. 2, pp. 7–53. Novosibirsk 1975
- Alekseev, A.S., Tsybul'chik, G.M.: On the relation of inverse problems of wave propagation theory and problems of visualization of wave fields (in Russian). Dokl. Akad. Nauk SSSR **242**, 1030–1033, 1978
- Alekseev, A.S., Zherniak, G.F.: Determination of inhomogeneous medium structure by the dynamic visualization method. Expanded Abstracts of the 53rd SEG Meeting, Las Vegas, USA, 615–618, 1983
- Claerbout, J.F.: Fundamentals of geophysical data processing. New York: McGraw-Hill Book Co. 1976
- Clayton, R.W., Stolt, R.H.: A Born-WKBJ inversion method for acoustic reflection data. Geophysics **46**, 1559–1567, 1981
- Kosloff, D.D., Baysal, E.: Migration with the full acoustic wave equation. Geophysics **48**, 677–687, 1983
- Raz, S.: Direct reconstruction of velocity and density profiles from scattered field data. Geophysics **46**, 832–836, 1981
- Voronin, V.V.: Numerical solution of a 2-D-problem of the elastic wave diffraction on an elastic body (in Russian). Preprint of Novosibirsk Computing Center, 26 p., 1978
- Zapreev, A.S.: On the possibility of determining the values of some functionals of a diffracting object located in a stationary magnetic field (in Russian). Preprint of Novosibirsk Computing Center-64, 25 p., 1977
- Zapreev, A.S., Cheverda, V.A.: On some inverse problems for wave equation. In: Mathematical methods for solving forward and inverse problems in geophysics (in Russian), pp. 39–54. Novosibirsk 1981
- Zherniak, G.F.: On determining the structure of a medium using a multifold seismic observation system (in Russian). Preprint of Novosibirsk Computing Center 363, 1982 (English translation: In: SEP-35, University of Stanford, USA, 201–221, 1983)
- Zherniak, G.F.: The continuation of wave field in spatially inhomogeneous media. In: Inverse problems and interpretation of geophysical observations (in Russian), pp. 14–29. Novosibirsk 1983
- Zherniak, G.F.: CDP and wave migration transformations in the problems for determination of a medium's structure (in Russian). Geology and Geophysics **3**, 105–116, 1984

Received September 15, 1984; revised version April 20, 1985
Accepted April 25, 1985

The resolution of the Graefenberg array for earthquake locations in the eastern Mediterranean

R.E. Habermann¹, R. Kind², M. Wyss³ and W. Stammler⁴

¹ School of Geophysical Sciences, Georgia Inst. of Technology, Atlanta, GA 30332, USA

² Seismologisches Zentralobservatorium, Krankenhaus Str. 1, 8520 Erlangen, Federal Republic of Germany

³ CIRES/Dept. of Geological Sciences, University of Colorado, Boulder, CO 80309, USA

⁴ Institut für Nachrichtentechnik, Universität Erlangen, Federal Republic of Germany

Abstract. One hope of making short-term predictions of upcoming mainshocks lies in the rapid location and identification of foreshocks. An advantage of the Graefenberg digital seismic array (GRF) for this task is that seismic signals are available for analysis at the array in real time. In this paper we ask the question: Is the location accuracy of the Graefenberg array in the Hellenic Arc high enough to allow foreshock recognition? We found that the delay patterns for events located throughout the Hellenic Arc were not sufficiently different to allow locating events using these patterns. We also examined waveform cross-correlation techniques and found the location errors to be too large (200–300 km). These errors are due to complexity in the upper mantle. The best technique for locating the events combines the azimuth estimate from the GRF array with the arrival time difference between P and P_4 (the P wave reflected off the 400-km discontinuity). The distance estimate from this travel-time difference has errors of about 60 km. The azimuth estimate has errors of about 3° which yields location errors of about 100 km. These errors are of the order of rupture zone sizes in the Hellenic Arc and, therefore, the Graefenberg array may be useful for approximate monitoring of seismic activity in real time. The techniques involved in determining the locations are, however, difficult for an automated location system.

Key words: Arrays – Earthquake location – Epicentre resolution

Introduction

Several regions of the Hellenic Arc south of Greece are presently experiencing seismic quiescence. Such quiescence has been observed before several large earthquakes in other subduction zones and, therefore, these regions may be in the late preparation stage for such events. It is important to monitor these regions for short-term precursors which may occur before these events. Foreshocks are one of the most commonly observed short-term precursors. The real-time analysis capability of the Graefenberg seismic array in West Germany provides a tool for monitoring a region for foreshocks. In this paper we investigate the location capability of this array in the Hellenic Arc to determine if it is good enough to use the array for foreshock monitoring.

Foreshocks generally occur only during several days to several weeks prior to the mainshock (e.g. Jones and Molnar 1978). This means that one must recognize the foreshocks as soon as possible to allow time for difficult decisions and meaningful warnings. The first step in this recognition is the detection and location of the events.

Wyss and Baer (1981) presented a detailed study of long- and intermediate-term seismicity patterns in the Hellenic Arc south of Greece. They concluded that most of the arc could be considered a seismic gap of the first or second type (highest seismic potential) and that a small section of the arc had no clear history of large earthquakes. They also examined temporal seismicity patterns throughout the arc and found two regions which were presently experiencing seismic quiescence. In the long-, intermediate- and short-term prediction framework, therefore, two segments of the Hellenic Arc have reached the intermediate stage and should be monitored for possible short-term precursors.

In the circum-Aegean area about 27% of mainshocks with $m_b \geq 5.5$ have foreshocks (swarmlike sequences excluded). Wong and Wyss (1984) showed that these foreshock sequences have a high level of spatial and temporal clustering. This observation suggests that foreshock sequences in the study area may be detected and identified in real time. Therefore, we investigated the possibility of using the Graefenberg digital seismograph array for foreshock monitoring in the circum-Aegean area. The primary advantage that the GRF data have over local data is that they can be analysed in real time which is not yet done with local data. The GRF array may, therefore, provide an important source of rapidly available information to local researchers. The source dimensions of the largest ruptures in the Hellenic Arc are expected to be about 100 km (Wyss and Baer 1981). Therefore, the GRF location capability must achieve an accuracy better than 100 km to monitor potential source volumes.

The array and the data

The Graefenberg array consists of 13 vertical and 6 horizontal Wielandt broadband seismometers located near Erlangen in southeast West Germany (Harjes and Seidl 1978; Seidl and Kind 1982; Fig. 1). The distance from the stations to the Hellenic Arc varies between 1,500 and 1,700 km. Digital seismic data are transmitted continuously from the seismographs to the central site in Erlangen where they

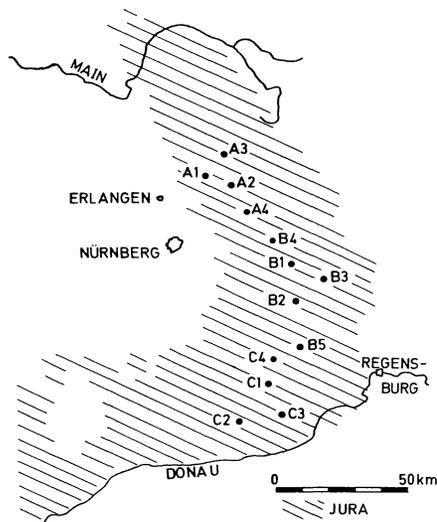


Fig. 1. Map of the GRF array stations. All stations have vertical Weilandt seismometers. The stations used in this study included A1–4 and B1–3

are recorded on magnetic tape. Generally the signals are examined within one day. This speed of analysis is crucial for the foreshock recognition problem. Faster analysis is possible, if necessary.

An advantage of using the Graefenberg array for this study is that one can use array techniques for phase recognition and event location. The most common array technique for event location is beam forming. In beam forming the array is focused into a particular region by applying time delays appropriate to a plane wave arriving from that region to the seismic traces. After these delays have been applied, the traces are summed and coherent plane waves from that region are amplified. By finding the best beam for a given event, one has located that event in terms of its azimuth and slowness.

Several features of Greek seismograms recorded at Graefenberg noted by Rademacher et al. (1983) are important for the work reported here. Greek events clearly show two major phases in the *P*-wave group at Graefenberg. The direct *P* wave has small amplitudes and is difficult to recognize in many cases. The second phase has mostly high amplitudes in the northern part of the array with smaller amplitudes to the south, causing large differences in the waveform across the array. This second phase has been interpreted by Rademacher et al. (1983) as a reflection from the 400-km discontinuity and is termed *P4*.

It is difficult to use *P* waves from Greek earthquakes recorded at GRF for accurate epicentre calculations because of their small amplitude. In many cases the phase would be missed and picking the generally emergent, and rather long-period, first arrival could be subject to large reading errors. We used the *P4* phase for array processing. To avoid difficulties in waveform correlation, we considered only data from seven of the instruments in the northern part of the array (Fig. 1) where the amplitude of this phase was largest and the waveforms most consistent. We considered only events which occurred when all of these stations were operating to avoid problems related to changing numbers of stations. We excluded station B4 which is in the northern part of the array because this station had been down during unusually many events. The *P4* phase is com-

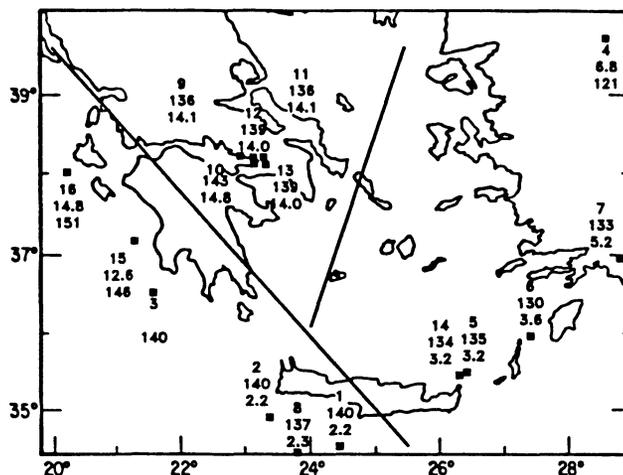


Fig. 2. Epicentre map showing the locations of the events used in this study. Each epicentre is labelled by event number, azimuth to the events from the centre of the array and the time difference between the arrival of the direct *P* wave and *P4* at station B3. This time is a function of distance, decreasing toward the south. The *P* arrival for event no. 3 could not be measured at GRF; therefore, this event was excluded from further study

plicated by interference from depth phases as well as other reflections and converted phases. These interfering phases often travel with different slownesses so they merge with and split from one another as they travel across the array. These complications make it difficult to pick the same phase consistently at all stations. For this reason we used only events which were large enough to be well recorded on broadband data.

Events from various locations in the Hellenic Arc were examined. The locations of these events are shown in Fig. 2 and listed in Table 1 (the distances and azimuths are calculated from array station A1). The pattern of arrival times of the *P4* phase across the array was determined by comparing waveforms of each event at different array stations and by comparing waveforms for different events at the same stations. The arrival time differences between all stations and a reference station (A1) were computed to form an empirical delay set which corresponds to the PDE location of the event being considered. Our hypothesis was that delays from new events with unknown locations could be automatically compared to the delays for the calibration events to provide a quick estimate of their location. The first step in determining if this would be possible was to check for regional consistency of the observed delay patterns.

Regional consistency of delays

The calibration events we chose are distributed fairly uniformly in space along most of the Hellenic Arc from the Ionian Sea to Turkey (Fig. 2). By examining the delay sets we determined that three regions showed fairly consistent delay patterns. These are divided by straight lines in Fig. 2 into West, Corinth and East, and the delay patterns for each of the three regions are shown in Fig. 3. A linear trend corresponding to an average velocity of *P4* across the array of 9.6 km/s was removed from the delays. Thus, the delays

Table 1. Greek earthquakes

No.	Date	PDE			Azimuth		Distance		m_b
		Lat. N	Long. E	P_4-P	GRF	PDE	GRF	PDE	
1	15. 5.79	34.530	24.437	2.2	140	143	19.65	19.55	5.6
2	18. 5.79	34.909	23.351	2.2	140	145	19.65	18.71	4.9
3	27. 5.79	36.544	21.563		140	147		16.37	4.9
4	18. 7.79	39.672	28.660	6.8	121	122	17.98	17.23	5.2
5	23. 7.79	35.483	26.322	3.2	135	137	19.28	19.53	5.2
6	22. 8.79	35.946	27.417	3.6	130	134	19.14	19.66	5.3
7	4.10.80	36.937	28.847	5.2	133	129	18.56	19.58	4.9
8	10. 2.81	34.379	23.779	2.3	137	145	19.61	19.41	4.2
9	24. 2.81	38.222	22.934	14.1	136	140	15.32	15.34	5.9
10	25. 2.81	38.125	23.141	14.8	143	139	15.07	15.52	5.6
11	4. 3.81	38.209	23.288	14.1	136	139	15.32	15.52	6.0
12	5. 3.81	38.207	23.129	14.0	139	139	15.36	15.44	5.1
13	7. 3.81	38.186	23.320	14.0	139	139	15.36	15.55	5.5
14	1. 6.81	35.445	26.307	3.2	134	137	19.28	19.55	5.1
15	22. 6.82	37.160	21.273	12.6	146	147	15.85	15.64	5.1
16	17. 1.83	38.026	20.228	14.8	151	148	15.07	14.29	6.1

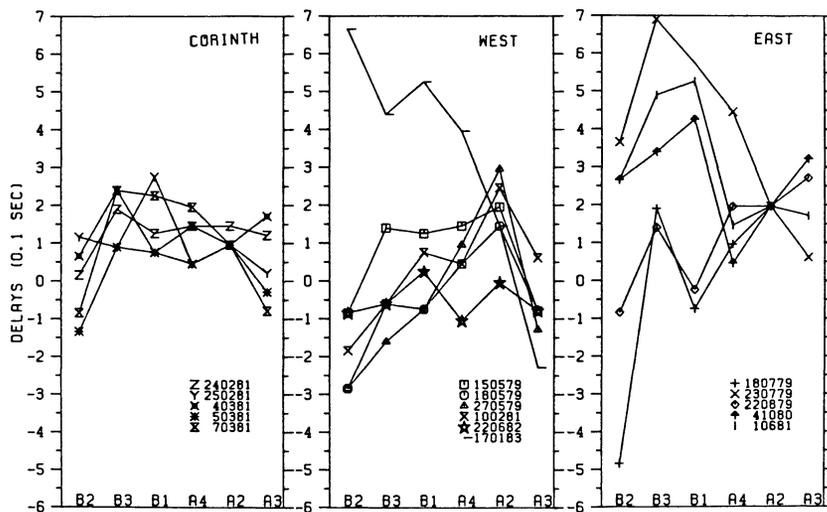


Fig. 3. Delays at the Graefenberg array stations relative to station A1 for events in three subregions of the Hellenic Arc. A constant depending on distance to the events from the array (distance/9.6 km/s) was subtracted from all delays to remove a linear trend. Note how the delays are fairly consistent in the Corinth and West regions

for an event which moved out across the array at this velocity would plot as a straight line across the plot at delay = 0.0.

The Corinth group is made up of five events which occurred during February and March, 1981 (Table 1). The average distance between these events is 18 km. The standard deviations for the delays at a given station in this group are all less than 0.1 s (see Fig. 3).

The West region includes six events from the Ionian Sea and Crete. This group shows surprisingly consistent delays considering that the average distance between the events is 264 km. The exception is the event of January 17, 1983 in the Ionian Sea. The delays for this event are clearly higher than for the others in this group in the B sub-array. The explanation for this difference is unknown. This event is large and well recorded, but the P_4 phase is long-period and very complex. This resulted in a questionable correlation of the waveforms across the array. Perhaps this complexity reflects complexity in the larger source of this event. If this earthquake is excluded, the standard deviations for delays in this group varies from 0.07 to 0.12 s. This variation is essentially the same as that observed for the Corinth group. This observation suggests that the vari-

ance of delays is only weakly dependent on the distance between events in this distance range.

The East group includes five events from east of Crete and Turkey. These events cover a large region, like the West group, and their delays are not consistent. The standard deviations in this region vary from 0.0 to 0.4 s.

Two of the three regions show consistent delay patterns. However, the patterns are not sufficiently different to allow one to place an event with an unknown location unambiguously in any one region on the basis of its delay pattern alone. Therefore, we conclude that the delay patterns do not provide the location capability needed.

Techniques using additional information

We clearly need to use some technique for locating the events which considers more than the delays. The goal is to find the delay pattern from the set of delays for the calibration events (Fig. 3 and Table 1) which most closely fits an event with an unknown location. In order to test these techniques we treated the calibration events as if their locations were unknown. We picked one of the calibration

delay sets and tested all of the calibration events against it to see if we could find the event whose delays we had picked. If this event could be recognized, we considered the test a success. If it could not, we considered the distance between the event picked by the test and the actual event as a measure of the location error of the technique.

First we tried a delay and sum technique based on the assumption that the correct delay pattern should be the one which maximizes the energy in a trace formed by summing the signals from all stations after they have been delayed according to that pattern. In order to test this technique, the signals for each event from all stations were delayed by the amounts determined for one of the calibration events (Fig. 3) and summed to form a sum trace. The integral of the squared sum trace was then computed over a time window including the phase used in identifying the delays to determine the total energy for each event. The same calculation was done for the calibration event whose delay set was being used to delay the other events and, because that calibration event determined the delays, we would expect it to show the highest energy of all of the events. In 5 of the 16 cases the calibration event did have the highest energy using this technique (as expected). In the other cases the distance between the event with the highest energy and the correct event varied from 17 to 510 km (average = 286 ± 158 km). This technique, therefore, does not provide satisfactory results.

We next calculated single lag correlation coefficients for all seven traces by multiplying together the delayed traces and summing the product trace over a time window which included the phase used to determine the delays. This technique was more successful than the summing technique at recognizing the correct event (9 of the 16 cases). Yet, in the other cases the distance between the correct event and the event found was again very high (263 ± 213 km).

In summary, the distances between the correct events and other events with high sums or correlation coefficients averaged between 250 and 300 km. This number provides an estimate of the location errors expected if these techniques are used for quick locations. These errors are too large to be acceptable.

A possible solution

In the previous section we discussed three techniques for quick event locations using delays observed across the Graefenberg array. These techniques do not provide the resolution that we need in this work. In this section we describe a distance and azimuth measurement which may allow reasonably quick locations of events in the Hellenic Arc.

Rademacher et al. (1983) made the observation that the *P*-wave group observed at Graefenberg for events from Greece characteristically contained two major phases. The first is the direct *P* wave and the second is a *P* wave reflected from the 400-km discontinuity (*P4*). These phases travel across the array with different average slowness, 13.6 and 11.5 s/deg (corresponding to 8.1 and 9.6 km/s) and the time difference between the two phases decreases as the sources move farther away (see Rademacher et al. for examples). This time difference, therefore, provides a measure of the distance to the earthquakes. We used the time difference between the first arrival of the *P* wave and the first maximum of *P4*, rather than the *P4* arrival time, because the

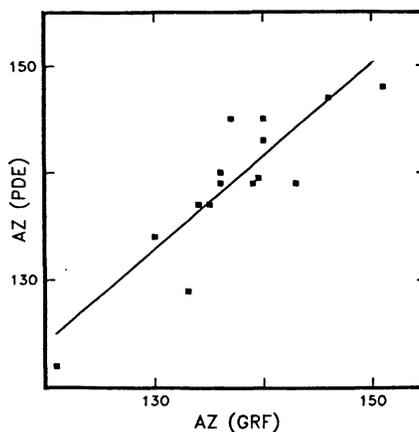


Fig. 4. Azimuth from the centre of the Graefenberg array to the epicentres listed in the PDE [Az(PDE)] versus the azimuth estimated by the GRF data alone [Az(GRF)]. The least-squares fit through the data was obtained minimizing the dependent variable only. The standard deviation of the latter is $\pm 3.25^\circ$

first arrival of *P4* is sometimes hard to recognize. The errors in these measurements are probably a few tenths of a second (provided that the correct phase is picked), which is small relative to the magnitude of the variations. The time differences at station B3 for the calibration events we examined are given in Fig. 2. They range from 2 s near Crete to 14 s near Corinth. This suggests that if the time differences can be measured even with only 1-s accuracy, distances accurate to less than 100 km can be measured. This is clearly better than the other techniques we described.

The second part of the earthquake location we need is the azimuth. The stations we used at Graefenberg were oriented along a line with an azimuth of N 140° E, pointing directly at the region we studied. This configuration is the worst possible for azimuth determinations. The azimuths determined from a least-squares fit to a plane wave crossing the array have errors of between 2° and 6° . The azimuths we determined are shown in Fig. 2 with the epicentres and difference times. They show the expected pattern of smaller azimuths (near 120°) for Turkey, the low 130s for the area between Rhodes and Crete, the upper 130s near Corinth and the 140s near Crete and in the Ionian Sea.

The correlation between the Graefenberg azimuths and those calculated on the basis of the PDE locations is shown in Fig. 4. Although only the northern part of the array was used, the standard deviation from the least-squares fit is not large ($\pm 3.25^\circ$). However, at the epicentral distances in question, this translates to a mislocation of approximately 100 km. The error introduced through uncertainties in the transformation of (*P4*-*P*) into distance is smaller. Figure 5 shows the epicentral distance (based on the PDE location) versus (*P4*-*P*). The straight-line fit in Fig. 5 was obtained by minimizing the dependent variable only and is described by the equation

$$\text{Distance (in degrees)} = 20.45 - 0.36 (P4 - P) \text{ (in seconds).}$$

Based on this relationship one could calculate distance for future events based on the observed (*P4*-*P*) value. The standard deviation for the distance estimate is $\pm 0.53^\circ$, corresponding to a location error of about 60 km.

Originally it was our intention to use only *P4* and not *P* for the location because *P* is usually weak at GRF. Now, with the method just described, we need the arrival time of *P* at one station. This is certainly a disadvantage for

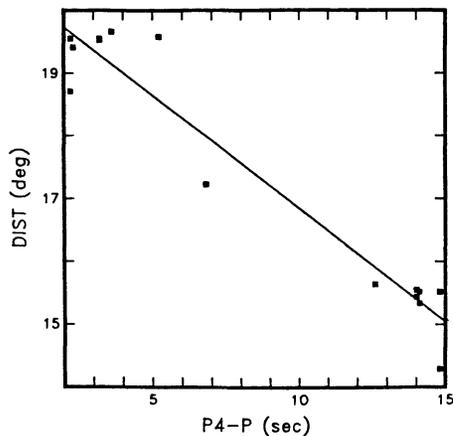


Fig. 5. Distance between the PDE location and Graefenberg, versus the time difference between the reflection off the 400-km discontinuity, P_4 , and the direct P phase. The straight line is a least-squares fit assuming a linear correlation between the two parameters for the relatively short distance range. The standard deviation in the dependent variable is $\pm 0.54^\circ = \pm 60$ km

weak events, but it is still much better than using all the P delays, which standard techniques do.

Discussions and conclusions

In this paper we examined the question of whether the location capability of the Graefenberg array was high enough to allow the use of the array for foreshock recognition in the eastern Mediterranean. The similarity of delays for events from all around the Hellenic Arc, the lack of an increase in the delay scatter in the West region relative to the Corinth region and the difficulty of picking out the calibration events using waveform correlations based on observed delays for those events all indicate that the locations determined using conventional array location techniques are not accurate enough for the foreshock problem. We found that the errors were between 200 and 300 km using such techniques.

Two parameters are used to describe locations by the Graefenberg array: slowness and azimuth. The configuration of stations we used is the worst possible for azimuth determinations and the best possible for slowness determinations for Greece. In spite of this, we found the azimuth resolution to be better than the slowness resolution. This is because the slowness in this distance range is controlled by complex vertical structure in the upper mantle and varies only slightly with distance. Azimuth, on the other hand, does not depend on vertical earth structure and can be measured well in any distance range.

Ironically, the same earth structure which makes it difficult to use slowness to determine distance provides the reflection (P_4) which we use for our distance estimate. If this reflection did not occur it would be very difficult to determine the distance to these events. Using the difference between the P_4 and the P arrival, epicentral distances can be estimated rapidly with an expected error of ± 60 km.

In conclusion we offer a technique for determining quick locations of earthquakes in the Hellenic Arc, using the time difference between the direct P wave and the reflection off the 400-km discontinuity and an azimuth estimate from the Graefenberg array. The error in the distance measure-

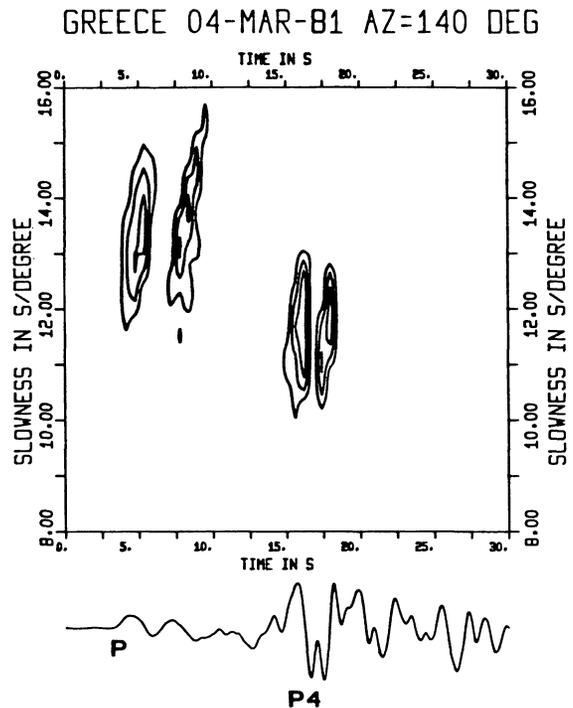


Fig. 6. Slowness analysis of one Greek earthquake. The slowness is measured as a function of time using all array data. The azimuth is fixed. Only the contour lines from the peaks of the phases P and P_4 are plotted. The different slowness of P and P_4 can easily be seen. This kind of data could be used to measure the time difference between P_4 and P and the azimuth on-line

ment may be as small as 60 km. The error in azimuth is about $\pm 3.3^\circ$ (100 km). This resolution is comparable to the size of rupture zones in the Hellenic Arc and may be sufficient for preliminary determinations of foreshock locations.

A number of problems limit the usefulness of this technique. First, it depends on the correct identification of the P and P_4 phases. As mentioned above, the P phase from events in the eastern Mediterranean is weak at GRF, but it can usually be identified. Late P phases, like P_4 , can have a number of causes including surface reflections, source complexity and reflections from upper-mantle discontinuities. Large location errors could occur if some other phase is mistakenly identified as P_4 . The waveforms of Greek events at GRF are very distinctive, primarily because of the different slownesses for the two main phases, so it is unlikely that such a misidentification would occur. These problems can, therefore, be overcome by an experienced analyst at GRF. However, our original goal of developing a foreshock recognition tool which could run automatically could not be achieved. A second problem is that we could only study events which were large enough to be well recorded at GRF. This includes most events with $m_b \geq 4.7-4.9$ (see Table 1). We must hope that foreshock sequences which occur will contain enough events of this size to be clearly recognized as anomalous.

It is possible that a more sophisticated array technique may provide a satisfactory solution to the problem. Figure 6 shows a slowness analysis of a Greek event using all array data. The slowness is plotted as a function of time with fixed azimuth. Only two dB contour lines from the peaks of the two phases P and P_4 are plotted. The

slowness difference of the two phases is clearly visible. It seems possible to measure the time difference between P_4 and P and the azimuth to the event from data like that shown in Fig. 6. This might provide an on-line location of Greek earthquakes accurate enough to identify events from the quiescent regions of the arc. Extensive computations are required for such a method, which works in the frequency and wave-number domain. Unfortunately, the computing facilities of the Graefenberg array are not sufficient for doing this analysis in real time at the present.

The possibility of using the Graefenberg array to monitor other regions of the world remains good. Our present results indicate that monitoring regions in Europe may be difficult or impossible unless the local structure provides some aid, like the P_4 phase used in this work. The resolution may improve with increasing distance when the slowness-distance relationship becomes more reliable and the signal shape is clearer.

Acknowledgements. This research was supported by the Deutsche Forschungsgemeinschaft (German Research Association) and the U.S. National Science Foundation, grant nos. EAR-8212116 and INT-8121311. We wish to thank Sonja Faber for reading the manuscript and Winfried Hanka for doing the slowness analysis. R.E.H.

would like to thank the entire GRF group and especially the Aichele family for making his stay in Erlangen so pleasant.

References

- Harjes, H.-P., Seidl, D.: Digital recording and analysis of broadband seismic data of the Graefenberg (GRF) Array. *J. Geophys.* **44**, 511–523, 1978
- Jones, L.M., Molnar, P.: Some characteristics of foreshocks and their possible relationship to earthquake prediction and premonitory slip on faults. *J. Geophys. Res.* **84**, 3596, 1978
- Rademacher, H., Odom, R.I., Kind, R.: The upper mantle structure under south-east Europe derived from GRF broadband records of Greek earthquakes. *J. Geophys.* **52**, 7–13, 1983
- Seidl, D., Kind, R.: Das instrumentelle und seismologische Konzept des Graefenberg Array. *Geol. Jahrb.* **E23**, 207–220, 1982
- Wong, K.C., Wyss, M.: Clustering of foreshocks and preshocks in the Circum-Aegean region. *Earthquake Prediction Research 1984* (in press)
- Wyss, M., Baer, M., Earthquake hazard in the Hellenic arc. In: *Earthquake prediction: An international review*, D.W. Simpson and P.G. Richards, eds.: pp 153–172. AGU, Washington, D.C. 1981

Received April 4, 1984; revised version May 23, 1985

Accepted July 16, 1985

Extremal inversion of vertical displacements, Long Valley Caldera, California 1982/1983

D.W. Vasco

Center for Computational Seismology, Lawrence Berkeley Laboratory and Department of Geology and Geophysics,
The University of California, Berkeley, CA 94720, USA

Abstract. Vertical displacement data from August 1982 and August 1983 leveling surveys in Long Valley Caldera are examined relative to a 1975 datum. These uplifts are hypothesized to be due to the inflation of a magma chamber of arbitrary shape at depth. Using extremal inversion techniques, which allow for uncertainties associated with random survey error, bounds on the depth to the top of the body and on the location of the edges of the body are produced. These bounds are unique horizontal and vertical limits on any possible volume source satisfying the data. The bounds indicate that any volume source satisfying the 1975–1982 leveling data must lie, in part, at or above 12 km. For the 1975–1983 displacement field, some volume change must have occurred at or above 11 km. The east-west bounds on the source have not changed from 1982 to 1983, requiring volume change east of 118.93°W and west of 118.90°W . However, the north-south bounds, which require a portion of the model to lie north of 37.65°N and south of 37.67°N , have widened one grid element to the north during this interval. These estimates are independent of both Poisson's ratio and the exact boundaries of the region modeled.

Key words: Positivity constraints – Inverse problem – Uplift

Introduction

Long Valley Caldera has been the site of recent permanent and seismic displacements. Repeated leveling surveys between 1975 and August 1983 within the caldera have measured up to 0.4 m of vertical displacement. These displacements have contributed to the hypothesis that a magma chamber still exists beneath the caldera and that this chamber has reinflated to some extent. This notion is compatible with recent moment tensor inversions of seismic data (Julian, 1983) and *P*-wave delay-time inversions (Steeles and Iyer, 1976). Similarly, a study of seismic attenuation within the caldera (Sanders and Ryall, 1983) suggests the presence of a “region of molten or partially molten magma”.

If one accepts the possibility of a magma body at depth, it is possible to invert the uplift data for parameters of the causative body such as the depth and the volume change. For example, Savage and Clark (1982) inverted the 1982 displacement data of the survey line along Highway 395. Assuming a point source, these authors produced

an estimate of the source depth as well as the volume change. Similarly, Castle et al. (1984) inverted the 1983 vertical displacement data along this line for estimates of the same parameters. Recently other models have been proposed. Savage and Cockerham (1984) were able to reproduce observed horizontal and vertical surface deformation reasonably well using two separate dike injection models. The first model consists of a single dike that dips 30° northward beginning at a depth of 8 km and extending to about 12 km in depth. The second model is similar to the first with the addition of a dike extending vertically from the top of the dipping intrusion to within 3 km of the surface. Right lateral slip was needed in both models in order to satisfy the horizontal displacement data. Recently, Rundle and Whitcomb (1984) proposed an additional model. In their model, deformation is attributed to the inflation of two spherical magma chambers; one at a depth of 5 km located 1.5 km west of station Casa (Fig. 1), the other 9 km deep about 5.5 km north-northwest of Casa. All of the above models fit the data reasonably well.

Given the deformation data alone, there is no reason to prefer one model over another. Even when including other information such as gravity or magnetotelluric data in the inversion, some ambiguity will remain in the description of the source. Therefore any proposed model must be viewed critically. Answers to the question “What magma body has produced these displacements?” are seldom unequivocally found. Definite answers are more forthcoming if one asks “How do the data constrain the range of possible models?” One way to answer this is to examine all the models which fit the data and determine properties common to all these models. However, this is a laborious task. There is a method available which allows one to find bounds or limits on certain properties of the models. Limits are placed on model properties such that all models satisfying the data must have properties within these bounds. Such limits are important in allowing one to assess the ambiguity present in the data set. It is for this reason that I have chosen to examine the bounds which the 1975–1982 and 1975–1983 leveling data place on the vertical and lateral extent of a proposed magma body under Long Valley Caldera. Using the method of extremal inversion (Parker, 1975; Sabatier, 1977a, b, c). I derive unique bounds on certain properties of the assumed source. Specifically, the bounds constrain the depth and horizontal extent of the perturbing body. Full nonsymmetric, three-dimensional bodies are allowed and random leveling errors are incorporated into the inversion procedure.

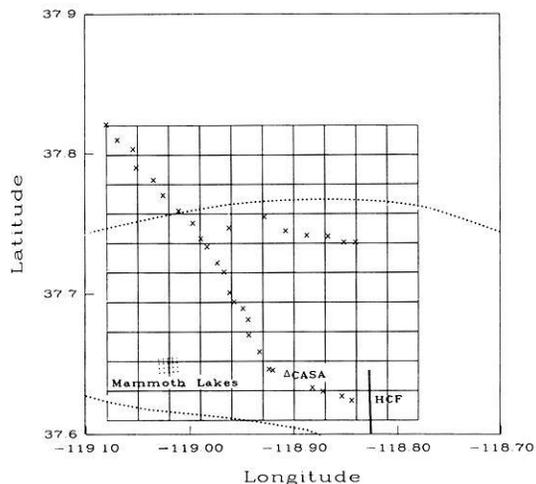


Fig. 1. Map of the Long Valley Caldera region. The caldera is denoted by the *dotted line*. The town of Mammoth Lakes is shown. The Hilton Creek fault is denoted by the *heavy black line* labeled HCF. *x*'s represent the leveling stations used in the study. Station Casa is the *labeled triangle*. Lee Vining is off the *upper left hand corner* of the map. The discretization of the region used in the inversion is shown

The 1982 and 1983 leveling surveys

Leveling surveys within Long Valley Caldera were run along Highway 395 in 1932, 1954, 1975, 1980, 1982 and 1983 (Castle et al., 1984). In addition, surveys were run along various access roads in the area. Figure 1 shows the stations used in the inversion. These are not all of the available data; some data near to and east of the Hilton Creek fault were not used in the inversion. Movement along this fault occurred in May 1980 and displacements, not associated with volume change, would adversely affect the inversion. Furthermore, trilateration data measuring horizontal length changes of survey lines within the caldera were not examined. These data would provide additional constraints on the model parameters.

The early surveys prior to and including the 1975 leveling line detected little or no uplift. However, between 1975 and 1982, up to 0.25 m of uplift occurred. Subsequent surveys in August 1982 and August 1983 detected 0.35 m and 0.40 m of maximum uplift, respectively. This suggests that one may take the 1975 elevation as a baseline with which to measure the changes occurring in the 1975/1982 and 1975/1983 intervals. These elevation changes are shown in Fig. 2 projected onto an east-west trending plane. In Fig. 3 the data are projected onto a north-south trending plane. The data shown are a portion of the leveling lines which extend along Highway 395 from the northwest to the Hilton Creek fault in the southeast. In addition, a second line of data extending approximately east-west was included in the inversion. One assumption made in the production of this uplift profile is that the southern end of the leveling line has remained stable with respect to Lee Vining in the north (Castle et al., 1984), which permitted one to treat the Long Valley system as if it were isolated from the surrounding region. Furthermore, Castle et al. (1984) also argue that only random errors are significant in the data, i.e. systematic deviations were shown to be negligible.

