

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

Volume 56 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authorities – tacitly or explicitly – in the laboratories where the work was carried out and that, if accepted, it will not be published elsewhere in the same form, in either the same or another language, without the consent of the copyright holders.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

Appel, E. 121	Herres, N. 174	Märcz, F. 47	Schurr, K. 1
	Hinzen, K.-G. 147	Mazzoni, P. 153	Siemes, H. 174
Becker, H. 1	Hroudá, F. 174		Soffel, H.C. 1, 121
Berckhemer, H. 9		Nass, H.U. 34	Tayton, J.W. 106
			Ulug, A. 9
Cassinis, R. 153	Jelinek, V. 58	Papadopoulos, K. 201	
Cormier, V.F. 20	Johnson, B.D. 106	Prodehl, C. 69	Varvoglis, H. 201
			Veevers, J.J. 106
Fahr, H.J. 34		Rabbel, W. 81	Vidberg, H.J. 183
	Koch, M. 160	Ranzoni, A. 153	
Gajewski, D. 69	Krummel, H. 147	Riska, D.O. 183	Wolf, D. 24, 100
			Zschau, J. 81
Hansen, L. 106		Samson, J.C. 133	
Hennig-Michaeli, C. 174	Lalmani 53	Schulz, R. 192	

Subject Index

Archaeomagnetism

Archaeomagnetic study of medieval fire-places at Mannheim-Wallstadt and ovens from Herrenchiemsee (southern Germany) and the problem of magnetic refraction (Schurr, K., et al.) 1

Aurora

Large-scale studies of Pi-2's associated with auroral breakups (Samson, J.C.) 133

Auroral electrojets

Large-scale studies of Pi-2's associated with auroral breakups (Samson, J.C.) 133

Auroral Physics

Stochastic ion acceleration by coherent electrostatic waves (Varvoglis, H., Papadopoulos, K.) 201

Australia

Magnetic expression of the continent-ocean boundary between the western margin of Australia and the eastern Indian Ocean (Veevers, J.J., et al.) 106

Book reviews

Book reviews 66
Book reviews 146
Book reviews 211

D.C. resistivity sounding

The method of integral equation in the direct current resistivity method and its accuracy (Schulz, R.) 192

Earthquakes

Source orientation from grid test and synthetic seismograms and an application to the Ibbenbüren earthquake of July 1981 (Hinzen, K.-G., Krummel, H.) 147

Active seismic layers and crustal structure in some Italian regions (Cassinis, R., et al.) 153

Nonlinear inversion of local seismic travel times for the simultaneous determination of the 3D-velocity structure and hypocentres – application to the seismic zone Vrancea (Koch, M.) 160

Earth's crust

Crustal structure beneath the Swabian Jura, SW Germany, from seismic refraction investigations (Gajewski, D., Prodehl, C.) 69

Active seismic layers and crustal structure in some Italian regions (Cassinis, R., et al.) 153

Nonlinear inversion of local seismic travel times for the simultaneous determination of the 3D-velocity structure and hypocentres – application to the seismic zone Vrancea (Koch, M.) 160

Earth's mantle

Frequency dependence of *Q* for seismic body waves in the Earth's mantle (Ulug, A., Berckhemer, H.) 9

Nonlinear inversion of local seismic travel times for the simultaneous determination of the 3D-velocity structure and hypocentres – application to the seismic zone Vrancea (Koch, M.) 160

Earth's Magnetosphere

Some aspects of whistler duct lifetimes at low latitudes (Lalmani) 53

Electromagnetic sounding

The inverse scattering problem for reflection of electromagnetic dipole radiation from Earth with vertical variation (Vidberg, H.J., Riska, D.O.) 183

Fault-plane solutions

Source orientation from grid test and synthetic seismograms and an applica-

tion to the Ibbenbüren earthquake of July 1981 (Hinzen, K.-G., Krummel, H.) 147

Geodynamics

Static deformations and gravity changes at the earth's surface due to atmospheric loading (Rabbel, W., Zschau, J.) 81

The normal modes of a uniform compressible Maxwell half-space (Wolf, D.) 100

Geomagnetic activity

Increased ionospheric absorption connected with Pc 1 pulsations after geomagnetic storms (Märcz, F.) 47

Geomagnetic pulsations

Increased ionospheric absorption connected with Pc 1 pulsations after geomagnetic storms (Märcz, F.) 47

Large-scale studies of Pi-2's associated with auroral breakups (Samson, J.C.) 133

Geomagnetic substorms

Large-scale studies of Pi-2's associated with auroral breakups (Samson, J.C.) 133

Global deformation of earth

Static deformations and gravity changes at the earth's surface due to atmospheric loading (Rabbel, W., Zschau, J.) 81

The normal modes of a uniform, compressible Maxwell half-space (Wolf, D.) 100

Gravity variations

Static deformations and gravity changes at the earth's surface due to atmospheric loading (Rabbel, W., Zschau, J.) 81

Indian Ocean

Magnetic expression of the continent-ocean boundary between the western margin of Australia and the eastern Indian Ocean (Veevers, J.J., et al.) 106

In memoriam

Hans Closs (1907–1982) 61
Herbert Flathe (1919–1983) 63

Ion acceleration

Stochastic ion acceleration by coherent electrostatic waves (Varvoglis, H., Papadopoulos, K.) 201

Isostasy

The relaxation of spherical and flat Maxwell Earth models and effects due to the presence of the lithosphere (Wolf, D.) 24

Italy

Active seismic layers and crustal structure in some Italian regions (Cassinis, R., et al.) 153

Lithosphere

The relaxation of spherical and flat Maxwell Earth models and effects due to the presence of the lithosphere (Wolf, D.) 24

Nonlinear inversion

Nonlinear inversion of local seismic travel times for the simultaneous determination of the 3D-velocity structure and hypocentres – application to the seismic zone Vrancea (Koch, M.) 160

Normal modes

The normal modes of a uniform, compressible Maxwell half-space (Wolf, D.) 100

Oceanic magnetic anomalies

Magnetic expression of the continent-ocean boundary between the western margin of Australia and the eastern Indian Ocean (Veevers, J.J., et al.) 106

Planetary atmospheres

Plasma-gas interactions in planetary atmospheres and their relevance for the terrestrial hydrogen budget (Nass, H.U., Fahr, H.J.) 34

Q

Frequency dependence of *Q* for seismic body waves in the Earth's mantle (Ulug, A., Berckhemer, H.) 9

Ray tracing

The polarization of S waves in a heterogeneous isotropic Earth model (Cormier, V.F.) 20

Refraction Seismology

Crustal structure beneath the Swabian Jura, SW Germany, from seismic refraction investigations (Gajewski, D., Prodehl, C.) 69

Rock magnetism

On a mixed quadratic invariant of the magnetic susceptibility tensor (Jelinek, V.) 58

Domain state of Ti-rich titanomagnetites deduced from domain structure observations and susceptibility measurements (Appel, E., Soffel, H.C.) 121

The relationship between the magnetic anisotropy and the *c*-axis fabric in a massive hematite ore (Hrouda, F., et al.) 174

Romania

Nonlinear inversion of local seismic travel times for the simultaneous determination of the 3D-velocity structure and hypocentres – application to the seismic zone Vrancea (Koch, M.) 160

Sea floor spreading

Magnetic expression of the continent-ocean boundary between the western margin of Australia and the eastern Indian Ocean (Veevers, J.J., et al.) 106

Seismic body waves

Frequency dependence of *Q* for seismic body waves in the Earth's mantle (Ulug, A., Berckhemer, H.) 9

The polarization of S waves in a heterogeneous isotropic Earth model (Cormier, V.F.) 20

Seismicity

Active seismic layers and crustal structure in some Italian regions (Cassinis, R., et al.) 153

Seismology

Nonlinear inversion of local seismic travel times for the simultaneous determination of the 3D-velocity structure and hypocentres – application to the seismic zone Vrancea (Koch, M.) 160

Synthetic seismograms

Source orientation from grid test and synthetic seismograms and an application to the Ibbenbüren earthquake of July 1981 (Hinzen, K.-G., Krummel, H.) 147

Whistlers

Some aspects of whistler duct lifetimes at low latitudes (Lalmani) 53

Urach geothermal anomaly

Crustal structure beneath the Swabian Jura, SW Germany, from seismic refraction investigations (Gajewski, D., Prodehl, C.) 69

Journal of Geophysics Zeitschrift für Geophysik

Volume 56 Number 1 1984

Original investigations

K. Schurr, H. Becker, H.C. Soffel

Archaeomagnetic study of medieval fireplaces at Mannheim-Wallstadt and ovens from Herrenchiemsee (southern Germany) and the problem of magnetic refraction 1

A. Ulug, H. Berckhemer

Frequency dependence of Q for seismic body waves in the Earth's mantle 9

V.F. Cormier

The polarization of S waves in a heterogeneous isotropic Earth model 20

D. Wolf

The relaxation of spherical and flat Maxwell Earth models and effects due to the presence of the lithosphere 24

H.U. Nass, H.J. Fahr

Plasma-gas interactions in planetary atmospheres and their relevance for the terrestrial hydrogen budget 34

F. Märckz

Increased ionospheric absorption connected with Pc 1 pulsations after geomagnetic storms 47

Lalmani

Some aspects of whistler duct lifetimes at low latitudes 53

Short communication

V. Jelinek

On a mixed quadratic invariant of the magnetic susceptibility tensor 58

In memoriam

Hans Closs (1907–1982) 61

Herbert Flathe (1919–1983) 63

Book reviews 66

82 Nat. 2148
24137

Indexed in Current Contents

Evaluated for Physics Briefs and INKA-PHYS data base

1-3, 7, 1

64

x, 501

20. Dez. 1984
Niedersächsische Staats-
Universitätsbibliothek
Göttingen



Springer International

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 (publisher, 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 resident in the Federal Republic of Germany or in West Berlin can apply to Verwertungsgesellschaft 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
G. Sternberg
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, 1984
Springer-Verlag GmbH & Co. KG
D-1000 Berlin 33
Printed in Germany



Springer International

Archaeomagnetic study of medieval fireplaces at Mannheim-Wallstadt and ovens from Herrenchiemsee (southern Germany) and the problem of magnetic refraction

K. Schurr¹, H. Becker², and H.C. Soffel¹

¹ Institut für Allgemeine und Angewandte Geophysik, Universität München, Theresienstr. 41, D-8000 München 2, Federal Republic of Germany

² Bayerisches Landesamt für Denkmalpflege, Arabellastrasse, D-8000 München, Federal Republic of Germany

Abstract. Ten fireplaces from Mannheim-Wallstadt (MW) and four ovens from Herrenchiemsee (HC) were studied with regard to their archaeomagnetic properties. The ages of the medieval structures were only poorly known and an age dating was intended with the help of the standard curves for the declination and inclination of the geomagnetic field of the past 2000 years, as set up by Thellier (1981) for France. The stability of the NRM was tested with Thellier's test. Of the ten MW fireplaces, only one passed the test, whereas of the four ovens of HC, all passed the test but one of them had to be excluded because of secondary displacements within the structure. The MW fireplace could be dated to 670–700 A.D., in agreement with other archaeological age determinations for the fireplaces. The ovens from HC yielded an age between 1100 and 1170 A.D.

The ovens showed the effect of magnetic refraction. Model calculations carried out on circular ring structures require a susceptibility of the oven material of the order of 0.5 SI units to explain the observed effect. The presently observed mean susceptibility at room temperature is only around 5×10^{-3} SI units. However, heating of the material at 550°C for 2 h in a reducing environment was able to increase the susceptibility by a factor of 20. This effect is explained by the reduction of secondary iron oxides and hydroxides to magnetite. Another increase by a factor of about 1.5 is obtained from the Hopkinson effect (increase of susceptibility with temperature) at the blocking temperature of magnetite. Combining all effects, it is possible to attain values for the susceptibility at the blocking temperature of magnetite as high as 0.2 SI units, which is the order of magnitude required for the explanation of the observed effect of magnetic refraction.

This effect demands that special considerations be made during archaeomagnetic sampling from archaeological ovens or kilns.

Key words: Archaeomagnetism – Archaeomagnetic age dating – Magnetic refraction – Susceptibility

Introduction

The initial purpose of the study was an attempt to date four medieval ovens from Herrenchiemsee and ten fireplaces near Mannheim-Wallstadt (both in southern Ger-

Table 1. Geographic coordinates, number *N* of archaeological structures per site and their estimated ages

Locality	(° E)	(° N)	<i>N</i>	Age
Mannheim-Wallstadt (MW)	8.5	49.5	10	First–third century or fifth–seventh century A.D.
Herrenchiemsee (HC)	12.4	47.9	4	Medieval

many) using the archaeomagnetic reference curve for France, published by Thellier (1981) for inclination and declination (Fig. 8). Table 1 shows the geographic coordinates of the sampling localities, the number of sites and their possible ages based on archaeological arguments.

The remnants of the fireplaces from Mannheim-Wallstadt were cast into large blocks of plaster of Paris (about 40 cm thick and wide, 100 cm long) and removed from the ground after orientation with respect to the horizontal plane and geographic north (using a sun compass). Figure 1 shows a sketch map of the position of the fireplaces on the archaeological site.