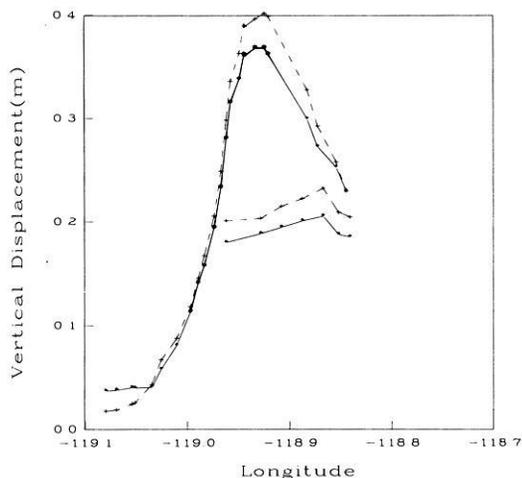


Fig. 2. The vertical displacement data from the August 1982 survey (*solid line*) and the 1983 survey (*dashed line*) projected onto the east-west axis. The data are shown relative to a 1975 baseline

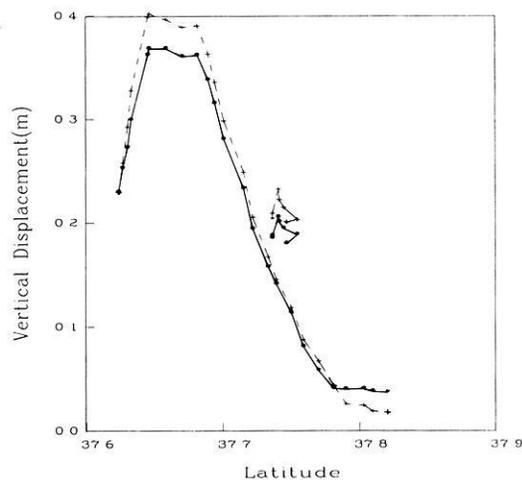


Fig. 3. The vertical displacement data from the August 1982 survey (*solid line*) and the 1983 survey (*dashed line*) projected onto the north-south axis. Again, these data are relative to the 1975 data

The method of extremal inversion

As mentioned previously, a variety of models have been proposed for the magma body giving rise to the observed displacements. Unfortunately, the data do not allow for discrimination among the various models. For this reason an alternative approach was taken. Properties common to all models fitting the data were searched for. Specifically, bounds on the depth and the horizontal location of the source were found. The method of extremal inversion was used to derive these bounds. This is a versatile technique and one that enables the inversion of all static displacement data, horizontal as well as vertical, to give unique bounds on properties of the source model. The only assumptions made are that the fractional volume change is of one sign and that the region is homogeneous and may be characterized by a single Poisson's ratio. The technique is discussed by Parker (1975), Sabatier (1977a, b), Safon et al. (1977) and Rietsch (1978). The adaptation of this method to the inversion of static earth displacements is given in Vasco and Johnson (1985).

As an introduction to the method consider the discrete case. The region of interest is divided into N blocks. Each block is capable of undergoing some fractional volume change $\Delta\vartheta$. A linear system of equations relates the fractional volume changes undergone by the blocks to the M displacements m_i measured at the surface,

$$\sum_{n=1}^N k_{i,n} \Delta\vartheta_n = m_i \quad i=1,2,3,\dots,M. \quad (1)$$

There is also the requirement that all fractional volume changes are of one sign, in this case they are non-negative,

$$\Delta\vartheta_n \geq 0 \quad n=1,2,3,\dots,N.$$

Here $\Delta\vartheta_n$ denotes the fractional volume change within the subregion ω_n , m_i denotes the measured displacement and $k_{i,n}$ gives the deformation at the i th station due to a unit fractional volume change within the n th block. $k_{i,n}$ is given by the point source response at the i th station integrated over the n th block,

$$k_{i,n} = \int_{\omega_n} K_i(\xi) dV(\xi).$$

Here $dV(\xi)$ denotes a volume element of the n th block and $K_i(\xi)$ denotes the integrand relating the displacement at station x^i to the fractional volume change at a point ξ . For vertical displacements,

$$K_i(\xi) = \frac{1}{3\pi} (\nu+1) \frac{\xi_3}{S^3}$$

where

$$S = \sqrt{(x_1^i - \xi_1)^2 + (x_2^i - \xi_2)^2 + \xi_3^2}.$$

ν is Poisson's ratio for the half space.

The general discrete extremal linear inversion problem is to find the extremum, minimum or maximum, of a generalized moment

$$A = \sum_{n=1}^N \alpha_n \Delta\vartheta_n \quad (2)$$

subject to the constraint that Eq. (1) is satisfied, i.e. the data are satisfied. The α_n in the above equation are constants. The generalized moment A may represent some physically significant property depending on the possible choices of the constants α_n . For example, choosing

$$\alpha_n = |\xi - \xi_0|^k \Delta V_n,$$

where ΔV_n is the volume of the ω_n subregion, will produce a bound on the k th-order moment about the point ξ_0 . For $k=0$ a bound on the total volume change in the region is produced,

$$A = \sum_{n=1}^N \Delta V_n \Delta\vartheta_n.$$

For $k=1$,

$$A = \sum_{n=1}^N |\xi - \xi_0| \Delta V_n \Delta\vartheta_n$$

the moment of the body about the point ξ_0 is given. Volume changes further from ξ_0 are given greater importance than changes closer to this point. Hence, A in this case can be

considered as a measure of the compactness of the body much like the moment of inertia.

Another possible choice of α_n and the one used in the following application to the Long Valley uplift data is

$$\alpha_n(\xi) = \begin{cases} 1 & \omega_n \text{ in } U \\ 0 & \omega_n \text{ not in } U \end{cases} \quad (3)$$

where U is a region of interest, that is some subset of the N blocks. Using such α_n , A represents the total fractional volume change in the region U . Such a choice can be used to bound the extent of the causative body in the following manner. If the lower bound of the linear functional A is zero then there exists at least one possible body which satisfies the data but has no volume change in the specified region U . On the other hand, if the smallest value of A is not zero then some volume change must have occurred in U .

The method of extremal inversion may be modified to treat data containing random errors ε_i . In this case one is interested in minimizing the moment (2) subject to the inequality constraints

$$m_i - \varepsilon_i \leq \sum_{n=1}^N k_{i,n} \Delta\vartheta_n \leq m_i + \varepsilon_i \quad i=1,2,3,\dots,M \quad (4)$$

$$\Delta\vartheta_n \geq 0 \quad n=1,2,3,\dots,N.$$

This problem may be transformed into one of the form (1) (Hadley, 1962). The values of ε_i can be estimated for the above surveys. It has been argued (Castle et al., 1984) that the errors in the leveling survey are principally random errors described by the standard deviation

$$\sigma_i = \gamma L_i^{\frac{1}{2}}$$

where γ is a constant and L_i is the distance between the i th station and the base bench mark. For the single-run first-order leveling surveys of 1982 and 1983, $\gamma = 2.0 \text{ mm/km}^{\frac{1}{2}}$. For the double-run first-order leveling survey of 1975, $\gamma = 1.5 \text{ mm/km}^{\frac{1}{2}}$. The lower precision value $\gamma = 2.0 \text{ mm/km}^{\frac{1}{2}}$ was considered as a measure of the error in the 1975–1982 and 1975–1983 data. Assuming a Gaussian distribution of errors, the 95% confidence intervals for $m_i (\pm 2\sigma_i)$ were incorporated into the inversion.

Application of extremal inversion to the leveling data of Long Valley Caldera

The method of extremal inversion was applied to the data set discussed above. The aim was to determine bounds on the depth to the top of the body and on the location of the east-west and north-south boundaries of all possible magma bodies fitting the data. First the region of interest was divided into 15 layers and each layer was divided into 100 horizontal blocks. This resulted in 1500 blocks, each of 1 km height, 2.65 km east-west length, and 2.39 km north-south length. The initial volume of each block was 6.33 km^3 . The range of possible models then is represented by all the possible combinations of fractional volume changes within the blocks. Given this space of possible models and a desire to produce bounds on the depth and horizontal boundaries of acceptable models, it is necessary to define an appropriate generalized moment. I have chosen to use α_n as defined in Eq. (3).

First examine the bound on the depth to the top of

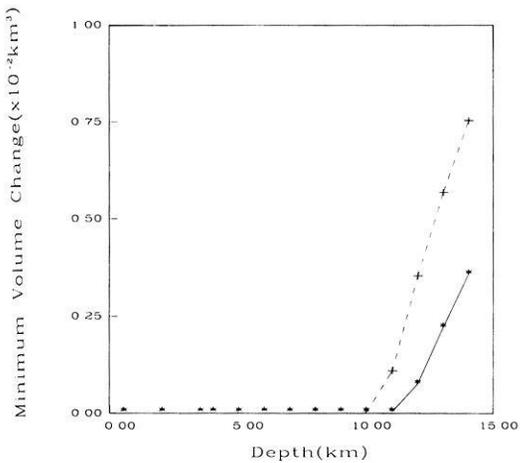


Fig. 4. Vertical depth bounds derived from the leveling data in Fig. 2. Shown here is the minimum volume change which must occur above the given depth. The bounds derived from the 1982 data are denoted by a *solid line*, while the 1983 bounds are denoted by a *dashed line*

the magma body. Consider a horizontal plane which lies at a depth h_1 . Define U to be the region between h_1 and the free surface and find the solution $\Delta \varrho_n$, $n=1, 2, \dots, N$ which minimizes the total fractional volume change in the region U given by the functional A while still satisfying the constraints (4). Now move to a greater depth h_2 and repeat the process. For each depth (h_1, h_2, \dots) one has a particular minimum value of A . Plotting these particular minimum values of A derived for the various regions with lower boundaries given by h_i against the depth h_i results in the curves in Fig. 4. The least upper bound on the depth of the body is given by the depth of the first point where the volume change is nonzero, for this is the shallowest depth above which some volume change is required in order to satisfy the data. If the lower boundary of the region U extends down to or deeper than this point, then some volume change is required in U . The lower bound on the required volume change is given by the ordinate. As can be seen in Fig. 4, in order to satisfy the 1982 leveling data some volume change must have occurred above 12 km. For the 1983 survey the bound is 11 km.

The method of extremal inversion was also used to produce horizontal bounds on the body. This was done in the same way as for the vertical bounds. A plane perpendicular to a specified direction defines a region U to the right or to the left of the plane. The minimum volume change in region U is sought and the plane is then shifted to a new position. The results are shown in Figs. 5 and 6 for east-west and north-south directions, respectively. Here, both right and left bounds are shown. In Fig. 5, one can see that there has been essentially no change in the east-west bounds between August 1982 and August 1983. However, there has been a change in the north-south limits to the body. The bound for the 1983 data has moved one grid element (2.39 km) to the north.

Discussion and conclusions

Extremal inversion techniques were able to produce depth bounds on a proposed magma body within Long Valley

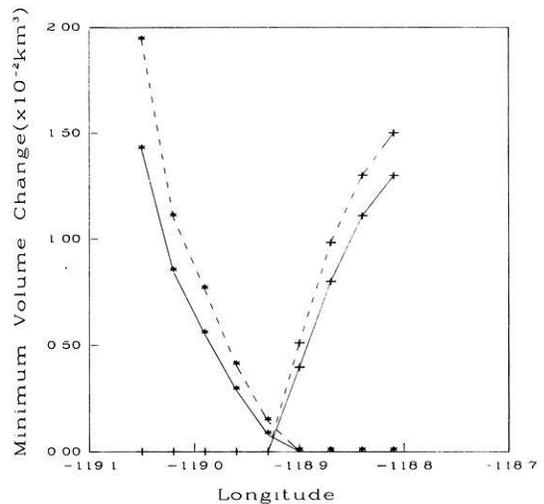


Fig. 5. East-west (longitudinal) bounds derived from the leveling data. This displays the minimum volume change which must occur to the east of the points* and to the west of the points+. The bounds on the 1982 data are indicated by *solid lines*, while the 1983 bounds are given by *dashed lines*

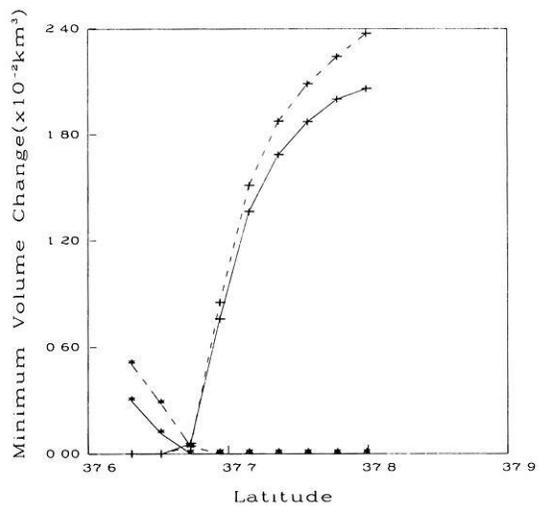


Fig. 6. North-south bounds (latitudinal) derived from the leveling data. This displays the minimum volume change which must occur to the north of the points* and to the south of the points+. The bounds on the 1982 data are indicated by *solid lines*, while the 1983 bounds are given by *dashed lines*

Caldera. The August 1982 leveling data require a volume change above a depth of 12 km while the August 1983 data require some volume change above a depth of 11 km. The horizontal bounds are for the most part unchanged, the only difference is the northward extension of the north-south bounds by one grid element. The significance of these results lies in what they indicate about the range of models that may fit the vertical leveling data. The sole conclusion one may make about the depth to the top of any supposed magma chamber is that it must be less than or equal to 12 km in 1982 and less than or equal to 11 km in 1983. As for the horizontal bounds (Figs. 5 and 6), in the case of the longitudinal or east-west bounds the only requirement is that volume change occur east of 118.93° W and west of 118.90° W. Similarly, any model satisfying the data for both 1982 and 1983 must lie north of 37.65° N and

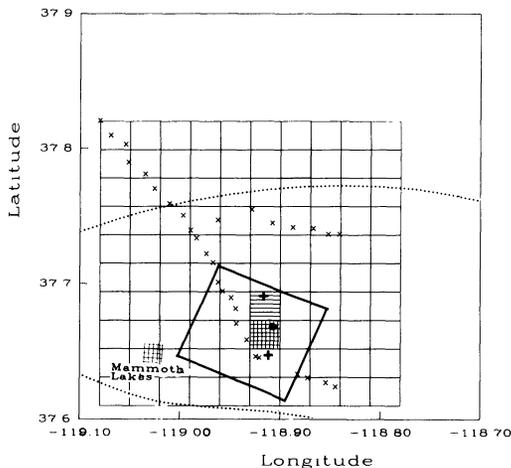


Fig. 7. Map of the Long Valley/Mono Craters region showing the 1982/1983 survey lines used. Stations are denoted by 'x's. The horizontal discretization of the region is indicated. The town of Mammoth Lakes is shown, as is the caldera boundary (dotted line). The horizontal bounds for 1975-1982 are shown (vertical hatches), as are the bounds for 1975-1983 (horizontal hatches). The point source model of Castle et al. (1984) is denoted by the dot. The two magma chambers of Rundle and Whitcomb (1984) are denoted by crosses. The projection of the dipping dike model of Savage and Cockerham (1984) on to the surface is the black rectangle

south of 37.67° N. The 1983 north-south bounds differ in that the northernmost bound has moved outward to 37.69° N. For the discretization given it is not possible to constrain the width of the source body. If one assumes that a homogenous body has given rise to the data then the body must lie, in part, between 118.93° W and 118.90° W and between 37.65° N and 37.67° N. So the geometrical constraints on possible models satisfying the data have been clearly laid out. It should be noted that these are necessary and not sufficient constraints. Any model satisfying the data must include some volume change in the region described above, but a model with volume change in the region does not necessarily satisfy the data. Finally, the bounds derived are not merely the properties of a point source in the given model space (discretization). This is because it is the requirement that the models fit the data within two standard deviations which determines the bounds.

A number of models have been proposed to explain the vertical and horizontal deformation (Rundle and Whitcomb, 1984; Savage and Cockerham, 1984; Castle et al., 1984). Though the models do differ in detail they seem to share some common properties. All models require inflation of a magma reservoir beneath the resurgent dome to fit the vertical displacement data. It is interesting to compare the above models with the bounds placed on the range of possible models by the method of extremal inversion. First consider the point source model of Castle et al. (1984). For the 1975-1983 vertical deformation data their model lies within the specified extremal bounds (Fig. 7). Similarly, the models of dipping dike intrusions of Savage and Cockerham (1984) satisfy the horizontal bounds (Fig. 7). Finally, consider the model of Rundle and Whitcomb (1984) which involves the inflation of two spherical magma chambers, one at a depth of 5 km located 1.5 km west of leveling station Casa and the other 9 km deep about 5.5 km north-

northwest of Casa. Their two magma chambers taken separately do not satisfy the horizontal bounds derived above. However, because the body is not a single body it is not required to lie in the region shown in Fig. 7, the intersection of the horizontal bounds. Non-convex or multiply connected bodies can satisfy the bounds without having volume change occur within the region in Fig. 7. The depths to the model of Rundle and Whitcomb of 5 and 9 km also satisfy the restriction that some or all of the magma intrusion occur at or above 12 km in the 1975-1982 interval and at or above 11 km in the 1975-1983 interval. Furthermore, the models of Castle et al. (1984) and Savage and Cockerham (1984) also satisfy the depth bounds derived. Therefore, the extremal bounds encompass three recently proposed models. However, the importance of the extremal bounds is not in judging proposed models. The importance lies in what the method states about the limits of the vertical displacement data in determining the location and shape of a model. The best one can say with the given data set, for the chosen parameters, is that volume change must have occurred somewhere in the rectangle defined by the latitude, longitude and depth bounds if the body is assumed to be a single convex body. If multiple or non-convex bodies are allowed, then one can merely say that the bodies must be distributed such that all of the individual bounds are satisfied.

In addition to the vertical and horizontal displacement data, the models are constrained by gravitational and magnetic field changes, teleseismic *P*-wave residuals (Steeple and Iyer, 1976) and *S*-wave attenuation data (Sanders, 1984). One might hope that comparisons could be made among the various data sets. Extremal inversion can provide one model-independent way to accomplish this. Extremal inversion techniques have been developed for gravity and magnetic (Safon et al., 1977), temperature (Huestis, 1979) and electro-magnetic induction (Weidelt, 1981) problems as well as for static displacements. The bounds derived from each of the above data sets can be compared. The data set which most tightly constrains some model property such as the depth to the top of the magma body can be determined. So the effectiveness of each data set in constraining the range of possible models becomes clear. This allows a more realistic assessment of the constraints on the body giving rise to the data sets.

It must be pointed out that the analysis was somewhat simplified. It was assumed that all fractional volume changes were positive. This excludes local deflation and assumes that the source of the material causing the expansion was sufficiently removed from the stations. Also, the interpretation of the bounds depends on assumptions of the form of the body, i.e. if it is convex or non-convex. A Poisson's ratio of 0.25 was assumed for the whole caldera. But, as can be shown, Poisson's ratio does not affect the depth estimate itself, though it does affect the minimum volume change estimates. Also, although it passed nearby, the survey did not traverse the region of maximum uplift. Hence, the depth bounds are slightly deeper than necessary but are still valid and unique for the given data set. A homogeneous halfspace was assumed in the calculations. Jovanovich et al. (1974) noted the effects of layering upon displacements. They also presented integral relationships between displacement and volume change which would allow one to invert the leveling data while accounting for overlying structure. Layers of high rigidity, such as thick lava flows,

tend to reduce and broaden surface flexure resulting in an overly cautious depth bound and a greater minimum width estimate. Finally, the volume change was assumed to have taken place in a specified region. That region was then discretized. By changing the boundaries of the volume considered, it was found that the extremal bounds are insensitive to the exact extent of the region. For example, vertical bounds were calculated for regions with total depth extents of 15, 20 and 25 km. The minimum volume changes at depth were identical for each of the regions. In order to estimate the depth or width bounds correctly it is only necessary for the region considered to encompass the boundary between the area of zero volume change and the area of nonzero volume change. However, the exact location of the bounds depends on the discretization; finer divisions of the region will give better bounds. One is only limited by computational expense in deciding on a discretization. Given these caveats, I believe that this is a robust technique which has produced meaningful bounds on the source volume in Long Valley Caldera.

Acknowledgements. I would like to thank Dr. J.C. Savage for supplying the uplift data. I would also like to thank Paul LeGros for the helpful criticism and encouragement.

References

- Castle, R.O., Estrem, J.E., Savage, J.C.: Uplift across Long Valley Caldera, California. *J. Geophys. Res.* **89**, 11507–11515, 1984
- Hadley, G.: *Linear Programming*, 1st edn. Reading, Massachusetts: Addison-Wesley 1962
- Huestis, S.P.: Extremal temperature bounds from surface gradient measurements. *Geophys. J.R. Astron. Soc.* **58**, 249–260, 1979
- Jovanovich, D.B., Husseini, M.I., Chinnery, M.A.: Elastic dislocations in a layered half-space I. Basic theory and numerical methods. *Geophys. J.R. Astron. Soc.* **39**, 205–217, 1974
- Julian, B.R.: Evidence for dyke intrusion earthquake mechanisms near Long Valley caldera, California. *Nature* **303**, 323–324, 1983
- Parker, R.L.: The theory of ideal bodies for gravity interpretation. *Geophys. J.R. Astron. Soc.* **42**, 315–334, 1975
- Rietsch, E.: Extreme models from the maximum entropy formulation of inverse problems. *J. Geophys.* **44**, 273–275, 1978
- Rundle, J.B., Whitcomb, J.H.: A model for deformation in Long Valley, California, 1980–1983. *J. Geophys. Res.* **89**, 9371–9380, 1984
- Sabatier, P.C.: Positivity constraints in linear inverse problems: I. General theory. *Geophys. J.R. Astron. Soc.* **48**, 415–441, 1977a
- Sabatier, P.C.: Positivity constraints in linear inverse problems: II. Applications. *Geophys. J.R. Astron. Soc.* **48**, 443–459, 1977b
- Sabatier, P.C.: On geophysical inverse problems and constraints. *J. Geophys.* **43**, 115–137, 1977c
- Safon, C., Vasseur, G., Cuen, M.: Some applications of linear programming to the inverse gravity problem. *Geophysics* **42**, 1215–1229, 1977
- Sanders, C.O.: Location and configuration of magma bodies beneath Long Valley, California, determined from anomalous earthquake signals. *J. Geophys. Res.* **89**, 8287–8302, 1984
- Sanders, C.O., Ryall, F.: Geometry of magma bodies beneath Long Valley, determined from anomalous earthquake signals. *Geophys. Res. Lett.* **10**, 690–692, 1983
- Savage, J.C., Clark, M.M.: Magmatic resurgence in Long Valley caldera, California: Possible cause of the 1980 Mammoth Lakes earthquakes. *Science* **217**, 531–533, 1982
- Savage, J.C., Cockerham, R.S.: Earthquake swarm in Long Valley Caldera, California, January 1983: Evidence for dike inflation. *J. Geophys. Res.* **89**, 8315–8324, 1984
- Steeple, D.W., Iyer, H.M.: Low-velocity zone under Long Valley as determined from teleseismic events. *J. Geophys. Res.* **81**, 849–860, 1976
- Vasco, D.W., Johnson, L.R.: Extremal inversion of static earth displacements due to volume sources. *Geophys. J.R. Astron. Soc.* **80**, 223–239, 1985
- Weidelt, P.: Extremal models for electromagnetic induction in two-dimensional perfect conductors. *J. Geophys.* **49**, 217–225, 1981

Received March 11, 1985; revised version July 16, 1985

Accepted July 22, 1985

In-situ permeability from non-dilatational soil deformation caused by groundwater pumping – a case study

H.-J. Kümpel¹, G. Lohr²

¹ Dalhousie University, Halifax, Canada; now at University of Kiel, Federal Republic of Germany

² University of Kiel, Federal Republic of Germany

Abstract. Short-term disturbances in the recording of a sensitive borehole tiltmeter are found to be generated by groundwater pumping at 120 m distance. Biot's consolidation theory for elastic porous media is applied to simulate the physical process involved. Results from various finite-element calculations for axisymmetric conditions are obtained. The findings are:

(1) The observed tilt disturbances can be explained as being provoked by elastic consolidation phenomena.

(2) Non-dilatational tilt and strain deformations can be measured outside the region where the pore pressure has declined due to pumping.

(3) The time variation of the deformations constrains the in-situ permeability of the aquifer within some volume around the screen of the pumped well.

The fundamentals of this case study may be useful to hydrologists and to those investigators who set up high-resolution tilt and/or strain measurements for purposes of geodynamic interest.

Key words: In-situ permeability – Soil deformation – Pump tests – Aquifer parameters – Tilt and strain measurements

Introduction

Methods to estimate aquifer parameters from pressure variations in a well have been described by different authors. Some of them investigated well tides from aquifer dilatation caused by earth tides or ocean loading tides (Bredehoeft, 1967; Robinson and Bell, 1971; Rhoads and Robinson, 1979), or considered well level undulations due to tidal gravity changes on the aquifer's overburden (Morland and Donaldson, 1984). Others analysed well pressure fluctuations due to volumetric strains caused by seismic disturbances (Cooper et al., 1965). These pressure variations are reported to be significant only if measured in confined aquifers. Soil parameters from artificially forced well level oscillations have been obtained by Krauss (1974). Unlike the methods relying on natural forcing functions, the latter technique is equally applicable to unconfined aquifers.

In many cases, satisfactory estimates can be made for pore fluid viscosities, temperatures, densities, compressibilities and shear moduli of formations. The quantities derived from well pressure variations are therefore used to estimate

the in-situ porosity and permeability of the aquifer. Under favourable conditions, the penetration depth related to the various methods extends from some ten to hundred metres, depending on the periods of the pressure variations and on the transmissivity of the aquifer (Bodvarsson, 1970; Krauss, 1974; Varga, 1976).

We present a case study showing how in-situ permeability of an unconfined aquifer is estimated from soil deformations caused by pumping. Since the deformations are sensed by a tiltmeter which is installed in a borehole situated away from the pumped well, mainly the non-dilatational part of the total deformation field is observed. The volume of the aquifer derived by this technique is probably larger than for the methods mentioned above.

The physical process involved is described by Biot's widely used consolidation theory (Biot, 1941). A load, when applied to a porous elastic medium, yields instantaneous deformation by compression of both the porous skeleton, called the matrix, and the pore filling. Further deformation of the medium arises from body forces; namely, from internal pore pressure gradients that are stimulated by the compression. These forces vary with time since the pressure gradients decrease through diffusion until a new steady-state equilibrium is achieved. The time variation of the process is mainly governed by the permeability of the medium and the viscosity of the pore filling. Zschau (1979) applied Biot's theory to show that meteorological air pressure variations can account for systematic tilt anomalies that he observed several tens of metres below the slope of a hill.

In our case, the deformation of the aquifer results from the pore pressure gradient that is stimulated at the screen of the pumped well. The tilt amplitudes measured are 5–10 times smaller than the earth tidal tilt amplitude. This might be the reason why soil engineers and hydrologists have not, so far, considered the method as a possible tool to determine aquifer parameters, although the use of high-precision instruments will not always be an inevitable necessity. The findings also have implications for the implementation and interpretation of tilt and strain measurements in projects of geodynamic interest.

Observations

The Institute of Geophysics, Kiel, runs a tiltmeter station in the northern Federal Republic of Germany, close to the village of Medelby (Fig. 1). Former purposes of the site, built up in 1977, have been loading investigations in the

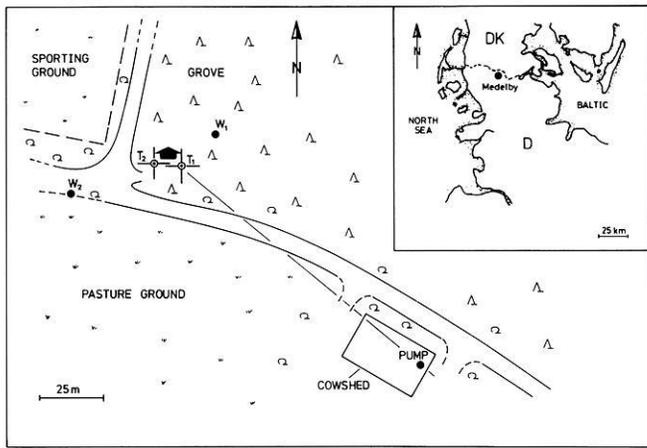


Fig. 1. Location map of the village of Medelby (DK = Denmark, D = Federal Republic of Germany) and of the tiltmeter site to the west of Medelby (T_1 , T_2 = tiltmeter boreholes with registration hut; W_1 , W_2 = groundwater observation wells)

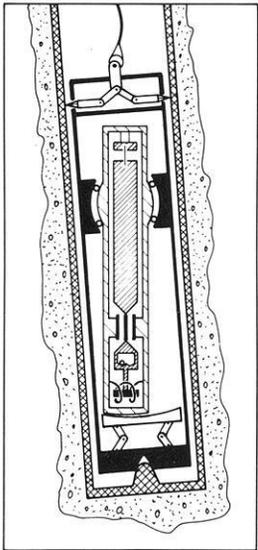


Fig. 2. Schematic cross-section of the major components of an Askania borehole tiltmeter with sensor pendulum (*heavy hatching*), pendulum holder (*light hatching*) and instrument casing (*black*). Elements of the sensor pendulum, from top to bottom, are: suspension strips, capacitance transducer, calibration chamber, and feedback magnet (inside coil). The lower end of the borehole ends in a cemented-in steel casing (*cross hatching*) that holds the instrument. When in operation, only the sensor pendulum hangs vertically

environments of the North Sea and the Baltic (Kümpel, 1982).

The tiltmeter we use is a continuously recording Askania-Gezeitenbohrlochpendel vertical pendulum type (Fig. 2). The top of the casing of the pendulum is clamped to the borehole by spring-loaded studs, its bottom resides on a stainless steel cone. Since the borehole is not vertical, the sensor pendulum hangs obliquely in the casing with respect to the instrument axis. To allow for a greater obliqueness of the borehole and for a wider measuring range without loss of sensitivity, the sensor pendulum is suspended in a pendulum holder which itself hangs in a nearly vertical position. Prior to operation of the tiltmeter, the pendulum holder is fixed to the casing by raising a

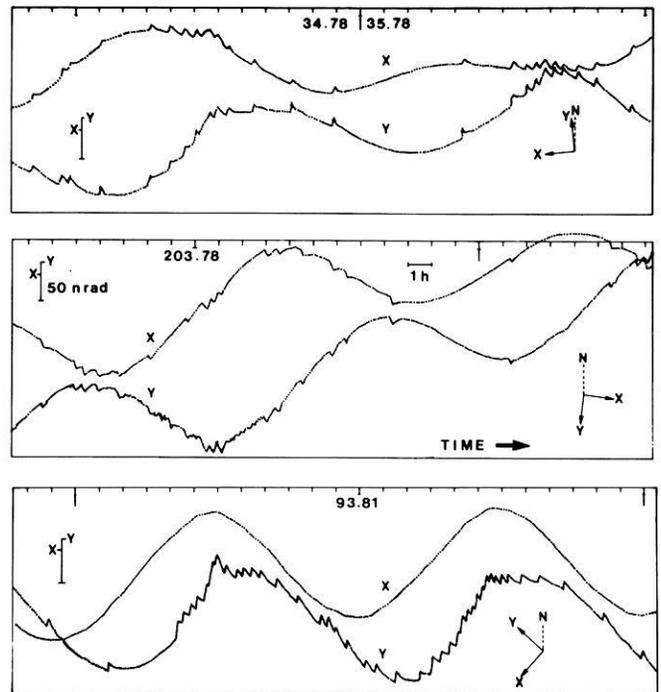


Fig. 3. Tilt disturbances 30 m below surface from pumping activities at 120 m distance. Graphs are tiltmeter recordings with components x and y at different azimuths (as indicated). The lower end of the tiltmeter borehole is always dragged towards the pump at 130° E of N; thus, on day number 93 in 1981 (*lower graph*), y is deflected towards 310° E of N with no disturbances in x . Long periodic undulations are tides

base plate. Tilt of the sensor pendulum with respect to the pendulum holder is sensed by a capacitance transducer. Calibration is achieved by forcing a steel ball to hop from one of two notches to the other within a small chamber. A feedback system made of coils and magnets provides active damping and raises the linear range of the instrument. The 60-cm-long sensor pendulum is suitable for measuring tilt in two perpendicular axes because of suspension strips that allow movements in any direction. Within the diurnal tidal bands, the instrumental resolution is 1 nrad (≈ 0.2 msec).

Due to microseismicity, the noise level at the site varies between 2 and 6 nrad. Besides the tidal signal and strong rainfall-induced tilt anomalies, we observe short-term tilt disturbances of 15–20 nrad amplitude (Fig. 3). Occurrence of these disturbances coincides with groundwater pumping for agricultural use at 120 m distance. The events become visible on the tilt recording within 30 s after activation of the pump, as was proven by monitoring the switch times of the pump. The time of response of the tiltmeter to pumping could even be shorter since the tilt signal is low pass-filtered with a cutoff period of 20 s. Without filter, a potentially immediate tilt response is masked by microseismicity.

A single event has a sawtooth-like shape with its recovery lasting longer than its stimulation. The sign of the disturbances is clearly independent of the azimuthal orientation of the tiltmeter. For seven different orientations (three are shown in Fig. 3), the maximum tilt effect occurs in a north-west direction. Moreover, the effect is independent of the individual tiltmeter borehole. The upper recording

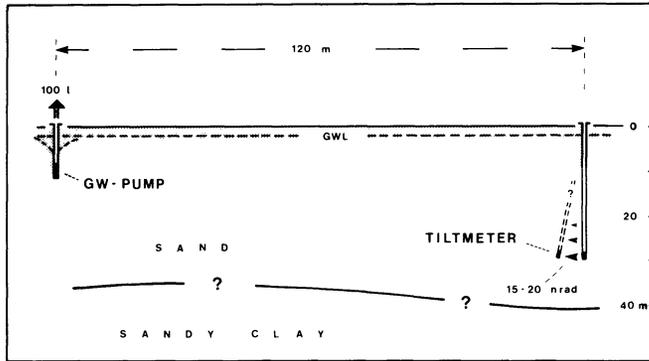


Fig. 4. Sketch of relevant features and proportions determining tilt disturbances at the Medelby site, northern FRG. Depth of intersection between sand and sandy clay layer is estimated from borehole information several hundred metres apart

in Fig. 3 is observed in hole T_1 , the lower two are observed in hole T_2 (see Fig. 1). Both holes are 30 m deep; T_1 is PVC-cased, T_2 is iron-cased. As the top of the tiltmeter is clamped to the borehole at 28.5 m depth, the non-Newtonian part of the tilt signal indicates the change in inclination of a vertical element between 28.5 and 30 m depth.

Figure 4 is a section of the site along the line T_1 - pump in Fig. 1. The well and the tiltmeter boreholes (only one is shown) have been drilled into a sand layer. Depending on the season, the depth of the groundwater level varies between 1.5 and 3 m. The well extends 9 m into the aquifer. Only the lower 3 m are perforated to allow the water to enter through the well screen. The diameter of the well is about 20 cm (8 in.) including the filter, made of pebbles. The diameter of the tiltmeter borehole is also 8 in. The productivity of the pump is approximately 20 l/min. Some 100 l are pumped during a typical pumping cycle. Whenever the pump draws groundwater, the lower end of the tiltmeter borehole is dragged towards the pumped well located to the south-east of the tiltmeter, thereby generating a tilt signal that is strongest in the opposite azimuth. Theoretically, the tilt events might also result from deflections of the plumb-line towards the north-west with the tiltmeter borehole remaining in a fixed position. However, simple calculation shows that the Newtonian effect of lifting a 100-kg mass over a height of 5 m at 120 m distance deflects the vertical by some negligible 10^{-7} nrad only. Within 30 m distance of the tiltmeter boreholes, we operate two continuously running water-table gauges of 5 mm resolution each (W_1 and W_2 in Fig. 1). Neither of them monitors a lowering of the water-table due to the pumping.

From other holes drilled in the vicinity of the site, it is known that (a) the sand is partially intersected by thin clay and clay shale layers and (b) somewhere below 30–40 m the sand changes its character to sandy clay. A detailed stratigraphy is not available. Several samples from the upper 2 m of the sand layer have been subjected to sieve analysis. The sand is well sorted, the median grain diameter is 0.2 mm, the clay fraction being less than 5%. Intrinsic permeability for a single sample from 1-m depth has been repeatedly measured in a laboratory flow experiment. The values obtained range from 5×10^{-8} to $22 \times 10^{-8} \text{cm}^2$, in agreement with Davis (1969) who reports permeabilities from 10×10^{-8} to $22 \times 10^{-8} \text{cm}^2$ for five sands of this size. Applying classical draw-down relations for wells in unconfined aquifers (e.g. Brown et al., 1972), we calculate in-situ

permeabilities in the range 1.5×10^{-8} – $6 \times 10^{-8} \text{cm}^2$ when taking into account that the well only partially penetrates the aquifer.

Formulation of model calculations

The relation between elastic deformation and pore pressure is described by the theory of consolidation for porous media (Biot, 1941). If axial symmetry holds, the governing system of differential equations is (Verruijt, 1969):

$$\mu \nabla^2 u - \mu \frac{u}{r^2} + (\lambda + \mu) \frac{\partial \varepsilon}{\partial r} - \alpha \frac{\partial \sigma}{\partial r} = 0$$

$$\mu \nabla^2 w + (\lambda + \mu) \frac{\partial \varepsilon}{\partial z} - \alpha \frac{\partial \sigma}{\partial z} = 0 \quad (1)$$

$$\frac{\alpha}{\lambda + \mu} \frac{\partial}{\partial t} \left(\frac{\sigma_r + \sigma_z}{2} \right) + \left(1 - \frac{\lambda}{\lambda + \mu} \right) \frac{\alpha}{r} \frac{\partial u}{\partial t} + \left(\frac{\alpha^2}{\lambda + \mu} + S \right) \frac{\partial \sigma}{\partial t} = \frac{k}{\eta} \nabla^2 \sigma$$

where the variables u (= radial displacement), w (= vertical displacement), σ (=excess pore pressure, with respect to the hydrostatic pressure) are functions of the coordinates r (= radius), z (= height) and t (= time). All forces are in equilibrium for t being negative, r is zero for the centre of the pore pressure disturbance, z is negative below the surface where the depth is positive.

$$\nabla^2 = \frac{\partial^2}{\partial r^2} + \frac{1}{r} \frac{\partial}{\partial r} + \frac{\partial^2}{\partial z^2} \quad \text{is the Laplacian operator,}$$

$$\varepsilon = \frac{\partial u}{\partial r} + \frac{u}{r} + \frac{\partial w}{\partial z}$$

is the volume strain (positive for compression),

$(\sigma_r + \sigma_z)/2$ is the average total stress.

Soil parameters are:

μ = shear modulus,

λ = Lamé constant,

$\alpha = 1 - c_g/c_m$ with c_g = grain compressibility and $c_m = (\lambda + 2\mu/3)^{-1}$ = compressibility of the porous matrix (after Nur and Byerlee, 1971),

$S = \rho \cdot c_f + (1 - \rho)c_m$, i.e. the hydraulic capacity where ρ is the volume porosity and c_f is the compressibility of the fluid filling the pore (after Bodvarsson, 1970),

k = intrinsic permeability,

η = dynamic viscosity of the pore fluid.

Following the notation of Davis and De Wiest (1966), $S_S = \delta \cdot g \cdot S$ is the specific storage of the formation where δ is the density of the pore fluid and g is the gravitational acceleration, and $K = k \cdot \delta \cdot g / \eta$ is the hydraulic conductivity. μ , λ , c_m are bulk formations constants for quasi-static, i.e. drained conditions as compared to undrained conditions that hold for elastic wave propagation.

Equation (1) is based on infinitesimal strains, elastic deformation (i.e. linear reversible stress-strain relations), soil that is saturated with pore fluid, the validity of Darcy's law (i.e. no turbulent flow) and the absence of inertia forces. The first two equations result from introducing Hooke's generalized law for elastic porous media into the well-known equilibrium conditions. Considerations on fluid flow according to Darcy's law and on the elastic storage capacity

of porous media under pressure lead to the third equation. Because the change in pore pressure σ is relevant for each part of Eq. (1), the system has to be solved simultaneously. Since the third equation is not exactly analogous to any heat conduction problem that has been analysed, analytical solutions are rare and become complex, even for simple geometric configurations.

Numerous consolidation problems have instead been solved using the finite-element technique. The program package available to us (IMSL, 1983) allows a straightforward conversion of the problem into programming code when not more than one time-derivative term per equation is encountered. Two modifications of Eq. (1) are introduced to meet this condition. The first modification neglects any time variation in the average total stress term. The effect of this simplification on various geometrical configurations has been analysed by Christian and Boehmer (1970). They did not find severe divergences with respect to the complete models if external loads remain constant with time, which also holds in our case. Yet, our numerical solutions are applicable only if other geometric and parametric simplifications within the model are as coarse approximations to reality as the assumption of constant average total stress is.

The second modification substitutes σ by a new variable

$$\Phi = A \cdot u/r + B \cdot \sigma,$$

which alters Eq. (1) into:

$$\begin{aligned} \mu \nabla^2 u - \mu \frac{u}{r^2} + (\lambda + \mu) \frac{\partial \varepsilon}{\partial r} - \frac{\alpha}{B} \left[\frac{\partial \Phi}{\partial r} + \frac{A}{r} \left(\frac{u}{r} - \frac{\partial u}{\partial r} \right) \right] &= 0 \\ \mu \nabla^2 w + (\lambda + \mu) \frac{\partial \varepsilon}{\partial z} - \frac{\alpha}{B} \left(\frac{\partial \Phi}{\partial z} - \frac{A}{r} \frac{\partial u}{\partial z} \right) &= 0 \quad (2) \\ \nabla^2 \Phi - \frac{A}{r} \left(\nabla^2 u - \frac{2}{r} \frac{\partial u}{\partial r} + \frac{u}{r^2} \right) &= \frac{B\eta}{k} \frac{\partial \Phi}{\partial t} \\ \text{with } A &= \left(1 - \frac{\lambda}{\lambda + \mu} \right) \alpha \quad \text{and } B = \frac{\alpha^2}{\lambda + \mu} + S. \end{aligned}$$

Tilt ψ of a vertical element at radius r , depth $z = (z_1 + z_2)/2$, and time t is obtained from

$$\tan^{-1} \psi(r, z, t) = \frac{u(r, z_2, t) - u(r, z_1, t)}{z_1 - z_2} \quad (3)$$

if $|z_2 - z_1|$ is taken much larger than vertical displacements. By this convention, ψ is positive when the upper end of the element is closer to the axis than the lower end. Respective relations hold for strain of vertical elements and for tilt and strain of horizontal elements.

Figure 5 shows the spatial division of the model into 96 triangular elements. Along a line segment combining two nodes, quadratic fits are exerted on the variables. The thin, drained layer above the water-table is omitted which leads to somewhat reduced depths of the well and of the borehole. The load of the drawn groundwater that increases atop the pump during a pumping cycle before it is supplied to a system of pipes also is neglected. The saturated sand layer and the sandy clay layer are characterized by their parameters E , ν , ρ and k , where $E = 2\mu(1 + \nu)$ is the Young's modulus of the formation and $\nu = 0.5 \lambda (\lambda + \mu)^{-1}$ is its Poisson's ratio. The difference in grain compressibility c_g between the two layers is of minor importance only.

The upper part of the sandy clay layer is modelled by a sequence of thin elements. This has been found necessary

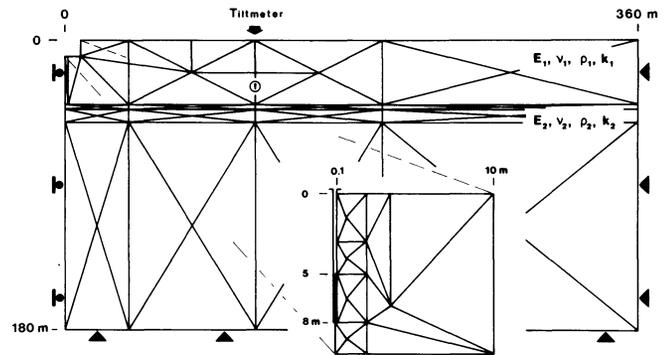


Fig. 5. Finite-element grid for modelling tilt disturbances in two-layered r -, z -plane (axial symmetry). The pump is located close to the upper left corner (see enlarged section). A pore pressure deficit is simulated between 5 and 8 m depth on the left boundary

to simulate the pore pressure diffusion across a layer intersection with a permeability contrast as is assumed here. The accuracy of the model for the net diffusion process is better than 10% when compared to analytical solutions. A further grid refinement is arranged close to the pump where the highest gradients in the variables u , w , σ are encountered.

As for the boundary conditions; displacements and pore pressure variations are prohibited at the lower boundary surface and at the outer radius of the model ($r = 360$ m), only vertical displacements and pressure changes are tolerated for the zero-radius. We disregard changes in the size of the well caused by the pumping, nor do we consider the elastic properties of the tiltmeter borehole as being different from those in its vicinity. Initially, displacements and excess pore pressure are assumed to be zero everywhere. The pumping is induced by a step function: at $t = 0$ s, a constant pore pressure deficit of -200 hPa is applied at the well screen. Eq. (2) is then successively evaluated at times $t = 2^n$ s ($n = 1, 2, \dots, 10$). Stability of the numerical results, in the sense of Booker and Small (1975), is achieved.

Computational results

Various models are tested. Parameters listed in Table 1 label the standard model. The static Young's moduli are deduced from a shallow seismic experiment at the site using theoretical and empirical relations between seismic velocities and static moduli for unconsolidated sediments (Ohkubo and Terasaki, 1977). In one of these relations, a Poisson's ratio of $\nu = 0.25$ is taken; we apply a more realistic value of $\nu = 0.45$ instead. The true static moduli might vary by up to $\pm 50\%$ of the values given in Table 1. Increasing E_1 , E_2 by the same factor, mainly reduces the displacement amplitudes by the reciprocal factor. For k_1 , the standard model adopts a value of $1.5 \times 10^{-8} \text{ cm}^2$. Effects due to different permeabilities may be easily traced from the results, as will be shown. Values of other parameters in Table 1 are taken from the literature in accordance with the parameters already known [ν_1 , ν_2 : Hamilton (1971); ρ_1 , ρ_2 , k_2 : Davis (1969); c_{g1} , c_{g2} : Simmons and Brace (1965); η , c_f : Kuchling (1971)].

Figure 6 shows how the excess pore pressure propagates in space and time. As expected, the pressure propagates faster for higher permeable sands. The time dependence

Table 1. Soil parameters of standard model

	z m	E dyne · cm ⁻²	ν	ρ %	k cm ²	c_g cm ² · dyne ⁻¹
Sand layer	(1)	0–32	2.0×10^9	0.47	45	15×10^{-9}
Sandy clay layer	(2)	32–180	3.33×10^9	0.45	63	4×10^{-11}
Pore fluid (water): $c_f = 5 \cdot 10^{-11}$ cm ² · dyne ⁻¹ , $\eta = 0.0165$ g · s ⁻¹ · cm ⁻¹ (20° C)						

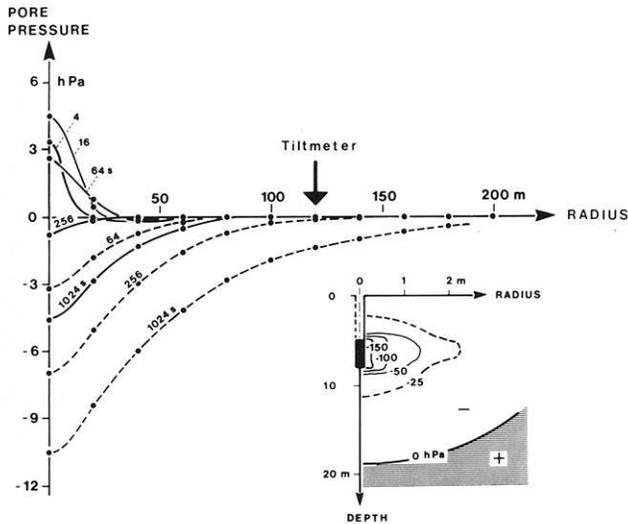


Fig. 6. Migration of pore pressure deficit after model calculations. Main diagram: change in pore pressure is plotted for depth 20 m, parameter is time in seconds from pump activation; solid and broken lines apply for permeabilities $15 \cdot 10^{-9}$ cm² (standard model) and $15 \cdot 10^{-8}$ cm² in sand layer. Inserted diagram: parameter is pore pressure change in hecto Pascals for standard model, time is 64 s from pump activation

of the process can be normalized when k/η is replaced by the dimensionless quantity $k \cdot t/L$ where L is some characteristic length of the geometry of the experiment.

The initial increase of the pore pressure for $r < 30$ m requires some further remarks. Although plotted for the standard model only, the increase occurs equally if other parameters are chosen. The phenomenon has been found in many types of consolidation problems, both theoretically and experimentally (e.g. Cryer, 1963; Verruijt, 1969), and can be explained from the elastic behaviour of the aquifer. When the pore pressure decreases at the well screen, mobile water is attracted from the aquifer. The loss of water yields a decrease in soil volume. The contraction is seen as a negative load by the surroundings. Loading deformation propagates with seismic velocity, i.e. much faster than the deficit of the pore pressure diffuses. As a consequence, soil particles displace towards the negative load, inwardly generating a compression of the soil. Since the pore fluid opposes volume compression, the pore pressure increases in areas where it has not yet declined due to diffusion (see also inserted diagram in Fig. 6). Within the first 1,024 s, the standard model predicts the pore pressure at the location of the tiltmeter to vary within the range of the numerical noise, i.e. by less than 0.1 hPa. Applying a k_1 -value 10 times higher already yields a pore pressure drop of 1.4 hPa after that time.

Clearly, significant soil displacements appear earlier at the tiltmeter location than any pressure drop. Radially hori-

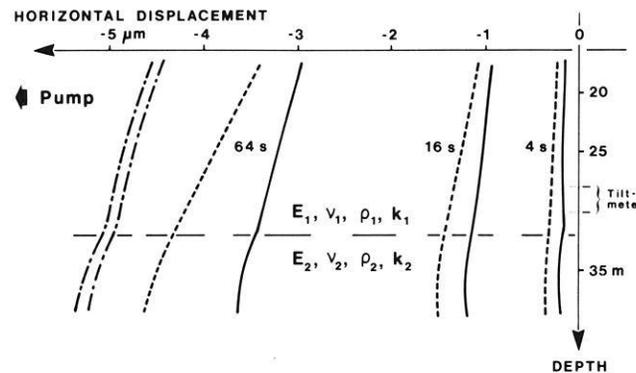


Fig. 7. Horizontal displacements for $r = 120$ m in the vicinity of the tiltmeter after model calculations. *Indexed lines*: parameter is time in seconds from pump activation, *solid lines* for standard model except $k_1 = 15 \times 10^{-8}$ cm², *broken lines* for same model except $\nu_2 = 0.47$, i.e. no contrast in Poisson's ratio between two layers. *Hatched dotted lines*: radial decrease of displacements, *left curve* for $r = 120$ m, *right curve* for $r = 122$ m (standard model, 1,024 s from pump activation)

zontal displacements for some of the models calculated are plotted in Fig. 7. It is noticeable how the displacements are influenced by the contrast in Poisson's ratio of the two layers. As time progresses, the displacements increase more rapidly if there is no contrast. The depth of the layer intersection is of minor influence as long as the tiltmeter borehole is completely in the upper layer. Some 5–10 m below the intersection, the displacements decrease with depth, which is partly due to the clamped boundary at 180 m depth. The radial decrease of displacements shown in Fig. 7 corresponds to horizontal strains of -5×10^{-8} ; respective vertical strains are positive and amount to 2×10^{-8} . Accordingly, the soil around the tiltmeter is found to be radially stretched; namely, more than twice as much as compressed vertically.

Referring to the observed tilt disturbances, the most essential results are summarized in Fig. 8. Herein the pump is simulated to run during a time span of 5 min. Tilt means the change in inclination of a vertical element, which – at the location of the tiltmeter – hardly differs from the change in inclination of a horizontal element. The tilt response following the shutdown of the pump is obtained by subtracting the inverse tilt response – shifted by 5 min – from the response for the persistently running pump (dotted lines). Computational results and observations are compatible within the shaded area. The vertical width of this area is estimated from uncertainties in both the amplitude of the pore pressure deficit at the well screen and the values of some soil parameters. Note that the observed tilt signal does not allow tilt variations shorter than 20 s to be resolved.

The initial 25-s section of curve A is identical to the

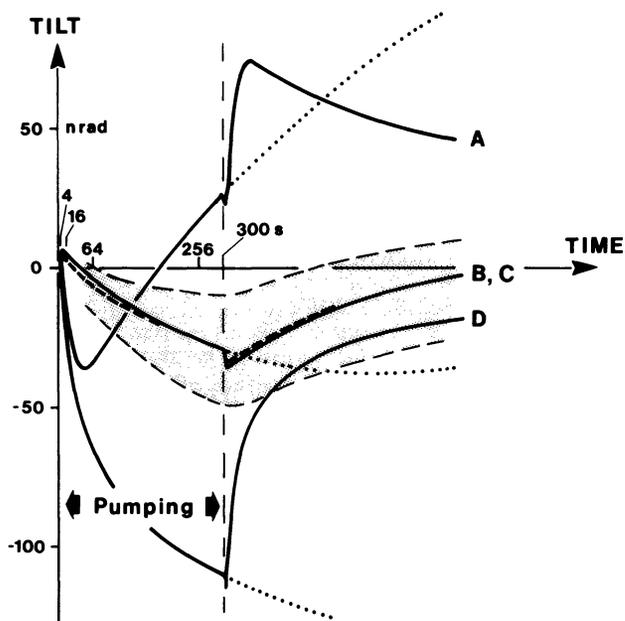


Fig. 8. Tilt response at location of tiltmeter due to pumping cycle lasting 5 min (linear time scale). A – standard model, except $k_1 = 15 \times 10^{-8} \text{ cm}^2$; B – standard model (solid line); C – standard model, except $k_2 = 15 \times 10^{-9} \text{ cm}^2$ (i.e. no permeability contrast between two layers, broken line); D – standard model, except $k_1 = 15 \times 10^{-8} \text{ cm}^2$ and $\nu_2 = 0.47$ (i.e. no contrast in Poisson's ratio between two layers); shaded area: consistent with observations

initial 300-s section of curve B except for a linear stretching of B due to a permeability 10 times lower. Curve D differs from A only in an absence of Poisson's ratio contrast between the two layers. Obviously, the deformation becomes sensitive to the parameter ν after some characteristic time, i.e. when the pore pressure starts to drop significantly at half the distance between the well and the tiltmeter borehole (Fig. 6). Neither curve A nor D fits the observations, nor do other models that are based on k_1 values as high as $15 \times 10^{-8} \text{ cm}^2$. If the permeability of the aquifer was better than $5 \times 10^{-8} \text{ cm}^2$, 5–10 times stronger tilt disturbances should have been recorded within the first minute after pump activation.

The tilt response of the standard model is consistent with the observations. This does not confirm the correctness of all the parameters applied. The permeability of the sandy clay layer, for example, is not constrained by any model within the first 300 s (see curve C). Other parameters, including the geometry of the standard model, may also vary considerably without destroying the compatibility condition. Still, the bulk in-situ permeability of the upper 30–40 m ranges among the parameters best constrained by the time variation of the tilt response.

So far, little has been said about the size of the region around the well that provokes the deformation at the location of the tiltmeter, i.e. the region to which we refer the permeability estimates. The source for the external deformation is the volume that decreases due to internal pore pressure drop. Naturally, if pumping continues, the size of this volume grows. Its shape is mainly determined by the shape of the well screen and variations in permeability around the well. Due to the diffusion process, the volume is not limited by a clearly defined boundary but can instead be described through concentric surfaces of equal pressure

drop (see also Fig. 6) or, in other words, through isobars normalized to the hydrostatic pressure.

As for the size of the volume that is mostly efficient for the external deformation, there are two effects that oppose each other. Since the gradient of the pore pressure is a body force in Eq. (1), those parts of the volume that enclose the highest gradients seem to be mostly efficient. With cylindrical or spherical (but not plane strain) diffusion processes, the gradient of the diffusion variable increases towards the centre of the disturbance, regardless of how long the disturbance has already been active (Carslaw and Jaeger, 1959). This proves the innermost volume around the well screen to be mostly efficient for the external deformation.

An isobar of low percentage pressure drop is effective, however, in that it is both more expanded and closer to external points of deformation measurements than is a more central isobar of higher percentage pressure drop. Indeed, the computational results show that the effects due to the latter reasoning predominate those due to the former. The instantaneous deformation following the pump activation is provoked by the highest pressure gradient ever occurring around the well. Still, the deformation increases substantially when the volume of the pore pressure drop expands, thereby lowering the pressure gradient at the well screen.

Somewhat arbitrarily, we define the 5% pressure drop isobar as the one limiting the effective volume. Then, in our case, the standard model yields a radial penetration depth of 15 m after 5 min of pumping and a vertical penetration depth of 8 m.

Conclusions

There are two major implications from the findings, provided the approximations made are tenable.

First, bulk in-situ permeabilities of aquifers may be estimated from ground deformations that are caused by pumping. Since elastic deformation is provoked beyond the region of declined pore pressure, indications for the transmissivity of the soil are available up to distances where this deformation can be measured. The penetration depth of the method exceeds that of those methods that rely on well pressure fluctuations because it is controlled by the duration of the stimulated pore pressure disturbance. Most characteristic for the in-situ permeability is the time function of the deformation. Experiment expenses might be reduced to reasonable amounts if strong pressure signals are generated in the well so that less sensitive instruments fulfil the deformation measurements. A detailed description of tiltmeters and strainmeters used for geodynamic purposes has been given by Agnew (1985). Pumping could also be replaced by fluid injection, yielding a change in the sign of the soil displacements. In many cases it will be sufficient to install a tiltmeter or a strainmeter in a shallow borehole. Surface deformation due to temperature variation or wind stress can be separated from the deformation signal caused by pumping because the switch times of the pump are well known. Signal enhancement can be achieved when cyclic pulse tests are set up.

Second, the consolidation theory seems to be applicable to certain deformation phenomena observed with highly sensitive instruments (like earth tide meters). In particular, rainfall-induced water-table variations are believed to produce the most annoying noise in geodynamic signals (e.g.

Wood and King, 1977; Kämpel, 1982). Modelling these generally local phenomena improves the understanding of tilt and strain recordings that are obtained in tectonic regions for the purpose of earthquake prediction, for instance.

Acknowledgements. We owe our sincere thanks to Chris Beaumont, John Peters and Jochen Zschau for many fruitful discussions, to Michael Huszak who did laboratory experiments on soil samples and to Kathleen Helbig who made numerous helpful comments on the manuscript. Hans-J. Kämpel received a Killam postdoctoral fellowship at Dalhousie University in Halifax, Canada, where all the model calculations were done. The German Research Foundation (DFG) supported the tiltmeter experiment at Medelby. The comments of two unknown reviewers are gratefully acknowledged.

References

- Agnew, D.C.: Strainmeters and tiltmeters. Reviews of Geophysics, submitted 1985
- Biot, M.A.: General theory of three-dimensional consolidation. *J. Appl. Phys.* **12**, 155–164, 1941
- Bodvarsson, G.: Confined fluids as strainmeters. *J. Geophys. Res.* **75**, 2711–2718, 1970
- Booker, J.R., Small, J.C.: An investigation of the stability of numerical solutions of Biot's equations of consolidation. *Int. J. Solids Structures* **11**, 907–917, 1975
- Bredhoeft, J.D.: Response of well-aquifer systems to earth tides. *J. Geophys. Res.* **72**, 3075–3087, 1967
- Brown, R.H., Konoplyantsev, A.A., Ineson, J., Kovalevsky, V.S.: Groundwater studies. Studies and reports in hydrology, 7, Paris (UNESCO), 1972
- Carslaw, H.S., Jaeger, J.C.: Conduction of heat in solids. Oxford: Clarendon Press 1959
- Christian, J.T., Boehmer, J.W.: Plane strain consolidation by finite elements. *J. Soil Mech. and Found. Div.*, ASCE **96**, 1435–1457, 1970
- Cooper, H.H. Jr., Bredhoeft, J.D., Papadopoulos, I.S., Bennett, R.R.: The response of well aquifer systems to seismic waves. *J. Geophys. Res.* **70**, 3915–3926, 1965
- Cryer, C.W.: A comparison of the three-dimensional consolidation theories of Biot and Terzaghi. *Q. J. Mech. Appl. Math.* **16**, 401–412, 1963
- Davis, S.N.: Porosity and permeability of natural materials. In: Flow through porous media, R.J.M. De Wiest, ed.: pp. 54–89, New York: Academic Press 1969
- Davis, S.N., De Wiest, R.J.M.: Hydrogeology. New York: John Wiley and Sons 1966
- Hamilton, E.L.: Elastic properties of marine sediments. *J. Geophys. Res.* **76**, 579–604, 1971
- IMSL, Inc.: Twodep, 5th ed. Houston: IMSL, Inc. 1983
- Krauss, I.: Die Bestimmung der Transmissivität von Grundwasserleitern aus dem Einschwingverhalten des Brunnen-Grundwasserleitersystems. *J. Geophys.* **40**, 381–400, 1974
- Kuchling, H.: Physik, Formeln und Gesetze. Köln: Buch- und Zeit-Verlagsgesellschaft mbH 1971
- Kämpel, H.-J.: Neigungsmessungen zwischen Hydrologie und Ozeanographie. Ph.D. thesis, Univ. Kiel, 1982
- Morland, L.W., Donaldson, E.C.: Correlation of porosity and permeability of reservoirs with well oscillations induced by earth tides. *Geophys. J. R. Astron. Soc.* **79**, 705–725, 1984
- Nur, A., Byerlee, J.D.: An exact effective stress law for elastic deformation of rocks with fluids. *J. Geophys. Res.* **76**, 6414–6419, 1971
- Ohkubo, T., Terasaki, A.: Physical property and seismic velocity of rock. OYO-Technical Note 22, translation from: Soil and Foundation **19**, 7, 1971 (in Japanese), 1977
- Rhoads, G.H. Jr., Robinson, E.S.: Determination of aquifer parameters from well tides. *J. Geophys. Res.* **84**, 6071–6082, 1979
- Robinson, E.S., Bell, R.I.: Tides in confined well-aquifer systems. *J. Geophys. Res.* **76**, 1857–1869, 1971
- Simmons, G., Brace, W.F.: Comparison of static and dynamic measurements of compressibility of rocks. *J. Geophys. Res.* **70**, 391–398, 1965
- Varga, P.: Investigation of earth tides by observing dilatational variations of the water table. *Bull. d'Inform. Marrées Terrestres*, P. Melchior (ed) **74**, 4319–4332, 1976
- Verruijt, A.: Elastic storage of aquifers. In: Flow through porous media, R.J. M. De Wiest, ed.: pp 331–376, New York: Academic Press 1969
- Wood, M.D., King, N.E.: Relation between earthquake, weather, and soil tilt. *Science* **197**, 154–156, 1977
- Zschau, J.: Air pressure induced tilt in porous media. In: Proc. 8th Int. Symp. Earth Tides, Bonn 1977, M. Bonatz, P. Melchior eds.: pp. 418–433, Bonn: 1979

Received March 27, 1985; Revised version July 8, 1985

Accepted July 17, 1985

Construction of conductance bounds from magnetotelluric impedances

P. Weidelt

Institut für Geophysik und Meteorologie, Technische Universität, Mendelssohnstraße 3, D-3300 Braunschweig, Federal Republic of Germany

Abstract. Whereas any finite set of impedance data does not constrain the electrical conductivity $\sigma(z)$ at a fixed level z in a 1D-model, the conductance function $S(z_2)$ as the depth-integrated conductivity from the surface to the depth z_2 will be constrained. Assuming only the non-negativity of $\sigma(z)$, it is shown that for a given depth z_2 the models generating the lower and upper bound of $S(z_2)$ consist of a sequence of thin sheets. The determination of the positions of the thin sheets and their conductances leads to a system of nonlinear equations. As a limitation the present approach requires the existence of a model, which exactly fits the data. The structure of the external models as a function of z_2 is discussed in examples with a small number of frequencies. Moreover, it is shown that any set of complex 1D impedances for M frequencies can be represented by a partial fraction expansion involving not more than $2M$ (positive) constants. For exactly $2M$ constants there are two complementary representations related to the lower and upper bound of $S(z_2)$.

For the simple one-frequency case, a more general extremal problem is briefly considered, where the admitted conductivities are constrained by a priori bounds $\sigma_-(z)$ and $\sigma_+(z)$ such that $\sigma_-(z) \leq \sigma(z) \leq \sigma_+(z)$. In this case, the extremal models for $S(z_2)$ consist of a sequence of sections with alternating conductivities $\sigma_-(z)$ and $\sigma_+(z)$. The sharpening of conductance bounds by incorporating a priori information is illustrated by an example.

Key words: Electromagnetic induction – Inverse problem – Extremal models

1. Introduction

The 1D-magnetotelluric inverse problem is known to be ill-posed and thus allows the construction of a whole set of satisfactory conductivity models from a given real data set. The COPROD study of Jones (1980) provides a good illustration of this fact. At present there are two attempts to overcome the problem: either the inversion is stabilized by incorporating known or assumed properties of the conductivity structure as a priori constraints, or one may try to extract geophysically useful properties pertaining to the whole

class of conductivity models consistent with the data. In principle, the latter problem can be approached either by exploring the whole space of feasible models by Monte Carlo techniques or by explicitly constructing the model, which extremizes the geophysically interesting property. Firm foundations for the use of the Monte Carlo method as a tool for geophysical inversion were laid by Anderssen and Seneta (1971, 1972) and the method of parameter extremization in geophysical inverse problems was pioneered by Parker (1972, 1974, 1975).

Any set of magnetotelluric impedances for a finite number of frequencies does not impose bounds on the conductivity $\sigma(z)$ at any fixed depth z . At this depth, either a thin insulating sheet or a sheet of unbounded conductivity, but of finite conductance (conductivity thickness product), may exist. However, conductivity averages over a finite depth range will, in general, be constrained by the data, provided the field penetrates down to this depth. This has been exemplified in detail by Oldenburg (1983), who constructed bounds on conductivity averages by linearizing the nonlinear problem. The existence of bounds for conductivity averages or for the simpler conductance function

$$S(z_2) = \int_0^{z_2} \sigma(z) dz \quad (1.1)$$

reflects the fact that the inverse problem for $S(z_2)$ is well-posed (V.I. Dmitriev, private communication).

The present study is centered on the computation of bounds for $S(z_2)$, imposing apart from the non-negativity condition $\sigma(z) \geq 0$ no further constraints on the conductivity. Any model consistent with the data must lie within these bounds. The problem under consideration resembles the problem of discovering extremal models of linear functionals of the density from a truncated set of eigenfrequencies of an elastic string, as treated by Barcilon (1979), Barcilon and Turchetti (1979), and Sabatier (1979). Briefly addressed is also the more general problem of computing bounds for conductivity averages rather than for $S(z_2)$.

We shall consider only for the simple one-frequency case the construction of bounds for $S(z_2)$ when $\sigma(z)$ is constrained by a priori bounds $\sigma_-(z)$ and $\sigma_+(z)$ such that $\sigma_-(z) \leq \sigma(z) \leq \sigma_+(z)$.

Contrary to the pragmatic approach of Oldenburg (1983), who applies his approximate method to a large number of frequencies, attention is confined in this very preliminary study to the exact extremal models for a small number of impedances. We also have to assume that there is a model that fits the data exactly. This restriction, however, may be dropped in subsequent work, thus extending the range of applicability to real inconsistent data sets. Only in connection with the COPROD study are approximate extremal models for many real data considered.

Although the present problem is one of the simplest extremal problems in electromagnetic induction, the nonlinearity introduces a great deal of complexity and leaves the treatment still in an experimental stage. Another relatively simple extremal problem in electromagnetic induction can be solved for two-dimensional perfect conductors (Weidelt, 1981).

The main part of the paper consists of three sections. The general structure of the extremal models is derived in Sect. 2. Then a detailed discussion and illustration of unconstrained extremal models is given in Sect. 3, and the concluding Sect. 4 is devoted to simple extremal models with a priori constraints on conductivity. The paper has two appendices, where in particular Appendix B contains the proof of an impedance representation theorem, to which we have to appeal in Sect. 3.

2. Necessary conditions for extremal models

Assuming SI units, a time factor $e^{i\omega t}$, a 1D-conductivity structure $\sigma(z)$, and neglecting displacement currents, Maxwell's equations reduce for a quasi-uniform incident magnetic field in the y -direction to

$$E'_x(z, \omega) = -i\omega\mu_0 H_y(z, \omega), \quad H'_y(z, \omega) = -\sigma(z)E_x(z, \omega), \quad (2.1)$$

implying

$$f''(z, \omega) = i\omega\mu_0 \sigma(z) f(z, \omega), \quad (2.2)$$

where $f(z) := E_x(z, \omega)$. In the sequel we use the transfer function c introduced by Schmucker (1970). Its theoretical value for given $\sigma(z)$ at a set of M frequencies ω_j , $j=1, \dots, M$ is defined as

$$c_j[\sigma] = \frac{E_x(0, \omega_j)}{i\omega_j\mu_0 H_y(-0, \omega_j)} = -\frac{f(0, \omega_j)}{f'(-0, \omega_j)}, \quad (2.3)$$

where $f(z)$ is a solution of Eq. (2.2), with $f'(z) \rightarrow 0$ for $z \rightarrow \infty$. In Eq. (2.3), provision is made for a possible discontinuity of H_y or f' at $z=0$ due to the presence of a thin conducting sheet. The M complex data c_j , which are assumed for the present to be exact, correspond to the functionals $c_j[\sigma]$. Then the problem of interpretation consists in finding at least one model $\sigma(z)$ such that $c_j[\sigma] = c_j$, $j=1, \dots, M$. When there is no risk of confusing functionals and data, $[\sigma]$ is omitted.

Within the class of models fitting the finite data set we are interested in those two models, which minimize and maximize either for a given depth range $z_1 \leq z \leq z_2$ the conductivity average

$$\bar{\sigma}(z_1, z_2) = \frac{1}{z_2 - z_1} \int_{z_1}^{z_2} \sigma(z) dz \quad (2.4a)$$

or simpler for a given $z_2 > 0$ the conductance function

$$S(z_2) = \int_0^{z_2} \sigma(z) dz. \quad (2.4b)$$

These four cases are equivalent to the problem of minimizing the objective function

$$Q[\sigma] = \int_0^{\infty} w(z) \sigma(z) dz \quad (2.5)$$

with the weights

$$w(z) = \begin{cases} 0, & 0 \leq z < z_1 \quad \text{and} \quad z > z_2 \\ 1, & z_1 \leq z \leq z_2 \quad \text{for} \quad Q = \bar{\sigma}_{\min} \\ -1, & z_1 \leq z \leq z_2 \quad \text{for} \quad Q = -\bar{\sigma}_{\max} \end{cases} \quad (2.6a)$$

or

$$w(z) = \begin{cases} 1, & 0 \leq z \leq z_2 \quad \text{for} \quad Q = S_{\min} \\ -1, & 0 \leq z \leq z_2 \quad \text{for} \quad Q = -S_{\max} \\ 0, & z > z_2 \end{cases} \quad (2.6b)$$

Assuming that $\sigma(z)$ is only constrained by the non-negativity condition, we are faced with the nonlinear programming problem of minimizing $Q[\sigma]$ subject to the constraints

$$c_j[\sigma] = c_j, \quad j=1, \dots, M \quad (2.7a)$$

and

$$\sigma(z) \geq 0. \quad (2.7b)$$

The equality constraints (2.7a) render the problem nonlinear. The Lagrange function is

$$L[\sigma] = Q[\sigma] + \text{Re} \sum_{j=1}^M \lambda_j \{c_j[\sigma] - c_j\} - \int_0^{\infty} \mu(z) \sigma(z) dz, \quad (2.8)$$

where the M complex Lagrangian multipliers λ_j enforce the equality constraints, whereas the non-negative function $\mu(z)$ takes the inequality constraint into account and satisfies

$$\mu(z) = 0, \quad \sigma(z) > 0; \quad \mu(z) \geq 0, \quad \sigma(z) = 0 \quad (2.9)$$

(e.g., Pearson, 1974, p. 1113). The formally introduced Lagrangian multipliers λ_j can be interpreted as the sensitivity of the minimum value Q_0 of the objective functional $Q[\sigma]$ to changes in the constraining data $c_j = g_j - ih_j$ (e.g., Pearson, 1974, p. 1118):

$$\text{Re} \lambda_i = -\frac{\partial Q_0}{\partial g_i}, \quad \text{Im} \lambda_j = -\frac{\partial Q_0}{\partial h_j}. \quad (2.10)$$

Hence, the λ_j allow immediately a rough estimate of the influence of data errors. Also $\mu(z)$ can be interpreted as the sensitivity of Q_0 by changing the lower bound of $\sigma(z)$ in the neighborhood of z to a positive value: let $\sigma(\zeta) \geq \tilde{\sigma}_-(\zeta)$ in $z - \Delta z/2 \leq \zeta \leq z + \Delta z/2$ and let

$$\Delta \tau = \int_{z - \Delta z/2}^{z + \Delta z/2} \tilde{\sigma}_-(\zeta) d\zeta.$$

Then we obtain in the limit $\Delta z \rightarrow 0$

$$\mu(z) = \frac{\partial Q_0}{\partial \tau}, \quad (2.11)$$

i.e. Q_0 is not affected, if the constraint is not binding ($\sigma(z) > 0$, $\mu(z) = 0$) and Q_0 will not decrease for a binding constraint ($\sigma(z) = 0$, $\mu(z) \geq 0$). These results, of course, were expected.