The ovens from Herrenchiemsee originally had a shape like that shown schematically in Fig. 2. Oriented samples were taken from the floor and from the walls of the structures wherever suitable material was preserved. Figure 3 shows the spatial distribution of the four ovens on the archaeological site as well as the position of the samples within each oven. From the four ovens at Herrenchiemsee, only the ovens HC I, II and IV were in a more or less undisturbed position. The floor of oven HC III showed later internal mechanical movements. Therefore, this material was only used for rock magnetic studies.

Preparation of specimens

From the huge blocks obtained from the fireplaces at Mannheim-Wallstadt and from the samples from Herrenchiemsee, smaller specimens were cut and embedded in cubes of plaster of Paris with a side length of 6 cm for a convenient measurement with a spinner magnetometer (see next section). Table 2a shows the number of samples and specimens obtained from each of the ten fireplaces at Mannheim-Wallstadt.

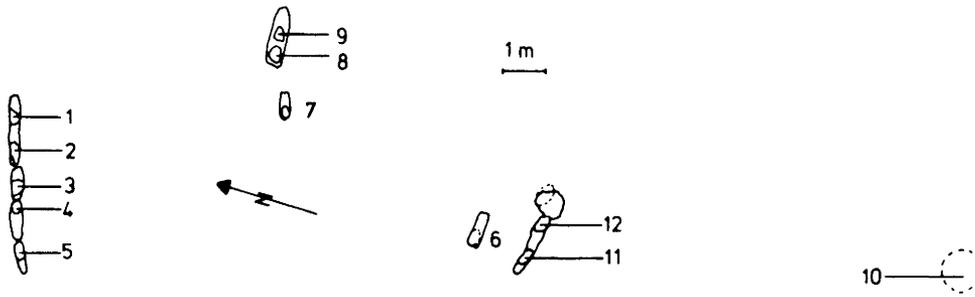


Fig. 1. Sketch map of the positions of the fireplaces at Mannheim-Wallstadt. Numbers refer to structure members MW1–MW12 (see also Table 2)

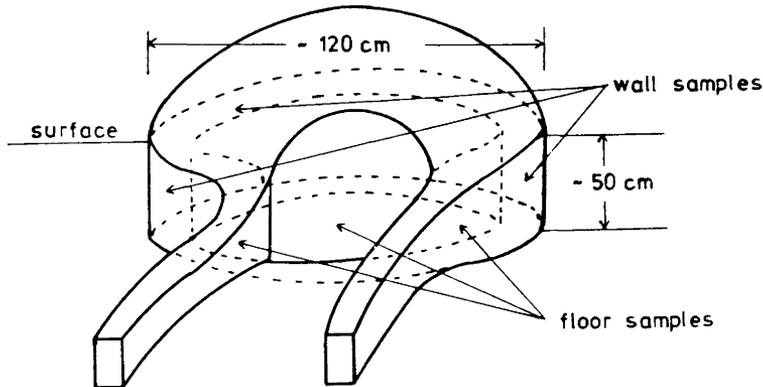


Fig. 2. Schematic sketch of the shape and dimensions of the ovens found at Herrenchiemsee

The distribution of the samples and specimens within the ovens at Herrenchiemsee is illustrated in Fig. 3. (Details can also be taken from Table 2b.) It illustrates the position of the samples (in a schematic way) in the wall or the floor of the ovens. The arrow with the full line represents the present magnetic north direction in the sampling area, while the arrow with the dashed line is the archaeonorth direction (see next section). The little arrows marked within the samples are the archaeoremanence directions determined for each specimen. For detail refer to section on magnetic refraction.

Archaeomagnetic measurements

The natural remanent magnetization (NRM) of the specimens was measured with a computerized big-sample flux-gate spinner magnetometer. The stability of NRM was checked with the Thellier test (Thellier and Thellier, 1959) by comparing the vertical Z -components, Z_1 and Z_2 , of the specimens before and after, respectively, a 4-week storage in the geomagnetic field in the reversed position. The viscosity coefficient v is determined by $v = \frac{Z_1 - Z_2}{Z_1 + Z_2}$. Furthermore, the directional shift of NRM was computed for the Thellier test.

Table 3 shows the mean viscosity coefficients \bar{v} for a selection of specimens of the fireplaces from Mannheim-Wallstadt and the oven specimens from Herrenchiemsee. For the fireplaces at Mannheim-Wallstadt, all viscosity coefficients are greater than 1% with the exception of a large and stable group of specimens from MW 9 with $\bar{v} = 0.22$. Similar small \bar{v} values ($\bar{v} < 0.5\%$) could be found for all ovens at Herrenchiemsee. The NRM of all structures with $\bar{v} < 0.5$ was taken as the archaeodirection. The results in terms of mean NRM directions and their statistical parameters are summarized in Table 4.

Magnetic refraction

In Fig. 3 the directions of the horizontal components of NRM of the individual specimens are shown as small arrows. The horizontal direction of the archaeofield (arrow with dashed line) is the average of the individual directions in the floor specimens. However, there is a systematic deviation of inclination and declination from the mean archaeofield direction in the wall specimens as well as a deviation only of inclination in the floor specimens. These deviations are dependent upon the sample positions (azimuth θ and distance r from the centre) within the structures. Figure 4 shows the variation of the inclination with distance from the centre, for three ovens. In all cases, the inclinations in the centre of the oven floor are significantly lower than in the walls. Effects of the same kind in kilns have already been described by Harold (1960). He interpreted them as a result of the cooling of the kiln and associated mechanical displacements (kiln wall fall-out). Experiments by Weaver (1962) and Aitken and Hawley (1971) indicated, however, that mechanical processes within the cooling kilns could not account for the effects. They discussed magnetic refraction as a possible interpretation.

Magnetic refraction describes the refraction of magnetic field lines at an interface where the magnetic susceptibility χ (or permeability $\mu = \chi + 1$) changes sharply. The relation between the angles of the incident field line (α_1) and the transmitted field line (α_2), and the permeabilities μ_1 and μ_2 is:

$$\tan \alpha_1 / \tan \alpha_2 = \mu_1 / \mu_2.$$

The magnetic refraction in a circular ring in vacuum with inner diameter a and outer diameter b with permeability μ has been described by Jackson (1975). A modification of the formulae with special application to the shape of the ovens can be taken from Schurr (1983). The alignment

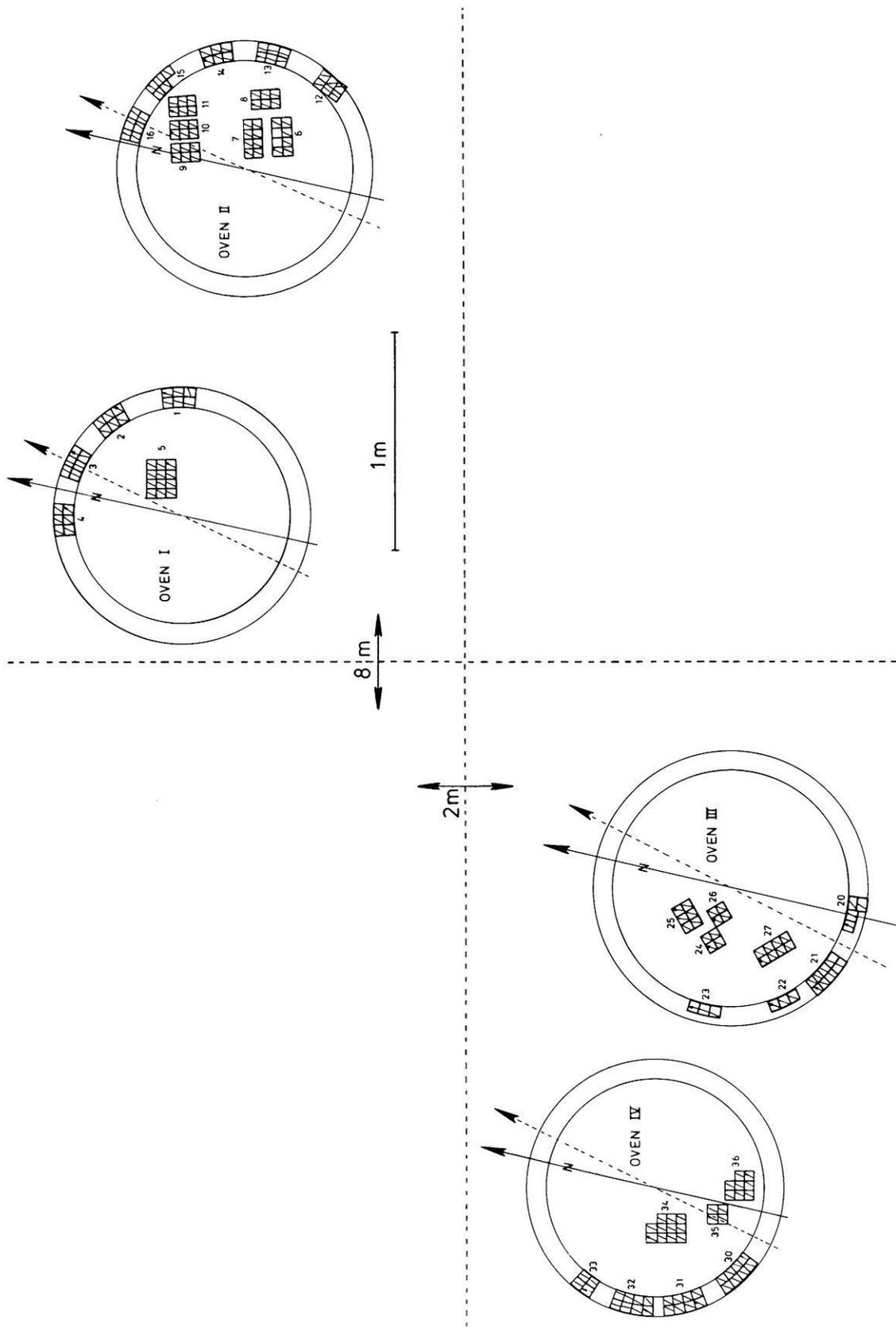


Fig. 3. Sketch map of the positions of the four ovens at Herrenchiemsee and distribution of the samples in the four ovens. The outer ring represents the oven wall, the inner area the oven floor. *Large arrows*: with *full line*: present orientation of the magnetic meridian; with *dashed line*: archaeo-orientation of the magnetic meridian. *Small arrows* and *thin lines* within each specimen: direction of the horizontal component of the stable natural remanence

Table 2. Number N of samples and number n of specimens obtained from each archaeological structure

a	Locality	Structure	N	n		
	Mannheim-Wallstadt	MW1	18	36		
		MW3	6	6		
		MW4	6	6		
		MW5	6	12		
		MW6	11	11		
		MW7	18	18		
		MW9	23	23		
		MW10	9	9		
		MW11	12	19		
		MW12	12	12		
		b	Locality	Structure	N	n
			Herrenchiemsee	HC I, floor	1	12
HC I, wall	4			24		
HC II, floor	6			38		
HC II, wall	5			29		
HC III, floor	4			22		
HC III, wall	4			18		
HC IV, floor	3			23		
HC IV, wall	4			28		

Table 3. Mean viscosity coefficients \bar{v} , in percent, of the fireplaces at Mannheim-Wallstadt and the ovens at Herrenchiemsee. n = number of specimen

Structure	n	\bar{v}	Remarks
MW1	36	8.96	
MW2	6	37.98	
MW4	6	1.78	
MW5	12	40.96	
MW6	11	12.77	
MW7	18	9.53	
MW9	23	2.89	All
	4	15.58	Very viscous group
	19	0.22	Stable group
MW10	9	4.55	
MW11	19	41.51	
MW12	12	5.09	
HC I, floor	12	0.38	
HC I, wall	24	0.27	
HC II, floor	8	0.26	
HC II, wall	8	0.38	
HC III, floor	8	0.27	
HC III, wall	8	0.42	
HC IV, floor	8	0.28	
HC IV, wall	8	0.38	

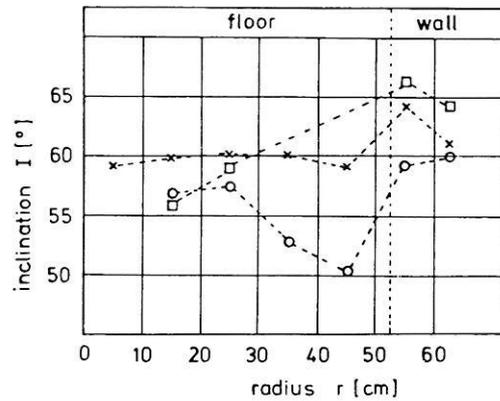
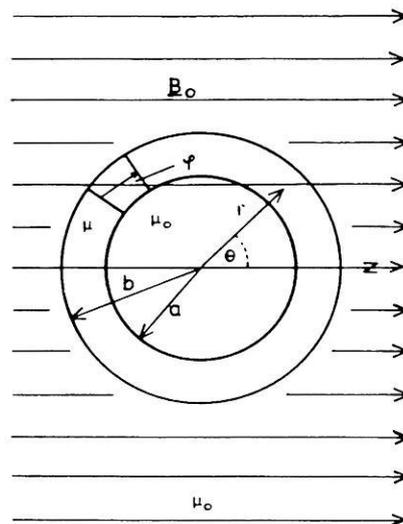
of the field lines within the circular ring is similar to that of Fig. 5 and has a distribution with the azimuth θ as shown in Fig. 6. The effect depends mainly on the susceptibility of the circular ring and, to some extent, also on the geometry of the ring. The curves shown in Fig. 6 have been computed for values of the radius of 57.5 cm (curves 2 and 4) and 62.5 cm (curves 1 and 3), with values for the susceptibility of 0.5 and 0.1 SI units, respectively.