As a *necessary* condition for an extremum of Q , the first variation of L with respect to σ has to vanish. This yields

$$w(z) + \operatorname{Re} \sum_{j=1}^M \lambda_j F_j(z) - \mu(z) = 0, \quad (2.12)$$

where $F_j(z)$ is the Fréchet derivative of the functional $c_j[\sigma]$, defined by

$$\delta c_j[\sigma] = \int_0^\infty F_j(z) \delta \sigma(z) dz \quad (2.13)$$

with

$$F_j(z) = -i\omega\mu_0 f_j^2(z) \quad (2.14)$$

(e.g., Parker, 1977), where $f_j(z) := f(z, \omega_j)$ is the solution of Eq. (2.2) with $f_j'(-0) = 1$ and $f_j'(z) \rightarrow 0$ for $z \rightarrow \infty$. The lowest-order differential equations satisfied by $F_j(z)$ are

$$2F_j F_j'' = 4i\omega_j \mu_0 \sigma F_j^2 + (F_j')^2, \quad (2.15a)$$

$$F_j''' = 4i\omega_j \mu_0 \sqrt{\sigma} (\sqrt{\sigma} F_j)'. \quad (2.15b)$$

On using Eq. (2.12), a function $D(z)$ is defined as

$$D(z) := w(z) + \operatorname{Re} \sum_{j=1}^M \lambda_j F_j(z) = \mu(z). \quad (2.16)$$

Now assume that in some interval $a < z < b$, completely inside an interval where $w(z)$ is constant, the conductivity $\sigma(z)$ is positive, i.e., $\sigma(z) > 0$ for $z \in (a, b)$. Then from Eq. (2.9) $\mu(z) = 0$ for $z \in (a, b)$ and Eq. (2.12) reads

$$D(z) = 0, \quad z \in (a, b). \quad (2.17)$$

In Appendix A it is shown that for downward diffusing fields $F_j(z)$ this equation has no solution, except for the trivial solution $\lambda_j = 0$ if $w(z) = 0$ for $z > a$. For $M = 1$ this can be verified easily: since Eq. (2.17) holds in a whole interval, it can be differentiated an arbitrary number of times. In particular, the first two derivatives at $z \in (a, b)$ read (skipping the subscript 1)

$$D' = -2\operatorname{Re} \{2F/c\} = 0,$$

$$D'' = +\operatorname{Re} \{\lambda F(i\omega\mu_0\sigma + 1/c^2)\} = 0$$

with $c = c(z) = -f(z)/f'(z) = g - ih$, $g > 0$, $h > 0$ for a downward diffusing field. The above equations can be considered as two linear homogeneous equations for $\operatorname{Re}(\lambda F)$ and $\operatorname{Im}(\lambda F)$, admitting a nontrivial solution only for a vanishing system determinant Δ . However,

$$|\Delta| = 4(h + \omega\mu_0\sigma g|c|^2)/|c|^4 > 0.$$

The linear independence of the Fréchet derivatives has the important consequence that the extremal models cannot comprise conductivity sections, where $\sigma(z)$ is positive over a *finite* interval. Therefore, the extremal

models consist of a sequence of insulating layers ($\sigma = 0$) and thin conducting sheets, where $\sigma(z)$ is positive only at an isolated point.

The problem of model construction consists in finding for a given impedance set a sequence of thin sheets and a set of complex Lagrangian multipliers such that the model fits the data and the function $D(z)$ is non-negative everywhere and in particular vanishes at the positions of the thin sheets, i.e.,

$$D(z) \geq 0; \quad D(z) = 0 \quad \text{for } \sigma(z) > 0. \quad (2.18)$$

This is clearly a nonlinear problem, since both the data functionals and the Fréchet derivatives depend nonlinearly on the positions and conductances of the thin sheets. Also, it is not yet clear from the outset how many sheets have to be considered. (Semi-empirical rules derived from experience with a small number of frequencies are listed in Sect. 3.6.) Concerning the non-negativity of $D(z)$, closer examination in Sect. 3.3 shows that for $z < z_2$ we have $D(z) > 0$, for $\sigma(z) = 0$, and $D(z) = 0$ for $\sigma(z) > 0$, whereas $D(z)$ vanishes identically below the first sheet occurring in $z > z_2$.

In a more general approach, the incorporation of a priori bounds $\sigma_-(z)$ and $\sigma_+(z)$ such that $\sigma_-(z) \leq \sigma(z) \leq \sigma_+(z)$ can be achieved by replacing the last term in the Lagrange function Eq. (2.8) by

$$-\int_0^\infty [\mu_+(z)\{\sigma_+(z) - \sigma(z)\} + \mu_-(z)\{\sigma(z) - \sigma_-(z)\}] dz,$$

where $\mu_\pm(z) \geq 0$. Then the definition of $D(z)$ in analogy to Eq. (2.16) is

$$D(z) = \mu_-(z) - \mu_+(z). \quad (2.19)$$

$D(z)$ vanishes, if $\sigma(z)$ attains neither its lower nor its upper bound. For the one-frequency case it was shown above that $D(z)$ cannot vanish in a finite interval. Therefore, in this case the extremal models consist of a stack of sections with alternating conductivities $\sigma_-(z)$ and $\sigma_+(z)$, satisfying

$$\left. \begin{array}{l} \sigma(z) = \sigma_+(z) \\ \sigma(z) = \sigma_-(z) \end{array} \right\} \quad \text{for } D(z) \begin{cases} \leq 0 \\ \geq 0 \end{cases}. \quad (2.20)$$

Jumps between σ_- and σ_+ occur, where $D(z)$ changes sign. The generalization to M frequencies must still be done, but it can be anticipated from the approximate extremal models presented by Oldenburg (1983) in his Fig. 7 that also in this case only the conductivities σ_- and σ_+ will occur.

3. Extremal models for the unconstrained conductance function

3.1 Formulas for a series of thin sheets

The discussion is started with the unconstrained extremal models requiring only the assumption $\sigma(z) \geq 0$, and attention is confined to the conductance function in Eq. (1.1)

$$S(z_2) = \int_0^{z_2} \sigma(z) dz \quad (3.1)$$

rather than considering conductivity averages. In the unconstrained case the extremal models reduce to a series of thin sheets. Models of this kind have been identified previously by Parker (1980) as giving the best fit to any real data set.

Formulas for the treatment of a series of thin sheets are briefly summarized. Assume a stack of K sheets with conductance τ_k at depth ζ_k , i.e.,

$$\sigma(z) = \sum_{k=1}^K \tau_k \delta(z - \zeta_k).$$

The inter-sheet separations are $d_k = \zeta_k - \zeta_{k-1}$, $k = 2, \dots, K$. Then the solution $f(z)$ of (2.2) varies linearly between two sheets, is continuous at $z = \zeta_k$, but shows the discontinuous slope

$$f'(\zeta_k + 0) - f'(\zeta_k - 0) = i\omega\mu_0\tau_k f(\zeta_k), \quad k = 1, \dots, K \quad (3.2)$$

$f(z)$ is constant in $z \geq \zeta_K$, i.e. $f'(\zeta_K + 0) = 0$. The theoretical transfer function $c[\sigma] = -f(0)/f'(-0)$ is obtained as $c[\sigma] = c_1 + \zeta_1$ with $c_k := -f(\zeta_k)/f'(\zeta_k - 0)$ recursively from

$$c_k = \frac{c_{k+1} + d_{k+1}}{1 + i\omega\mu_0\tau_k(c_{k+1} + d_{k+1})}, \quad k = K-1, \dots, 1, \quad (3.3)$$

starting with $c_K = 1/(i\omega\mu_0\tau_K)$. (In this section the subscript k on c denotes the value of c at $z = \zeta_k - 0$, whereas in a different context the subscript j specifies the particular frequency ω_j of c at $z = -0$.) The value of f at $z = \zeta_k$, normalized to $f'(-0) = 1$ is

$$f(\zeta_1) = -c_1, \quad f(\zeta_k) = f(\zeta_1) \prod_{n=2}^k \frac{c_n}{c_n + d_n}, \quad k > 1, \quad (3.4)$$

and the values of $f'(\zeta_k - 0) = f'(\zeta_{k-1} + 0)$ are determined from

$$f'(\zeta_1 - 0) = 1, \quad f'(\zeta_k - 0) = \prod_{n=1}^{k-1} (1 - i\omega\mu_0\tau_n c_n), \quad k > 1. \quad (3.5)$$

The partial derivatives of c with respect to the model parameters τ_k and ζ_k are obtained by means of Eqs. (2.13) and (2.14):

$$\frac{\partial c[\sigma]}{\partial \tau_k} = F(\zeta_k) = -i\omega\mu_0 f^2(\zeta_k), \quad (3.6a)$$

$$\frac{\partial c[\sigma]}{\partial \zeta_k} = \tau_k \bar{F}'(\zeta_k) = [f'(\zeta_k - 0)]^2 - [f'(\zeta_k + 0)]^2, \quad k = 1, \dots, K-1 \quad (3.6b)$$

$$\frac{\partial c[\sigma]}{\partial \zeta_K} = \frac{1}{2} \tau_K F'(\zeta_K - 0) = [f'(\zeta_K - 0)]^2, \quad (3.6c)$$

where

$$\bar{F}'(\zeta_k) = \frac{1}{2} [F'(\zeta_k - 0) + F'(\zeta_k + 0)]. \quad (3.7)$$

3.2 An impedance representation theorem

The theoretical transfer function of a 1D conductivity distribution admits the spectral expansion

$$c[\sigma] = a_0 + \int_0^\infty \frac{a(\lambda)d\lambda}{\lambda + i\omega}, \quad a_0 \geq 0, \quad a(\lambda) \geq 0, \quad (3.8)$$

where $a(\lambda)$ is a generalized function to include both the discrete and continuous part of the spectrum (Weidelt, 1972; Parker, 1980; Parker and Whaler, 1981). As an example, the stack of thin sheets considered in the previous section has a finite discrete spectrum and leads to the representation

$$c[\sigma] = \zeta_1 + \sum_{k=1}^{K-1} \frac{a_k}{b_k + i\omega} + \frac{a_K}{i\omega}, \quad a_k, b_k > 0; \quad \zeta_1, a_K \geq 0. \quad (3.9)$$

Let the theoretical impedances $c_j[\sigma]$ for M distinct frequencies ω_j be given by

$$c_j[\sigma] = a_0 + \sum_{n=1}^N \frac{a_n}{b_n + i\omega_j} \quad (3.10)$$

with $a_n, b_n > 0$, $n = 1, \dots, N-1$; $a_0, b_N \geq 0$, where the b_n , $n = 1, \dots, N$ are distinct. [This form rather than Eq. (3.8) is chosen for ease of presentation only, for equivalent integral analogs based on (3.8) exist for all formulas involving the summation over n .] Then for $N \geq M$ $c_j[\sigma]$ allows the two representations

$$I: c_j[\sigma] = \sum_{m=1}^M \frac{A_m}{B_m + i\omega_j}, \quad (3.11a)$$

$$II: c_j[\sigma] = \bar{A}_0 + \sum_{m=1}^{M-1} \frac{\bar{A}_m}{\bar{B}_m + i\omega_j} + \frac{\bar{A}_M}{i\omega_j}, \quad (3.11b)$$

where all constants $A_m, B_m, \bar{A}_0, \bar{A}_m, \bar{B}_m$, and \bar{A}_M are positive. For $N < M$ there is no representation other than Eq. (3.10). [See also Parker (1980).] The proofs are given in Appendix B.

Model I is a series of $M+1$ thin sheets: the first sheet is at $z=0$ and the last sheet has an infinite conductance. Complementary model II consists of M thin sheets: the first sheet lies at $z = \bar{A}_0 > 0$ and the last sheet has a bounded conductance. Between models I and II is a duality relationship in the sense that they can be transformed into each other by replacing a sheet of finite (infinite) conductance by an insulating layer of finite (infinite) thickness, and vice versa.

Assuming the ordering $b_n > b_{n+1}$, $\bar{B}_m > \bar{B}_{m+1}$, we infer from (B-19, 20), (B-24), and (B-26) the inequalities

$$\sum_{m=1}^M (A_m/B_m) \leq a_0 + \sum_{n=1}^N (a_n/b_n), \quad (3.12a)$$

$$\sum_{m=1}^M A_m \leq \sum_{n=1}^N a_n \quad (a_0 = 0), \quad (3.12b)$$

$$\bar{A}_0 \geq a_0, \quad (3.12c)$$

$$\bar{A}_M \geq a_N \quad (b_N = 0). \quad (3.12d)$$

These inequalities express the extremal properties of models I and II, with the following physical interpretation:

a) The depth of a perfect conductor (if present) is given by

$$z_\infty = \lim_{\omega \rightarrow 0} c(\omega). \quad (3.13a)$$

Hence Eq. (3.12a) implies that I is the model with the *shallowest* perfect conductor consistent with the data. Its position marks “the limiting depth below which nothing can be learnt about the conductivity” (Parker, 1982).

b) The conductance of a surface sheet, the presence of which requires $a_0=0$, is given by

$$S(+0) = \lim_{\omega \rightarrow \infty} \frac{1}{i\omega\mu_0 c(\omega)}. \quad (3.13b)$$

Hence Eq. (3.12b) says that model I has the *greatest* surface conductance.

c) The depth to the first conductor is

$$z_0 = \lim_{\omega \rightarrow \infty} c(\omega). \quad (3.13c)$$

Therefore according to Eq. (3.12c), model II is the conductivity structure with the *deepest* top.

d) The total conductance in the absence of a perfect conductor ($b_N=0$) is

$$S(\infty) = \lim_{\omega \rightarrow 0} \frac{1}{i\omega\mu_0 c(\omega)}. \quad (3.13d)$$

Hence, Eq. (3.12d) implies that model II has the *least* total conductance of all models fitting the data.

The presentation of Eqs. (3.11a, b) also clarifies a situation described by Parker (1980): the best fit to a set of M complex measured impedances by an expansion (3.10) is obtained when using quadratic programming. This yields a representation with at most $N=2M$ terms. In the case $N>M$ (which may occur at least for small M), the expansion (3.10) would require more constants than available independent data. The above representation theorem overcomes this unsatisfactory situation by compressing Eqs. (3.10) to (3.11a) or (3.11b) with exactly $2M$ constants. If the quadratic programming procedure requires $N \geq M$ terms, the best fitting model is nonunique.

For a given representation (3.10), the condensed versions (3.11a, b) can be obtained as follows: for model I, one first either has to solve the M -dimensional nonlinear system (B-13) for the M constants B_m , $m=1, \dots, M$, or one determines B_m alternatively as the M roots of the polynomial (B-9) with $K=M$, $q_k^{(0)} \equiv q_k$, and $q_M=1$, where the coefficients q_k , $k=1, \dots, M-1$ are the solution of the linear system (B-8b). With the knowledge of B_m the constants A_m (or G_m) are immediately obtained from (B-16) or (B-18). Similarly, for model II the constants \bar{B}_m , $m=1, \dots, M-1$, are obtained, for instance, as the solution of the first $M-1$ equations of the system (B-22) (note $b_1 \geq \bar{B}_m \geq b_N$); A_0 then follows from the M -th equation of (B-22) or from (B-24), and finally \bar{A}_m (or \bar{G}_m) is given by (B-25, 26). For the solution of the various systems of nonlinear equations considered in this paper, Brown's algorithm turned out to be extremely useful, as it does not require the provision of partial derivatives. A FORTRAN program is published in Brown (1973). After the determination of the constants in Eqs. (3.11a, b), the parameters of the conductivity models are found on using the stable Rutishauser algorithm as described by Parker and Whaler (1981).

The above considerations refer to theoretical or synthetic data, for which the existence of representations (3.8) or (3.10) is granted. As mentioned above, real data can be approximated by an expansion (3.10) by quadratic programming. However, the construction of extremal models for the “cleaned” data is meaningful only for $N \geq M$, since for $N < M$ those modified data can be interpreted by only one model.

3.3 General properties of the extremal models for $S(z_2)$

In the absence of a priori constraints, the extremal models for $S(z_2)$ [Eq. (3.1)] consist of a sequence of thin sheets. The necessary extremal conditions (2.18) provide no information about the required number of sheets as a function of both z_2 and the number M of frequencies. However, the extremal properties of models I and II considered in the previous section give the first hint. Let $S_{\max}(z_2)$ denote the upper bound of $S(z_2)$ and let $\bar{z}_\infty = \min z_\infty = \sum (A_m/B_m)$. Then the extremal models for both $S_{\max}(+0)$ and $S_{\max}(\bar{z}_\infty)$ are given by the $(M+1)$ -sheet model I [extremal properties a) and b)]. Similarly, let $S_{\min}(z_2)$ signify the lower bound of $S(z_2)$ and $\bar{z}_0 = \max z_0 = A_0$. Then $S_{\min}(z_2)=0$ for $z_2 \leq \bar{z}_0$ and the M -sheet model II is the extremal for both $S_{\min}(\bar{z}_0)$ and $S_{\min}(\infty)$ [extremal properties c) and d)]. Thus, $S_{\max}(z_2)$ and $S_{\min}(z_2)$ have models I and II, respectively, both at the beginning and end. Although it has not been proved generally, numerical experiments have shown that $S_{\max}(S_{\min})$ returns to model I (II) also at intermediate depths $z_2 = \zeta_k + 0$ ($z_2 = \zeta_k - 0$), where ζ_k is the position of a thin sheet in model I (II). Hence these models form the backbone in the evolution of the extremal models for S_{\max} and S_{\min} as a function of z_2 .

To study the continuous deformation of the extremal models for varying z_2 , the necessary condition (2.18) is discussed in more detail. In the unconstrained case the sections with $\sigma(z) > 0$, implying $D(z)=0$, shrink to one point. This means that $D(z)$ must be either identically zero in the neighborhood of this point or it must show a “double zero” at the position $z=\zeta_k$ of a thin sheet *not coinciding with* $z=z_2$. Since $D(z)$ may have a discontinuous slope at $z=\zeta_k$, the necessary conditions demand

$$D(\zeta_k)=0, \quad D'(\zeta_k-0)+D'(\zeta_k+0)=0, \quad (3.14a, b)$$

or using (2.16) and (3.7)

$$w(\zeta_k) + \operatorname{Re} \sum_{j=1}^M \lambda_j F_j(\zeta_k) = 0, \quad (3.15a)$$

$$\operatorname{Re} \sum_{j=1}^M \lambda_j \bar{F}_j(\zeta_k) = 0 \quad (3.15b)$$

with $w(z)$ defined in (2.6b). Multiplying (3.15b) by τ_k , the conditions (3.15a, b) are, on account of (3.6a-c), equivalent to

$$\operatorname{Re} \sum_{j=1}^M \lambda_j \frac{\partial c_j[\sigma]}{\partial \tau_k} = -w(\zeta_k), \quad (3.16a)$$

$$\operatorname{Re} \sum_{j=1}^M \lambda_j \frac{\partial c_j[\sigma]}{\partial \zeta_k} = 0. \quad (3.16b)$$

Assuming a model with K thin sheets, the condition (3.16b) also holds for an ultimate perfect conductor $\tau_K = \infty$ at $z = \zeta_K$, where $F_j(\zeta_K) = 0 = F'_j(\zeta_K - 0)$, but

$$\lim_{\tau_K \rightarrow \infty} \frac{1}{2} \tau_K F'_j(\zeta_K - 0) = \frac{\partial c_j[\sigma]}{\partial \zeta_K} = [f'_j(\zeta_K - 0)]^2 \neq 0.$$

Now it is shown that the extremal model for $S_{\max}(z_2)$ always has a conducting sheet at $z_2 - 0$, which is just included in the range of integration, whereas $S_{\min}(z_2)$ has a conducting sheet at $z_2 + 0$, which is just excluded. It is assumed that z_2 lies between sheet $p-1$ and sheet p , i.e., $\zeta_{p-1} < z_2 < \zeta_p$. In between, sheets $D(z)$ can vary at most as a second degree polynomial since $f_j(z)$ is a linear function in z . The discontinuity of $w(z)$ between ζ_{p-1} and ζ_p introduces in this range in addition a discontinuity in $D(z)$, but not in $D'(z)$:

$$D(z_2 + 0) - D(z_2 - 0) = \begin{cases} +1 & \text{for } S_{\max}(z_2) \\ -1 & \text{for } S_{\min}(z_2) \end{cases} \quad (3.17)$$

$$D'(z_2 + 0) - D'(z_2 - 0) = 0. \quad (3.18)$$

First, it is shown that the conditions (3.14a, b) in connection with the quadratic variation of $D(z)$ lead to the conclusion that $D(z) \equiv 0$ for $z \geq \zeta_p$. From $f_j(z) = \text{const.}$ for $z \geq \zeta_K$ and $D(\zeta_K) = 0$ follows $D(z) \equiv 0$ for $z \geq \zeta_K$. With $D'(\zeta_K + 0) = 0$ it is inferred from Eq. (3.14b) that also $D'(\zeta_K - 0) = 0$. Hence for $p < K$ we have $D(z) = A_K(z - \zeta_K)^2$ in $\zeta_{K-1} \leq z \leq \zeta_K$, where $D(z) \geq 0$ implies $A_K \geq 0$. However, $A_K = 0$, since $D(\zeta_{K-1}) = 0$ on account of Eq. (3.14a). Therefore, $D(z) \equiv 0$ in $\zeta_{K-1} \leq z \leq \zeta_K$. Repeating the arguments, it is found that $D(z) \equiv 0$ for $z \geq \zeta_p$. In $\zeta_p > z \geq z_2 + 0$ we again have $D(z) = A_p(z - \zeta_p)^2$, $A_p \geq 0$. With reference to Eq. (3.18), $D'(z)$ is a continuous linear function in $\zeta_{p-1} \leq z \leq \zeta_p$ with $D'(\zeta_p - 0) = 0$, $D''(\zeta_p - 0) \geq 0$, implying $D'(\zeta_{p-1} + 0) \leq 0$. On the other hand, if $\zeta_{p-1} + 0 < z_2 < \zeta_p - 0$, then the two conditions $D(\zeta_{p-1}) = 0$ and $D(z) \geq 0$ would require $D'(\zeta_{p-1} + 0) \geq 0$. This does not contradict the preceding result only if $D'(\zeta_{p-1} + 0) = 0$, implying $A_p = 0$ and $D(z) \equiv 0$ in $\zeta_{p-1} \leq z \leq \zeta_p$ [because of the continuity of $D'(z)$]. The condition (3.17), however, requires that $D(z) \neq 0$ in this interval. The remedy is to take either $z_2 = \zeta_{p-1} + 0$ or $z_2 = \zeta_p - 0$. In this case a thin sheet lies at the preassigned depth $z_2 \pm 0$, which is no longer a freely variable parameter and hence is not to be included in the necessary conditions (3.14b)–(3.16b). In view of Eq. (3.17), the case $z_2 = \zeta_{p-1} + 0$ applies to S_{\max} with

$$D(z_2 - 0) = D(\zeta_{p-1}) = 0, \quad D(z_2 + 0) = 1, \quad (3.19a)$$

and the case $z_2 = \zeta_p - 0$ pertains to S_{\min} :

$$D(z_2 + 0) = D(\zeta_p) = 0, \quad D(z_2 - 0) = 1. \quad (3.19b)$$

Since the evolution of the extremal models for varying z_2 is quite involved in the general case, it will be studied in detail only by two relatively simple examples for one and two frequencies.

3.4 Extremal models for one frequency

The data for the frequency $\omega_1 =: \omega$ is $c_1 =: c$ with

$$c = g - ih = |c|e^{-i\psi}, \quad (3.20)$$

where $|c|$ is related to the apparent resistivity ρ_a and apparent conductivity σ_a by

$$\rho_a = \omega \mu_0 |c|^2, \quad \sigma_a = 1/\rho_a, \quad (3.21)$$

and ψ is the complement to the phase φ between the electric and magnetic field, $\varphi + \psi = 90^\circ$. The data is consistent with a 1D model if $g \geq 0$, $h \geq 0$, or equivalently $0 \leq \psi \leq 90^\circ$. The results of the last section have already shed some light on the structure of the extremal models. The associated models I and II [Eq. (3.11a, b)] using the notation of Sects. 3.1 and 3.2 are

$$\text{I: } c[\sigma] = \frac{A_1}{B_1 + i\omega} = \frac{\zeta_2}{1 + i\omega \mu_0 \tau_1 \zeta_2}, \quad (3.22a)$$

$$\text{II: } c[\sigma] = \bar{A}_0 + \frac{\bar{A}_1}{i\omega} = \zeta_1 + \frac{1}{i\omega \mu_0 \tau_1}. \quad (3.22b)$$

Equating with Eq. (3.20), it is found that model I consists of a thin sheet with conductance $\tau_1 = |c| \sigma_a \sin \psi$ at $z = 0$ and a second sheet with $\tau_2 = \infty$ at $z_\infty = |c| \sec \psi$, whereas model II consists of a single sheet $\tau_1 = |c| \sigma_a / \sin \psi$ at $z_0 = |c| \cos \psi$. Hence,

$$S_{\max}(+0) = |c| \sigma_a \sin \psi, \quad S_{\max}(z_2) = \infty, \quad z_2 \geq |c| \sec \psi, \quad (3.23a)$$

$$S_{\min}(\infty) = |c| \sigma_a / \sin \psi, \quad S_{\min}(z_2) = 0, \quad z_2 \leq |c| \cos \psi. \quad (3.23b)$$

The models of the start and end are identical, with the difference, however, that at the end the (lower) sheet is included in S . Now the evolution of the models between these limits has to be studied. Starting with S_{\max} , it is assumed that the extremal model for $z_2 > 0$ evolves from that of $z_2 = +0$ by moving the surface sheet to $\zeta_1 = z_2 - 0 > 0$. Then also the second sheet with $\tau_2 = \infty$ will move. The two free parameters of the model are τ_1 and ζ_2 , whereas ζ_1 and τ_2 are preassigned. From Eqs. (2.7a) and (3.16a, b) results the system

$$c[\sigma] = c, \quad (3.24a)$$

$$\text{Re} \left\{ \lambda \frac{\partial c[\sigma]}{\partial \tau_1} \right\} = 1, \quad (3.24b)$$

$$\text{Re} \left\{ \lambda \frac{\partial c[\sigma]}{\partial \zeta_2} \right\} = 0 \quad (3.24c)$$

($\lambda_1 =: \lambda$), where the partial derivatives of $c[\sigma]$ are determined from (3.6a–c). The system (3.24a–c) consists of four real equations for the four real unknowns τ_1 , ζ_2 , $\text{Re } \lambda$, and $\text{Im } \lambda$. In practice, first Eq. (3.24a) is solved for τ_1 and ζ_2 to obtain $S_{\max}(z_2) = \tau_1$, and then Eq. (3.34b, c) is solved for the sensitivity measure λ , which is required, for instance, in the discriminant function (2.16)

$$D(z) = w(z) + \text{Re} \{ \lambda F(z) \}$$

with $F(z) := F_1(z)$, as defined in Eq. (2.14). This non-negative function has to vanish at $z = \zeta_1$ and $z = \zeta_2$, where $D(\zeta_1) = 0$ is equivalent to Eq. (3.24b) and $D(\zeta_2) = 0$ is satisfied because of $F(\zeta_2) = 0$. With the present two-sheet model the condition $D(z) \geq 0$ can be satisfied only for sufficiently small z_2 . This is a consequence of $D(\zeta_1) = D(z_2 - 0) = 0$ and the fact that $D(z)$ is a second-degree polynomial in $0 \leq z \leq z_2$, which for some limiting

value $z_2 = z_a$ vanishes at $z=0$ and thus signals the emergence of a third sheet at $z=0$. To determine z_a , the system (3.24a-c) is augmented by the condition $D(0) = 0$. This condition and (3.24b, c) are linear in the two unknowns $\text{Re } \lambda$ and $\text{Im } \lambda$ and are compatible only if linearly dependent. Using $F(\zeta_1) = \partial c / \partial \tau_1$, the compatibility condition is

$$\text{Im} \left\{ [F(\zeta_1) - F(0)] / \frac{\partial c[\sigma]}{\partial \zeta_2} \right\} = 0.$$

Expressing F and $c[\sigma]$ in terms of τ_1 , $\zeta_1 (= z_2 - 0)$, and ζ_2 , this condition reads

$$(\zeta_1 - \zeta_2)^2 (1 + \omega^2 \mu_0^2 \tau_1^2 \zeta_1^2) = \zeta_2^2,$$

and writing τ_1 and ζ_2 in terms of the data, we end up with a cubic equation for $z_a = \zeta_1$ as limiting value of z_2 :

$$z_a^3 - 4g z_a^2 + (2g^2 + 3|c|^2) z_a - 2g|c|^2 = 0. \quad (3.25)$$

Of interest is the root z_a in $0 \leq z_a \leq g = |c| \cos \psi$.

Renumbering the sheet parameters in the three-sheet problem, the determination of $S_{\max}(z_2) = \tau_1 + \tau_2$ for $z_2 > z_a$ requires the determination of τ_1 , τ_2 , and ζ_3 from the set of equations

$$c[\sigma] = c, \quad (3.26a)$$

$$\text{Re} \left\{ \lambda \frac{\partial c[\sigma]}{\partial \tau_1} \right\} = 1, \quad (3.26b)$$

$$\text{Re} \left\{ \lambda \frac{\partial c[\sigma]}{\partial \tau_2} \right\} = 1, \quad (3.26c)$$

$$\text{Re} \left\{ \lambda \frac{\partial c[\sigma]}{\partial \zeta_3} \right\} = 0. \quad (3.26d)$$

In practice, first τ_1 , τ_2 , and ζ_3 are computed from (3.26a) and the compatibility condition

$$\text{Im} \left\{ \left[\frac{\partial c[\sigma]}{\partial \tau_2} - \frac{\partial c[\sigma]}{\partial \tau_1} \right] / \frac{\partial c[\sigma]}{\partial \zeta_3} \right\} = 0,$$

and λ is obtained in turn from two of the three equations (3.26b-d).

With the emergence of the third sheet, the perfectly conducting ultimate sheet, which was first moving downwards, is moving upwards again and merges for $z_2 \rightarrow z_\infty = |c| \sec \psi$ with the sheet at $z_2 = 0$, which is getting increasingly conductive. The conductance of the surface sheet has meantime increased to $|c| \sigma_a \sin \psi$, and we return for $z_2 = z_\infty$ to the conductivity model for $z_2 = +0$.

The extremal models for $S_{\min}(z_2)$ evolve similarly. As mentioned earlier, $S_{\min}(z_2) = 0$ for $z_2 \leq g = |c| \cos \psi$. For $z_2 > g$ we try a model consisting of a surface sheet of conductance τ_1 at $z = \zeta_1 = 0$ and a second sheet of conductance τ_2 at $z = \zeta_2 = z_2 + 0$. The unknowns are τ_1 and τ_2 . The system of equations analogous to (3.24a-c) is

$$c[\sigma] = c, \quad (3.27a)$$

$$\text{Re} \left\{ \lambda \frac{\partial c[\sigma]}{\partial \tau_1} \right\} = -1, \quad (3.27b)$$

$$\text{Re} \left\{ \lambda \frac{\partial c[\sigma]}{\partial \tau_2} \right\} = 0. \quad (3.27c)$$

Again, Eq. (3.27a) is sufficient to determine τ_1 and τ_2 . The two-sheet model with the fixed surface sheet has to be modified, when ζ_1 can also be considered as a variable, i.e. if, in addition to Eq. (3.27a-c), the following holds

$$\text{Re} \left\{ \lambda \frac{\partial c[\sigma]}{\partial \zeta_1} \Big|_{\zeta_1=0} \right\} = 0. \quad (3.27d)$$

The augmented system (3.27a-d) allows the determination of the limiting value $z_2 = z_i$, where the surface sheet starts moving. The condition of linear dependence of (3.27b-d) leads to

$$\text{Im} \left\{ \frac{\partial c[\sigma]}{\partial \tau_2} / \frac{\partial c[\sigma]}{\partial \zeta_1} \Big|_{\zeta_1=0} \right\} = 0,$$

or

$$\tau_1 + 2\tau_2 = \omega^2 \mu_0^2 \tau_1 \tau_2^2 \zeta_2^2.$$

The expression of τ_1 and τ_2 in terms of the data then yields again a cubic equation for $\zeta_2 = z_i$:

$$4g^2 z_i^3 - 4g(|c|^2 - g^2) z_i^2 + |c|^2 (|c|^2 - 4g^2) z_i - |c|^4 g = 0. \quad (3.28)$$

Of interest is the solution $z_i \geq g$ with $z_i \rightarrow |c|^2 / (2g)$ for $z_i \rightarrow \infty$. For $z_2 > z_i$ the system (3.27a-c) is augmented by

$$\text{Re} \left\{ \lambda \frac{\partial c[\sigma]}{\partial \zeta_1} \right\} = 0$$

and the unknowns τ_1 , τ_2 , and ζ_1 are determined from Eq. (3.27a) and

$$\text{Im} \left\{ \frac{\partial c[\sigma]}{\partial \tau_2} / \frac{\partial c[\sigma]}{\partial \zeta_1} \right\} = 0.$$

The surface sheet, starting moving for $z_2 = z_i$, approaches $z = g$ for $z_2 \rightarrow \infty$ and attains the conductance $\tau_1 = |c| \sigma_a / \sin \psi$. In this limit the lower sheet disappears at infinity, and we return to the starting one-sheet model.

Figure 1 shows $z_a(\psi)$, $z_i(\psi)$, and $z_\infty(\psi)$. It also sketches the type of extremal models for the various ranges of z_2 . The depth $z^* = g = |c| \cos \psi$ is the "center of gravity" of the induced currents. Unless there is only one sheet, conductors must be present both above and below z^* . The extremal models are not unique for S_{\max} in the range $z_2 > z_\infty$ ($S_{\max} = \infty$) and for S_{\min} in the range $0 \leq z_2 < z_0 = g$ ($S_{\min} = 0$). The position of conducting sheets for $\psi = 45^\circ$ is depicted in the left part of Fig. 6 as function of z_2 .

Figure 2 displays the curves bounding $S(z_2)$ for different phases ψ . S is normalized by the conductance $|c| \sigma_a = 1 / (\omega \mu_0 |c|)$ [Eq. (3.21)]. All models compatible with the two-data problem must fall into the shaded areas. These areas get narrow both for $\psi \rightarrow 0^\circ$ and $\psi \rightarrow 90^\circ$, reflecting the fact that these two limiting phases can be interpreted by one model only, consisting, respectively, of a perfect conductor at $z = |c|$ and a thin sheet of conductance $|c| \sigma_a$ at $z = 0$. The limiting values of S_{\max} and S_{\min} as a function of ψ are given in Eq. (3.23a, b).

As a particular feature of Fig. 2, note that the shaded areas for ψ and $90^\circ - \psi$ are mirror images obtained

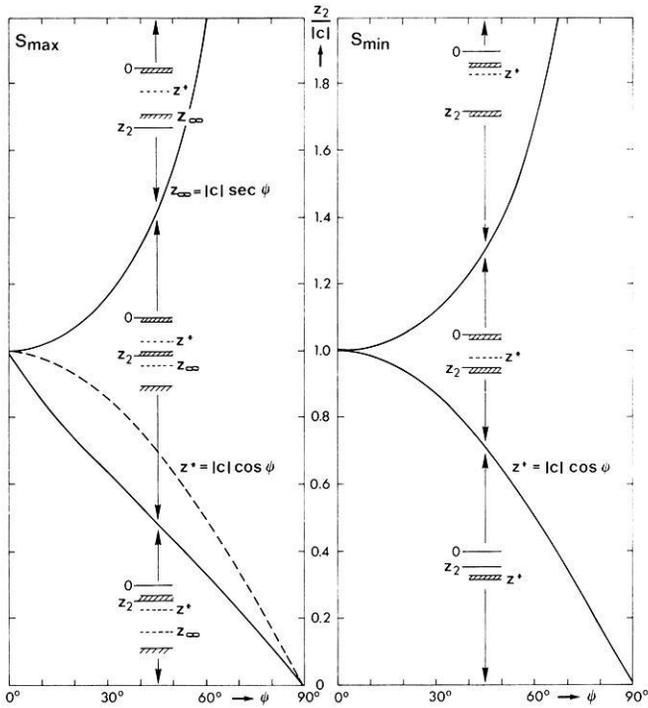


Fig. 1. The different types of extremal models in the unconstrained one-frequency case as function of the phase ψ . In S_{\max} a surface sheet emerges for $z_2 = z_a(\psi)$ (lower solid curve at left) and in S_{\min} the surface sheet is detached for $z_2 = z_i(\psi)$ (upper curve at right)

by reflection at the main diagonal. This is an expression of the duality transformation (Weidelt, 1972), which transforms Eq. (2.2) by means of

$$\bar{z} := \int_0^z \sigma(\zeta) / \sigma_0 d\zeta, \quad \bar{\sigma}(\bar{z}) := \sigma_0^2 / \sigma(z), \quad (3.29a, b)$$

$$\bar{f}(\bar{z}) := f'(z) \quad (3.29c)$$

into

$$\bar{f}''(\bar{z}) = i\omega\mu_0\bar{\sigma}(\bar{z})\bar{f}(\bar{z}) \quad (3.30)$$

with

$$\bar{c}(\omega) = -\bar{f}'(0) / \bar{f}(0) = 1 / [i\omega\mu_0\sigma_0 c(\omega)], \quad (3.31)$$

where $\sigma_0 > 0$ is an arbitrary reference conductivity. Equations (3.29a, b) and (3.31) imply

$$S(z) = \sigma_0 \bar{z}, \quad \bar{S}(\bar{z}) = \sigma_0 z, \quad \psi + \bar{\psi} = 90^\circ$$

and admit the following interpretation. When determining for given z_2 and ψ the bounds $S_{\max}(z_2)$ and $S_{\min}(z_2)$, one automatically also solves the dual problem consisting in determining for $\bar{\psi} = 90^\circ - \psi$ and $\bar{S}_2 = \sigma_0 z_2$ the depth bounds $\bar{z}_{\max}(\bar{S}_2)$ and $\bar{z}_{\min}(\bar{S}_2)$, to which $\bar{\sigma}(\bar{z})$ has to be integrated to reach the given conductance \bar{S}_2 under the most adverse and favorable conditions. The congruence of the shaded areas results from the particular scaling on exploiting $|c|\sigma_a = |\bar{c}|\sigma_0$, $|\bar{c}|\bar{\sigma}_a = |c|\sigma_0$ [Eq. (3.31)]. The symmetry between conductance and depth breaks down, if a priori constraints are imposed on conductivity (cf. Fig. 7).

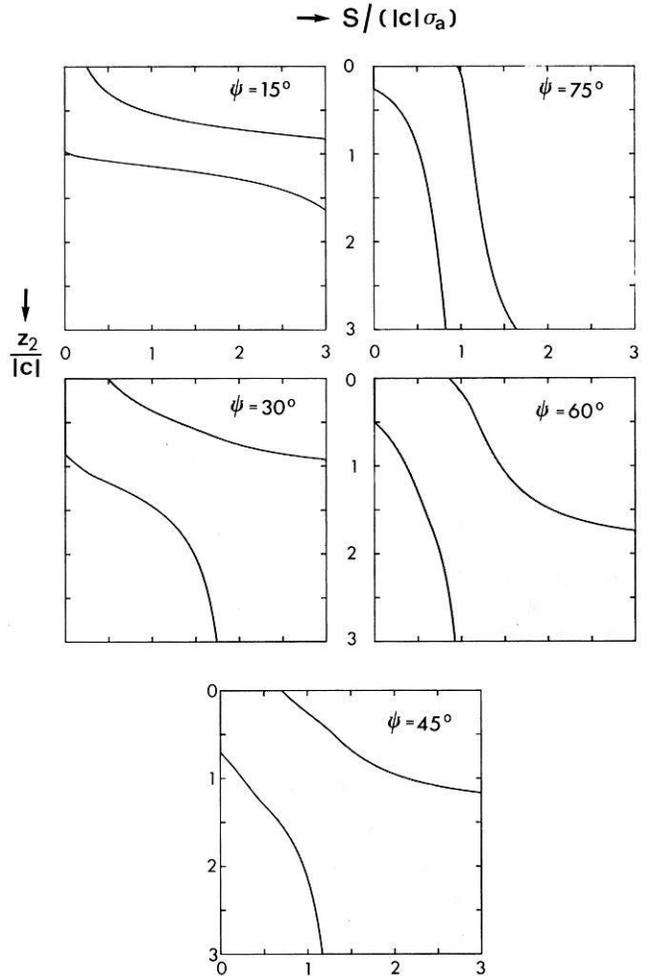


Fig. 2. The bounds for $S(z_2)$ in the unconstrained one-frequency case for different phases ψ . The mirror symmetry of the shaded domains for ψ and $90^\circ - \psi$ is explained in the text

3.5 Extremal models for two frequencies

For more than one frequency various different situations may occur in the evolution of the extremal models, and the general discussion becomes cumbersome (or even impossible). For this reason, attention is confined in this section to the exemplary study of a real data set for $M=2$ frequencies, consisting of the estimates of the transfer function c for the first and fourth Sq harmonic for Europe as given by Schmucker (1984):

$$1 \text{ cpd: } c_1 = (575 - i260) \text{ km}$$

$$4 \text{ cpd: } c_2 = (290 - i275) \text{ km.}$$

By quadratic programming on using the program NNLS of Lawson and Hanson (1974) [as suggested by Parker (1980)], it is found that the data can be represented *exactly* by the series (3.10) with the maximum number of $N=2M=4$ terms. The condensation of this series according to Eq. (3.11a, b) then yields the two canonical models

$$\begin{aligned} \zeta_1 &= 0, & \tau_1 &= 4.09466 \cdot 10^3 \text{ S} \\ \text{I: } \zeta_2 &= 5.23675 \cdot 10^5 \text{ m}, & \tau_2 &= 3.77586 \cdot 10^4 \text{ S} \\ & \zeta_3 = 7.83023 \cdot 10^5 \text{ m}, & \tau_3 &= \infty \\ \text{II: } \zeta_1 &= 1.05947 \cdot 10^5 \text{ m}, & \tau_1 &= 6.66233 \cdot 10^3 \text{ S} \\ & \zeta_2 = 6.99673 \cdot 10^5 \text{ m}, & \tau_2 &= 1.01835 \cdot 10^5 \text{ S} \end{aligned}$$

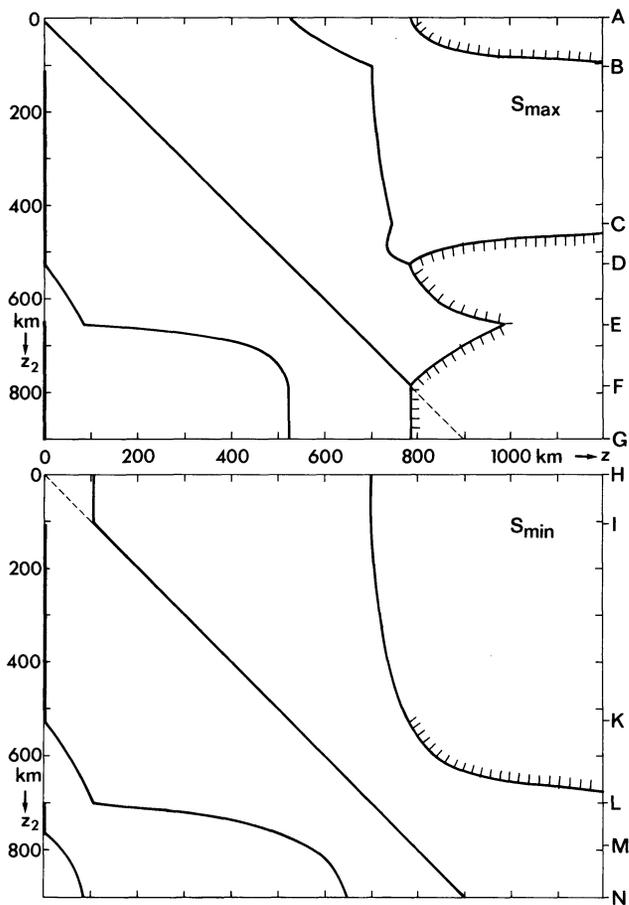


Fig. 3. The position z of conducting sheets as function of z_2 in the extremal models for the unconstrained two-frequency problem (first and fourth S_q harmonic)

(giving for a possible check of the algorithm three to four more digits than warranted by the accuracy of the data). $S_{max}(z_2)$ has model I not only at the start and end, but returns to it also at the intermediate depth $z_2 \approx 524$ km. Similarly, $S_{min}(z_2)$ also returns to model II

at $z_2 \approx 700$ km. The position z of conducting sheets (abscissa) as a function of z_2 (ordinate) is shown in Fig. 3 for S_{max} and S_{min} . The ticks mark perfectly conducting sheets. This figure provides an idea of the nonlinearities in the problem by showing emerging, coalescing, disappearing, and reappearing sheets.

The continuous deformation of the start model to an identical end model is performed, on the one hand, by the sheets emerging at $z=0$ (level B, E, I, and L) and getting detached at D, K, and M, and on the other by the lower sheets either coalescing at D and F with the ultimate perfect conductor (S_{max}), or moving at L and below N to infinity (S_{min}).

Model I is realized at A, D, K, and below F, whereas model II occurs at B, I, L, and below N for $z_2 \rightarrow \infty$. The models for S_{max} in the ranges BC and DE are identical with those of S_{min} in the same range of z_2 ; S_{max} and S_{min} differ only by the conductance of the sheet at z_2 . In the missing range CD all necessary conditions imposed on S_{max} can be satisfied on taking the same model as for S_{min} , but it turns out that choosing the four-sheet model for S_{max} with the reappearing perfect conductor produces in this range slightly greater values. This underlines the fact that the conditions derived in Sect. 2 are necessary but not sufficient.

Figure 4 displays the resulting bounds for the two frequencies, along with the bounds obtained by each frequency separately. In the present example, joint consideration of the two frequencies improves spectacularly the lower bound on S , whereas the upper bound curve deviates only slightly from the curve consisting of the smaller of the values of S_{max} obtained by single-frequency interpretation. At level $z_2 = \zeta_2$ of model II, the values of S_{min} for 1 cpd and 1 & 4 cpd do not appear to be strictly equal, but differ only by 0.01%. The marks at the right margin indicate the depth of the shallowest perfect conductor, moving from 551 km for 4 cpd to 692 km for 1 cpd and to 783 km for 1 & 4 cpd.

Finally, it is noted that at the distinguished levels $z_2 \approx 106, 524,$ and 700 km, where the slope of the bounding curves changes discontinuously, the Lag-

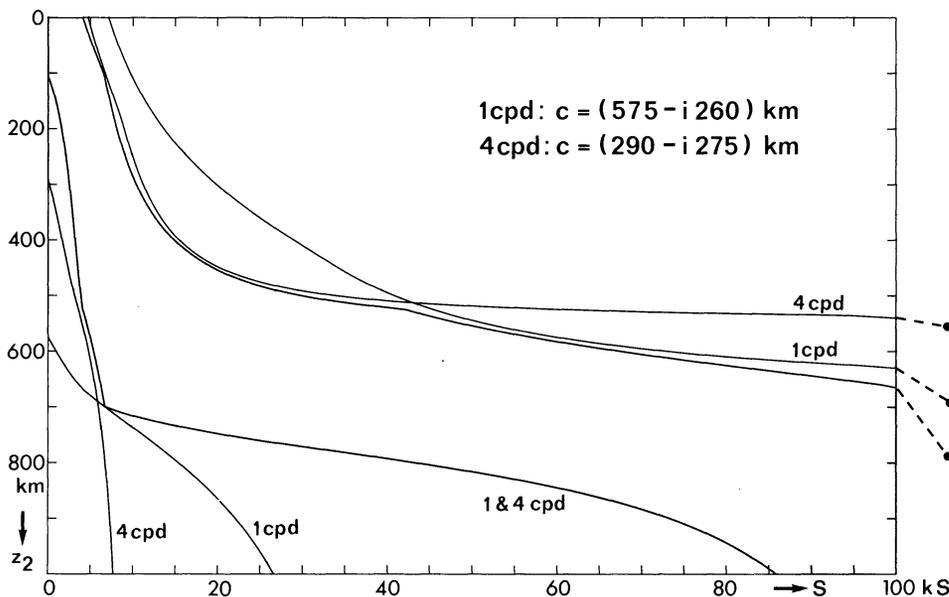


Fig. 4. Bounds for $S(z_2)$ in the unconstrained two-frequency problem. All models fitting these data must fall into the shaded area. The joint consideration of 1 and 4 cpd considerably improves the bound for S_{min} and shows that all models must have a good conductor not deeper than 700-800 km

rangian multipliers λ_1 and λ_2 are also discontinuous, meaning that near these levels the relation (2.10) for an estimate of the influence of data errors has to be applied with caution (cf. also case C1 in Sect. 3.6).

3.6 Extremal models for many frequencies

As a basic limitation of this study, it is assumed that there is a 1D model, which exactly fits the data set. This restriction may be dropped in subsequent work, where the equality constraints (2.7a) are replaced by inequality constraints that demand only a fit of the data within a suitable multiple of the standard deviation. With the present tools, however, there are two approximate ways to handle inconsistent data:

a) The 1D information (3.10) is extracted by quadratic programming; the cleaned data allow a reasonable construction of extremal models only if the number M of frequencies does not exceed the number N of terms, since for $M > N$ only one model exists.

b) Extremal models are constructed for different consistent subsets of the data, and the bounds on $S(z_2)$ for the whole set are estimated by taking the greatest of the lower bounds for each z_2 and the smallest of the upper bounds. In the simplest case a consistent subset comprises the data for just one frequency, requiring only $\text{Re } c \geq 0, \text{Im } c \leq 0$.

Generalizing the experience gained from a small number of frequencies, it appears that in the M -frequency case the extremal models belong to one of the six categories:

A1: The model consists of $M+1$ thin sheets with $\tau_{M+1} = \infty$ and with one sheet fixed at depth z_2 . There remain $2M$ free model parameters, which can be determined from the $2M$ real data. If required, the M -complex multipliers λ_j are obtained by solving the linear system of order $2M$ resulting from the necessary conditions (3.16a, b) for the $2M$ free model parameters (examples in Fig. 3: $\overline{AB}, \overline{DE}, \overline{KL}$).

A2: This case differs from A1 only in preassigning $\zeta_1 = 0$ rather than $\tau_{M+1} = \infty$ (examples: $\overline{BC}, \overline{IK}, \overline{LM}$).

B1: $M+2$ sheets with preassigned $\zeta_1 = 0, \tau_{M+2} = \infty$, and one sheet at fixed depth z_2 . There remain $2M+1$ free model parameters, for which there are $2M+1$ equations of type (3.16a, b), linear in the $2M$ real quantities $\text{Re } \lambda_j$ and $\text{Im } \lambda_j$. As a compatibility condition, the determinant of order $2M+1$ formed by the matrix of the linear system and the right-hand side of (3.16a, b) as $(2M+1)$ st column has to vanish. This equation in connection with the $2M$ data furnishes $2M+1$ nonlinear equations for the $2M+1$ free-model parameters. If required, the multipliers λ_j can finally be obtained by $2M$ of the linear equations (omitting an equation with the right-hand side vanishing). Examples are $\overline{CD}, \overline{EF}$.

B2: $M+1$ sheets of finite conductance with one sheet fixed at $z = z_2$. There are again $2M+1$ free-model parameters, which are determined in analogy to B1 (example: \overline{MN}).

C1: $M+1$ sheets with $\zeta_1 = 0, \tau_{M+1} = \infty$ (=canonical model I), z_2 happens to coincide with the position of one of the sheets. The $2M$ -model parameters (includ-

ing z_2) can be determined from the $2M$ data. However, the multipliers λ_j remain undefined, since Eqs. (3.16a, b) yield only $2M-1$ real equations (z_2 fixed!) for the M complex λ_j . In fact, the λ_j are discontinuous for these special values of z_2 , yielding different sets for slightly smaller and greater z_2 (examples: A, D, F, K).

C2: M sheets of finite conductance (=canonical model II), z_2 coincides with one of the sheets (cf. C1 for further discussion; examples: B, I, L).

The models A1 and A2, B1 and B2, C1 and C2 are dual in the sense of Sect. 3.2.

The model type for varying z_2 may have to be changed for the following reasons:

a) The system of nonlinear equations no longer has a solution (e.g., because the deepest perfectly conducting sheet disappears or reappears at infinite depth or merges with another sheet, or the conductance of the deepest sheet increases from finite values to infinity).

b) A surface sheet emerges. According to Eqs. (3.14a), (2.14), and (2.3), the condition is

$$D(0) = w(0) - \text{Re} \sum_{j=1}^M i \omega_j \mu_0 \lambda_j c_j^2 = 0,$$

whereas $D(0) > 0$ in the absence of a surface sheet.

c) The surface sheet becomes detached and moves downwards. This happens when ζ_1 becomes a freely varying parameter, implying according to (3.14b), (2.14), (2.3), and (3.2) that

$$D'(-0) + D'(0) = \text{Re} \sum_{j=1}^M 2i \omega_j \mu_0 c_j \lambda_j (2 - i \omega_j \mu_0 \tau_1 c_j) = 0.$$

This section concludes with an example for approximate extremal models from an inconsistent data set. We choose the COPROD data of Jones (1980), which have been inverted by different authors using a variety of techniques. The shaded area in Fig. 5 was determined by the approximate method b) from 11 consistent subsets comprising a single frequency. The figure also displays as $S(z_2)$ -curves the various results of inversion. These models freely use the space allotted them and occasionally even transgress it because of the approximate nature of bounds and models. This example shows that the approximate method b), which for consistent data tends to produce conservative bounds (cf. S_{\min} in Fig. 4), can even yield too narrow bounds for inconsistent data, where the range of acceptable models is broader because in fitting the data different frequencies (or frequency bands) can be emphasized. Also, inconsistencies in the data may lead to unreasonable restrictions in the feasible (shaded) area. For this reason the four longest periods contained in the COPROD data have been omitted when constructing the bounds.

4. Extremal models for the constrained conductance function

Section 3 was devoted to the unconstrained extremal models with $\sigma_-(z) \equiv 0, \sigma_+(z) \equiv \infty$. In this section we shall consider as a simple example for constrained extremal models the one-frequency problem of Sect. 3.4, where now the range of admitted conductivities is

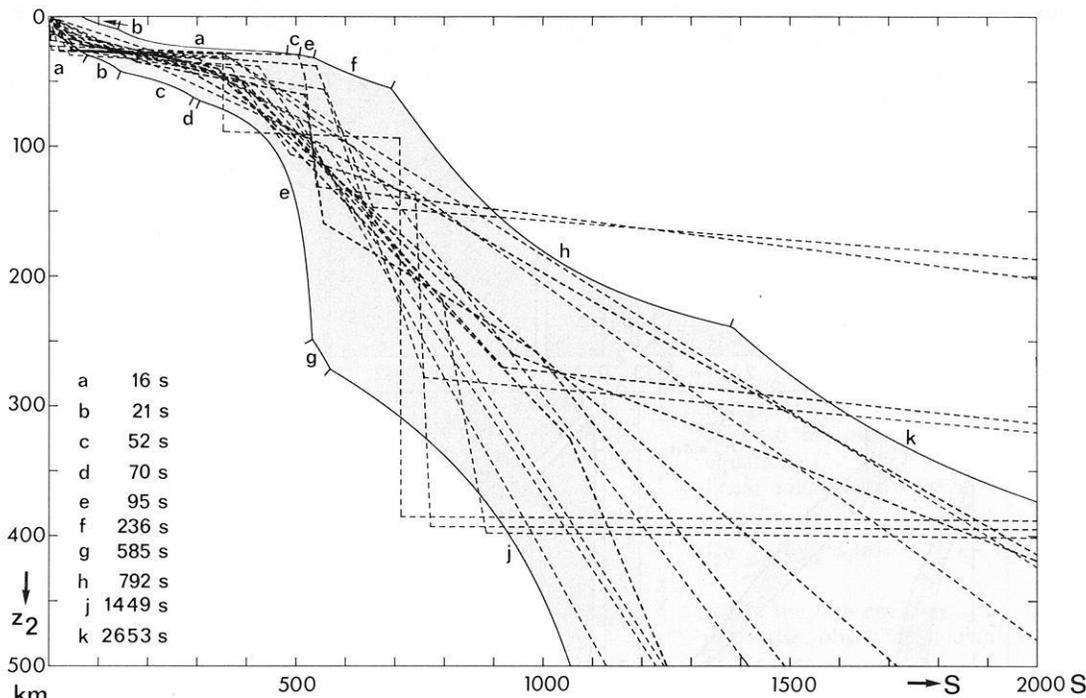


Fig. 5. Approximate bounds on $S(z_2)$ as derived from the 11 shortest periods of the COPROD data; only the 10 periods listed have contributed to the bounds. In general, longer periods should be binding at greater depth. Due to inconsistent data the longer period of 21 s (b) is bounding S_{\max} at shallower depth than the shorter period of 16 s (a). Also displayed as $S(z_2)$ curves are the results of the various interpreters

bounded by $\sigma_+(z) \equiv \sigma_+$, where σ_+ is a positive constant. In particular, the structure of the extremal models is studied when assigning to σ_+ a varying multiple α of the apparent conductivity σ_a [cf. (3.21)], i.e., $\sigma_+ = \alpha\sigma_a$. In this particular case, the extremal models consist of a sequence of uniform layers with $\sigma(z) = \sigma_+$ and insulators with $\sigma(z) = 0$. According to Eq. (2.20), the necessary extremal conditions are $D(z) \geq 0$ for $\sigma(z) = 0$ and $D(z) \leq 0$ for $\sigma(z) = \sigma_+$, with layer boundaries at positions where $D(z)$ changes sign.

Before considering more details, it might be useful to discuss first the structure of the extremal models as displayed in Fig. 6 for $\psi = 45^\circ$ and various values of α . The position of the thin sheets (abscissa) as a function of z_2 (ordinate) in the unconstrained case $\alpha = \infty$ is shown at the left. The ticks mark again perfect conductors. The main structural units of these models are clearly discernible in the following models of constrained conductivity, where the layers with $\sigma = \sigma_+$ are shaded. As expected from experience with the unconstrained case, S_{\min} generally shows a conducting layer starting at $z = z_2$, which is just excluded from the range of integration, whereas S_{\max} in general has a conducting layer, which ends at z_2 and is just included. There are two notable exceptions (dashed diagonal). For sufficiently small z_2 , the extremal models demand only that $\sigma = 0$ for S_{\min} and $\sigma = \sigma_+$ for S_{\max} in $0 \leq z \leq z_2$, which can be reached in different ways. When z_2 exceeds a certain limit, however, there will only be one model. This first unique model is then also shown for smaller z_2 . This situation is comprised in the extremal condition (2.20) since, in this range of z_2 , S_{\min} and S_{\max} do not depend on the data, i.e., $\lambda = 0$ [Eq. (2.10)] and

$D(z) = w(z)$, which is non-negative for S_{\min} and non-positive for S_{\max} . The other exception occurs for S_{\max} in a certain range below the point, where the two conducting layers merge into a thick conductor having $z = z_2$ as interior point. In this case $D(z_2 - 0) < -1$ and $D(z_2 + 0) = D(z_2 - 0) + 1 < 0$, i.e., $D(z)$ does not change sign at $z = z_2$.

For smaller values of α , the conducting layers necessarily become thicker, and for $\alpha = 1$ the extremal models for S_{\max} are just uniform half-space models. Therefore, the case of $\alpha = 1.1$, for which narrow non-conductive channels still occur, is considered. Due to the overshoot phenomenon of the apparent resistivity curve, it is possible to construct extremal models even for $\alpha < 1$, meaning that the true conductivity is smaller everywhere than the apparent conductivity. It is easily verified that for the considered phase $\psi = 45^\circ$ the smallest value of α is $\alpha_{\min} = \tanh^2(\pi/2) = 0.8412$, which corresponds to a surface layer of conductivity $\alpha_{\min}\sigma_a$, thickness $(\pi|c|/\sqrt{2}) \coth(\pi/2) = 2.4221|c|$, and an insulator below. For the smallest possible conductivity the models for S_{\min} and S_{\max} coincide and are independent of z_2 .

The resulting bounding curves of $S(z_2)$ for selected values of α (curve parameter) are shown in Fig. 7. A bounded value of α mostly affects S_{\max} , whereas S_{\min} is only influenced for relatively small α . The dashed lines refer to the smallest possible value of α , for which only one model exists. For $0 \leq \psi < 45^\circ$ it consists of a slab at finite depth, and for $45^\circ \leq \psi \leq 90^\circ$ it is a surface slab.

The actual construction of the extremal models is similar to that in the unconstrained case. If $\sigma_-(z)$ and $\sigma_+(z)$ are independent of z , the conductivity models

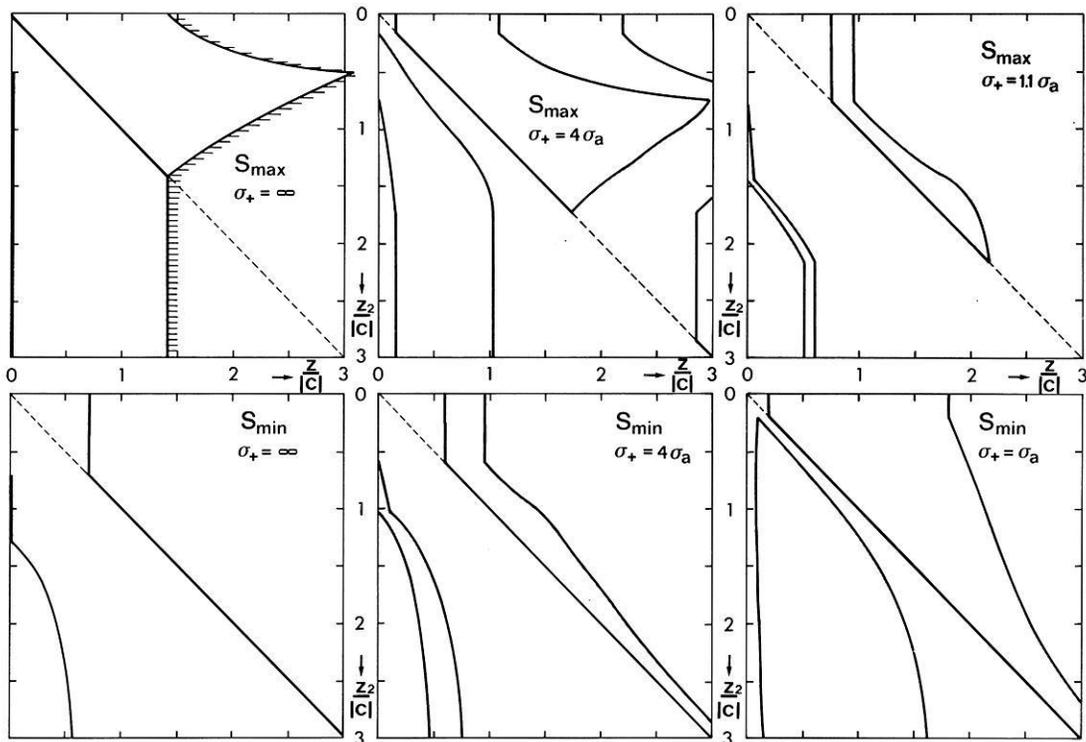


Fig. 6. Structure of the unconstrained (*left*) and constrained (*center and right*) one-frequency extremal models for $\psi = 45^\circ$. For a given z_2 (*ordinate*) the unconstrained models are presented only by the position of the conducting thin sheets (without specifying their conductance), whereas the constrained models ($\sigma_+ < \infty$) are completely characterized by specifying the position of the conducting layers (*shaded*)

consist of a sequence of uniform layers. The pertinent formulae are briefly summarized. Let K uniform layers exist, with the top at $z = \zeta_k$ ($\zeta_1 = 0$), and layer conductivities σ_k , $k = 1, \dots, K$. With $\gamma_k^2 := i\omega\mu_0\sigma_k$ and $d_k := \zeta_{k+1} - \zeta_k$ as thickness of layer k , $k = 1, \dots, K - 1$, the theoretical response $c[\sigma] = c_1$ is recursively determined from

$$c_k = \frac{1}{\gamma_k} \frac{\gamma_k c_{k+1} + \tanh(\gamma_k \alpha d_k)}{1 + \gamma_k c_{k+1} \tanh(\gamma_k d_k)}, \quad k = K - 1, \dots, 1 \quad (4.1)$$

starting with $c_K = 1/\gamma_K$. [Note again that in this context the subscript k on c refers to the level ζ_k , whereas in other applications the subscript specifies the frequency.] The Fréchet derivative $F(z) = -i\omega\mu_0 f^2(z)$ at $z = \zeta_k$ is determined from

$$f(\zeta_1) = f(0) = -c[\sigma],$$

$$f(\zeta_k) = f(\zeta_1) \prod_{n=1}^{k-1} \frac{\gamma_n + 1/c_n}{\gamma_n + 1/c_{n+1}} e^{-\gamma_n d_n}, \quad k \geq 2. \quad (4.2)$$

The necessary conditions (2.20) require that at a discontinuity $z = \zeta_k$, different from 0 and z_2 , $D(z)$ changes sign, i.e.,

$$D(\zeta_k) = w(\zeta_k) + \text{Re} \{ \lambda F(\zeta_k) \} = 0, \quad \zeta_k \neq 0, z_2, \quad (4.3)$$

$D'(\zeta_k) \neq 0$. At $z = z_2$, the conductivity is only discontinuous if $D(z_2 - 0) \cdot D(z_2 + 0) < 0$. The condition (4.3) for each *unknown* discontinuity level, together with the data $c[\sigma] = c$, obviously provide the correct number of equations to compute the positions of the unknown discontinuities and λ . The examples in Fig. 6 show that there

will be up to four unknown levels for the one-frequency case. The dimension of the resulting nonlinear system can again be reduced to four by eliminating λ on replacing the four equations (4.3), which are linear in λ , by two compatibility conditions. These conditions demand that any two of the four (3×3) subdeterminants of the augmented (4×3) matrix, i.e., $w(\zeta_k)$ as third column, have to vanish.

All extremal models (except the half-space for S_{\max} , $\psi = 45^\circ$, $\alpha = 1$) terminate with an insulator after a finite number of layers. This pattern changes, when the a priori bound $\sigma_- > 0$ is imposed, because $w(z)$ vanishes for $z > z_2$, and in this range $D(z) = \text{Re} \{ \lambda F(z) \}$ will show a kind of damped oscillations rather than tend to a constant. As a consequence, there will be an infinite number of layers, which below a certain level, however, will have an insignificant influence on the actual bounds.

5. Conclusion

The explicit examples in Sects. 3 and 4 have demonstrated the possibility of constructing rigorous bounds for $S(z_2)$. But the examples have also shown that the consideration of only a few frequencies yields pessimistic bounds, which might not be very useful in the geophysical application. The situation will even be worse if bounds for spatial averages of the conductivity are constructed rather than for the conductance. Therefore, the joint interpretation of a larger number of frequencies appears to be mandatory. The handling of

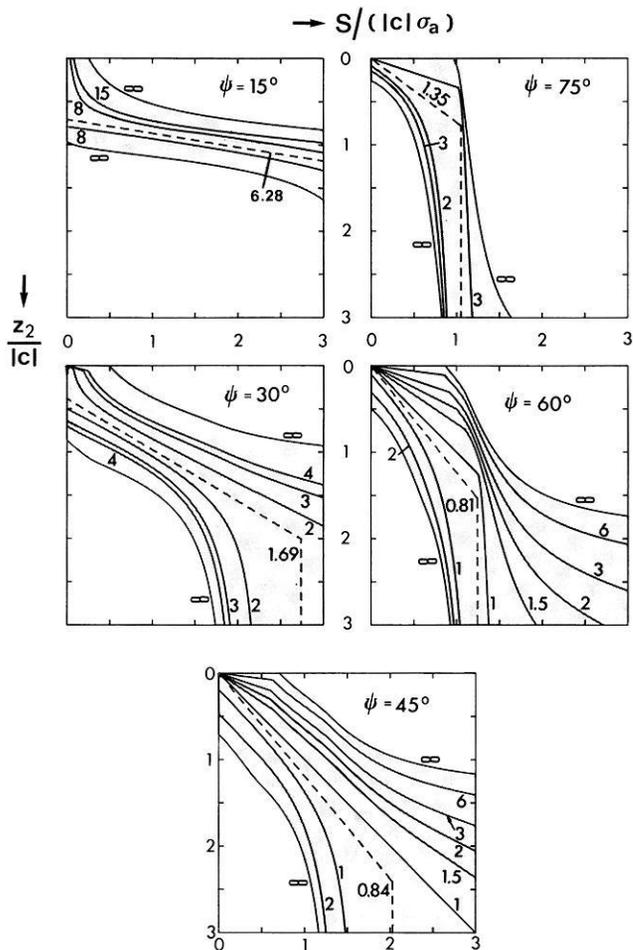


Fig. 7. Bounds on $S(z_2)$ in the one-frequency case constrained by $\sigma(z) \leq \alpha \sigma_a$, where α is the curve parameter. The limiting case $\alpha = \infty$ is already shown in Fig. 2. For the smallest possible value of α there is only one model depicted by the dashed curve

many consistent data will not pose serious problems, but an extension of the theory is still necessary to deal with the most interesting case of many real data that are generally inconsistent.

Acknowledgements. The author is greatly indebted to one of the referees for pointing out a substantial error in an earlier version of Appendix A. Moreover, the author wishes to thank Prof. H.-J. Dürbaum for providing the opportunity to carry out a major part of the present study at the Federal Institute for Geosciences and Natural Resources (BGR) at Hannover. Many thanks also to Dr. A.G. Jones for making available the results of the COPROD study prior to publication.

Appendix A: Linear independence of the Fréchet kernels

The simple structure of the extremal models is a result of the assertion that there is no set of M complex constants λ_j , not all equal zero, so that (2.17) is satisfied in any interval where $w(z)$ is constant and $\sigma(z)$ is positive. The functions $F_j(z) \sim f_j^2(z)$ have to present a downward diffusing field with $F_j'(z) \rightarrow 0$ for $z \rightarrow z_{\max}$, where z_{\max} is either infinity or the depth to a perfect conductor. This qualification will eliminate the nontrivial solutions of (2.17), which may exist for a particular choice of $\sigma(z)$.

At the beginning it is assumed that at least one interval

exists where (2.17) is satisfied and that $a < z < b$ is the deepest. Then (2.17) reads

$$D(z) = w(z) + \operatorname{Re} \sum_{j=1}^M \phi_j(z) = 0, \quad z \in (a, b) \quad (\text{A-1})$$

with $\phi_j(z) := \lambda_j F_j(z)$. For $a < z < b$, the conductivity $\sigma(z)$ is positive. Let $\sigma^{(p)}(b)$, $p \geq 0$ be the first nonvanishing derivative of $\sigma(z)$ at $z = b - 0$. In $z > b$, a set of K thin sheets of finite conductance may exist, possibly terminated by an additional perfect conductor. For ease of presentation the case $K = 0$ and $K > 0$ are considered separately.

$K = 0$

Let the possible perfect conductor be at $z = b + H$. Differentiate (A-1) $2M$ ($p+2$) times and evaluate the derivatives at $z = b - 0$ on using $\phi_j'(b) = -(2/H) \phi_j(b)$ and the differential equations (2.15a, b). The result is a set of $2M$ homogeneous linear equations for $\operatorname{Re} \phi_j(b)$ and $\operatorname{Im} \phi_j(b)$

$$\operatorname{Re} \sum_{j=1}^M (i\omega_j)^n \phi_j(b) = 0, \quad n = 1, \dots, 2M, \quad (\text{A-2})$$

where the n -th equation is obtained from the $n(p+2)$ -th order derivative, observing that this derivative of $\phi_j(z)$ at $z = b - 0$ is of the type

$$\phi_j(b) \sum_{m=0}^n \gamma_m (i\omega_j)^m$$

with real frequency-independent coefficients γ_m and $\gamma_n \neq 0$. The set (A-2) disintegrates into two uncoupled systems for $\omega_j \operatorname{Im} \phi_j$ and $\omega_j^2 \operatorname{Re} \phi_j$, both having as system determinant the nonvanishing Vandermonde determinant (e.g., Smirnov 1964, p. 21)

$$\Delta = \begin{vmatrix} 1 & \dots & 1 \\ \omega_1^2 & \dots & \omega_M^2 \\ \vdots & & \vdots \\ \omega_1^{2M-2} & \dots & \omega_M^{2M-2} \end{vmatrix} = \prod_{1 \leq r < s \leq M} (\omega_s^2 - \omega_r^2) \neq 0. \quad (\text{A-3})$$

Hence, (A-2) allows only the trivial solution $\phi_j(b) = 0$, $j = 1, \dots, M$. This solution is compatible only with $w(b-0) = 0$ and either a perfect conductor at $z = b$ or $b = \infty$, because the modulus of ϕ_j cannot increase with depth. In the first case also $\phi_j'(b) = 0$, but $\phi_j''(b) \neq 0$. Repeating the above arguments for $\phi_j''(b)$ by considering the derivatives of order $n(p+4)$, $n = 1, \dots, 2M$, it follows that also $\phi_j''(b) = 0$, implying $\phi_j(z) \equiv 0$ in $a < z < b$. In the case $b = \infty$ differentiate (A-1) three times, use (2.15b), divide by $4\mu_0 \sqrt{\sigma}$, and integrate to obtain

$$\operatorname{Re} \sum_{j=1}^M i\omega_j \sqrt{\sigma(z)} \phi_j(z) = \gamma, \quad z > a, \quad (\text{A-4})$$

where γ is a constant. By comparing with power laws $\sigma(z) = Az^m$ leading to Bessel function solutions (or appealing to a bounded energy dissipation), it is always found that

$$\lim_{z \rightarrow \infty} \sqrt{\sigma(z)} \phi_j(z) = 0,$$

which implies that $\gamma = 0$. Dividing (A-4) by $\sqrt{\sigma}$ and applying the same operations another $(2M-1)$ times, we end up again with the system (A-2), $\phi_j(b)$ being replaced by $\phi_j(z)$. Hence, $\phi_j(z) \equiv 0$ in $z \geq a$. This shows that for $K = 0$, no solution $\phi_j(z) \neq 0$ of (A-1) can be found.

$K > 0$

Below $z = b$ let there exist K thin sheets of finite conductance τ_k at $z = \zeta_k$, $k = 1, \dots, K$ and define $d_1 := \zeta_1 - b \geq 0$, $d_k := \zeta_k - \zeta_{k-1} > 0$, $k = 2, \dots, K$. In addition, there may be a perfect

conductor at $z = \zeta_K + H$. Separately considered are the cases $w(z) \equiv 0$ in $z \geq b$ and $w(z) \not\equiv 0$ in $z \geq b$.