The actually observed variations of the declination from the mean archaeodirection, dependent on the azimuthal positions of the specimens in the oven walls, have also been plotted in Fig. 6. Despite a considerable scatter of the exper-

Table 4. Mean stable NRM directions of the archaeological structures, subdivided into floor and wall specimens

Structure	I (°)	D (°E)	N	α_{95} (°)	Standard deviations		k
					I (°)	D (°)	
MW9	71.04	356.53	18	0.83	0.28	1.12	1,700
HC I, floor	57.92	13.85	12	1.58	0.81	0.49	750
HC I, wall	65.42	10.68	24	1.71	0.89	2.56	310
HC II, floor	59.88	11.08	38	0.78	0.32	0.50	880
HC II, wall	63.08	9.55	29	2.53	1.34	3.18	110
HC III, floor	63.38	2.79	22	1.22	0.39	1.25	720
HC III, wall	65.76	10.69	18	1.12	0.38	1.15	1,100
HC IV, floor	54.12	15.29	23	1.43	0.75	0.44	450
HC IV, wall	59.81	14.24	28	1.60	0.56	1.35	290

I : inclination; D : declination; N : number of specimens; α_{95} radius of cone of confidence; k : precision parameter

**Fig. 4.** Variation of the mean inclination of floor and wall specimens with distance r from the centre, for three ovens. The inclination of the wall specimens is significantly higher than that of floor specimens**Fig. 5.** Alignment of the field lines within a circular ring of permeability μ situated in vacuum (permeability μ_0). r : radius of the ring. θ : azimuth of a volume element of the ring. φ : angle between the field direction in the volume element and the external homogeneous field. (Modified from Bergmann-Schaefer, 1974)

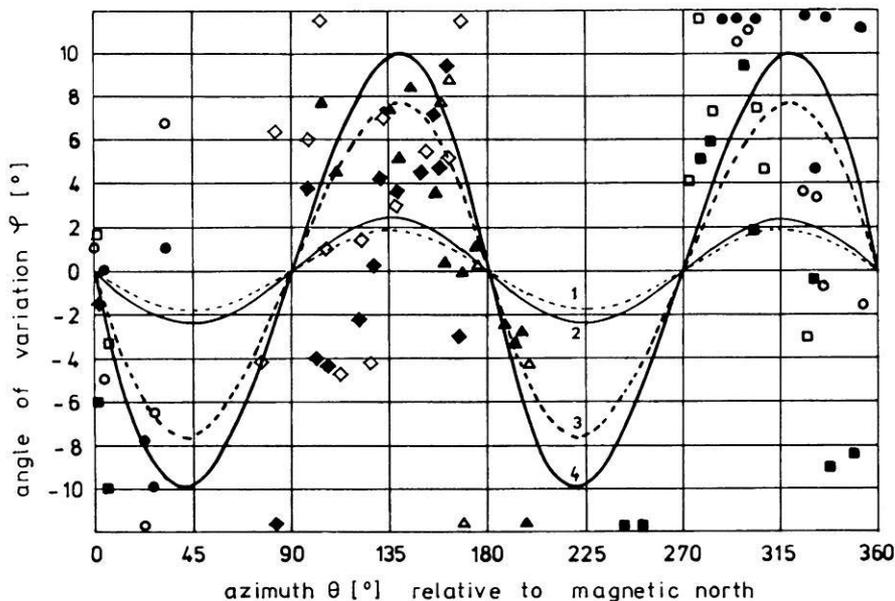


Fig. 6. Variation of the orientation φ of the field lines within a circular ring of radius r and susceptibility χ with azimuth θ . For definition of φ , r and θ , see legend of Fig. 5. The sine curves are computed for the following parameters: 1; $r=62.5$ cm and $\chi=0.1$ SI units. 2; $r=57.5$ cm and $\chi=0.1$ SI units. 3; $r=62.5$ cm and $\chi=0.5$ SI units. 4; $r=57.5$ cm and $\chi=0.5$ SI units. The open symbols refer to $r=57.5$ cm, the closed symbols refer to $r=62.5$ cm. Circles: HC I; squares: HC II; triangles: HC III; diamonds: HC IV

imental data, it can be seen that the data follow, in general, the predicted azimuthal distribution. Unfortunately, the state of conservation of the ovens did not allow sampling of all azimuthal positions of all ovens.

Rock magnetic studies

As seen from Fig. 6, a susceptibility of the order of 0.5 SI units of the oven material is required to explain the observed effects. The carriers of remanence of baked earths have been studied in detail by Le Borgne (1955, 1960a, 1960b, 1965) Mullins (1974) and Graham and Scollar (1976). X-ray powder studies were only possible for the material of the fireplaces at Mannheim-Wallstadt. Here, enough ore grains could be extracted with a hand magnet. Measurements on material from seven fireplaces gave a mean cubic lattice constant of 8.376 \AA (± 0.019) indicating that the extracted magnetic phase is mainly magnetite with, however, some modifications towards maghemite (Rumble 1976). Unfortunately, the magnetic phases of the oven material was too fine-grained for extraction.

The Curie temperatures could be determined for the MW samples as well as for HC samples (for abbreviations see Table 1). They are around $T_c=580\text{--}590$ °C, which is higher than T_c for pure magnetite (580 °C) and also indicates some degree of maghemitization.

The susceptibility of specimens from HC I, II and IV was measured with a 1 kHz bridge. The mean value for 12 specimens was 4.8×10^{-3} SI units, with maximum values as high as 1.4×10^{-2} . The anisotropy of susceptibility was also measured. Aitken et al. (1981) showed that the anisotropy of susceptibility, which is indicative of an alignment of the magnetic minerals, may affect archaeomagnetic data. However, we only found anisotropies of the order of less than 2.5%.

These mean susceptibility values are of course not sufficient to explain the observed effects of magnetic refraction as illustrated in Fig. 6. However, it should be considered that the thermoremanent magnetization was formed in the oven walls and floors under special conditions. Firstly, it was certainly formed at high temperatures, not much below

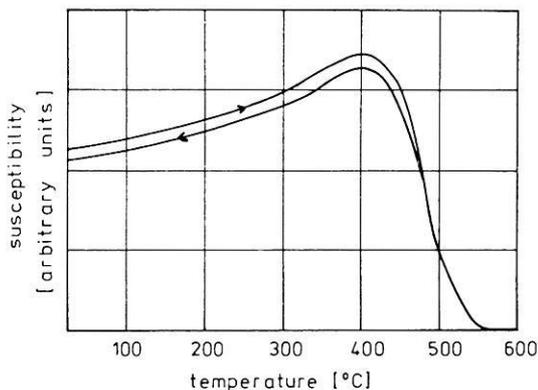


Fig. 7. Hopkinson effect showing the increase of susceptibility with temperature, with a maximum below the Curie temperature. Susceptibility given in arbitrary units

the Curie temperature of magnetite. This implies that the susceptibility of the material at the blocking temperature of the magnetite should be considered rather than that measured at room temperature. Secondly, during firing, there was a reducing atmosphere within the oven which probably favoured the formation of magnetite from most of the iron present in the kiln material. Maghemite may have formed later on, during cooling in an oxidizing environment. Weathering of the entire structure in the soil may also have caused the formation of iron hydroxides like Goethite.

We therefore measured the susceptibility of oven wall and oven floor material as a function of temperature. Figure 7 shows such a measurement, which indeed showed an increase at elevated temperatures (Hopkinson-effect), but the enhancement was not larger than a factor of 1.5.

Mullins (1974) and Graham and Scollar (1976) showed that the susceptibility of archaeological material could be increased considerably by heating the material in a reducing atmosphere together with organic matter. We did similar experiments by mixing powdered material from MW as well as HC with wheat flour (5% by weight) and heating in a nitrogen atmosphere for 2 h at 550 °C. The results are summarized in Table 5. They show that in some cases the

Table 5. Enhancement of natural susceptibility χ_o by heating of oven and fireplace materials for 2 h at 550° C in a reducing environment (χ_h). For experimental details see text. χ_o and χ_h are given in 10^{-3} SI-units

Sample	χ_o	χ_h	χ_h/χ_o
MW 15/6+	1.60	20.30	15.8
MW 2/4			
MW 15/12	5.80	13.30	2.3
HC III, wall	0.60	11.50	19.2
HC I, wall	0.83	15.90	19.2
HC II, wall	3.25	15.20	4.7
HC IV, wall	1.55	13.70	8.8

susceptibility could be enhanced by a factor of almost 20, while in some cases only a small increase by a factor of 2 could be obtained. The enhancements are especially large in those cases where the initial natural susceptibility χ_o was very low ($<3 \times 10^{-3}$ SI units). The low values of χ_o may have been caused by iron hydroxides and hematite as the

most abundant iron-containing minerals, while in the cases with high χ_o values magnetite may already have been present in the natural samples. The reduction of the material may be responsible for an enhancement of χ by as much as a factor of 20. This could enhance the measured mean susceptibility of 5×10^{-3} SI units up to about 0.1 SI units.

The total amount of iron present in the material (assuming that it is all transformed into magnetite in a reducing environment) should determine an upper limit for the susceptibility of the material. The total iron content was determined for four representative specimens by wet chemical analysis. A surprisingly uniform iron content was obtained, from which a Fe_3O_4 content of $(2.43 \pm 0.03)\%$ was calculated. With an initial mass susceptibility χ_{gi} of magnetite of 2.5×10^{-3} SI units (Landolt-Börnstein 1970) and the observed density of 2 g/cm^3 for the material, we obtained a value of 0.12 SI units as an upper limit for the volume susceptibility χ_o at room temperature. Due to the Hopkinson effect, the volume susceptibility near the blocking temperature will be of the order of 0.2 SI units. This is the required range to cause the observed variation of declination in most of the wall samples from Herrenchiemsee.

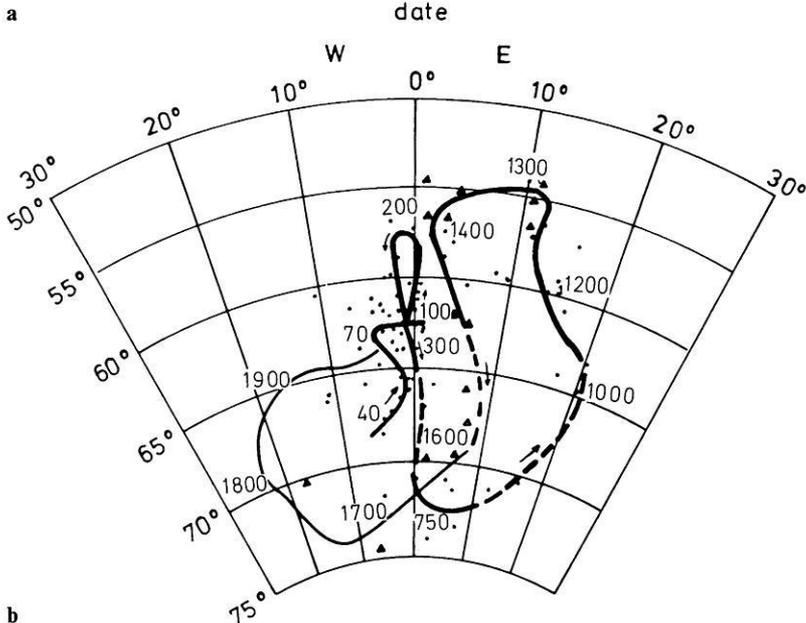
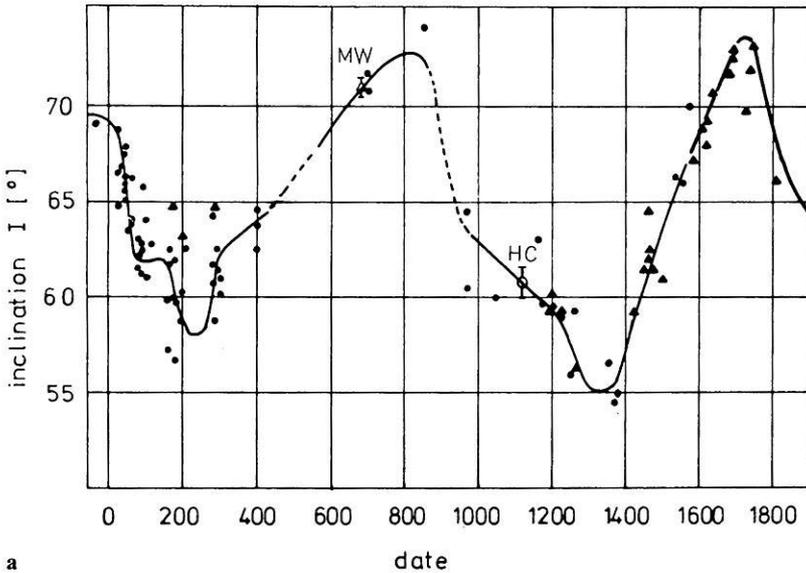


Fig. 8a and b. Standard curves for the archaeosecular variation in France compiled by Thellier (1981) for the last 2000 years. The values for the declination and inclination have been reduced to the geographical coordinates at Paris. **a** Inclination. *Heavy full line* from 1550 A.D. to present: direct measurements in Paris. *Dots*: data from archaeological structures. *Triangles*: data from bricks and tiles. *Full line*: reliable parts of the curve. *Dashed lines*: uncertain parts. *Open circle*: mean inclination value for the three ovens at Herrenchiemsee. *Open triangle*: mean inclination value for the fireplace MW 9 from Mannheim-Wallstadt. Error bars are also shown. **b** Declination and inclination, combined. Numbers are years A.D. *Thin curve* from 1600 A.D. to present: direct measurements in Paris. *Heavy lines*: interpolated curve from reliable archaeomagnetic data. *Dashed heavy lines*: uncertain parts of the curve

Age of the archaeological structures

The updated standard curves for the variation of I and D have been published by Thellier (1981) and they are shown in Fig. 8. For the age determination, the mean of all specimens (walls and floors) was calculated for three of the four kilns at Herrenchiessee and for the fireplace number 9 at Mannheim-Wallstadt. The data are summarized in Table 6 together with the statistical parameters.