$\alpha)$ $w(z) \equiv 0$ in $z \geq b$

In this case the necessary conditions (3.15a, b) have to be satisfied at all K sheets and lead to the $2K$ equations

$$\operatorname{Re} \sum_{j=1}^M e_{2K-n}(\omega_j) \phi_j(\zeta_K) = 0, \quad n=0, \dots, 2K-1. \quad (\text{A-5})$$

Working from $z = \zeta_K$ upwards and starting with $e_{2K}=1$, the functions $e_{2K-n}(\omega)$ are, on account of Sect. 3.1 [in particular Eqs. (3.2), (3.4), and (3.5)], recursively defined as

$$\begin{aligned} e_{2k-1} &= e_{2k}(i\omega\mu_0\tau_k - 2/c_k), \\ e_{2k-2} &= e_{2k}(1 + d_k/c_k)^2 \end{aligned} \quad (\text{A-6})$$

with $k=K, \dots, 1$ and the c_k are obtained by the recurrence relation (3.3), starting with $c_K = H/(1 + i\omega\mu_0\tau_K H)$. Examining the structure of $e_{2K-n}(\omega)$, it is easily found that

$$e_{2K-n}(\omega_j) = \sum_{m=0}^n \bar{\gamma}_m(i\omega_j)^m, \quad n=0, \dots, 2K-1 \quad (\text{A-7})$$

where $\bar{\gamma}_m$ is again real and frequency independent with $\bar{\gamma}_n \neq 0$. Hence by linear combination (A-5) is equivalent to

$$\operatorname{Re} \sum_{j=1}^M (i\omega_j)^n \phi_j(\zeta_K) = 0, \quad n=0, \dots, 2K-1. \quad (\text{A-8})$$

For $K \geq M$ (A-8) implies already $\phi_j(\zeta_K) = 0$; for $K < M$ the missing linearly independent equations are obtained from the higher derivatives of $D(z)$ at $z = b - 0$. First, let $d_1 = \zeta_1 - b > 0$. Then the three quantities

$$\begin{aligned} \phi_j(b)/\phi_j(\zeta_K), \quad \phi'_j(b)/\phi_j(\zeta_K), \quad \text{and} \\ \phi'_j(b)/\{\phi_j(b)\phi_j(\zeta_K)\} \end{aligned} \quad (\text{A-9})$$

are polynomials in $i\omega_j$ of exact degree $2K$. After evaluating the derivatives of $\phi_j(z)$ of order $n(p+2)$, $n=1, \dots, 2M-2K-1$ at $z = b - 0$ on, using the preceding result in connection with (A-8) and (2.15a, b), the validity of (A-8) is extended up to powers $2M-1$. As a consequence, $\phi_j(\zeta_K) = 0$, implying in fact $\phi_j(z) \equiv 0$. For $d_1 = 0$ the first two terms of (A-9) are only of degree $2K-2$ and $2K-1$, but the same result is obtained by considering the derivatives of order $-p+n(p+2)$, $n=1, \dots, 2M-2K$.

$\beta)$ $w(z) \not\equiv 0$ in $z \geq b$

According to Sect. 3.3 the discontinuity of $w(z)$ coincides with the position of a sheet. If this is the sheet $k=1$, then (cf. Sect. 3.3)

$$\begin{aligned} D(\zeta_1 - 0) = 1, \quad D'(\zeta_1) = 0 \quad \text{for } S_{\min}, \\ D(\zeta_1 - 0) = 0, \quad D'(\zeta_1) < 0 \quad \text{for } S_{\max} \end{aligned} \quad (\text{A-10})$$

However, $D(b) = 0$, $D'(b) = 0$ implies $D(z) = A(z-b)^2$ in $b \leq z \leq \zeta_1$, which for $d_1 \geq 0$ is incompatible with both cases of (A-10). If the discontinuity occurs at level $k > 1$, then

$$D(\zeta_1) = 0, \quad D'(\zeta_1 - 0) < 0$$

which again cannot be met by any choice of A for $d_1 \geq 0$.

Summarizing the above results, in the unconstrained case no downward diffusing field solution satisfying (2.17) can be found for $\sigma(z) > 0$ in $a < z < b$. The assumption $\sigma(z) > 0$ is essential, since for $\sigma(z) \geq 0$ non-trivial diffusing solutions of (2.17) exist, e.g., $D(z) \equiv 0$ for $z > z_2$ in the case of S_{\min} (cf. Sect. 3.3) without $\phi_j(z)$ vanishing identically.

Appendix B: Equivalent partial fraction expansions

Let for M distinct frequencies ω_j the theoretical transfer function $c_j[\sigma]$ of a layered ground be represented by

$$c_j[\sigma] = a_0 + \sum_{n=1}^N \frac{a_n}{b_n + i\omega_j}, \quad j=1, \dots, M, \quad (\text{B-1}) \quad = (3.10)$$

where the b_n are distinct (otherwise N can be reduced). Assume the ordering $b_n > b_{n+1}$. Then a_0 and b_N have to be non-negative, all other constants are strictly positive. The following statement will be proved:

a) For $N \geq M$ exist two (condensed) representations

$$c_j[\sigma] = \sum_{m=1}^M \frac{A_m}{B_m + i\omega_j}, \quad (\text{B-2}) \quad = (3.11a)$$

$$c_j[\sigma] = \bar{A}_0 + \sum_{n=1}^{M-1} \frac{\bar{A}_n}{\bar{B}_n + i\omega_j} + \frac{\bar{A}_M}{i\omega_j}, \quad (\text{B-3}) \quad = (3.11b)$$

$j=1, \dots, M$. Excluding only the identity $M=N$ with $a_0 = b_N = 0$, implying also $\bar{A}_0 = \bar{B}_N = 0$, the representations (B-2) and (B-3) consist each of $2M$ positive constants corresponding to the $2M$ data. There is no representation with less than $2M$ constants, i.e., B_m and \bar{B}_m are distinct.

b) For $N < M$ the representation (B-1) is unique, i.e., no alternative partial fraction expansion exists.

The proof is concentrated on the representation (B-2), and at the end only the necessary modifications for (B-3) are stated. At the outset it is assumed that there is an expansion of type (B-2) with $K \leq M$ complex terms. Then it has to be shown that the nonlinear system of $2M$ equations

$$\sum_{m=1}^K \frac{A_m B_m}{B_m^2 + \omega_j^2} = a_0 + \sum_{n=1}^N \frac{a_n b_n}{b_n^2 + \omega_j^2} \quad \left. \vphantom{\sum_{m=1}^K} \right\} j=1, \dots, M \quad (\text{B-4a})$$

$$\sum_{m=1}^K \frac{A_m}{B_m^2 + \omega_j^2} = \sum_{n=1}^N \frac{a_n}{b_n^2 + \omega_j^2} \quad \left. \vphantom{\sum_{m=1}^K} \right\} j=1, \dots, M \quad (\text{B-4b})$$

has $2K$ positive solutions $A_m, B_m, m=1, \dots, K$. First, we recall the elementary partial fraction decomposition

$$x^{2k} \left/ \prod_{j=1}^M (x^2 + \omega_j^2) \right. = \sum_{j=1}^M \frac{\alpha_{jk}}{x^2 + \omega_j^2}, \quad k=0, \dots, M-1 \quad (\text{B-5})$$

with

$$\alpha_{jk} = (-\omega_j^2)^k \left/ \prod_{\substack{l=1 \\ l \neq j}}^M (\omega_l^2 - \omega_j^2) \right.,$$

then multiply (B-4a, b) for $k=0, \dots, M-1$ by α_{jk} , sum over j , identify x on the left-hand side with B_m , on the right-hand side with b_n , and obtain by (B-5) the new set of $2M$ equations

$$\sum_{m=1}^K G_m B_m^l = a_0 \delta_{l, 2M-1} + \sum_{n=1}^N g_n b_n^l, \quad l=0, \dots, 2M-1, \quad (\text{B-6})$$

where δ_{ik} is the Kronecker symbol and

$$G_m := A_m \left/ \prod_{j=1}^M (B_m^2 + \omega_j^2) \right., \quad g_n := a_n \left/ \prod_{j=1}^M (b_n^2 + \omega_j^2) \right. \quad (\text{B-7})$$

The equations for odd and even l result from (B-4a) and (B-4b), respectively. The quantities G_m rather than A_m may now be considered as unknowns. In the derivation of the first term on the right-hand side it has been observed that by virtue of (B-5)

$$\sum_{j=1}^M \alpha_{jk} = \lim_{x \rightarrow \infty} x^{2k+2} \left/ \prod_{j=1}^M (x^2 + \omega_j^2) \right. = \delta_{k, M-1}, \quad k \leq M-1.$$

The $2M$ equations (B-6) are linear in the K unknowns G_m . In order for there to be a solution, any $K+1$ equations of (B-6) have to be linearly dependent. Taking, for instance, the $K+1$ equations from $l=i$ to $l=i+K$, $i=0, \dots, 2M-K-1$, then for each i there has to be a set of coefficients $q_k^{(i)}$, $k=0, \dots, K$, not all equal to zero, such that

$$\sum_{k=0}^K q_k^{(i)} B_m^{i+k} = 0, \quad m=1, \dots, K, \quad (\text{B-8a})$$

$$a_0 q_K^{(i)} \delta_{i, 2M-K-1} + \sum_{k=0}^K q_k^{(i)} \sum_{n=1}^N g_n b_n^{i+k} = 0. \quad (\text{B-8b})$$

Since we are searching for positive B_m , (B-8a) may be divided by B_m^i to yield

$$\sum_{n=0}^K q_k^{(i)} B_m^k = 0, \quad m=1, \dots, K. \quad (\text{B-8c})$$

$q_k^{(i)}$ is independent of m . Hence, each B_m satisfies the same algebraic K -th order equation, i.e., by the fundamental theorem of algebra the existence of the set $q_k^{(i)}$ implies the existence of solutions B_m being the K roots z of

$$\sum_{k=0}^K q_k^{(i)} z^k = 0. \quad (\text{B-9})$$

On the other hand, the roots define the coefficients $q_k^{(i)}$ uniquely apart from a scaling factor. Therefore, the $q_k^{(i)}$ are in fact independent of i , $q_k^{(i)} = q_k$. Normalizing by $q_K=1$, the remaining K coefficients q_k have to be determined from the $2M-K$ linear equations (B-8b). First, let $K=M$. Then the system determinant

$$\Delta_M = \det \left\{ \sum_{n=1}^N g_n b_n^{k+i} \right\}, \quad i, k=0, \dots, M-1 \quad (\text{B-10})$$

can be expressed for $N \geq M$ as

$$\Delta_M = \sum \left\{ \prod_{k=1}^M g_{n_k} \right\} \cdot \begin{vmatrix} 1 & b_{n_1} & \dots & b_{n_1}^{M-1} \\ \vdots & \vdots & & \vdots \\ 1 & b_{n_M} & \dots & b_{n_M}^{M-1} \end{vmatrix}^2, \quad (\text{B-11})$$

where the summation extends over all M -tupels n_k , $k=1, \dots, M$ with $1 \leq n_1 < \dots < n_M \leq N$ (e.g., Smirnov, 1964, p. 28). The determinants in (B-11) are again Vandermonde determinants, which are generally defined as $V = \det \{x_j^{k-1}\}$, $j, k=1, \dots, M$ and are given by (e.g., Smirnov, 1964, p. 21)

$$V = \prod_{1 \leq p < q \leq M} (x_q - x_p). \quad (\text{B-12})$$

All terms in the sum (B-11) are positive by virtue of the assumptions $N \geq M$ and b_n distinct. Hence, $\Delta_M > 0$ and the coefficients q_k can be determined uniquely. In the case $N < M$, however, $\Delta_M = 0$, as is seen by formally adding in (B-10) $M-N$ terms with $g_{N+1} = \dots = g_M = 0$ and using (B-11). In this case there is no set of coefficients q_k , implying that (B-1) cannot be modified.

The case of $K < M \leq N$ must still be discussed, in which the linear system (B-8b) consists of $2M-K = K+2(M-K)$ equations for only K unknowns. A necessary condition for the existence of a solution is the linear dependence of the first $K+1$ equations, which is equivalent to the statement that $\Delta_{K+1} = 0$. However, (B-11) with $M=K+1$ shows that $\Delta_{K+1} > 0$. Hence, for $K < M$ the first $K+1$ equations are already linearly independent and there is no solution q_k . Therefore, in the sequel only the case $K=M$, $N \geq M$ has to be considered.

Having established the existence of the set B_m as the roots of (B-9) with $K=M$, we have to show that B_m is real and positive. Without exploiting (B-9) any further, we return to (B-8b, c). In order for there to be a solution $q_k^{(i)} \equiv q_k$, $k=0, \dots, M$

for this homogeneous linear system, the determinant has to vanish, implying

$$a_0 \delta_{i, M-1} \begin{vmatrix} 1 & B_1 & \dots & B_1^{M-1} \\ \vdots & \vdots & & \vdots \\ 1 & B_M & \dots & B_M^{M-1} \end{vmatrix} + \sum_{n=1}^N g_n b_n^i \begin{vmatrix} 1 & B_1 & \dots & B_1^M \\ \vdots & \vdots & & \vdots \\ 1 & B_M & \dots & B_M^M \\ 1 & b_n & \dots & b_n^M \end{vmatrix} = 0,$$

$i=0, \dots, M-1$. The above Vandermonde determinants are easily evaluated by (B-12) and yield after much cancellation since the B_m are distinct

$$a_0 \delta_{i, M-1} + \sum_{n=1}^N g_n b_n^i \prod_{m=1}^M (b_n - B_m) = 0, \quad i=0, \dots, M-1, \quad (\text{B-13})$$

or equivalently

$$a_0 \gamma_{M-1} + \sum_{n=1}^N g_n p_{M-1}(b_n) \prod_{m=1}^M (b_n - B_m) = 0, \quad (\text{B-14})$$

where $p_{M-1}(b_n)$ is an arbitrary polynomial in b_n of degree not higher than $M-1$ with γ_{M-1} as coefficient of b_n^{M-1} . The M equations (B-13) form a nonlinear system for B_m . The existence of a solution set B_m is granted from (B-8c) and (B-9). To show the positivity of the particular element B_k , $k=1, \dots, M$, we take

$$p_{M-1}(b_n) = \prod_{m=1}^M (b_n - B_m^*), \quad \gamma_{M-1} = 1,$$

where prime and asterisk denote, respectively, the omission of the factor $m=k$ and the complex conjugate, and obtain from (B-14)

$$a_0 + \sum_{n=1}^N g_n (b_n - B_k) \prod_{m=1}^M |b_n - B_m|^2 = 0. \quad (\text{B-15})$$

This shows that B_k is real and positive.

Next it is proved that also A_k [or G_k , cf. Eq. (B-7)] is positive. The solution of the first M equations of (B-6) for G_k yields by Cramer's rule on exploiting, again the simple properties of the resulting Vandermonde determinants

$$G_k = \sum_{n=1}^N g_n \prod_{m=1}^M \frac{b_n - B_m}{B_k - B_m}. \quad (\text{B-16})$$

The definite sign of G_k is not yet obvious. However, expanding the product by

$$\begin{aligned} \prod_{m=1}^M (B_k - B_m) &= \prod_{m=1}^M [(b_n - B_m) - (b_n - B_k)] \\ &= (b_n - B_k) p_{M-1}(b_n) + \prod_{m=1}^M (b_n - B_m), \end{aligned} \quad (\text{B-17})$$

where p_{M-1} is, in fact, a polynomial of degree $M-2$, we obtain from (B-16) by means of (B-14) with $\gamma_{M-1} = 0$:

$$G_k = \sum_{n=1}^N g_n \prod_{m=1}^M \left(\frac{b_n - B_m}{B_k - B_m} \right)^2 > 0. \quad (\text{B-18})$$

This proves the existence of the representation (B-2).

Skipping details of derivation, we mention only two additional identities required in Sect. 3.2:

$$\sum_{m=1}^M (A_m - 2a_0 B_m) = \sum_{n=1}^N \left[a_n - g_n \prod_{m=1}^M (b_n - B_m)^2 \right], \quad (\text{B-19})$$

$$\begin{aligned} \sum_{m=1}^M (A_m/B_m) &= a_0 + \sum_{n=1}^N (a_n/b_n) \\ &\quad - \prod_{j=1}^M (\omega_j/B_j)^2 \cdot \left[a_0 + \sum_{n=1}^N (g_n/b_n) \prod_{m=1}^M (b_n - B_m)^2 \right]. \end{aligned} \quad (\text{B-20})$$

The existence of the representation (B-3) is proved similarly. The equivalents of (B-6), (B-13), and (B-14) are

$$\sum_{m=1}^{M-1} \bar{G}_m \bar{B}_m^l + \bar{G}_M \delta_{l0} = (a_0 - \bar{A}_0) \delta_{l, 2M-1} + \sum_{n=1}^N g_n b_n^l, \quad l=0, \dots, 2M-1, \quad (\text{B-21})$$

$$(a_0 - \bar{A}_0) \delta_{iM} + \sum_{n=1}^N g_n b_n^i \prod_{m=1}^{M-1} (b_n - \bar{B}_m) = 0, \quad i=1, \dots, M, \quad (\text{B-22})$$

$$(a_0 - \bar{A}_0) \gamma_{M-1} + \sum_{n=1}^N g_n p_{M-1}(b_n) b_n \prod_{m=1}^{M-1} (b_n - \bar{B}_m) = 0. \quad (\text{B-23})$$

The positivity of \bar{B}_k , $k=1, \dots, M-1$, is proved by taking

$$p_{M-1}(b_n) = \prod_{m=1}^{M-1} (b_n - \bar{B}_m^*), \quad \gamma_{M-1} = 0,$$

implying

$$\sum_{n=1}^N g_n (b_n - \bar{B}_k) \prod_{m=1}^{M-1} |b_n - \bar{B}_m|^2 = 0$$

and $b_1 \geq \bar{B}_k \geq b_N$. Choosing

$$p_{M-1}(b_n) = \prod_{m=1}^{M-1} (b_n - \bar{B}_m), \quad \gamma_{M-1} = 1$$

we obtain

$$\bar{A}_0 = a_0 + \sum_{n=1}^N g_n b_n \prod_{m=1}^{M-1} (b_n - \bar{B}_m)^2 > 0. \quad (\text{B-24})$$

Finally, \bar{G}_m is found by solving the first M equations of (B-21) for \bar{G}_m and using analogues of the identity (B-17):

$$\bar{G}_k \bar{B}_k = \prod_{n=1}^N g_n b_n \prod_{m=1}^{M-1} \left(\frac{b_n - \bar{B}_m}{\bar{B}_k - \bar{B}_m} \right)^2 > 0, \quad k=1, \dots, M-1, \quad (\text{B-25})$$

$$\bar{G}_M = \sum_{n=1}^N g_n \prod_{m=1}^{M-1} \left(\frac{b_n - \bar{B}_m}{\bar{B}_M} \right)^2 > 0. \quad (\text{B-26})$$

References

- Anderssen, R.S., Seneta, E.: A simple statistical estimation procedure for Monte Carlo inversion in geophysics. *Pure Appl. Geophys.*, **91**, 5-13, 1971
- Anderssen, R.S., Seneta, E.: A simple statistical estimation procedure for Monte Carlo inversion in geophysics. II: Efficiency and Hempel's paradox. *Pure Appl. Geophys.*, **96**, 5-14, 1972
- Barcilon, V.: Ideal solution of an inverse normal mode problem with finite spectral data. *Geophys. J.R. Astron. Soc.*, **56**, 399-408, 1979
- Barcilon, V., Turchetti, G.: On an inverse eigenvalue problem with truncated spectral data. In: *Inverse and improperly*

- posed problems in differential equations*, G. Anger, ed.: pp. 25-34, Berlin: Akademie-Verlag 1979
- Brown, K.M.: Computer oriented algorithms for solving systems of simultaneous nonlinear algebraic equations, G.D. Byrne, C.A. Hall, ed.: pp. 281-348. New York and London: Academic Press 1973
- Jones, A.G.: Report on the COPROD study. Paper read at the Fifth Workshop on Electromagnetic Induction in the Earth and Moon, Istanbul, August 18-23, 1980
- Lawson, C.L., Hanson, R.J.: *Solving least squares problems*. Englewood Cliffs: Prentice-Hall 1974
- Oldenburg, D.W.: Funnel functions in linear and non-linear appraisal. *J. Geophys. Res.*, **88**, 7387-7398, 1983
- Parker, R.L.: Inverse theory with grossly inadequate data. *Geophys. J.R. Astron. Soc.*, **29**, 123-138, 1972
- Parker, R.L.: Best bounds on density and depth from gravity data. *Geophysics*, **39**, 644-649, 1974
- Parker, R.L.: The theory of ideal bodies for gravity interpretations. *Geophys. J.R. Astron. Soc.*, **42**, 315-334, 1975
- Parker, R.L.: The Fréchet derivative for the one-dimensional electromagnetic induction problem. *Geophys. J.R. Astron. Soc.*, **49**, 543-547, 1977
- Parker, R.L.: The inverse problem of electromagnetic induction: existence and construction of solutions based on incomplete data. *J. Geophys. Res.*, **85**, 4421-4428, 1980
- Parker, R.L.: The existence of a region inaccessible to magnetotelluric sounding. *Geophys. J.R. Astron. Soc.*, **68**, 165-170, 1982
- Parker, R.L., Whaler, K.A.: Numerical methods for establishing solutions to the inverse problem of electromagnetic induction. *J. Geophys. Res.*, **86**, 9574-9584, 1981
- Pearson, C.E. [ed.]: *Handbook of applied mathematics*. New York: Van Nostrand Reinhold Co. 1974
- Sabatier, P.C.: On extremal solutions of Sturm-Liouville inverse problems. In: *Inverse and improperly posed problems in differential equations*, G. Anger, ed.: pp. 223-231. Berlin: Akademie-Verlag 1979
- Schmucker, U.: Anomalies of geomagnetic variations in the southwestern United States. *Bull. Scripps Inst. Oceanogr.* **13**, 1-165, 1970
- Schmucker, U.: EM-Übertragungsfunktionen für Sq. In: *Protokoll Kolloqu. „Erdmagnetische Tiefenforschung“ in Grafrath March 19-23, 1984*, J. Homilius and V. Haak, ed., Hannover: Niedersächsisches Landesamt für Bodenforschung 1984
- Smirnov, V.I.: *A course of higher mathematics*, Vol. III/1. Oxford [etc.]: Pergamon Press 1964
- Weidelt, P.: The inverse problem of geomagnetic induction. *Z. f. Geophys.*, **38**, 257-289, 1972.
- Weidelt, P.: Extremal models for electromagnetic induction in two-dimensional perfect conductors. *J. Geophys.*, **49**, 217-225, 1981

Received October 1, 1984; Revised August 1, 1985
Accepted August 1, 1985

The frequency characteristics of anomalous vertical fields observed in the British Isles

D. Beamish

British Geological Survey, Murchison House, West Mains Road, Edinburgh EH9 3LA. United Kingdom

Abstract. The frequency characteristics of anomalous vertical fields are presented and summarised in a consistent manner across a substantial area of the UK. The period range considered is from 10 to 10000 s. The results have been necessarily divided into coastal and non-coastal observations. An effective coast-effect has been observed at only one location, on the west coast of Ireland. At other coastal sites, individual frequency characteristics are observed that appear dependent on site location. The controlling regional influence of the deep ocean and shelf seas at periods in excess of 1000 s is noted and to some extent defined in the results obtained across a major portion of the UK land-mass.

When mainland results are compared on a regional basis, spatially consistent features emerge in both the real and imaginary induction arrows. The results demonstrate that the in-quadrature component of the anomalous vertical field, although often small (i.e. <0.1), can be determined with sufficient accuracy to provide information as characteristic and detailed as that of the in-phase component. When transfer function estimates are well-resolved, it is demonstrated that the frequency characteristics observed in the azimuth differences and phase of the anomalous vertical field can be used to characterise the response observed at a given location. In particular, the detailed behaviour of such parameters with frequency may indicate the degree of near-field 2-dimensionality. A characteristic period, predicted by simple 2-dimensional models, is clearly observed at a number of locations. It is pointed out that such a period is diagnostic of the structural parameters defining the geoelectric anomaly.

Key words: Electromagnetic induction – Anomalous vertical fields – Induction arrows – Frequency characteristics – British Isles – Observations – Models

Introduction

Estimates of the anomalous vertical field obtained from geomagnetic sounding experiments contain unique information on lateral contrasts in geoelectric structure. Such transfer function estimates are obtained as a function of frequency across a given bandwidth. The estimates then define frequency response curves at one or more locations. Over the past 15 years several geomagnetic deep sounding experiments conducted in the British Isles have been de-

scribed (Edwards et al., 1971; Hutton and Jones, 1980; Sik et al., 1981; Beamish and Banks, 1983). Across the same time interval data logging and analysis techniques have improved. Over the past 10 years the British Geological Survey has operated digitally recording rubidium and fluxgate magnetometers at a number of locations, in relation to a variety of field experiments. These data have been subjected to a common formatting and analysis technique that provides well-resolved transfer function estimates over the period range 10–10000 s. The results obtained define a set of characteristic frequency response curves from widely dispersed locations across Britain. The aim of this study is to summarise, present and comment on these results. It is hoped that the results presented will form a background to which further observations may be both compared and understood.

The spatially extensive data set considered by Edwards et al. (1971) covered the period range from 720–28800 s. The results obtained were interpreted on the basis of the regional influence exerted by the induced current systems in the ocean and shelf seas. Based mainly on results covering the period range from 2400–8640 s, the authors defined the six major current concentrations shown in Fig. 1, referred to as L1–L6. L1 was defined along the continental slope to the west of Ireland, current concentrations L2–L5 were located in loops within the shelf seas, while L6 defined a current concentration within the Southern Uplands of Scotland. Frequency-dependent effects were noted, particularly at sites in Ireland, as the period decreased and the relative strength of the current systems in the shelf seas increased. These long-period frequency characteristics were also observed in the analogue model study of the British Isles presented by Dosso et al. (1980). The authors considered scale model results equivalent to the period range 1200–7200 s and the results indicate that over this interval the anomalous vertical field may change by a factor of five or six between certain coastal locations, and at least by a factor of two or three between inland locations.

In the geomagnetic array study of Scotland (Sik et al., 1981; Kirkwood et al., 1981), transfer function results were presented down to a period of 300 s although no explicit use was made of their frequency characteristics. More recently, Banks and Beamish (1984) used the frequency characteristics observed across a dense network of sites in southern Scotland and northern England to investigate the mechanisms of local and regional induction. At periods greater than 2000 s, the transfer functions at the majority of sites

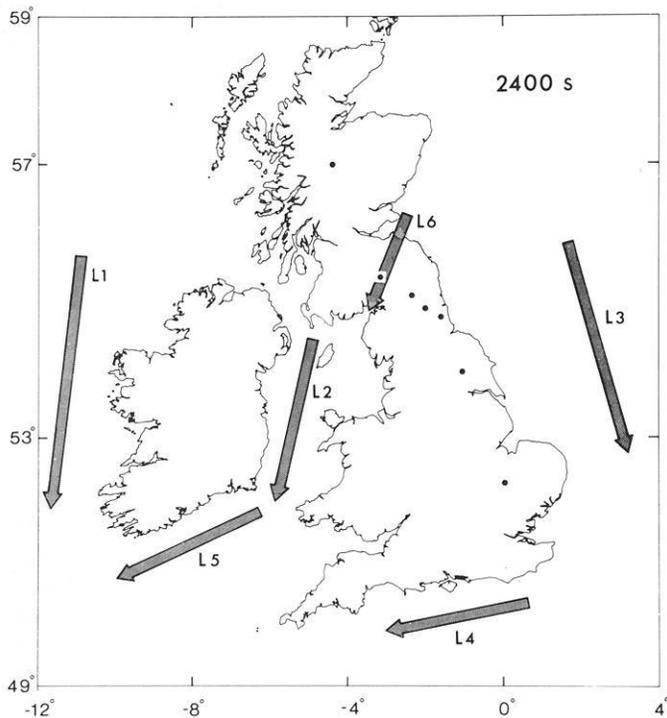


Fig. 1. Location of six current concentrations (L1–L6) in and around the British Isles, at a period of 2400 s. After Edwards et al. (1971)

are determined by current concentrations associated with the Atlantic ocean to the west and south-west. In the period range 400–2000 s, the transfer functions are determined by currents induced in a thin sheet of laterally varying conductance comprising shallow seas and both on- and off-shore sedimentary basins. At periods less than 200 s, the transfer functions are compatible with an induction process which is controlled by the local geological structure.

In the present study the results are necessarily divided into coastal and non-coastal observations. In each case the frequency response curves obtained are compared with the equivalent results obtained from simple models. General methods for the interpretation of anomalous fields using their frequency characteristics have been considered by Rokityansky (1975, 1982). The existence and properties of a characteristic period generated by simple geoelectric structures may be used to further characterise a frequency response curve. The accuracy and broadband nature of the response curves considered allows such characteristics to be examined in detail.

Transfer functions and their presentation

The data considered in the present study are single-site vertical-field transfer functions. These relate the vertical and horizontal components of the magnetic variation field as a function of frequency at a particular location. Although certain necessary assumptions must be made concerning single-site data (Schmucker, 1970; Banks, 1973), the transfer functions obtained from such data avoid the requirement for simultaneous operation of instruments and enable a basic response set to be generated using data col-

lected at different times. Such an advantage outweighs the limitations of the approach for the present comparative study. The geomagnetic data for this study were obtained from both fluxgate and rubidium vapour magnetometers, recording digitally on cassettes. The procedures used to determine the single-site transfer function (A , B) from such data are given by Beamish and Banks (1983). The method used provides band-averaged transfer function estimates over the three decades from 10 to 10000 s. These three decades are divided logarithmically into 18 overlapping period bands.

The magnitude of the transfer functions generally observed in the U.K. is small (e.g. $\ll 0.5$) and we require well-resolved estimates if detailed information is to be obtained. An important aspect of the present study is that the procedure adopted, together with the amount of data analysed, provides well-estimated transfer functions that are smooth functions of frequency. In particular, the phase information contained in the transfer function is accurately determined. The data have been subject to the same analysis procedure with the result that stable and accurate transfer functions can be presented and compared.

The single-site vertical-field transfer function consists of the complex pair of numbers (A , B) determined as a function of frequency or period band. The most common method of presentation is the induction arrow which allows the transfer function to be presented in terms of a magnitude and an azimuth which defines the normal to the strike of a local geoelectric lateral gradient giving rise to an anomalous concentration of current. Two such arrows are defined for vertical fields responding in-phase (real) and in-quadrature (imaginary) with the horizontal component with which the vertical field possesses maximum correlation. The magnitudes of the real and imaginary induction arrows are defined as

$$G_R = (A_R^2 + B_R^2)^{1/2}$$

$$G_I = (A_I^2 + B_I^2)^{1/2}, \quad (1)$$

where the subscripts R and I refer to the real and imaginary parts. The azimuths of the real and imaginary induction arrows are here defined as

$$\theta_R = \tan^{-1} - (B_R/A_R)$$

$$\theta_I = \tan^{-1} - (B_I/A_I) \quad (2)$$

clockwise from geomagnetic north. It can be noted that both real and imaginary induction arrows are reversed, in keeping with an implied time dependence of $\exp(i\omega t)$, (Lilley and Arora, 1982). The induction-arrow method of presentation has a simple interpretation only when the structure is 2-dimensional. In such circumstances, given suitable ratios of conductivity contrast and frequency, the arrows will be perpendicular to the conductivity contrast.

When the transfer function possesses an intermediate phase, the separation into real and imaginary components can be artificial. The maximum and minimum response function presentation of Banks and Ottey (1974) is aimed at separating from (A , B) that part of the transfer function that is compatible with a 2-dimensional assumption. The presentation again requires four parameters consisting of the maximum and minimum response magnitudes (G_M and G_L), together with the azimuth and phase (θ_M , ϕ_M) of the maximum response. The three maximum response param-

ters G_M , θ_M , ϕ_M , then define that part of the response that is compatible with a 2-dimensional assumption. An example of the eight parameters that can be obtained from the transfer function are shown in Fig. 4, as a function of frequency.

The frequency response of two simple models

Prior to the presentation and discussion of the observed frequency characteristics, it is worthwhile considering the frequency characteristics that can be derived from simple geoelectric models. When considering the transfer function data set across the U.K. it is necessary to subdivide the site locations into coastal and non-coastal (i.e. mainland) sites. In so doing, it is useful to consider two broad categories of models that have found applications in these two situations. To examine the frequency response that may be observed at coastal sites, a typical thin-sheet geoelectric model is considered. To examine the frequency response that may be observed at mainland sites, a typical 2-dimensional geoelectric model with an embedded anomaly is considered.

The geomagnetic coast effect, reviewed by Parkinson and Jones (1979), is commonly related to the edge, or near, effect produced by electromagnetic induction in an electrically thin, conducting sheet (or shell), taken to represent the ocean. The thin-sheet approximation requires two limiting conditions on the electrical thickness of the surface sheet (Schmucker, 1970, 1971). These conditions impose a lower bound on the variation periods for which the thin-sheet approximation may be used. For the deep (4 km) ocean, the lower period bound lies in the range 900–1000 s (Schmucker, 1971; Weaver, 1979). Solutions to a thin-sheet model can therefore only provide “one-sided” solutions at periods greater than the lower bound.

A number of analytic and numerical procedures have been developed to study the edge-effect of induced currents in thin-sheets (Hewson-Browne and Kendall, 1976; Fischer, 1979). We here consider a solution obtained using a matching technique suggested by Hewson-Browne and Kendall (1976) and applied by Quinney (1979). The model consists of a half-plane of finite conductivity (the ocean) underlain by, and insulated from, a region of perfect conductivity whose upper surface lies at a depth of $P/2$. P is the depth of the image currents induced in the perfect conductor which simulates the conducting Earth. The solutions for the electric field obtained by Quinney (1979) together with the resulting magnetic fields constitute universal curves, being a function of distance x from the ocean edge, and depend only on the parameter

$$\gamma = 0.5\omega\kappa\mu_0 P, \quad (3)$$

where ω is angular frequency, κ is the conductance of the surface half-plane (a constant), μ_0 is the permeability of free space and P is defined above.

Solutions to this problem using $\kappa = 16000$ mho (i.e. for an ocean of depth 4 km) and $P = 1000$ km have been obtained by D. Quinney for the period range corresponding to the observational bandwidth of the present study. The anomalous vertical field provided by the solution is divided by the total horizontal field to provide a direct comparison with observation. The in-phase and in-quadrature ratios, equivalent to the induction-arrow magnitudes G_R and G_I ,

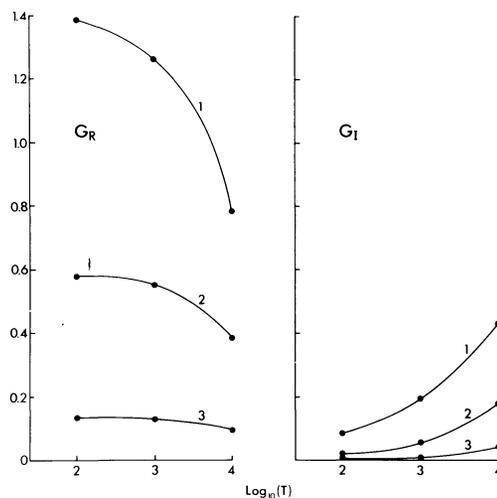


Fig. 2. Induction-arrow magnitudes as a function of period T (s) and distance (x) from the ocean edge, computed from a thin-sheet model. G_R is real magnitude and G_I is imaginary magnitude. Curve 1: $x=0.1 P=100$ km. Curve 2: $x=0.5 P=500$ km. Curve 3: $x=1.0 P=1000$ km

at three locations inland from the ocean edge are shown in Fig. 2. The ratios obtained for (curve 1) $x=0.1 P$ (100 km) can be considered a near-field effect while the ratios obtained for (curve 3) $x=1.0 P$ (1000 km) are a far-field effect of the current concentration at the ocean edge. The frequency characteristics observed for the ratios are similar to the corresponding results obtained for the Roden strip of finite width (Roden, 1964) presented by Edwards et al. (1971, Fig. 23). In the present case, the results have been obtained to much shorter periods. At short periods ($T < 1000$ s), the solution provides an upper bound in the in-phase response G_R and a lower bound in the in-quadrature response G_I . With increasing period ($T > 1000$ s), a characteristic asymptotic frequency response is observed which is a function of distance from the ocean edge. As noted previously, for a 4-km ocean the results obtained for a thin-sheet model apply for $T > 900$ s, the equality expressing the likely lower bound for quasi-static induction in the deep ocean. At shorter periods, as $T \rightarrow 0$, the induced tangential electric field within the ocean suffers rapid attenuation and the effects of self and mutual induction must be taken into account. The characteristics shown in Fig. 2 reveal that the far-field effect of deep ocean induction may be discernable in G_R (i.e. $G_R = 0.1$) as far as $x = 1000$ km.

The solutions provide the constraint that $G_R > G_I$ over the observational bandwidth. Perhaps the most characteristic feature of the response occurs at the lower bound for quasi-static induction. At this period, the anomalous vertical field will be a maximum and will be predominantly in-phase. An additional feature of the solutions obtained is that the ratios obtained are universal curves depending on the parameter γ . As can be seen from Eq. (3), with P a constant equivalence is obtained by retaining $\omega\kappa$ as a constant. We may, in fact, displace all curves one decade lower in period (i.e. to the left) for a sea of depth 0.4 km and the same conductivity. Such curves therefore allow a general assessment of the frequency characteristics of anomalous vertical fields in both the near- and far-fields. The results provided are inevitably one-sided and can only be used to obtain the asymptotic frequency response from a

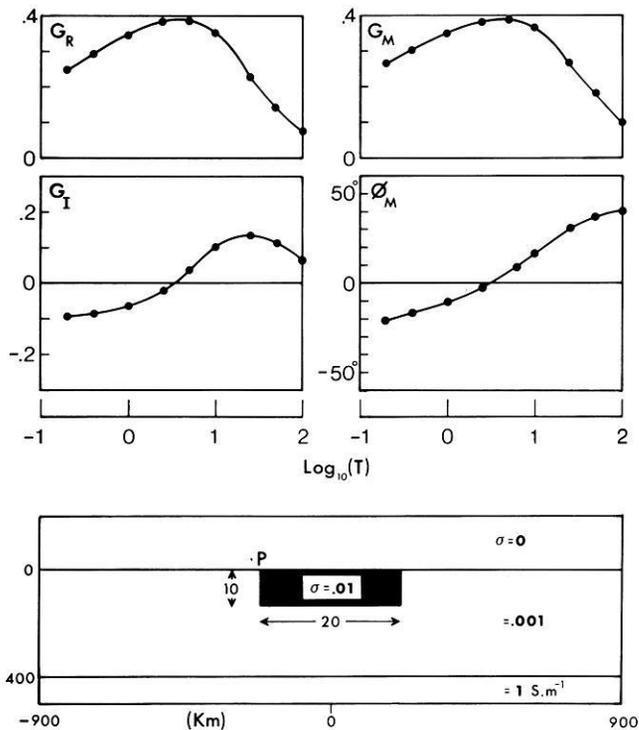


Fig. 3. Two-dimensional model of conducting anomaly (*lower diagram*). *Upper four diagrams* are frequency response curves of the anomalous vertical field computed at point P. G_R , G_I are the amplitudes of the real and imaginary induction arrows. G_M and ϕ_M are the magnitude and phase of the maximum response (see text)

lower period bound. The parameter specification of the model, however, makes it one of the simplest that can be applied in the study of the frequency dependence of the anomalous vertical field.

The thin-sheet models discussed above have a limited application when considering the frequency response of the lower two decades (10–1000 s) at sites in mainland U.K. To examine broader-band frequency characteristics we consider instead a simple 2-dimensional numerical model. For our present purposes the simplest model consists of an isolated conducting anomaly embedded in a half-space. Reviews of both analytic and numerical solutions to such problems are provided by Jones (1973), Kaufman and Keller (1981) and Rokityansky (1982). The model considered is defined in Fig. 3, solutions being obtained by a form of the finite difference method (e.g. Pascoe and Jones, 1972; Brewitt-Taylor and Weaver, 1976). The frequency dependence of the four parameters G_R , G_I , G_M and ϕ_M , evaluated at point P on the surface is shown in Fig. 3.

The main frequency characteristic of the anomalous vertical field derived from such models is a change of sign in the phase of the vertical field at the frequency where the maximum response is observed. This phase change is a general result of both analytic and numerical models (e.g. Rokityansky, 1982). The phase change can be observed in terms of the induction-arrow azimuths or possibly more directly in terms of the phase of the maximum response. For a 2-dimensional structure the real and imaginary induction-arrow azimuths are strictly parallel or antiparallel. Using the transfer function parameters defined previously, two frequency-dependent conditions are imposed on the parameters derived from the vertical-field transfer function. Let

T_0 be the period at which the maximum response is observed, then for $T < T_0$, the azimuths of the real and imaginary induction arrows are antiparallel (they differ by 180°) and the phase of the maximum response is negative. For $T > T_0$, the azimuths of the real and imaginary induction arrows are parallel and the phase of the maximum response is positive. At T_0 and for a strictly 2-dimensional structure, the phase of the maximum response is zero and the azimuth of either the real or imaginary induction arrow will change by 180° depending on whether the maximum is observed in G_R or G_I . Clearly, in assessing the degree of 2-dimensionality, a consideration of the difference $\theta_R - \theta_I$ should be a useful exercise, while the condition $\phi_M = 0$ provides a better assessment of T_0 .

The frequency response at coastal sites

Transfer function results over three decades are available at five coastal sites, four of these sites lie within the British Isles while the fifth site is on Faroe (FA). One of the sites, Valentia (VA), is situated on the west coast of Ireland and was considered by Parkinson (1962) to be a 'normal' coast site in that the vertical field is strongly influenced by the west-European continental shelf. The frequency response of the eight parameters that may be used to characterise the transfer function are shown in Fig. 4, for VA. The upper four curves define the magnitude and azimuth of the two induction arrows while the lower four curves define the maximum and minimum response parameters. It can be noted in Fig. 4 that since the response is largely in-phase, i.e. $G_R > G_I$, there is a high degree of equivalence between the upper and lower methods of presentation.

Our purpose here is to summarise the frequency dependence of these parameters at the five widely separated coastal sites shown in Fig. 5 in relation to the bathymetry of the continental margin. The coastal sites are denoted as FA, LE, DU, HA and VA. We first restrict our attention to the estimates obtained from the first five long-period bands of the analysis. These bands cover the period range 7200–900 s. The azimuthal information contained in the transfer function is best considered using θ_M , the azimuth of the maximum response considered in isolation from the additional complexities of amplitude and phase. The azimuths obtained for the period range 7200–900 s at the five coastal sites, together with four mainland sites shown for comparison, are displayed in Fig. 5. The sense of rotation with decreasing period is indicated by arrows.

It is quite evident from Fig. 5 that the long-period anomalous vertical fields are strongly influenced by induced current systems within both the deep ocean and shelf seas. Despite the lack of deep ocean between the west Shetland (LE) and Faroe (FA) platforms, a concentration of current exists in the shelf seas (<1 km) separating the two islands. It is also apparent that, while the azimuths at the coastal sites are relatively constrained over the bandwidth considered, the azimuths at three of the inland sites (ES, YO and CA) rotate through a large angle, from the NW to the NE quadrant, with decreasing period. The consistent rotation pattern displayed at the six sites LE, DU, LL, ES, YO and CA is indicative of a large-scale effect extending over a large part of mainland Britain and has been commented upon by Banks and Beamish (1984). At 10000 s the anomalous fields are dominated by the current system in the deep ocean to the west and SW. As the period de-

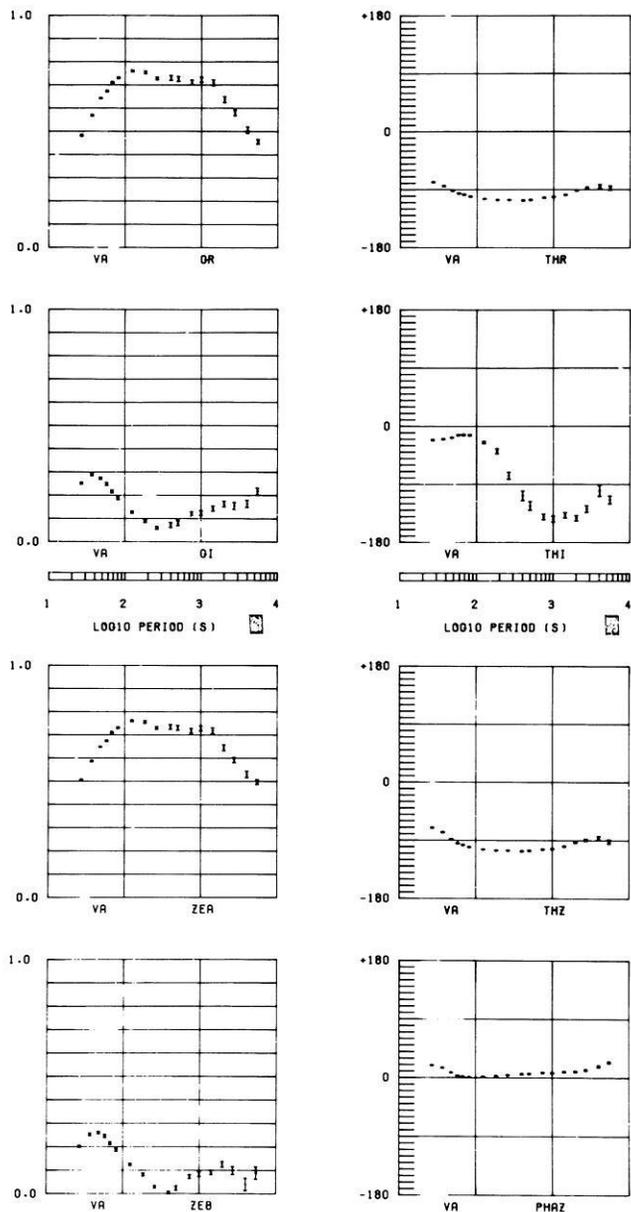


Fig. 4. Frequency response of the anomalous vertical field at Valentia VA. *Upper four diagrams* are induction-arrow parameters. G_R , THR : magnitude and azimuth of the real induction arrow. G_I , THI : magnitude and azimuth of the imaginary induction arrow. *Lower four diagrams* are the maximum/minimum response parameters. ZEA , ZEB : magnitude of the maximum and minimum response. THZ : azimuth of the maximum response. $PHAZ$: phase of the maximum response

creases, the relative importance of currents to the east in the North Sea increases causing the observed rotations.

We now consider the frequency response observed in the magnitude of the anomalous vertical field at the same five coastal sites. The frequency response is presented in terms of the induction-arrow amplitudes G_R and G_I . The results for VA are shown in Fig. 4 and the results for the other four sites are plotted on a common scale in Fig. 6. It is apparent that the frequency response curves observed at each site are highly individual. Before considering their general characteristics we must note from Fig. 5 the complex distribution of offshore bathymetric units which reflect

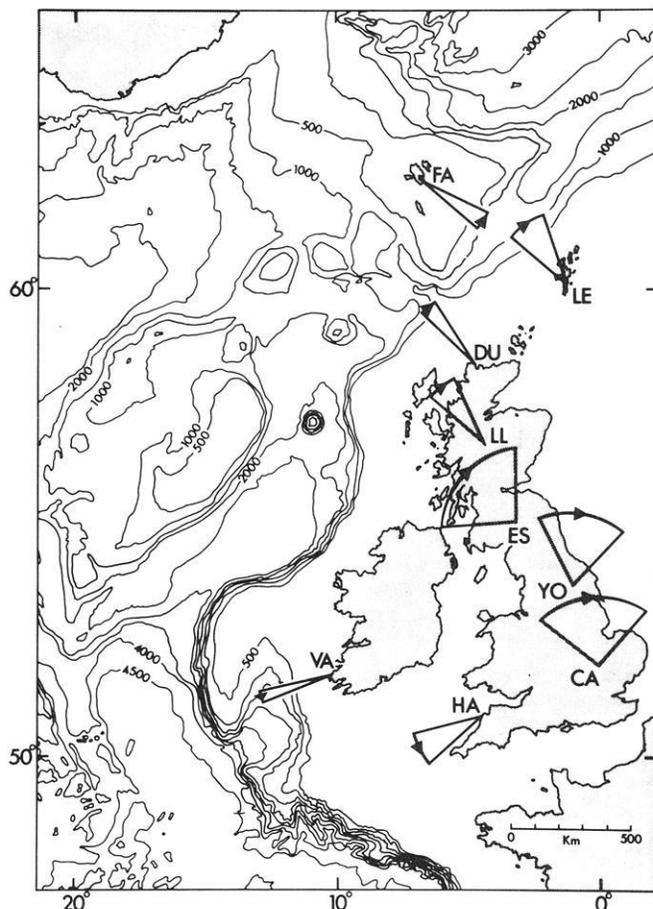


Fig. 5. Bathymetry of the west European continental margin contoured at an interval of 500 m. Also shown are the azimuths of the maximum response at 7200 and 900 s, at nine locations. The arrows denote the sense of rotation with decreasing period

the presence of both oceanic crust and continental fragments in the region (Laughton and Roberts, 1978). A relatively straightforward continental slope from deep ocean may only be defined to the SW of Britain, below a latitude of 55° N. In terms of the likely influence of induced currents in the deep ocean in the near-field (i.e. $x < 500$ km), the three northern sites (DU, LE and FA) must be viewed as distinct from the two southern sites (VA and HA).

At the two southern sites, VA and HA, the frequency response is characterised by the occurrence of a maximum response in the period range 100–1000 s. In the vicinity of the maximum, the imaginary response displays a minimum, giving rise to a maximum response that is largely in-phase. For these two sites we may then define a long-period ‘asymptotic’ response curve for periods in excess of 1000 s. The site VA on the west coast of Ireland is the only site within 500 km of the deep ocean. The long-period response from this site may therefore be compared with the results obtained from the thin-sheet model for an ocean of depth 4 km, discussed previously. The comparison, in terms of G_R and G_I , is shown in Fig. 7 for model results obtained at a distance of $x = 200$ km from the ocean edge. This distance roughly corresponds to the distance between the site and the shelf-sea margin. Given the simplicity of the model, a quantitative comparison is not warranted. However, the correct long-period asymptotic behaviour is clearly displayed by the observations. The level of corre-

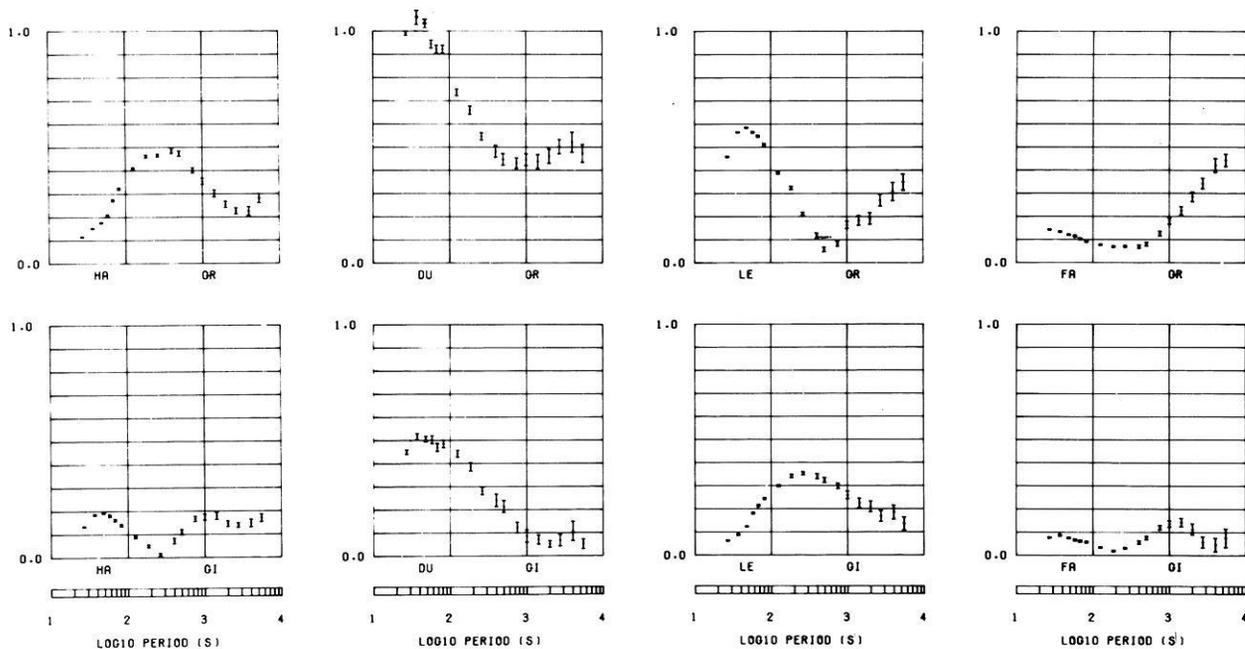


Fig. 6. Frequency response curves at four of the coastal sites shown in Fig. 5. G_R : magnitude of the real induction arrow. G_I : magnitude of the imaginary induction arrow

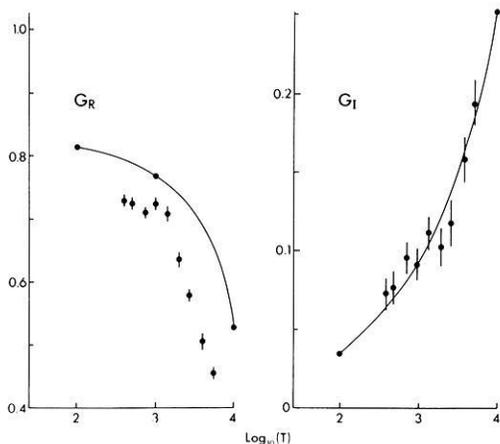


Fig. 7. Comparison of observed and calculated real (G_R) and imaginary (G_I) response magnitudes at VA. The *solid line* is calculated from a thin-sheet model of the deep ocean (4 km) at a distance of 200 km from the ocean edge

spondence shown in Fig. 7 is not displayed at any other of the coastal sites, as anticipated given the bathymetric information of Fig. 5.

Turning now to the response curves at the three northern coastal sites DU, LE and FA (Fig. 6), we observe very different and distinct frequency characteristics at each of the three locations. The only common feature appears to be a minimum in the G_R response in the period range 100–1000 s, at each of the three sites. The existence of a short-period (<100 s) maximum in G_R may indicate the influence of an induction process taking place in the shallow shelf seas. However, the asymptotic behaviour of G_I does not generally exhibit the predicted behaviour. The very different frequency response curves at periods less than 1000 s observed on the two shelf-sea islands of Shetland (LE) and Faroe (FA) is worthy of note.

The frequency characteristics at mainland sites

The single-site transfer functions within mainland Britain are far more numerous than at coastal sites. In order to examine spatially representative frequency characteristics it is necessary to subdivide the existing data set into regional data sets. Four representative regions have been chosen. The four regions cover that part of Britain lying between the Grampian Highlands of Scotland and northern England. To display each regional data set the simplest method of presentation is used, i.e. an overlay plot of individual results within each region. When the transfer function estimates are well-resolved this simple approach is adequate. The four parameters used for display are the magnitude and azimuth of the two induction arrows.

The four representative regions (R1–R4), together with the individual site locations, are shown in Fig. 8. It is worth noting that the regions cover two of the main structural elements of Caledonian Britain. The sites within region 1 (R1) lie within the Caledonian metamorphic foreland. The sites within the three regions to the south lie roughly within the Southern Uplands (R2), the Northumberland Basin (R3) and the Alston Block of northern England (R4). The combined results for each of the four regions are displayed in Fig. 9. The results are plotted on a common scale and define, quite clearly, the limiting bounds on the magnitude of the frequency response observed over a considerable area of northern Britain. The range of upper bounds in both real and imaginary magnitudes are lower by at least a factor of two when compared to the range of upper bounds observed at the coastal sites presented in Fig. 6.

Considering the complete frequency response curves for the four regions, it is apparent that the response characteristics of regions R2 and R3, the Southern Uplands and Northumberland Basin, exhibit common features. Regions R1 and R4 exhibit separate and distinct response characteristics to R2, R3. Within all four regions we observe an important frequency characteristic in the two azimuthal re-

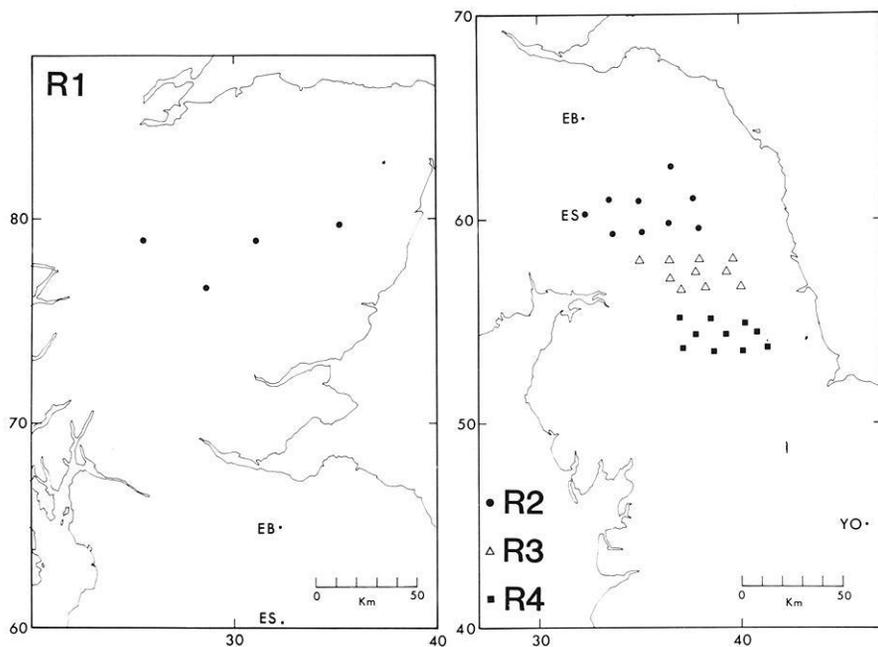


Fig. 8. Site locations within the four regions (R1–R4) referred to in the text. Coordinate numbers refer to the National Grid (ten units = 100 km). Geographical coordinates are provided in Fig. 5

sponses. With decreasing period, below 1000 s, the azimuths rotate to define local, rather than regional, strike directions. The rotation to predominantly local azimuths, increases with decreasing period. The obvious exception to this is θ_R in region 4 (R4) which is highly constrained. Region 1 is characterised from all other regions by its long-period (>1000 s) imaginary response G_I , θ_I . At periods in excess of 1000 s the imaginary azimuth appears consistent at the four widely separated sites. Again it is the imaginary response G_I , θ_I that characterises the frequency response in regions 2 and 3. In these two regions both the magnitude and azimuth of the imaginary response appear to be consistently constrained in the period range 100–1000 s. The response in region 4 is characterised from the above three regions by both the real and imaginary parts of the response. The upper and lower bounding values in the magnitude of G_R is quite apparent, as is the constraint on θ_R at periods less than 1000 s. A second characteristic feature of region 4 is the short-period (60 s) maximum observed in the imaginary component. The induced current systems responsible for the anomalous vertical field characteristics observed in region 4 are considered by Beamish and Banks (1983) and Banks and Beamish (1984).

Clearly the rotation patterns from the NW to the NE quadrant, observed across regions 1 and 4 at periods in excess of 1000 s, confirm the large-scale effect of thin-sheet currents in the deep ocean and shelf seas noted previously. The lack of such a consistent rotation in regions 2 and 3 is likely to be due to the degree of attenuation experienced by the anomalous vertical field across these two regions. With the exception of region 4, which is in the near-field of regional east-west current flow, the rotation to local azimuths with decreasing period below 1000 s is clearly observed. Broadly, the results presented support the hypothesis proposed by Banks and Beamish (1984) that the anomalous vertical field at mainland sites responds principally to *local* geological structure only at periods less than 200 s.

The frequency characteristics considered in detail

The frequency characteristics observed in Fig. 9 define a set of regionally representative anomalous vertical-field estimates, presented in terms of induction arrows. In order to consider the characteristics in more detail we next reconsider the general frequency behaviour obtained previously from the simple 2-dimensional model. For the isolated conducting anomaly of Fig. 3, the existence of a characteristic period T_0 was observed. At this period, a maximum anomalous vertical field is observed. Across such a period, the imaginary response together with the phase of the vertical field pass through zero and the azimuth difference ($\theta_R - \theta_I$) changes by 180° . As noted previously, such 2-dimensional behaviour is most readily assessed using the magnitudes of the induction arrows together with the azimuth difference ($\theta_R - \theta_I$) and the phase of the maximum response. Obviously, $(\theta_R - \theta_I) = 0$ defines parallel induction arrows while $(\theta_R - \theta_I) = 180^\circ$ defines antiparallel induction arrows. This method of presentation is displayed for a selection of four locations in Fig. 10a–d. The results for each of the four sites are now discussed in turn.

Figure 10a shows the frequency characteristics at York (YO, Fig. 5) near the eastern coast of England. The real magnitude (G_R) exhibits two maxima with associated non-zero minima in the imaginary magnitude (G_I). The lack of true zeros in G_I implies a departure from near-field 2-dimensionality. If we define two characteristic periods (T_0) at the position of the above maxima, we observe that the phase passes through zero from positive values for $T > T_0$ to negative values for $T < T_0$, as the model predicts. The behaviour of the azimuth difference, however, displays a smooth rotation through a large azimuth with frequency again indicating the absence of strong near-field 2-dimensionality.

The frequency characteristics observed in Fig. 10b at the coastal site of Hartland (HA, Fig. 5) are simpler to

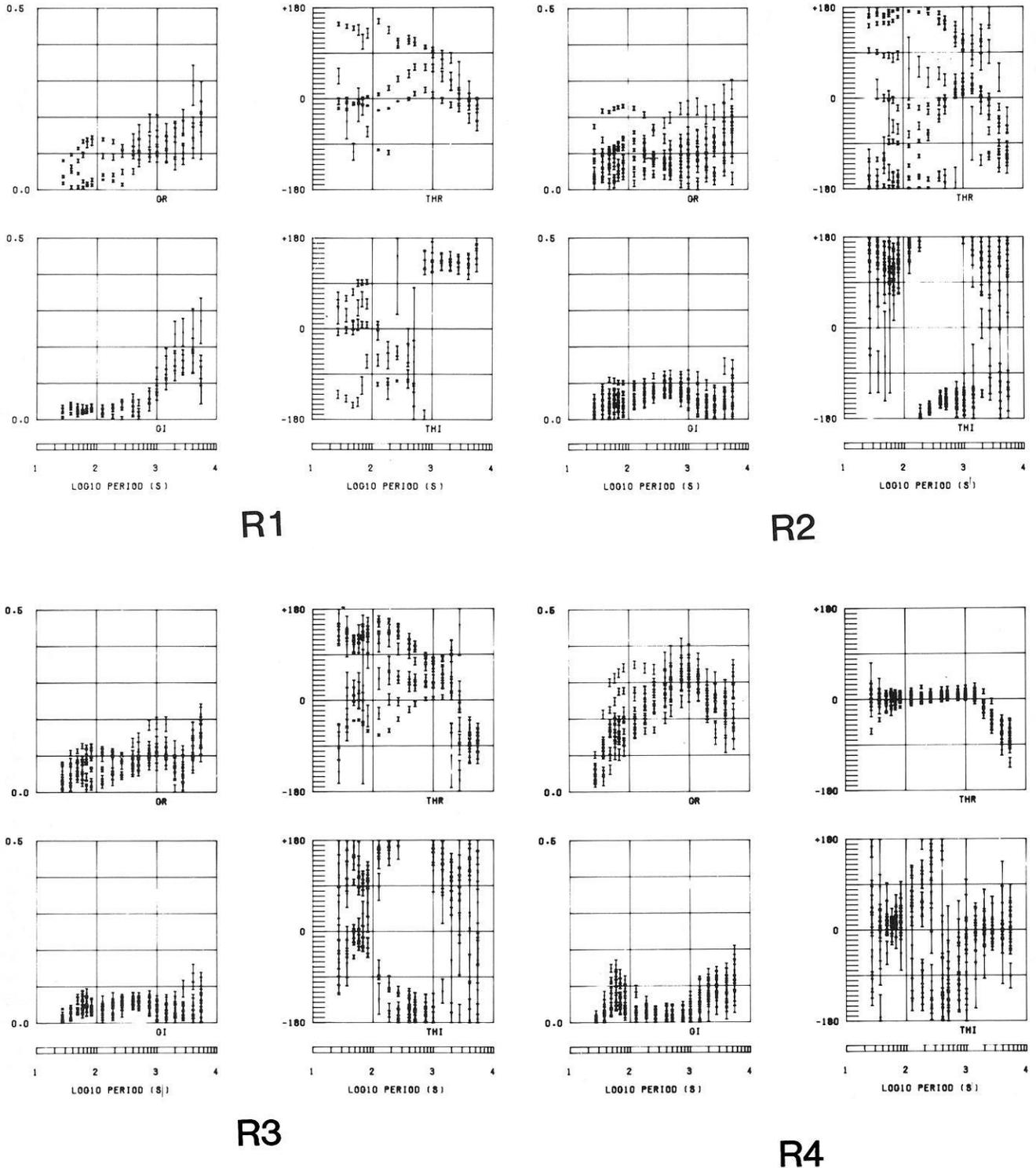


Fig. 9. Frequency response curves for the four regions R1–R4. Induction-arrow results overlaid. G_R , G_I : induction-arrow magnitudes. THR , THI : induction-arrow azimuths. (R1) 4 sites R1. (R2) 9 sites R2. (R4) 10 sites R3. (R3) 10 sites R4

interpret in that the location is truly coastal. A single maximum exists in the real magnitude with a corresponding true zero in the imaginary response. At periods just in excess of the characteristic period T_0 , the phase is positive and $(\theta_R - \theta_I)$ is zero. At periods slightly less than T_0 , the phase is negative and the azimuth difference approaches -180° . Thus, at periods close to T_0 , the characteristic behaviour of a simple 2-dimensional conducting anomaly is observed.

At periods removed from T_0 , the response departs from such simple behaviour which is not surprising given the complexity of the 'offshore conductor'.

The second two examples are taken from mainland locations and involve quite distinct 2-dimensional near-field effects. Figure 10c shows the frequency characteristics observed at Earlyburn (EB, Fig. 8) located in the vicinity of a major crustal Caledonian feature, the Southern Uplands

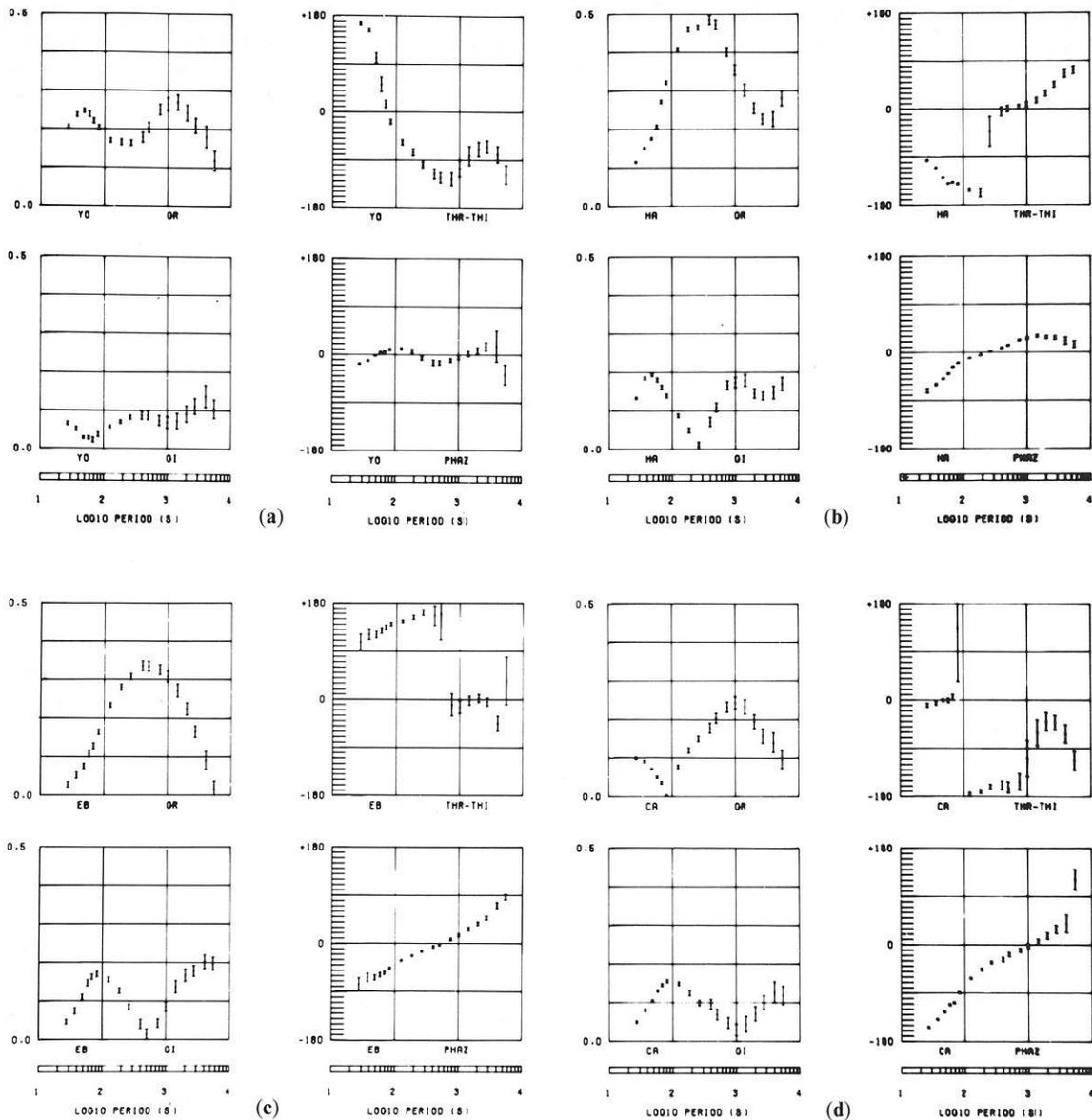


Fig. 10a-d. Frequency response curves at four locations. G_R , G_I : induction-arrow magnitudes. $THR-THI$: azimuth difference between real (G_R) and imaginary (G_I) induction arrows. $PHAZ$: phase of the maximum response. **a** Site YO. **b** Site HA. **c** Site EB. **d** Site CA

Fault. The distinct maximum observed is, in fact, the largest mainland response observed in the current data set. The features observed in all the parameters closely resemble the 2-dimensional effects of a single isolated conducting anomaly observed in the near-field. The 2-dimensional condition on the azimuth difference appears more consistent at longer rather than shorter periods.

Figure 10d shows the frequency characteristics observed at Cambridge (CA, Fig. 5). The long-period maximum at 1000 s owes its existence to off-shore induced currents as can be seen by comparing the response parameters observed in Fig. 10a and d and by noting the previously defined rotational properties at these sites at long periods. The interesting feature of Fig. 10d is the short-period maximum observed in the imaginary component and associated true zero in the real component. If we define a characteristic period T_0 , at this maximum in G_I , we note that the phase passes through -90° and the azimuth difference transfers from -180° to 0° with decreasing period. Such behaviour conforms to that predicted for an isolated 2-dimensional

resistive anomaly observed in the near-field. The short-period behaviour observed at this site is unique among the current data set.

Discussion

Following from the above work, it seems clear that in addition to providing information on the mechanisms of local and regional induction, the existence of a characteristic period observed within the broad-band frequency response curves goes some way to identifying the characteristic anomalous vertical field observed at a particular location. The use of the parameter pair consisting of phase and azimuth difference has been shown to be particularly effective in this regard. The existence of near- and far-field effects together with their interaction as a function of frequency will generally distort any simple concept of isolated 2-dimensionality. Despite the likely influence of such real Earth complexity, three of the frequency response curves presented in Fig. 10 display an adequate correspondence to

the behaviour anticipated from a single, isolated 2-dimensional anomaly. When such effects are observed it would be useful to establish some general relationships between the characteristic period observed and the anomaly from which they derive.

A large amount of Soviet work concerning the behaviour of anomalous magnetic field characteristics due to simple 2-dimensional structures is summarised by Rokityansky (1982). The type of models considered are insertion and graben surface structures together with immersed (i.e. buried) elliptical cylinders, the frequency response curves being calculated using analytical methods. It is clearly difficult, given the non-linearity of the problem, to obtain general results from even simple models. However, given the large number of models considered, Rokityansky (1982) considers general interpretational methods which may be applied in given situations. In particular, various estimation techniques to obtain the longitudinal conductivity ($G = a \cdot b \cdot \sigma$ mho·m), defined by the cross-sectional area ($a \cdot b$) of an anomalous body, are considered. For observations at a point at which the frequency response characteristics have been used to obtain the characteristic period T_0 , Rokityansky (1982, eq. 6.63) suggests the empirical relationship

$$G = 5 \times 10^4 (T_0)^{1.2} \quad (4)$$

may be used to estimate longitudinal conductivity. The above relationship is the most probable result with maximum discrepancies occurring at a period of 400 s. For buried anomalies and structure with finite horizontal dimensions, the estimate obtained from the above will be an underestimate. From the strongly 2-dimensional frequency characteristics observed at EB (Fig. 10c), T_0 is estimated as 700 s. Using the above, G is estimated as 129.74 mho·m. If we take the dimensional parameters for the Southern Uplands anomaly to be $a = 10$ km and $b = 100$ km, we obtain an estimate of the anomalous conductivity of 0.13 mho/m which is not unrealistic (Beamish, submitted for publication). When a characteristic period (T_0) is observed at a coastal site, Rokityansky (1982, eq. 7.1) suggests an empirical formula relates T_0 to the sea depth (h):

$$T_0 = 250 \cdot (h)^{1.7} \quad (5)$$

with T_0 in seconds and h in km. The characteristic period observed at the coastal site of HA (Fig. 10b) is estimated to be 200 s. Using the above relationship, the sea depth h is estimated to lie between 700 and 800 m. Such an estimate, bounded by the actual depths of shelf sea and deep ocean, seems reasonable. Since most of the published work concerns conductive (cf. resistive) anomalies, the characteristic short-period in-quadrature anomaly observed at CA (Fig. 10d) remains, for the present, an intriguing result.