Thellier did not accept that kilns show any systematic pattern of distortion (Thellier 1981). However, he always averaged the results of as many specimens as possible from widely spaced places within one archaeological structure and so minimized the effect of magnetic refraction. Therefore, a separate discussion of the directions of remanence of floor and wall samples was not regarded as appropriate, at present, for the archaeomagnetic age dating using Thellier's standard curves.

Although the standard curves have been reduced to the geographic coordinates of Paris, they can also be used for southern Germany with the application of a dipole correction, at least for the inclination values. However, the corrections are so small ($<1^\circ$) that they have been neglected. From the present dipole field, the differences in declination are of the order of 2° and have probably not been very much larger in the past 2000 years.

For the fireplace MW 9, with an inclination of 71° , there are several possible age intervals on the standard curve (Fig. 8a). The ages around 1,700 A.D. must be excluded because of the general archaeological situation. The observed declination value of 356.5° (see Fig. 8b) and the inclination (71°) make an age of between 670 and 700 A.D. for the last firing most likely, and not the other possible ages of around 50 B.C. (from an extrapolation of the curve of the inclination to earlier dates) or 900 A.D. The age of 670–700 A.D. is also likely from the archaeological dating of the entire group of fireplaces at Mannheim-Wallstadt by other methods, yielding ages of the fifth to the seventh century A.D. Reference structure nearest in direction: Thellier (1981) Table IV, No. 125: Huy (oven 2), age: approx. 700 A.D.

The remanence directions determined for the three ovens gave values for I between 57° and 63° , and values for D between 12° and 15° E. From the standard curves of Fig. 8, only the age interval between 950 and 1,250 A.D. is in agreement with the observed data set. Other age intervals, which could be possible from the inclination values alone (1,400–1,500 A.D. or 50–400 A.D.) can be excluded; the later alternative because of archaeological reasons, the earlier one because of the declination values.

The archaeological situation of the ovens at Herrenchiessee makes it likely that all were in use almost simultaneously, that is to say at least within one generation, and not over a time span of almost 300 years which we obtain if we use the determined mean remanence direction of each oven. For this reason we used the mean remanence direction of all three ovens ($I=60.6^\circ$ and $D=13.8^\circ$ E) for the archaeomagnetic age dating.

From the averaged standard curve of the inclination (Fig. 8a) we determine the interval between 1,100 and 1,170 A.D. as the probable time span for the last firing of the three ovens at Herrenchiessee. Reference structure nearest in direction: Thellier (1981) Table IV, No. 163: Plo-meur St. Saturnin, age: $1,050 \pm 90$ A.D.

Table 6. Mean stable NRM directions. For details see also legend of Table 4

Structure	I ($^\circ$)	D ($^\circ$ E)	N	α_{95} ($^\circ$)	k	Standard deviations	
						I ($^\circ$)	D ($^\circ$)
MW 9	71.04	356.53	18	0.83	1,700	0.28	1.12
HC I	63.43	12.30	36	1.92	160	0.90	1.70
HC II	61.26	10.59	67	1.21	210	0.49	1.04
HC IV	57.24	14.67	51	1.28	240	0.60	0.76
HC (all)	60.63	13.81	154	1.24	84	0.57	1.10

Discussion and conclusion

The observed effect of magnetic refraction seems to be a serious problem for the evaluation of the secular variation of the geomagnetic field by archaeomagnetic measurements and the setting up of standard curves for archaeomagnetic dating. Thellier (1981), and also in earlier papers, tried to avoid this by collecting as many specimens as possible from the archaeological sites hoping that the effects would be averaged out. As a general rule, one can say that the declination values are least affected in specimens taken from the centre of the floor of kilns or from places with azimuth angles 0° , 90° , 180° and 270° with respect to the magnetic meridian of circular structures. (For definition of the azimuth see Figs. 5 and 6.) The inclination values are most disturbed (too shallow) in the floor specimens. They are also disturbed in wall specimens (too steep) at azimuth positions 0° and 180° and are probably undisturbed in azimuth positions 90° and 270° . In poorly preserved structures it is often possible to take samples in only a few places. If they are not adequately placed over the entire structure, both D and I values can be biased in an unknown way. So far there are no possibilities of making proper corrections. As shown earlier the corrections depend not only on the geometry of the archaeological structures, but also on the rock magnetic properties of the material at the blocking temperature when the TRM was acquired. Further rock magnetic studies are planned with the aim of providing means to correct archaeomagnetic data for the effects of magnetic refraction.

Acknowledgements. The measurements were made in the Institute für Allgemeine und Angewandte Geophysik, University of Munich. We thank Professor Angenheister for his support. The financial support of the Deutsche Forschungsgemeinschaft and the Stiftung Volkswagenwerk is gratefully acknowledged.

References

- Aitken, M.J., Alcock, P.A., Bussell, G.D., Shaw, C.J.: Archaeomagnetic determination of the past geomagnetic intensity using ancient ceramics: allowance for anisotropy. *Archaeometry* **23**, 53–64, 1981
- Aitken, M.J., Hawley, H.N.: Archaeomagnetism: Evidence for magnetic refraction in kiln structures. *Archaeometry* **13**, 83–85, 1971
- Bergmann, L., Schaefer, C.: *Lehrbuch der Experimentalphysik II*, W. de Gruyter, Berlin, 1974
- Graham, I.D.G., Scollar, I.: Limitation on magnetic prospection in archaeology imposed by soil properties. *Archaeo-Physica* **6**, 1–126, 1976

- Harold, M.R.: Magnetic dating: kiln wall fall-out. *Archaeometry* **3**, 47–49, 1960
- Jackson, I.D.: *Classical electrodynamics*. J. Wiley and Sons, London, 1975
- Landolt-Börnstein: *Zahlenwerte und Funktionen aus Naturwissenschaften und Technik*, Vol. III, 4b, 1970
- Le Borgne, E.: Susceptibilité magnétique anormale du sol superficiel. *Ann. Géophys.* **11**, 399–419, 1955
- Le Borgne, E.: Influence du feu sur les propriétés magnétiques de sol et sur celles du schiste et du granite. *Ann. Géophys.* **16**, 159–196, 1960a
- Le Borgne, E.: Etude expérimentale du trainage magnétique dans le cas d'un ensemble de grains magnétiques très fins dispersés dans une substance non magnétique. *Ann. Géophys.* **16**, 445–503, 1960b
- Le Borgne, E.: Les propriétés magnétiques du sol. Application à la prospection des sites archéologiques. *Archaeo-Physica I*, Beihefte der Bonner Jahrbücher **15**, 1–20, 1965
- Mullins, C.E.: The magnetic properties of soil and their application to archaeological prospection. *Archaeo-Physica* **5**, 143–348, 1974
- Rumble, D. (ed): *Reviews in Mineralogy*, Vol. 3: Oxide Minerals. Min. Soc. of America, Short Course Notes, Southern Printing Comp., Blacksburg, Virginia, USA, 1976
- Schurr, K.: *Archäomagnetische Untersuchungen an Feuerstellen aus Mannheim-Wallstadt und Öfen aus Herrenchiemsee*. Unpubl. Thesis, Universität München, 1983
- Thellier, E., Thellier, O.: Sur l'intensité du champ magnétique terrestre dans le passé historique et géologique. *Ann. Géophys.* **15**, 285–376, 1959
- Thellier, E.: Sur la direction du champ magnétique terrestre, en France, durant les deux derniers millénaires. *Phys. Earth Planet. Inter.* **24**, 89–132, 1981
- Weaver, G.H.: Archaeomagnetic measurements on the Second Boston Experimental Kiln. *Archaeometry* **5**, 93–104, 1962

Received February 10, 1984; Revised June 1, 1984

Accepted June 15, 1984

Frequency dependence of Q for seismic body waves in the Earth's mantle

A. Ulug* and H. Berckhemer

Institut für Meteorologie und Geophysik, Feldbergstr. 47, 6000 Frankfurt, Federal Republic of Germany

Abstract. In this paper an attempt is made to determine the frequency dependence of Q in the Earth's mantle in the frequency range 0.03–1.5 Hz from the spectral ratio of teleseismic S - and P -waves. Digital broad-band data of 17 earthquakes at $40^\circ < \Delta < 90^\circ$ recorded at the Central Seismological Observatory of the Federal Republic of Germany at Erlangen were analysed. The method implies the following assumptions: frequency independence of the crustal transfer function, proportionality of $Q_P(f)$ and $Q_S(f)$, and proportionality of P - and S -source spectra. This last and most critical assumption was carefully investigated by kinematic and dynamic source models.

The calculated \bar{Q} -spectra for the individual events vary considerably but all have in common a general increasing trend with frequency which can best be represented by a power law $Q \sim f^\alpha$ with $0.25 < \alpha < 0.6$. A further increase in slope near 1 Hz suggests an absorption band corner with an upper cut-off relaxation time $\tau_m = 0.33 \pm 0.18$ s. The significance of the Q -spectra and their variability is estimated by manipulating semi-synthetic seismograms with different error-producing processes such as length and shape of the time window, superposition of noise, digital filter process and source spectra. It is concluded that none of these processes is able to destroy or to imitate the observed increasing trend of Q with frequency.

The results are compared with those from other seismological investigations and from laboratory experiments on mantle rocks at high temperature and in the seismic frequency band.

Key words: Anelastic wave attenuation in the Earth's mantle – Frequency dependence of Q – Spectral ratio method

1. Introduction

Amplitude and shape of seismic signals are, to a large extent, influenced by the anelastic attenuation in the earth. This physical property is most conveniently de-

scribed by the elastic quality factor Q . Detailed knowledge of Q in space and frequency is not only required as a crucial correction term for restoring the seismic source function from teleseismic signals or for generating realistic synthetic seismograms, but also provides intrinsic information on physical properties of the material traversed by the seismic waves. Anelasticity in mantle and core is most probably controlled by thermally activated solid state processes and, therefore, highly temperature dependent. Comparison of laboratory experiments, theory and field observations may lead to a better understanding of the physical processes governing anelasticity, and to temperature estimates in the earth.

Until a few years ago Q was generally treated by seismologists as a frequency independent quantity, chiefly because of the lack of more precise information. The constraints to assume some increase of Q with frequency, however, are growing. For the Earth's mantle the number of specific investigations with regard to the frequency dependence of Q is still very limited, e.g. Sipkin and Jordan (1979), Lundquist and Cormier (1980), Der et al. (1982). These and other studies are based on differences in attenuation derived from long- and short-period seismograms, but not from continuous spectra over a broader frequency band. The mode of variation of Q with frequency deduced from these data remains, therefore, ambiguous. Also, depending on the method applied, it is often difficult to separate spatial variations of Q from those in frequency (e.g. Der et al. 1982). This is true, in particular, for all Q determinations from surface waves.

The exclusive aim of the present study is to obtain Q -spectra estimates for mantle body waves. No attempt is made to determine absolute values or spatial variations of Q . The authors were motivated to this study by laboratory measurements of Q by Berckhemer et al. (1979, 1982 and unpublished results). These experiments, which surpassed the solidus temperature of peridotite and dunite, were considered to be of relevance to the anelasticity of the upper mantle, especially of the asthenosphere. Consequently, the Q -structure in the Earth's mantle was also investigated, in particular its possible frequency dependence, and both results compared.

A continuous spectrum of $Q(f)$ can only be determined if attenuation is measured over a sufficiently wide frequency band along a well specified ray path.

* Present address: Dokuz Eylül Üniversitesi, Institute of Marine Sciences and Technology, P.O. Box 478, Konak-Izmir, Turkey

Offprint requests to: H. Berckhemer

This requires knowledge of the seismic signal spectrum at both ends of the path, and of all factors influencing the amplitude spectrum along the path.

Since the frequency dependence of Q can not be expected to be very large it is insufficient to replace the source spectrum just by some simple theoretical model, as done for constant- Q estimates by several authors (e.g. Tsujiura, 1966; Okada et al., 1970; Burdick, 1978). Use of the spectral ratios of multiple ScS (Sipkin and Jordan, 1979) or ScS/ScP (Kanamori, 1967) eliminates, in principle, the problem of the source spectrum. However, these signals are typically of low signal/noise ratio, in particular at short periods, and are affected by the particular conditions of the core-mantle boundary zone (low- Q zone?).

2. Principles of the S/P spectral ratio method

In the present study we use the spectral ratio of S and P of a teleseismic event where both phases are recorded at the same station. This has the advantage of almost identical ray paths, high signal to noise ratio and large contrast in attenuation for both wave types. In fact the less attenuated P -signal serves as a kind of reference signal for the unknown source time function. The necessary assumptions to be made are stated below and will be subject to a critical discussion in Sect. 3.

To derive the basic equation for the determination of $Q(f)$ it is convenient to describe the transmission of the seismic signal from the source to the receiver in the frequency domain by a chain of linear filters:

$$A(f) = S(f) C(f) M(f, \Delta) \phi(f, R), \quad (1)$$

where $A(f)$ = seismogram spectrum,
 $S(f)$ = seismograph transfer function,
 $C(f)$ = transfer function of the receiver crust,
 $M(f, \Delta)$ = mantle transfer function,
 Δ = epicentral distance,
 $\phi(f, R)$ = source spectrum with
 R = source radiation characteristic.

All functions may be complex and, with the exception of $S(f)$, depend on the respective wave type. We are particularly interested in the mantle transfer function $M(f, \Delta)$ which consists of the frequency independent geometrical spreading factor $G(\Delta)$ and the frequency dependent anelastic absorption factor $\exp[-\int \gamma(f, r) ds]$

along the ray path s , where, for a spherically symmetrical earth, the absorption coefficient $\gamma(f, r)$ is a function of the geocentric distance r and the frequency f . Since γ is related to Q by

$$\gamma(f, r) \approx \pi f / [Q(f, r) v(r)], \quad (2)$$

with the propagation velocity $v(r)$ of the wave type considered, we obtain

$$M(f, r) = G(\Delta) \exp \left[-\pi f \int_s \frac{ds}{Q(f, r) v(r)} \right]. \quad (3)$$

Replacing the integral in Eq. (3) by the quotient t^* of travel time t and the mean value $\bar{Q}(f)$ along the ray

path s

$$\int_s \frac{ds}{Q(f, r) v(r)} = \frac{t}{\bar{Q}(f)} = t^*, \quad (4)$$

the spectral ratio of S - and P -wave recorded at the same station by the same instrument may be written according to Eqs. (1), (3) and (4) as

$$\frac{S(f)}{P(f)} = \frac{C_S(f)}{C_P(f)} \cdot \frac{G_S(\Delta)}{G_P(\Delta)} \cdot \exp \left\{ \pi f \left[\frac{t_P}{\bar{Q}_P(f)} - \frac{t_S}{\bar{Q}_S(f)} \right] \right\} \frac{\phi_S(f, R_S)}{\phi_P(f, R_P)}. \quad (5)$$

In order to solve Eq. (5) for $\bar{Q}_S(f)$ and $\bar{Q}_P(f)$ we introduce the following approximations and assumptions:

a) Approximative frequency independence of the crustal transfer functions $C_S(f)$ and $C_P(f)$ in the frequency range 0.01–3 Hz

$$C_S(f), C_P(f) \approx \text{const.}, C_S/C_P = n (\approx 1). \quad (5a)$$

b) Approximative equality of the ray divergence factors $G_S(\Delta)$ and $G_P(\Delta)$

$$G_S(\Delta) \approx G_P(\Delta). \quad (5b)$$

c) Approximative proportionality of $\bar{Q}_S(f)$ and $\bar{Q}_P(f)$. With the common assumption that anelasticity is confined to shear deformation,

$$\bar{Q}_P(f) = k \bar{Q}_S(f), \quad k = \frac{3}{4} (v_P^2/v_S^2) \approx \frac{3}{4} (t_S^2/t_P^2). \quad (5c)$$

d) Assumption of proportionality of the source spectra ϕ_S and ϕ_P . This is certainly correct for a point source but neglects the rupture propagation process of a source of finite dimensions:

$$\phi_S(f, R_S)/\phi_P(f, R_P) \approx R_S/R_P = g. \quad (5d)$$

Under these conditions Eq. (5) reads

$$\frac{S(f)}{P(f)} = m \exp \left[\pi f \frac{1}{\bar{Q}_S(f)} \left(\frac{t_P}{k} - t_S \right) \right] \quad (6)$$

$$\text{with } m = ng. \quad (6a)$$

Equation (6) may be resolved with respect to $\bar{Q}_S(f)$, and with Eq. (5c) one obtains

$$\bar{Q}_S(f) = \frac{\pi f}{\ln S(f) - \ln P(f) - \ln m} \frac{\frac{4}{3} t_P^3 - t_S^3}{t_S^2},$$

$$\bar{Q}_P(f) = \frac{3}{4} \frac{t_S^2}{t_P^2} \bar{Q}_S(f). \quad (7)$$

All quantities in Eq. (7), except the factor m , can be directly determined from the seismogram if the travel times are known. The direct determination of the factor m , which reflects essentially the ratio of the S - and P -radiation patterns in the direction of the ray from the source to the station, would require a very accurate and stable fault-plane solution for the event. This can be circumvented by an empirical determination of m using a standard value of \bar{Q}_S (or \bar{Q}_P) at a reference

frequency f . As pointed out already, we are only concerned with the frequency dependence of Q , not with its absolute level. Standard values of $t^*(\Delta) = \frac{t(\Delta)}{Q(\Delta)}$ are available from several Q_P - and Q_S -models. We have chosen t^* values from model SL1 (Anderson and Hart, 1978) and assigned them to a reference frequency $f_r = 0.1$ Hz, typical for the range at which SL1 was determined. Any other reasonable model would do it as well and simply shift the level of $Q(f)$ somewhat, without a significant influence on the frequency dependence of Q . From Eq. (7) it follows

$$\ln m = \ln \left[\frac{S(f)}{P(f)} \right] - \pi 0.1 t_S^*(\Delta) \frac{\frac{4}{3} t_P^3 - t_S^3}{t_S^3}, \quad (8)$$

$$f = 0.1 \text{ Hz}$$

3. Discussion on the assumptions involved

Crustal transfer function

The layered structure of the receiver crust acts as a seismic frequency filter. The ratio of surface motion and wave motion at the base of the crust, the crustal transfer function, has been calculated for many crustal structures and angles of incidence i by various authors. Kanamori (1967) has shown that, e.g. for a 4-layer crust and steep incidence ($i = 10^\circ - 20^\circ$), the ratio of the vertical surface motion and P -wave amplitude in the frequency range 0–1 Hz assumes an almost constant average level of 2.5 with variations of $\pm 20\%$, and similar for the ratio of the horizontal ground motion and SV amplitude. Leblanc (1967) has shown that the average of 20 transfer functions of different crustal models varies in the range $0.4 < f < 1.6$ Hz by less than ± 1 dB about a frequency independent mean value. This is considered a sufficient justification to approximate C_S and C_P by frequency independent mean values which do not have to be known explicitly if m is determined empirically.

Ray divergence factor

Since the ray divergence at a given epicentral distance is almost identical for P - and S -waves (e.g. Shimshoni and Ben Menahem, 1970), the ratio of both can safely be assumed to be unity. In any case ray divergence would not affect the frequency dependence of Q .

Relation between Q_S and Q_P

An isotropic solid having anelastic properties can conveniently be described by complex elastic moduli

$$M(f) = M'(f) + iM''(f). \quad (9)$$

Then, the corresponding elastic quality factor is defined as

$$Q(f) = M'(f)/M''(f). \quad (10)$$

The moduli governing the S - and P -wave propagation, respectively, are

$$M_S(f) = \mu'(f) + i\mu''(f), \quad (11a)$$

$$M_P(f) = K'(f) + \frac{4}{3}\mu'(f) + i[K''(f) + \frac{4}{3}\mu''(f)], \quad (11b)$$

and the corresponding Q -factors

$$Q_S(f) = \frac{\mu'(f)}{\mu''(f)}, \quad (12a)$$

$$Q_P(f) = \frac{K'(f) + \frac{4}{3}\mu'(f)}{K''(f) + \frac{4}{3}\mu''(f)}, \quad (12b)$$

or

$$\frac{Q_S(f)}{Q_P(f)} \approx \frac{v_S^2}{v_P^2} \left[\frac{4}{3} + K''(f)/\mu''(f) \right], \quad (13)$$

where, for $Q \gg 1$,

$$v_S^2 \approx \frac{\mu'}{\rho}, \quad v_P^2 \approx \frac{K' + \frac{4}{3}\mu'}{\rho}. \quad (14)$$

If anelasticity is entirely due to energy loss at shear deformation ($K'' = 0$), which is a widely adopted postulate and also used in the present paper, then we arrive at Eq. (5a). It is seen from Eq. (13) that proportionality of Q_S and Q_P exists as long as the velocity dispersion remains negligible. Even if $K''(f) \neq 0$ but is small relative to μ'' or varying with frequency in a similar way, proportionality of Q_S and Q_P may still be assumed a good approximation.

Seismic source spectra

The assumption of proportionality of ϕ_P and ϕ_S certainly needs careful and critical consideration. In a homogeneous elastic medium the far-field displacement $u(r, t)$ of a seismic point source with moment function $M_0(\tau)$ may be written in terms of radial (P) and tangential (S) displacement components:

$$u^P = \frac{1}{4\pi\rho} \frac{dM_0(\tau)}{d\tau} R_P(\vartheta, \lambda) G_P(\Delta), \quad (15a)$$

$$u^S = \frac{1}{4\pi\rho} \frac{dM_0(\tau)}{d\tau} R_S(\vartheta, \lambda) G_S(\Delta). \quad (15b)$$

From Eq. (15a and b) it is evident that for a point source u^S and u^P , and therefore also the respective source spectra ϕ_S and ϕ_P , are proportional, with R_S/R_P as the factor of proportionality. In our notation the radiation functions of a double couple point source are: for P

$$\text{(radial)} \quad R_P(\vartheta, \lambda) = \frac{1}{v_P^3} \sin 2\vartheta \cos \lambda,$$

for S

$$\text{meridional} \quad R_{Sm}(\vartheta, \lambda) = \frac{1}{v_S^3} \cos 2\vartheta \cos \lambda,$$

$$\text{(lateral)} \quad R_{Sl}(\vartheta, \lambda) = -\frac{1}{v_S^3} \cos \vartheta \sin \lambda, \quad (16)$$

where ϑ and λ are co-latitude and longitude on the focal sphere with respect to the fault plane and the slip vector. As an angular average one would expect to find $R_S/R_P = g \approx m$ to be of the order of $v_P^3/v_S^3 \approx 5$. Extreme