Acknowledgements. I would like to thank Doug Quinney for providing the solutions for the thin-sheet ocean model used in this study. This paper is published with the approval of the Director, British Geological Survey (NERC).

References

- Banks, R.J.: Data processing and interpretation in geomagnetic deep sounding. *Phys. Earth Planet. Inter.* **7**, 339–348, 1973
 Banks, R.J., Beamish, D.: Local and regional induction in the British Isles. *Geophys. J.R. Astron. Soc.* **79**, 539–553, 1984
 Banks, R.J., Ottey, P.: Geomagnetic deep sounding in and around the Kenya rift valley. *Geophys. J.R. Astron. Soc.* **36**, 321–335, 1974

- Beamish, D.: Deep crustal geoelectric structure under the Northumberland Basin. *Geophys. J.R. Astron. Soc.*, submitted, 1985
 Beamish, D., Banks, R.J.: Geomagnetic variation anomalies in northern England: processing and presentation of data from a non-simultaneous array. *Geophys. J.R. Astron. Soc.* **75**, 513–539, 1983
 Brewitt-Taylor, C.R., Weaver, J.T.: On the finite difference solution of two-dimensional induction problems. *Geophys. J.R. Astron. Soc.* **47**, 375–396, 1976
 Dosso, H.W., Nienaber, W., Hutton, V.R.S.: An analogue model study of electromagnetic induction in the British Isles Region. *Phys. Earth Planet. Inter.* **22**, 68–85, 1980
 Edwards, R.N., Law, L.K., White, A.: Geomagnetic variations in the British Isles and their relationship to electrical currents in the ocean and shallow seas. *Phil. Trans. R. Soc. Lond.* **A270**, 289–323, 1971
 Fischer, G.: Electromagnetic induction effects at an ocean. *Proc. IEEE*, **67**, 1050–1060, 1979
 Hewson-Browne, R.C., Kendall, P.C.: Magnetotelluric modelling and inversion in three-dimensions. *Acta Geodaet., Geophys. et Montanist.* **11**, 427–446, 1976
 Hutton, V.R.S., Jones, A.G.: Magnetovariational and magnetotelluric investigations in S. Scotland. *Adv. Earth Planet. Sci.* **9**, 141–150, 1980
 Jones, F.W.: Induction in laterally non-uniform conductors: theory and numerical models. *Phys. Earth Planet. Inter.* **7**, 282–293, 1973
 Kaufman, A.A., Keller, G.V.: The magnetotelluric sounding method. *Methods in geochemistry and geophysics.* **15**, Amsterdam-Oxford-New York: Elsevier 1981
 Kirkwood, S.C., Hutton, V.R.S., Sik, J.: A geomagnetic study of the Great Glen Fault. *Geophys. J.R. Astron. Soc.* **66**, 481–490, 1981
 Loughton, A.S., Roberts, D.G.: Morphology of the continental margin. *Phil. Trans. R. Soc. Lond.* **A290**, 75–85, 1978
 Lilley, F.E.M., Arora, B.R.: The sign convention for quadrature Parkinson arrows in geomagnetic induction studies. *Rev. Geophys. Space Phys.* **20**, 513–518, 1982
 Parkinson, W.D.: The influence of continents and oceans on geomagnetic variations. *Geophys. J.R. Astron. Soc.* **6**, 441–449, 1962
 Parkinson, W.D., Jones, F.W.: The geomagnetic coast effect. *Rev. Geophys. Space Phys.* **17**, 1999–2015, 1979
 Pascoe, L.J., Jones, F.W.: Boundary conditions and calculations of surface values for the general two-dimensional electromagnetic induction problem. *Geophys. J.R. Astron. Soc.* **27**, 179–193, 1972
 Quinney, D.A.: A note on computing coastal edge corrections for induced oceanic electric fields. *Geophys. J.R. Astron. Soc.* **56**, 119–126, 1979
 Roden, R.B.: The effect of an ocean on magnetic diurnal variations. *Geophys. J.R. Astron. Soc.* **8**, 375–388, 1964
 Rokityansky, I.I.: The interpretation of anomalous fields by using their frequency characteristics. *Phys. Earth Planet. Inter.* **10**, 271–281, 1975
 Rokityansky, I.I.: Geoelectromagnetic investigation of the Earth's crust and mantle. Berlin-Heidelberg-New York: Springer-Verlag 1982
 Schmucker, U.: Anomalies of geomagnetic variations in the south-western United States. *Bull. Scripps Inst. Oceanogr., Monograph.* **13**, 1970
 Schmucker, U.: Interpretation of induction anomalies above non-uniform surface layers. *Geophysics* **36**, 156–165, 1971
 Sik, J.M., Hutton, V.R.S., Dawes, G.J.K., Kirkwood, S.C.: A geomagnetic variation study of Scotland. *Geophys. J.R. Astron. Soc.* **66**, 491–512, 1981
 Weaver, J.T.: Electromagnetic induction in thin sheet conductivity anomalies at the surface of the Earth. *Proc. IEEE*, **67**, 1044–1050, 1979

Received March 1, 1985; Revised June 10, 1985

Accepted June 10, 1985

On the origin of the cusp field-aligned currents

L.C. Lee, J.R. Kan and S.-I. Akasofu

Geophysical Institute, University of Alaska, Fairbanks, Alaska 99775-0800, USA

Abstract. A complicated system of field-aligned currents is known to exist in the high-latitude region near noon, associated with the cusp. We suggest that the equatorward part of this system, referred to as the Region 1 field-aligned currents, is caused by the leakage of the field-aligned currents associated with the rotational discontinuities at the dayside magnetopause. The poleward part, referred to as the cusp field-aligned currents, is associated with the tail magnetopause. In this situation, it can be shown that the direction of the field-aligned currents at the magnetopause is controlled by the y -component of the interplanetary magnetic field (B_y). The poleward (equatorward) part of this field-aligned current system is found to flow out of (into) the northern polar ionosphere when $B_y > 0$ and into (out of) the northern ionosphere when $B_y < 0$. This current pattern reverses systematically in the southern polar ionosphere. Therefore, the suggested mechanism can explain qualitatively the observed changes of the cusp current systems. Further, the latitudinal width of the cusp field-aligned current system at the ionospheric altitude is estimated to be 100–400 km, consistent with observations.

Key words: Rotational discontinuity – Field-aligned currents

Introduction

Large-scale field-aligned currents or Birkeland currents in the cusp region have been reported by Iijima and Potemra (1976), McDiamid et al. (1978, 1979), Wilhelm et al. (1978), Iijima et al. (1978), Rostoker (1980), Doyle et al. (1981) and Bythrow et al. (1982). These cusp field-aligned currents are located between 78° and 81° invariant latitude and between 1000 MLT and 1400 MLT. From a data base that encompassed all orientations of the interplanetary magnetic field (IMF) vector, Iijima and Potemra (1976) found that these currents were determined as flowing into the ionosphere in the post-noon and away from the ionosphere in the pre-noon MLT hours. However, Doyle et al. (1981) showed that at any given time the cusp currents exist only on one side of the noon-midnight axis. In particular, they showed that when the B_y component of the interplanetary magnetic field (IMF) is negative ($B_y < 0$), the cusp field-aligned currents are found to flow mainly into the post-

noon sector of the northern polar ionosphere and flow out of the pre-noon sector of the southern polar ionosphere. When $B_y > 0$, the cusp field-aligned currents flow mainly out of the pre-noon sector of the northern polar ionosphere and into the post-noon sector of the southern polar ionosphere.

The earlier and later analyses of both the satellite and ground magnetic-field data over the northern and southern polar regions indicated the existence of a pair of field-aligned current sheets flowing in opposite directions, and also revealed a close correlation between the flow direction of the cusp field-aligned currents and the B_y component of the interplanetary magnetic field (McDiamid et al. 1978, 1979; Wilhelm et al. 1978; Iijima et al. 1978; Bythrow et al. 1982; Friis-Christensen et al. 1985). The equatorward part of the paired field-aligned current sheets is referred to as the Region 1 field-aligned currents, while the poleward part is referred to as the cusp field-aligned currents. In the northern hemisphere, the cusp field-aligned currents are observed to flow predominantly out of the ionosphere for $B_y > 0$ and into the ionosphere for $B_y < 0$. This current pattern reverses systematically in the southern hemisphere.

Lee and Kan (1979), D'Angelo (1980) and Primdahl and Spangselev (1981) suggested that part of the magnetopause currents may flow into and out of the cusp ionosphere. Lee and Kan (1979) studied the tangential discontinuity, in which B_n , the normal component of the magnetic field at the magnetopause, is zero, and found that the magnetopause currents have a significant field-aligned component. They suggested that part of the magnetopause field-aligned currents can be diverted to and closed through the polar ionosphere. However, the B_y dependence of the field-aligned current was not discussed in their study.

D'Angelo (1980) suggested that the north-south voltage difference across the magnetopause due to the presence of IMF B_y may drive the observed cusp field-aligned currents. Under the assumption of a closed magnetosphere, Primdahl and Spangselev (1981) suggested that the north-south component of the magnetopause current, which is generated to shield the IMF B_y , can be responsible for the cusp field-aligned currents. However, for a closed magnetosphere, it is easy to show on the basis of magnetic flux conservation that the field-aligned currents originating from the magnetopause would be distributed at the ionospheric altitude in a small region with a latitudinal width < 30 km, in contrast to the observed width of 100–400 km. Reiff et al. (1978) and Cowley (1981) suggested that the cusp field-

aligned currents are due to the closure of the ionospheric currents which are associated with the observed azimuthal plasma flows in the dayside “throat” region.

In this paper we suggest a different mechanism for the generation of the cusp field-aligned currents which can explain qualitatively some of the most crucial aspects of the cusp current system. In our model, the equatorward part of this field-aligned current system, referred to as the Region 1 currents, is caused by the leakage of the field-aligned currents associated with the rotational discontinuities at the dayside magnetopause and the poleward part, referred to as the cusp field-aligned currents, is generated in the tail magnetopause. The flow directions of the field-aligned currents at the magnetopause and of the corresponding cusp field-aligned currents are found to be sensitive to B_y of the IMF. We show also that the flow directions are consistent with the observed patterns for positive and negative B_y values. The latitudinal width of the cusp field-aligned currents is estimated to be 100–400 km. The effects of the field-aligned currents on the plasma flow will also be examined.

Generation of cusp field aligned currents

Magnetopause currents are generated and maintained due to the imposed transition of magnetic fields between the magnetosheath and the magnetosphere. Figure 1 illustrates two possible magnetic hodograms across the dayside magnetopause. The x axis is normal to the magnetopause and directed towards the magnetosheath, while the y and z axes are parallel to the magnetopause. The fields \mathbf{B}_{mt} and \mathbf{B}_{st} are the tangential components of the magnetic field in the magnetosphere and in the magnetosheath, respectively.

For simplicity, the magnetopause will be considered locally as one-dimensional in which all physical quantities depend only on x . The magnetic field at the magnetopause can be written as:

$$\mathbf{B}(x) = B_n \hat{\mathbf{x}} + \mathbf{B}_t(x) = B_x \hat{\mathbf{x}} + B_y(x) \hat{\mathbf{y}} + B_z(x) \hat{\mathbf{z}} \quad (1)$$

where B_n and \mathbf{B}_t are, respectively, the normal and tangential components of the magnetic field, and $\hat{\mathbf{x}}$, $\hat{\mathbf{y}}$ and $\hat{\mathbf{z}}$ are unit vectors. Since $\nabla \cdot \mathbf{B} = 0$, we have $\mathbf{B}_x = B_n = \text{constant}$. If $B_n = 0$, the magnetopause is closed and the magnetopause structure is a “tangential discontinuity.” If $B_n \neq 0$, the magnetopause is open and can be identified as a “rotational discontinuity.”

By substituting Eq. (1) into Ampere’s law, the field-aligned component and the perpendicular component of the magnetopause current density \mathbf{J} can be written as:

$$J_{\parallel}(x) \equiv \mathbf{J} \cdot \mathbf{B} / B = \mu_0 (B_t^2 / B) d\theta / dx \quad (2)$$

and

$$J_{\perp}(x) = \mu_0 [(dB_t/dx)^2 + (B_t B_n / B)^2 (d\theta/dx)^2]^{1/2} \quad (3)$$

where $\theta \equiv \tan^{-1}(B_y/B_z)$ and μ_0 is the permeability in free space.

It can be seen from Eq. (2) that the magnetopause current may have a field-aligned component if the magnetic field rotates across the magnetopause, $d\theta/dx \neq 0$. The flow direction of the field-aligned current depends on the sense of magnetic field rotation: (i) $J_{\parallel} > 0$ if $d\theta/dx > 0$, and (ii) $J_{\parallel} < 0$ if $d\theta/dx < 0$. In fact, the field-aligned component of the magnetopause currents is required to exist if the magnetic fields change direction across the magnetopause. The ex-

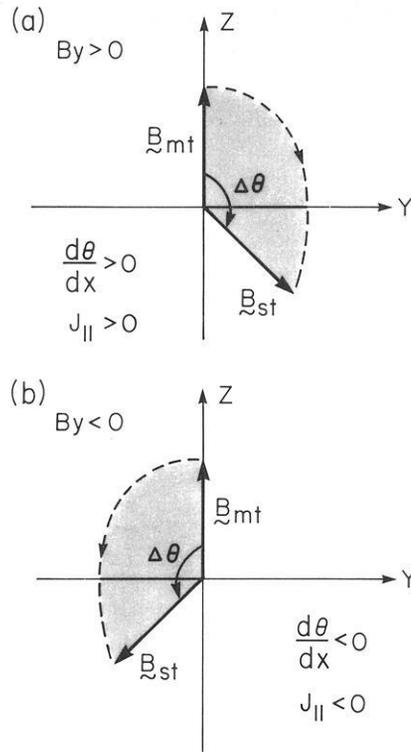


Fig. 1 a and b. Magnetic hodograms at the dayside magnetopause for (a) $B_y > 0$ and (b) $B_y < 0$. The x axis is perpendicular to the magnetopause plane (y - z plane). \mathbf{B}_{mt} and \mathbf{B}_{st} are the tangential components of the magnetic field in the magnetosphere and magnetosheath, respectively

istence of field-aligned currents at the magnetopause has been shown observationally (e.g. Kaufman and Cahill, 1977; Russell and Elphic, 1978; Sonnerup et al. 1981; Berchem and Russell, 1982) as well as theoretically (Lee and Kan, 1979, 1982; Wang and Sonnerup, 1984).

We propose that part of the field-aligned currents at the magnetopause will leak, along the geomagnetic field lines, to the polar ionosphere and produce the observed cusp field-aligned currents. The leaking of field-aligned currents from the dayside magnetopause produces the equatorward part (Region 1) of this system, while the leaking from the tail magnetopause leads to the poleward (cusp) part of the field-aligned currents.

It should be mentioned that a rotational discontinuity can be considered as a localized large-amplitude Alfvén wave, across which magnetic fields twist (Lee and Kan, 1982). Since the plasma and field conditions on the two sides of the magnetopause are observed to fluctuate (e.g. Russell and Elphic, 1978), it is likely that part of the highly twisted fields at the magnetopause may propagate as Alfvén waves to the ionosphere, carrying along the field-aligned currents. This is a physical interpretation of the leakage of the magnetopause currents to the ionosphere.

We now discuss the relationship between the flow direction of the field-aligned currents and the B_y component of the interplanetary magnetic field. For this purpose, we note first that the total angle of rotation ($\Delta\theta$) in the magnetic field across the rotational discontinuities at the magnetopause is generally observed to be smaller than or equal to 180° , i.e. $\Delta\theta \leq 180^\circ$. Berchem and Russell (1982) made an extensive study of the rotational discontinuities at the

magnetopause based on ISEE 1 and 2 data. They found that the sense of rotation in the magnetic field is controlled by the relative orientation of the magnetosheath and magnetospheric magnetic fields and that the sense of rotation is such as to minimize the total rotational angle $\Delta\theta$ ($\Delta\theta \leq 180^\circ$). Computer simulation of the rotational discontinuity ($B_n \leq 0$) by Swift and Lee (1983) also indicated that a rotational discontinuity with $\Delta\theta \leq 180^\circ$ is stable, while a rotational discontinuity with $\Delta\theta > 180^\circ$ is unstable and tends to evolve into a rotational discontinuity with a rotational angle $\Delta\theta < 180^\circ$. With these facts in mind, we examine field-aligned currents on the dayside and tail magnetopauses.

A) Dayside magnetopause field-aligned currents

We consider the B_y dependence of the field-aligned currents at the dayside magnetopause. For B_y (IMF) > 0 , we have $d\theta/dx > 0$ as illustrated in Fig. 1a since the rotational angle $\Delta\theta$ cannot be greater than 180° . It follows from Eq. (2) that $J_{\parallel} > 0$. In this case, the field-aligned current at the magnetopause is parallel to the magnetic field and flows into the dayside cusp region of the northern polar ionosphere. On the other hand, for B_y (IMF) < 0 , we have $d\theta/dx < 0$ and $J_{\parallel} < 0$ as shown in Fig. 1b. In this case, the field-aligned current is antiparallel to the geomagnetic field and flows out of the northern cusp ionosphere. We identify this field-aligned current as the equatorward part (Region 1) of the field-aligned currents observed by Iijima et al. (1978), McDiamid et al. (1978) and Wilhelm et al. (1978).

B) Tail magnetopause field-aligned currents

Next, we consider the B_y dependence of the field-aligned currents at the tail magnetopause which are responsible for the poleward part of the observed cusp field-aligned currents. By a similar argument as for the dayside magnetopause, it is found in the northern hemisphere that $J_{\parallel} < 0$ for $B_y > 0$, and $J_{\parallel} > 0$ for $B_y < 0$. Therefore, for $B_y > 0$ ($B_y < 0$), the field-aligned current generated at the northern flank of the tail magnetopause is antiparallel (parallel) to the geomagnetic field and flows out of (into) the northern polar ionosphere. The results are thus consistent with observations (McDiamid et al. 1978; Iijima et al. 1978; Wilhelm et al. 1978; Friis-Christensen et al. 1985). In the southern polar ionosphere, the current pattern reverses systematically.

Global pattern of the cusp field-aligned currents

A global distribution of the cusp field-aligned currents for $B_y > 0$ is schematically shown in Fig. 2, which provides a perspective view of the magnetic fields and currents at the magnetopause and in the cusp region of the northern hemisphere. The currents flow into the cusp ionosphere from the dayside magnetopause and flow away from the ionosphere to the tail magnetopause. It is known that for $B_y > 0$, the interplanetary magnetic flux penetrates into the Earth's magnetosphere mainly from the dawn side of the northern magnetosphere and leaves mainly from the dusk side of the southern magnetosphere (Cowley, 1981; Akasofu and Roederer 1984). Therefore, the cusp field-aligned current density in the northern hemisphere is larger in the pre-noon region than in the post-noon region as shown in Fig. 2. In the southern hemisphere the cusp field-aligned current pattern is reversed, i.e. for $B_y > 0$, currents flow into the

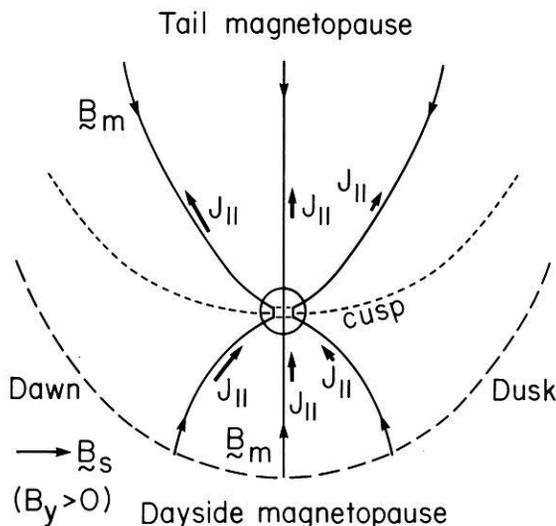


Fig. 2. A global distribution of the cusp field-aligned currents (J_{\parallel}) for $B_y > 0$. The view is from the dayside magnetosheath above the equatorial plane

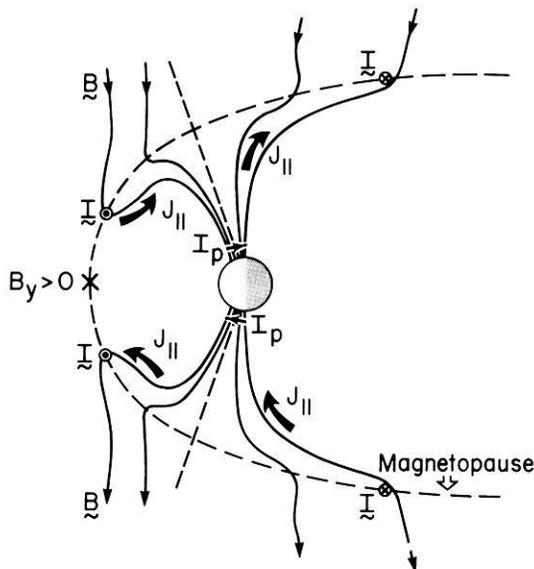


Fig. 3. The cusp current-flow pattern in the noon-midnight meridional plane for $B_y > 0$. J_{\parallel} is the field-aligned current, I_p is the Pedersen current in the ionosphere, and I is the magnetopause current

cusp ionosphere from the tail magnetopause and flow away from the ionosphere to the dayside magnetopause. The cusp field-aligned current density is expected to be larger in the post-noon region of the southern hemisphere than in the pre-noon region of the southern hemisphere.

For $B_y < 0$, the current patterns are systematically reversed and the cusp field-aligned currents flow predominantly in the post-noon (pre-noon) region of the northern (southern) hemisphere.

Note that, in our model, the Region 1 field-aligned current at noon flows toward (away from) the ionosphere in the northern hemisphere for $B_y > 0$ ($B_y < 0$). This result is consistent with the observations by Friis-Christensen et al. (1985).

The current-flow pattern in the noon-midnight meridional plane for $B_y > 0$ is shown in Fig. 3. The poleward and equatorward parts of the cusp field-aligned currents are con-

nected by the Pedersen current I_p in the ionosphere. The equatorward part of the cusp field-aligned current is closed by the dawn-to-dusk dayside magnetopause currents and the poleward part is closed by the dusk-to-dawn tail magnetopause currents. Note that the current closure at the dayside magnetopause in Fig. 3 is different from those presented by D'Angelo (1980) and Primdahl and Spangselev (1981). In their model, the equatorward part of cusp field-aligned currents in the southern hemisphere is directly connected to that in the northern hemisphere through the dayside magnetopause. In our model, on the other hand, the field-aligned currents in the southern hemisphere are not connected to those in the northern hemisphere through the dayside magnetopause.

The latitudinal width of the cusp field-aligned currents in the ionosphere can be estimated as follows. All of the dayside open magnetic field lines can carry the leaked field-aligned currents to the cusp ionosphere. On the other hand, since the anti-sunward boundary-layer plasma-flow velocities at the tail magnetopause are observed to be higher than the local Alfvén speed in the distant tail region with $x \leq -20 R_E$ (Akasofu et al., 1973; Rosenbauer et al. 1975), the Alfvén waves will not carry the leaked currents in that region to the ionosphere. The latitudinal width of the cusp field-aligned currents at the ionosphere altitude is just the latitudinal width of the dayside open field lines and near-earth open field lines ($x \geq -20 R_E$), which can then be estimated to be 100–400 km, consistent with observations, (Iijima and Potemra 1976; McDiamid et al. 1978, 1979; Wilhjelm et al. 1978; Friis-Christensen et al. 1985). It can also be estimated that a leakage of 5%–20% of the magnetopause currents is sufficient to account for the observed cusp field-aligned current intensity (integrated in the north-south direction) of 0.4 A/m, or the observed total field-aligned current of 2×10^5 A. One the other hand, if the cusp currents are due to the diversion of the north-south magnetopause currents associated with the closed magnetopause, almost all of the magnetopause currents must be diverted to the ionosphere in order to account for the total observed currents (Primdahl and Spangselev, 1981).

The injection of the magnetopause currents to the cusp ionosphere may generate the Pedersen currents, Hall currents and the plasma flows observed in the “throat” region (e.g. Banks et al. 1984). In our model, the strength of the cusp field-aligned currents and the induced plasma flow may depend on the ionospheric conductivity through the reflection of the incoming Alfvén waves which carry the field-aligned currents. For a higher conductivity in the cusp ionosphere, the ionospheric electric field and hence the plasma flow will be smaller; but the field-aligned current will be larger through the enhancement from reflected waves. Thus, our model predicts a smaller “throat” plasma flow and a larger cusp current in the summer hemisphere than in the winter hemisphere. This prediction should be tested by future observations.

Finally, we point out that the cusp currents in our model are different from the “Region O” currents proposed by Heikkila (1984), which are the closure currents of part of the Region 1 currents and flow between open and closed magnetic field lines. In our model, the cusp-region currents flow along open field lines and the equatorward part of the cusp-region currents contributes to the Region 1 currents.

In summary, we suggest that the observed cusp field-

aligned current system can be explained in terms of the leakage of the magnetopause currents associated with the rotational discontinuities at the dayside and tail magnetopause. For such a mechanism, the latitudinal width of the generated cusp field-aligned currents at the ionospheric altitude can be estimated to be of the order of 100–400 km, consistent with observations.

Acknowledgements. This work is supported by the NSF Grants ATM83-17454 and ATM83-12515 to the University of Alaska.

References

- Akasofu, S.-I., Hones, E.W. Jr., Bame, S.J., Asbridge, J.R., Lui, A.T.Y.: Magnetotail and boundary layer plasmas at a geocentric distance of $\sim 18 R_E$. Vela 5 and 6 observations. *J. Geophys. Res.* **78**, 7257–7274, 1973
- Akasofu, S.-I., Roederer, M.: Dependence of the polar cap geometry on the IMF. *Planet. Space Sci.* **32**, 111–115, 1984
- Banks, P.M., Araki, T., Clauer C.R., St. Maurice, J.P., Foster, J.C.: The interplanetary electric field, cleft currents and plasma convection in the polar caps. *Planet. Space Sci.* **32**, 1551–1557, 1984
- Berchem, J., Russell, C.T.: Magnetic field rotation through the magnetopause: ISEE 1 and 2 observations. *J. Geophys. Res.* **87**, 8139–8148, 1982
- Bythrow, P.F., Potemra, T.A., Hoffman R.A.: Observations of field-aligned currents Iparticles, and plasma drift in the polar cusps near solstice. *J. Geophys. Res.* **87**, 5131–5139, 1982
- Cowley, S.W.H.: Magnetospheric asymmetries associated with the y-component of the IMF. *Planet. Space Sci.* **29**, 79–96, 1981
- D'Angelo, N.: Field-aligned currents and large scale magnetospheric electric fields. *Ann. Geophys.* **36**, 31–39, 1980
- Doyle M.A., Rich, F.J. Burke, W.J., Smiddy, M.: Field-aligned currents and electric fields observed in the region of the dayside cusp. *J. Geophys. Res.* **86**, 5656–5664, 1981
- Friis-Christensen, E., Kamide, Y., Richmond, A.D., Matsushita, S.: IMF control of high latitude electric fields, ionospheric and Birkeland currents determined from Greenland magnetometer data. *J. Geophys. Res.* **90**, 1325–1338, 1985
- Heikkila, W.J.: Magnetospheric topology of fields and currents. In: *Magnetospheric currents*. T.A. Potemra ed.: p 208–222. Washington, D.C.: American Geophys. Union 1984
- Iijima, T., Fujii, R., Potemra, T.A., Saflekos, N.A.: Field-aligned currents in the south polar cusp and their relationship to the interplanetary magnetic field. *J. Geophys. Res.* **83**, 5595–5603, 1978
- Iijima, T., Potemra, T.A.: Field-aligned currents in the dayside cusp observed by Triad. *J. Geophys. Res.* **81**, 5971–5979, 1976
- Kaufman, R.L., Cahill, L.J., Jr.: The magnetopause at $5.2 R_E$ on August 4, 1972: Magnetopause shape and structure. *J. Geophys. Res.* **82**, 1573–1584, 1977
- Lee, L.C., Kan, J.R.: A unified kinetic model of the tangential magnetopause structure. *J. Geophys. Res.* **84**, 6417–6426, 1979
- Lee, L.C., Kan, J.R.: Structure of the magnetopause rotational discontinuity. *J. Geophys. Res.* **87**, 139–143, 1982
- McDiamid, I.B., Burrows, J.R., Wilson, M.D.: Magnetic field perturbations in the dayside cleft and their relationship to the IMF. *J. Geophys. Res.* **83**, 5753–5756, 1978
- McDiamid, I.B., Burrows, J.R., Wilson, M.D.: Large-scale magnetic field perturbations and particle measurements at 1400 km on the dayside. *J. Geophys. Res.* **84**, 1431–1441, 1979
- Primdahl, F., Spangselev, F.: Cusp region and auroral zone field-aligned currents. *Ann. Geophys.* **37**, 529–538, 1981
- Reiff P.H., Burch, J.L., Heelis, R.A.: Dayside auroral arcs and convection. *Geophys. Res. Lett.* **5**, 391–394, 1978
- Rosenbauer, H., Grünwaldt, H., Montgomery, M.D., Paschmann, G., Scokpe, N.: HEOS 2 plasma observations in the distant polar magnetosphere: The plasma mantle. *J. Geophys. Res.* **80**, 2723–2737, 1975

- Rostoker, G.: Magnetospheric and ionospheric currents in the polar cusp and their dependence on the B_y component of the interplanetary magnetic fields. *J. Geophys. Res.* **85**, 4167–4176, 1980
- Russell, C.T., Elphic, R.C.: Initial ISEE magnetometer results: magnetopause observations. *Space Sci. Rev.* **22**, 681–715, 1978
- Sonnerup, B.U.Ö., Paschmann, G., Papamastorakis, I., Scokopke, N., Haerendel, G., Bame, S.J., Asbridge, J.R., Gosling, J.T., Russell, C.T.: Evidence for magnetic field reconnection at the earth's magnetopause. *J. Geophys. Res.* **86**, 10049–10067, 1981
- Swift, D.W., Lee, L.C.: Rotation discontinuities and the structure of the magnetopause. *J. Geophys. Res.* **88**, 111–124, 1983
- Wang, D.J., Sonnerup, B.U.Ö.: Electrostatic structure of the rotational discontinuity II: Shock pair solutions. **27**, 2828–2834, 1984
- Wilhelm, J., Friis-Christensen, E., Potemra T.A.: The relationship between ionospheric and field-aligned currents in the dayside cusp. *J. Geophys. Res.* **83**, 5586–5594, 1978

Received April 16, 1985; Revised version June 17, 1985

Accepted June 27, 1985

Book Review

Intracontinental Fold Belts. Case studies in the Variscian Belt of Europe and the Damara Belt in Namibia. Edited by H. Martin and F.W. Eder. Springer-Verlag, Berlin Heidelberg New York Tokyo, 1983. XIV, 945 pp., 300 figs., 24 plates. DM 120,— ; approx. US \$ 47.70.

Die Herausgeber stellen mit diesem Sammelband zahlreicher Autoren das Ergebnis 11-jähriger interdisziplinärer Arbeiten des Sonderforschungsbereichs (SFB) 48 – Göttingen vor: „Entwicklung, Bestand und Eigenschaften der Erdkruste, insbesondere der Geosynklinalräume“.

Untersucht und einander gegenübergestellt wurden die beiden intrakontinentalen Orogene Varisziden in Mitteleuropa und Damara in Südwestafrika. Als zentrale Fragestellungen wurde betrachtet: Bildeten sich die zugehörigen Geosynklinalen auf kontinentaler Kruste oder an Kontinentalrändern?

Sind diese Orogene Folge der plattentektonischen Zyklen Wilson's mit Plattentrennung, sea-floor-spreading, Kontinent-Kontinent-Kollision? Müssen weitere dynamische Prozesse in Betracht gezogen werden? Schwerpunkt der Arbeiten lag in sedimentologischen und petrologischen Fragestellungen: Untersuchungen von Mélanges an Deckenfronten im Saxothuringikum und den Montagnes noires in Südfrankreich, Untersuchung von Turbiditfolgen und Metaplayasequenzen im Damara-Orogen, lateraler und vertikaler Änderungen der Gesteine und Gesteins-Chemismen des Rhenohertzynischen Beckens, Änderungen der Hauptliefergebiete der Sedimentfolgen. Dargestellt wurden Folgen von Mineralreaktionen während der Metamorphose und versucht P-T-Bedingungen der Regionalmetamorphose festzulegen. Weiter wurden Schmelzbedingungen der Granitoide des Damarasystems untersucht. Wichtige Beiträge sind Untersuchungen zur Strukturevolution und Mineralisation während des Geosynklinal-Orogen-Zyklus. Die Geochronologie lieferte Daten von Abkühlungsaltern im Damara-Orogen und ein Abkühlungsmodell. Der Höhepunkt der Metamorphose im Rhenohertzynikum wird datiert. Die Geophysik steuert zu diesen Arbeiten Modelle isostatischer Vertikalbewegungen und deren Einfluß auf die Wärmegeschichte und die Metamorphose im Varizikum bei. Außerdem werden seismische Profile und magnetotellurische Messungen im Varizikum diskutiert. Ergebnisse mehrerer refraktionsseismischer Profile im Damara-Orogen werden vorgestellt. Der Sammelband wendet sich mehr an Geologen und Petrologen als an den Geophysiker, trotzdem ist er ein wertvoller Beitrag interdisziplinärer geowissenschaftlicher Forschung vor allem im Vorfeld der europäischen Geotraverse und des kontinentalen Tiefbohrprogramms.

Die Strukturgeologie z.B. zeigt die Existenz von Decken im Rhenohertzynikum und Saxothuringikum in großräumigen Krustenverschuppungen entlang der südlichen Begrenzung dieser Zonen. Gerade diese Beiträge sind von besonderer Wichtigkeit auch für den Geophysiker, der sich in Probleme der Geologie und Petrologie der Kruste einlesen will.

So wird die Hypothese des „Crustal stacking“ zur Krustenevolution in Mitteleuropa ausführlich dargestellt: Große Überschiebungsbahnen scheinen nach refraktionsseismischen Profilen durch das Rhenohertzynikum und das Saxothuringikum als trennende Störungszonen tektonischer Großenheiten bis zum Mantel oder in die Unterkruste zu fungieren. Im südlichen Rhenohertzynikum und im nördlichen Saxothuringikum werden die intrakrustalen Diskontinuitäten als Späne von tieferer Kruste und der Krusten-Mantel-Übergangzone interpretiert. Während der kompressiven orogenen Phasen werden Grundgebirgs- und Sedimenteinheiten von ihrer Unterlage abgesichert und horizontal über zehn bis möglicherweise über 100 km weit verschoben, unter Bildung von Überschiebungs- und Stapleinheiten. Solche Crustal-Stacking-Prozesse könnten dann wesentlich zum Aufbau der Kruste während der variszischen Tektogenese beigetragen haben.

Für das frühe Saxothuringikum wird ein Riftmodell als Vorläufer der variszischen Geosynklinalen diskutiert. Eine Aufwölbung heißen asthenosphärischen Mantelmaterials führt zu Dehnungsstrukturen in der überlagernden Lithosphäre und wird durch Intrusionen von Schmelzen in magmatischen Gangsystemen kompensiert. CO₂ aus dem Mantel durchgast die Unterkruste und H₂O wird in seichtere Niveaus ausgetrieben. In der Unterkruste bilden sich Granulite. Basische Intrusiva werden in Pyriklasite (basische Granulite) transformiert. Die saueren Granulite der Unterkruste enthalten Quarz und die Unterkruste verhält sich deshalb duktil. Die Dehnungsbeanspruchung zeigt sich an Texturen der heute oberflächlich aufgeschlossenen Granulite. Die schichtige Einregelung der Granulite erzeugt eine Anisotropie der thermischen Leitfähigkeit.

Solche Diskussionsbeispiele zeigen, wie wichtig das interdisziplinäre Gespräch in den Geowissenschaften zur Erarbeitung von Modellen geworden ist. Die Anschaulichkeit der Darstellung in den Übersichtsarbeiten ist den Autoren gut gelungen, während viele der Einzeldarstellungen von sedimentologischem oder petrologischem Inhalt für den Geophysiker nur schwer zu lesen sind. Die Fülle der Gedanken, die in diesem Band vorgestellt werden, lohnen es jedoch allemal sich mit den Arbeiten auseinanderzusetzen.

Peter M. Sachs