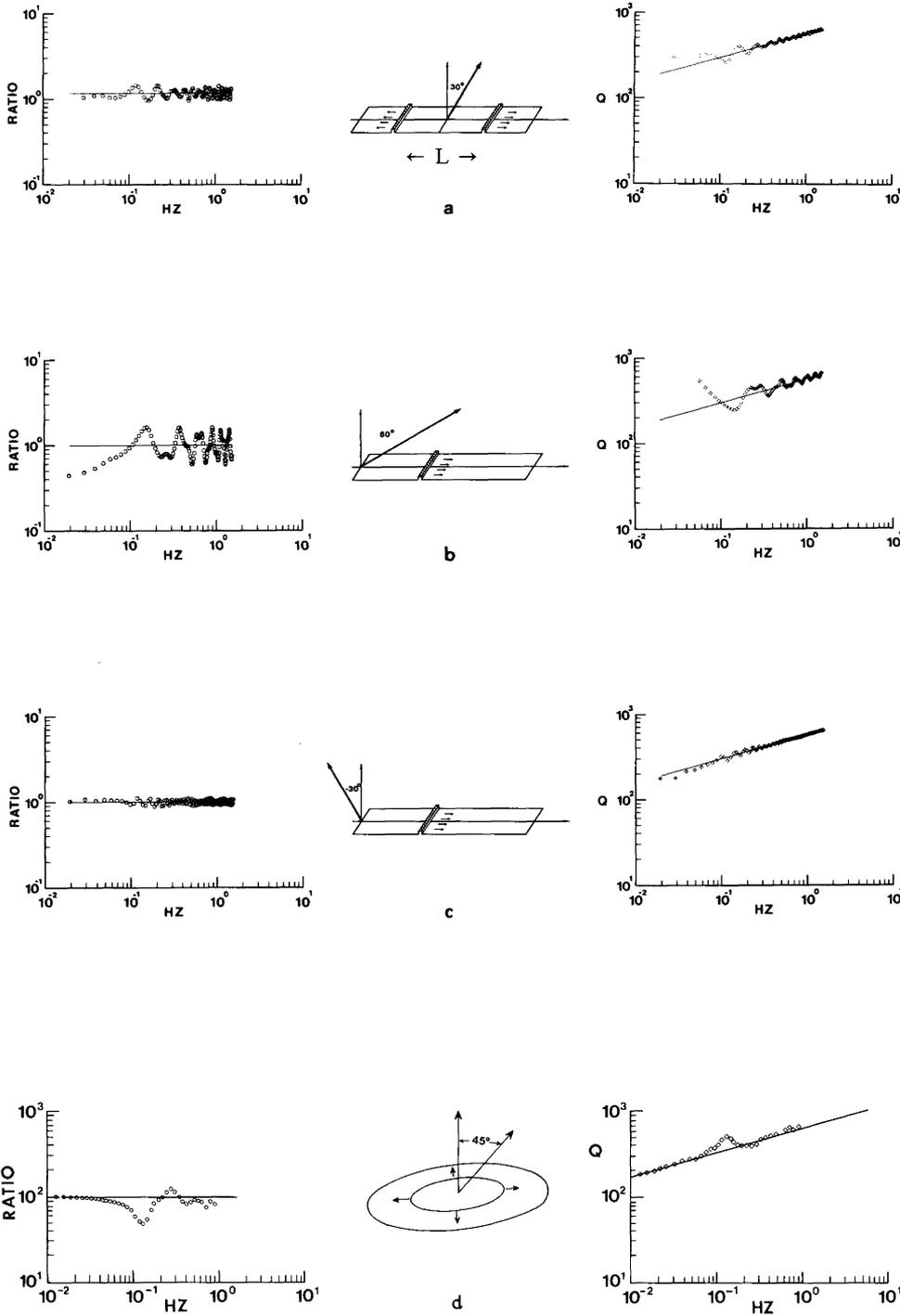


Fig. 1. Synthetic S/P -spectral ratios (*left*) and Q -spectra (*right*) for finite source models and different radiation angles (*middle*). Straight lines in the spectra correspond to a respective point source. **a-c:** Haskell type dislocation models **d:** Madariaga stress relaxation model. In the synthetic Q -spectra, absorption according to a power law in $Q(\alpha=0.276)$ is assumed

values are to be expected near nodal planes for $\vartheta \approx 0, \pi/4, \pi/2$ and $\lambda \approx 0, \pi/2$. The question is whether the proportionality of ϕ_S and ϕ_P is also a justified assumption for real earthquakes with an extended source. If we consider the corner frequencies $f^*(P)$ and $f^*(S)$ as representative quantities of the source spectra of P and S , respectively, it turns out that the information on $f^*(P)/f^*(S)$ found in literature is controversial. Following the critical review recently published by Hanks (1981), it is found that $f^*(P)/f^*(S) > 1$ in the majority of observational cases. However, some near-source observations in low-loss material show $f^*(P)/f^*(S) \lesssim 1$. Hanks concluded from numerical estimates that dif-

ferent absorption of P - and S -waves may account for some of the difference of $f^*(P)$ and $f^*(S)$ but not for all. This is also not to be expected for a source of finite dimensions and finite rupture velocity.

In order to gain some feeling for the possible influence of finiteness of the source on ϕ_S/ϕ_P in the frequency band under consideration, we have applied two different types of source models to produce synthetic source spectra.

A: One-dimensional unilateral and bilateral dislocation models of Haskell type with rupture propagation along an elongated rectangular fault of $L=50$ km. Rupture velocities $v_r=0.9v_S$ and $v_r=0.65v_S$ were used with

Table 1. List of earthquakes analysed

No.	Region name	Date	Origin time h:m:s	Coordinates:		Δ degrees	Depth km	Mag- nitude
				latitude φ°	longitude λ°			
1	Oaxaca-Mexico	29.11.78	19:52:53	17.0N-	96.0W	87.9	49	$M_s=7.8$
2	St. Elias-Alaska	28.02.79	21:27:08	60.6N-	141.6W	68.6	25	$M_b=6.8$
3	Okinawa	12.06.78	08:14:27	41.0N-	142.0E	79.8	40	$M_s=7.5$
4	Alaska	13.02.79	05:34:26	55.5N-	157.2W	74.8	24	$M_s=6.7$
5	Mexico	14.03.79	11:07:19	17.8N-	101.3W	90.3	59	$M_s=7.5$
6	N. Atlantic	22.04.79	09:50:18	33.0N-	39.7W	40.8	33	$M_b=5.7$
7	Kodiak-Alaska	20.05.79	08:14	56.7N-	156.7W	73.5	72	$M_s=6.7$
8	Tibet	20.05.79	22:59:16	32.0N-	79.0E	52.4	31	$M_s=5.5$
9	Iran	16.09.78	15:35:56	33.2N-	57.4E	37.6	33	$M_b=6.1$
10	U.S.S.R.	01.11.78	19:48:29	35.4N-	72.7E	46.2	33	$M_b=6.3$
11	Kuriles	06.12.78	14:02:03	44.7N-	146.4E	78.2	97	$M_b=6.5$
12	Honshu	07.03.78	02:48:46	32.0N-	137.4E	85.7	430	$M_b=7.0$
13	Tibet	18.11.77	05:20:11	33.0N-	89.0E	58.1	33	$M_s=5.9$
14	Taiwan	02.09.78	01:57:34	24.9N-	122.0E	83.7	115	$M_b=6.0$
15	Colombia	23.11.79	23:40:30	4.8N-	76.2W	84.7	108	$M_b=6.4$
16	Japan	18.01.81	18:17:24	38.7N-	142.8E	82.2	33	$M_b=6.2$
17	Japan	23.05.78	07:50:28	31.1N-	130.1E	83.1	161	$M_b=6.3$

$v_s=3.4$ km/s. Some typical examples of model calculations with $v_r=0.9v_s$ and arbitrary, but uniform rise-time along the fault are shown in Fig. 1a-c. Since this is only a consideration of secondary importance in the context of this paper, for details refer to Ulug (1983). The spectral ratios shown on the left of Fig. 1 are smoothed, as described in Sect. 5, and normalized by putting $R_S/R_P=1$. It is seen that the oscillations strongly depend on the radiation angle ϑ (Eq. 16). Radiation under steep angles to the rupture plane produces unimportant oscillations only (Fig. 1a and c), while comparatively large oscillations occur at low angles (Fig. 1b) due to the increasing Doppler effect. With $v_r=0.9v_s$, a rather unfavourable choice has been made in this respect. Negligible oscillations are found for a set of models with $v_r=0.65v_s$ for most radiation angles.

B: Madariaga stress relaxation models with constant rupture velocity. It may be argued that Haskell type models might lead to unrealistic results because in these models $f^*(P) \leq f^*(S)$ (Hanks, 1981). For that reason we have also used the more sophisticated source model of Madariaga (1976). It represents a dynamic solution of the wave radiation from a circular rupture front propagating with the finite rupture velocity v_r . As shown by Madariaga, generally, $f^*(P) \geq f^*(S)$. The diameter of the rupture area is chosen to be 30 km and, as in the previous models, $v_s=3.4$ km/s, $v_r=0.9v_s$. Figure 1d shows the smoothed ϕ_S/ϕ_P ratio for a radiation angle $\vartheta=45^\circ$. There is some disturbance in the spectral ratio slightly above the corner frequencies but very little in the lower and higher part of the spectrum.

In summary we may conclude that, despite the more or less pronounced oscillations, it is important to note that all spectral ratios follow a constant mean level throughout the frequency band which may be taken as a justification for our assumption $\phi_S/\phi_P \approx \text{const}$.

Most of the earthquakes used in this study occurred at normal depth. No contributions of waves generated by reflections and wave conversion at the surface near

the source have been taken into consideration in the models. It is to be expected that these secondary waves, if falling in the time window for spectral analysis, will create a tendency to moderate the differences between ϕ_S and ϕ_P . It is a basic idea of this study that even if an individual observation might have been made under unfavourable conditions a sufficient number of cases should give, in a statistical sense, a representative picture as regards $Q(f)$.

4. Selection and preparation of data

For this kind of study broad-band digital seismograms with high signal/noise ratio are required. Seventeen events of $M \geq 5.5$, recorded at the Central Seismological Observatory of the Federal Republic of Germany at Erlangen (station abbreviation GRF) at epicentral distances $40^\circ < \Delta < 90^\circ$, were selected. Care was taken for equally good quality of P and S signals, i.e. for sufficient distance from nodal planes. Of the 17 earthquakes, 4 had a focal depth of more than 100 km and one was deeper than 300 km. The list of events is given in Table 1.

The frequency response of the broad-band seismometers (Wielandt and Streckeisen, 1982) is flat with respect to ground velocity from 0.05–5 Hz. The dynamic range extends to 132 dB with a resolution of 66 dB. The sampling rate is 20 s^{-1} . P -wave spectra were computed from the vertical component. S -wave signals were separated into SV and SH components. For spectral analysis, only SH was used in order to eliminate PS and waves converted near the receiver site.

In order to get a satisfactory spectral estimate of the signal in question, the time window T was usually taken as $T \approx 3T_0$ (Abramovici, 1973), where T_0 is the longest period at which P and S still have a good signal/noise ratio. The window length had to be limited to 40–100 s to keep undesired phases like PL out. Be-

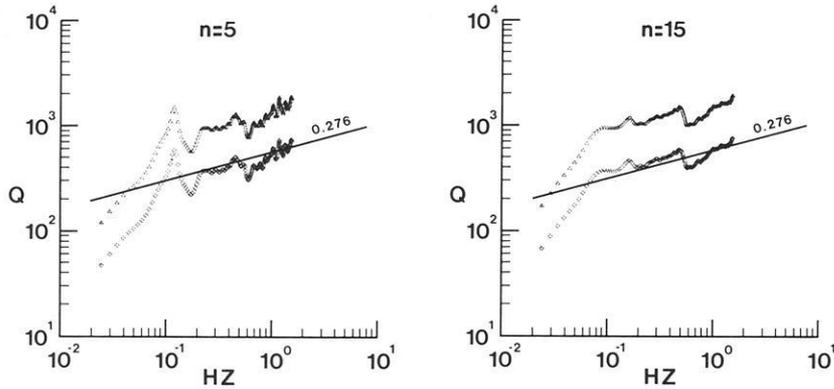


Fig. 2. Effect of smoothing of the spectra using the mean of $2n+1=11$ and 31 values, respectively

yond $\Delta \approx 60^\circ$, PcP and ScS overlap with P and S , respectively. Kurita (1969) has shown that, if PcP and P or ScS and S spectra are similar, PcP or ScS respectively would modulate the P or S spectrum. Since the PcP/P amplitude ratio is less than 0.2 for $\Delta > 60^\circ$ (Fraser and Chowdhury 1974) and ScS/S less than 0.5 for $\Delta > 70^\circ$ (Mitchell and Helmberger, 1973), the modulation of the P - and S -spectra by PcP or ScS will not be very severe. The superposition of direct P and S with near-source surface reflections can also not be avoided for shallow foci but is, in some way, even desirable, as mentioned at the end of Sect. 3. It may, however, also cause some modulation of the source spectrum.

The choice of the shape of the time window was also subject to a comparative investigation. It turned out, however, that in practical cases the difference in the final result for a simple box car window and a cosine window, or a combination of both, was negligible. Therefore, the box car window was used.

5. Data analysis and results

P - and S -signal spectra were calculated by fast Fourier transformation at 1024 equidistant frequencies. For sufficient resolution at lower frequencies and since the signal to noise ratio admits only an upper signal frequency limit of 1.5 Hz, the original data were reduced to a sampling rate of 5 s^{-1} corresponding to a Nyquist frequency of $f_v = 2.5 \text{ Hz}$ and a spectral resolution of $\Delta f \approx 0.005 \text{ Hz}$.

The spectra show oscillations and break-ins caused for reasons discussed earlier (see Fig. 3). To reduce these irregularities, the raw P - and S -spectra were smoothed taking the unweighted mean of the value in question and n neighbouring values to both sides, i.e. of $2n+1$ values. Figure 2 shows the result of the smoothing process on $Q(f)$ for $n=5$ and $n=15$. The general trend which reflects the frequency dependence of Q appears more clearly after smoothing and is not modified by the procedure. The subsequent results are obtained with $n=7$.

From the smoothed signal spectra the Q -spectra were calculated according to Eqs. (7) and (5c). The value of m was determined by inserting a standard t^* into Eq. (8). The full procedure is exemplified for the Kodiak-Alaska earthquake of 20 May 1979. Figure 3 shows the time functions of P , S and noise (some minutes before the P -onset), their raw and smoothed spec-

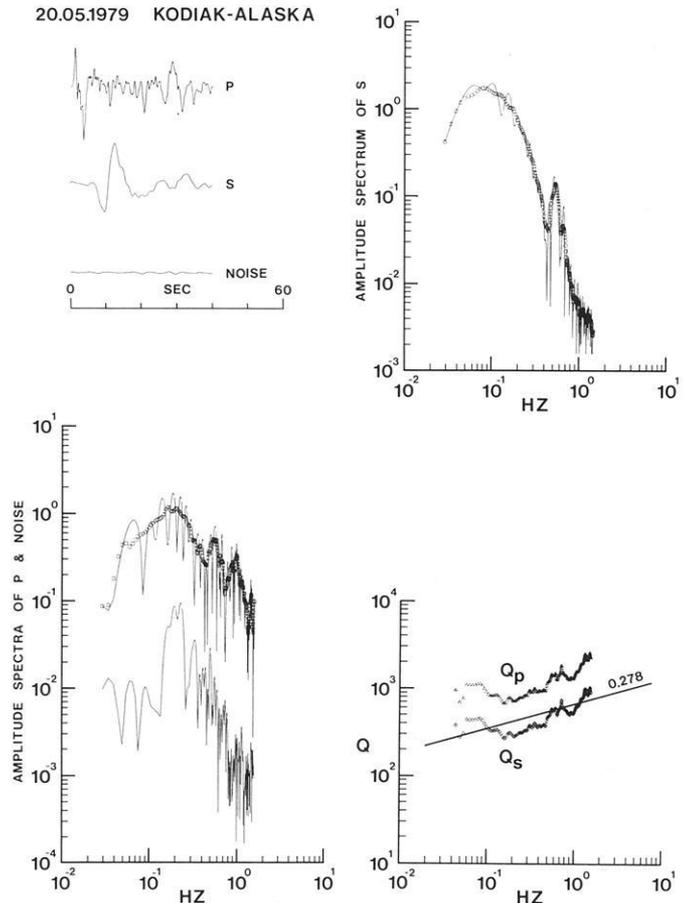


Fig. 3. Example of the numerical data treatment. P - and S -signal time functions, P , S and noise ground displacement Fourier spectra; Q_p and Q_s spectra of the earthquake of 20.05.1979 at Kodiak-Alaska, $M_s = 6.7$

tra and $Q_p(f)$ and $Q_s(f)$ in the frequency range 0.06–1.5 Hz. As a representative collection of the results obtained, the Q -spectra of 9 of the 17 earthquakes investigated are shown in Fig. 4. The mean of 15 Q -spectra, scaled at $f=1 \text{ Hz}$ (thick line), is given in Fig. 5, together with the standard deviation (dotted area) and two model curves. Although the spectra vary considerably in detail, they all have in common a generally increasing trend of Q with frequency, which may be expressed by a power law represented as a best fit straight line in the figures on bi-logarithmic scale. In Fig. 6 the α -values of all 17 events are plotted with

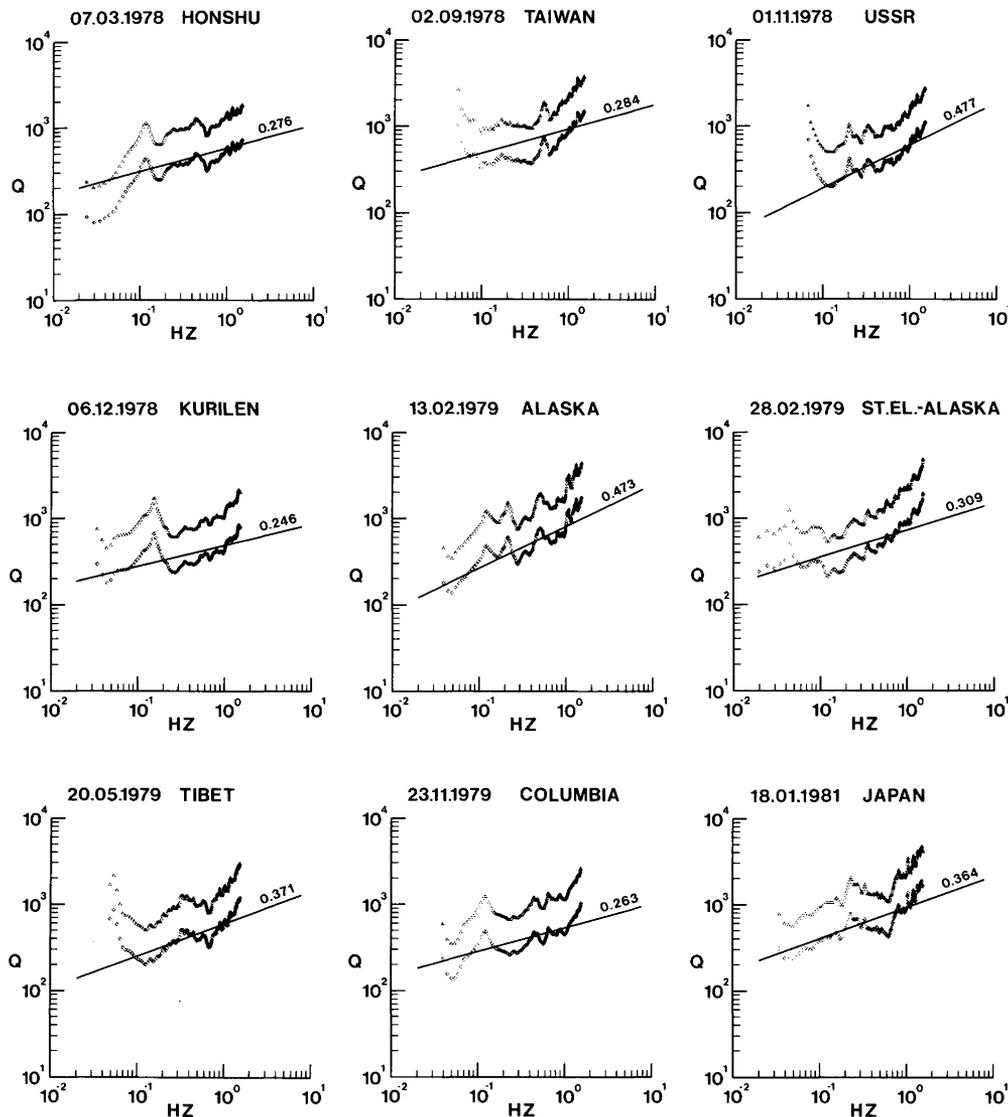


Fig. 4. $\bar{Q}(f)$ for P (upper curve)- and S (lower curve)-signals along the ray path for 9 of 17 events. Focal data are given in Table 1

rather conservative error bars versus epicentral distance. It may be concluded that α -values range from 0.25–0.6 with some tendency to increase with decreasing distance. Since signals at shorter distances have travelled relatively longer in the asthenosphere than those which penetrated very deep in the mantle, the data may indicate that absorption in the asthenosphere is represented by higher α -values.

In quite a number of cases, as seen from Figs. 4 and 5, a simple power law fails to describe properly the increasing slope above 1 Hz. As an alternative approach we have, therefore, attempted to fit the spectra by the high frequency end of an absorption band model with frequency independent relaxation density, as proposed by Liu et al. (1976) and applied, e.g. by Sipkin and Jordan (1979) and by Lundquist and Cormier (1980). The model curve b was calculated according to

$$Q(f) = Q_B \frac{\pi}{2} \left\{ \tan^{-1} \left[\frac{2\pi f(\tau_M - \tau_m)}{1 + 4\pi^2 f^2 \tau_M \tau_m} \right] \right\}^{-1} \quad (17)$$

(Kanamori and Anderson, 1977), where Q_B is the value in the absorption band and τ_M and τ_m are the upper and the lower cut-off relaxation times.

From fits of the individual Q -spectra with this model we obtained $\tau_m = 0.33 \pm 0.18$ s with no significant dependence on the epicentral distance. The spectral mean in Fig. 5 is fitted with $\tau_m = 0.4$ s.

6. Significance of the results

Before speculating on further features of the Q -spectra it is necessary to carefully examine factors which might have influenced the results obtained. For this purpose synthetic or semi-synthetic seismograms were used and manipulated with artificial disturbances.

Concerning the effect of the finiteness of the source on the Q -spectra, we make use of the model calculations of Sect. 3. On the right of Fig. 1 synthetic Q -spectra are shown corresponding to the respective source spectra ratio to the left. It has been presumed that anelastic absorption on the way from the source to

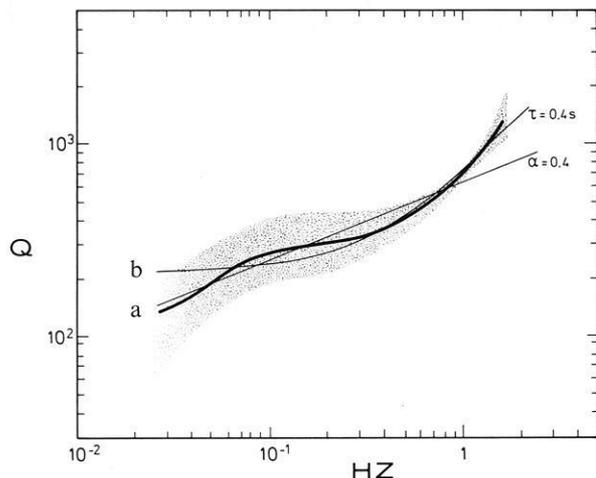


Fig. 5. Mean of 15 Q -spectra scaled at 1 Hz (thick line) and range of standard deviations (dotted area). Power law fit with $Q \approx f^{0.4}$ (a) and, alternatively, fit by a constant- Q absorption band model with cut-off relaxation time $\tau_m = 0.4$ s (b)

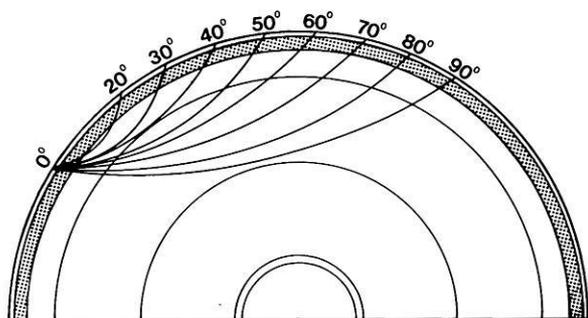
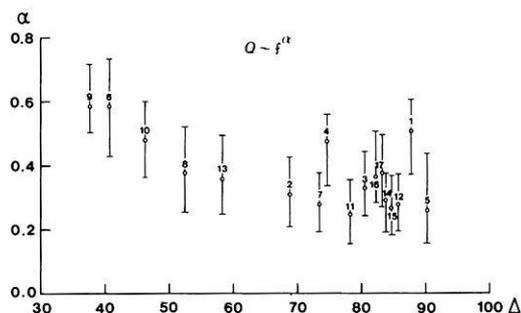


Fig. 6. α -values (with error bars) of $Q(f) \approx f^\alpha$ for all events analysed as function of epicentral distance Δ (degrees) and the corresponding mantle ray paths

the station follows a power law $Q \sim f^\alpha$ with $\alpha = 0.276$. The Q -spectrum for an ideal source ($\phi_s \sim \phi_p$) would then be the straight line shown on all diagrams. The case 1c follows the true Q -spectrum very closely and 1a and d sufficiently closely. In the unfavourable case 1b we see, in particular at the low frequency end, considerable deviations but still $Q(f)$ follows the straight line in its general trend. The low frequency behaviour and some irregularities of $Q(f)$ in the spectra of Fig. 4 might be explained by seismic source effects.

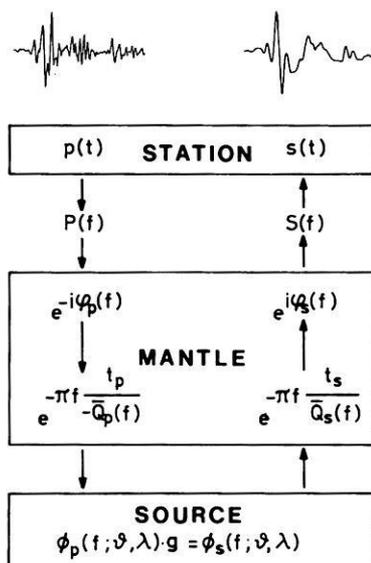


Fig. 7. Scheme for synthesizing an S -signal corresponding to an observed P -signal for a given $Q(f)$ model

The influence of other factors has been studied by means of a pair of semi-synthetic seismograms which are connected by an ideal power law absorption. To be most realistic we started from an actual P -seismogram and its Fourier transform. By inverse application of the absorption operator of an idealized anelastic mantle, the P -source spectrum ϕ_p was obtained. Assuming proportionality of ϕ_p and ϕ_s , the corresponding forward operation was executed to obtain the S -time function corresponding to the starting P -signal. As seen from the scheme in Fig. 7, the process takes into account an amplitude attenuation factor, $\exp[-\pi ft/\bar{Q}(f)]$, as well as absorption linked phase shift, $\varphi(f, t)$. Absorption was assumed to follow $\bar{Q}_p(f) = Af^\alpha$ and $\bar{Q}_s(f) = Bf^\alpha$, with $A = 1,345$, $B = 565$ and $\alpha = 0.276$. The parameters $g \approx m$, t_p , t_s were taken from the actual event to which the P -signal belonged.

The phase shift, $\varphi(f, t)$, is obtained from the approximate dispersion law for weak frequency dependence (Futterman, 1962; Kanamori and Anderson, 1977)

$$\bar{c}(f)/\bar{c}(f_0) \approx 1 + \frac{1}{\pi \bar{Q}(f)} \ln(f/f_0), \quad (18)$$

where $\bar{c}(f)$ is the mean phase velocity and $\bar{Q}(f)$ the corresponding Q value along the ray. $\bar{c}(f_0)$ may be equalized with the signal velocity $v = s/t$ for $f_0 = 1$ Hz. The phase shift, $\varphi(f)$, is related to $\bar{c}(f)$ by

$$\varphi(f) = 2\pi f [t - s/\bar{c}(f)]. \quad (19)$$

The pair of P and S signals thus obtained was manipulated in several ways. Of course $Q(f)$, determined according to Eqs. (7) and (5c), resulted in an almost straight line with the slope of 0.276, as expected. Small deviations are attributed to numerical noise (smoothing procedure). Perhaps the most important results from these numerical experiments are those concerning the influence of artificial noise. In Fig. 8a white noise (with respect to ground velocity) is superimposed on the signal. The Q -spectra show a marked increase at high

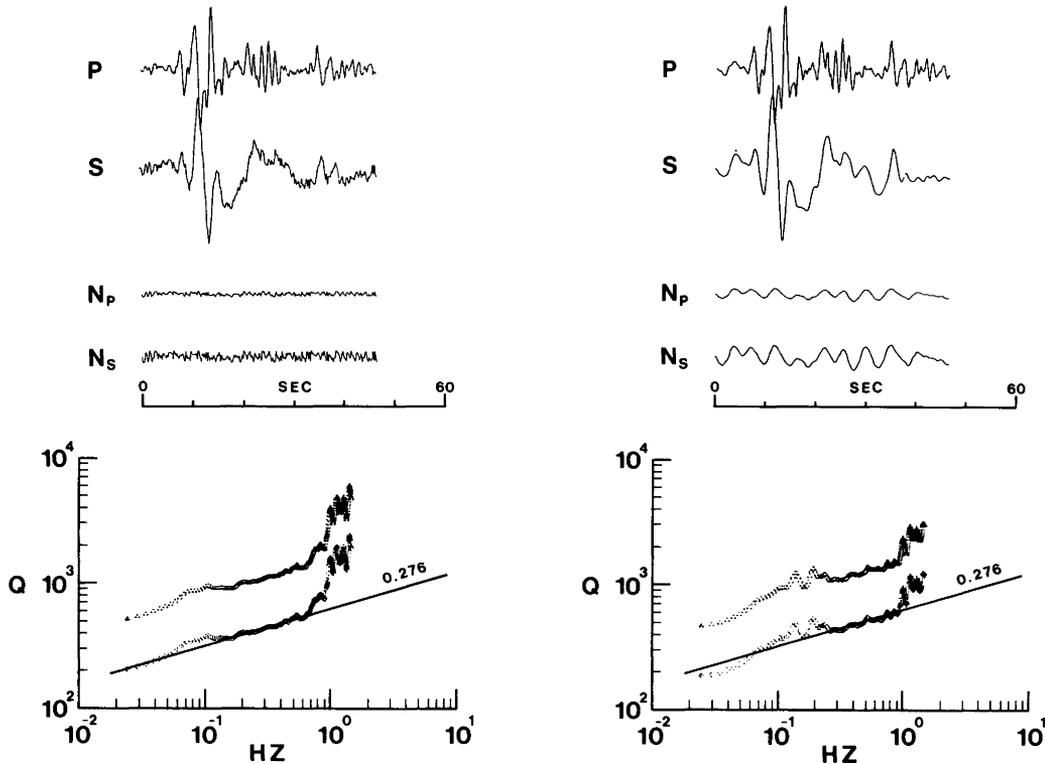


Fig. 8. Effect of noise, superimposed on a pair of semi-synthetic *P*- and *S*-signals, on the *Q*-spectra. *Left*: white noise; *Right*: coloured noise (central frequency 0.2 Hz)

frequencies where the signal amplitude is rapidly dropping (compare Fig. 3). However, the mean slope of 0.276 is still recognizable without difficulty. This may account, at least in part, for the rising high-frequency end of some of the *Q*-spectra in Fig. 4, and may also be of importance in judging results of other authors. In Fig. 8b the white noise is convolved with a wavelet of $T_0 = 5$ s in order to obtain "coloured" noise with a main frequency of 0.2 Hz. As expected, this tends to raise the *Q*-level in the frequency range around 0.2 Hz but *Q*(*f*) still follows the slope of 0.276 quite well. From this experiment we may conclude that noise superimposed on the signal may modify the *Q*-spectrum considerably, in particular at higher frequencies, but still the general trend remains preserved.

Other experiments carried out with our semi-synthetic pair of *P*- and *S*-signals may just briefly be mentioned. As already tested with real data (Sect. 4), the shape of the time window has little influence on *Q*(*f*). Variation of the length of the window by a factor of 2.5 modified *Q*(*f*) only at the low-frequency end.

We may conclude this section with the following statement, valid for the frequency range considered. Several factors are capable of modifying the *Q*-spectra in detail, but no conceivable process can imitate or destroy the observed general increasing trend of *Q*(*f*) which can most simply be expressed by a power law.

7. Discussion

Comparison with other seismic evidence

In this section we compare our results with those of other investigations. If only short- and long-period *Q*-

Table 2. Comparison of equivalent α -values in different investigations

Study	α	Frequency range (Hz)
<i>Whole earth</i>		
Anderson, Minster (1979)	0.2-0.4	10^{-8} - 10^{-2}
<i>Mantle</i>		
This study	0.25-0.6	0.03 - 1.5
Sipkin, Jordan (1979)	0.35	0.01 - 2
Der et al. (1982)	0.15	0.01 - 2
Sacks (1980)	0.15-0.25	0.001- 0.3
Clements (1982)	0.2 -0.5	0.1 - 1
Zschau (personal communication 1983)	0.4 -0.6	0.1 - 1
<i>Crust</i>		
Aki (1980)	0.6 -0.8	0.5 -25
Mitchell (1980)	0.3 -0.5	0.025- 1
Singh et al. (1982)	1.0	3 -25
Roecker et al. (1982)	0.5 -1.0	0.4 -48
Thouvenot (1983)	0.25	10 -25
<i>Laboratory data</i>		
Berckhemer et al. (1982)	0.25	0.003-30
Gueguen (personal communication 1982)	0.3	10^{-4} -10

values are available it is still possible to bring them into the form of a power law $Q \sim f^\alpha$. This has been attempted in Table 2.

Similarities in results for crust and mantle are probably just of formal character because the physical na-

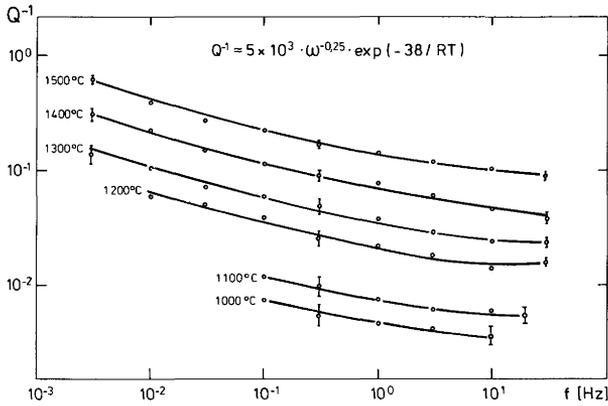


Fig. 9. Laboratory data on $Q^{-1}(f, T)$ for synthetic polycrystalline forsterite at shear deformation

ture of attenuation processes in the low-temperature, fractured and brittle crust, and in the high-temperature ductile mantle must be fundamentally different. The most specific results for the mantle, namely those of Sipkin and Jordan (1979) and ours, are in surprisingly good agreement although quite different techniques were used. The representation of the data by a power law is, however, not the interpretation given by Sipkin and Jordan. They applied relaxation band models with constant \bar{Q} inside the band (see Sect. 5) and got the closest fit for a lower cut-off relaxation time $\tau_m \approx 0.4$ s. There exists, however, a gap in their data in the period range $3 < T < 15$ s which leaves some freedom for the interpretation.

As already suggested in Sect. 5, our \bar{Q} -spectra can probably best be represented by a power law absorption band with a slope of $0.25 < \alpha < 0.4$ and a cut-off relaxation time $0.2 < \tau_m < 0.5$ s. This model is of the type proposed recently by Anderson and Given (1982) for absorption in the mantle. Their values of τ_m , however, are either considerably lower for the depth range $h < 500$ km or considerably higher for $500 < h < 2,300$ km. Of course it must be kept in mind that our \bar{Q} -values are weighted averages over the whole ray path, dominated by the absorption in the upper mantle. The effect of noise on the high-frequency end of the \bar{Q} -spectra, as discussed in Sect. 6, calls for care in drawing conclusions. In any case, values of $\alpha > 1$, appearing on some spectra, must be artefacts.

Comparison with laboratory data

Finally, let us come to the original aim of the study: the comparison of our laboratory data with Q in the Earth's mantle. Q_5^{-1} was measured in forced torsion oscillation experiments by the phase shift of stress and strain in the frequency range 3 mHz–30 Hz, at temperatures up to 1,500°C and under ambient pressure of 1 bar (Berckhemer et al. 1982). The maximum shear in the sample reached values of 3×10^{-5} , which was proven to be well within the limits of linearity. Polycrystalline synthetic forsterite, natural dunite and peridotite were used as candidate materials for the upper mantle. The $Q^{-1}(f)$ -spectra of dunite (Fig. 9), and similarly of polycrystalline forsterite, follow a power law with $\alpha \approx$

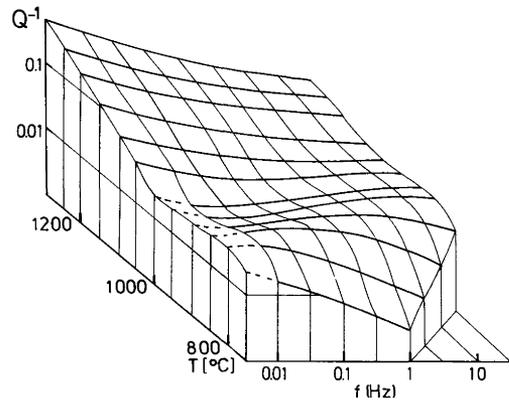


Fig. 10. Laboratory data on $Q^{-1}(f, T)$ for dry lherzoolite-peridotite (Balmuccia, zone of Ivrea, northern Italy) at shear deformation

-0.25 . Apparently the whole experimental frequency range lies inside the absorption band. This is in good agreement with the seismic results, at least for $0.03 < f < 1$ Hz. The laboratory experiments with lherzolite-periodotite (about 65% olivine, 30% pyroxenes) show a similar behaviour in the experimental frequency range for $T > 1,200^\circ\text{C}$ (Fig. 10). At lower temperature, however, a flat relaxation peak enters the field from the high frequency side. It appears to be the high-frequency corner of the absorption band. The slope of $Q^{-1}(f)$, however, is more gentle than -1 at the flank of the band, probably due to a smooth tapering-off of the relaxation density distribution. The laboratory experiments seem to reflect, at least in a qualitative way, the absorption properties of the mantle characterized by a weak power law inside the absorption band related to the so-called high-temperature background absorption (Anderson and Minster, 1979), and the high-frequency corner of the absorption band with a steeper slope.

The question is whether the increasing slope of the mantle Q -spectra above about 0.5 Hz (Figs. 4 and 5) may be related to the absorption band corner seen in the experimental Q -spectra of peridotite near 0.3 Hz at 850°C ? This is a problem of extrapolating laboratory data obtained at atmospheric pressure to asthenosphere conditions - since most of the mantle absorption takes place in the asthenosphere.

For thermally activated processes, the corner frequency $f_m \approx 1/\tau_m$ depends on temperature T and pressure p as $f_m = f_{n0} \exp[(A + PV^*)/RT]$, where A is the activation energy, V^* the activation volume of the respective material and process and R the universal gas constant. From the shift of f_m with T (Fig. 10), a value of $A \approx 550$ kJ/mol (130 kcal/mol) is derived. Values of the activation volume of minerals are very scarce due to the experimental difficulties of their determination. For creep in olivine, Karato and Ogawa (1982) obtained $V^* = 16 \pm 3$ cm³/mol. From dislocation recovery experiments by Kohlstedt et al. (1980), Karato (1981) derived $V^* = 19 \pm 2$ cm³/mol. From creep experiments on dunite (Ross et al. 1979), Kohlstedt et al. (1980) deduced $V^* = 18$ cm³/mol. Taking $V^* \approx 18$ cm³/mol as a mean and taking the continental geotherm of Ito and Kennedy (1967) we shall find the experimentally observed absorption band corner under asthenosphere

conditions: for $h=150$ km ($p=50$ kbar), $T=1,050$ °C at $f_m=0.7$ Hz; for $h=200$ km ($p=66$ kbar), $T=1,150$ °C at $f_m=3$ Hz. Indeed, these values compare fairly well with the trend in the mantle Q -spectra, which suggests a causal relation. This is, however, by no means a stringent proof, because only modest variations of temperature shift f_m by orders of magnitudes. f_m also turns out to be strongly depth dependent and the actually observed mantle Q corresponds to a weighted mean along the ray path. The latter problem was studied by Anderson and Given (1982) and by Lundquist and Cormier (1980).

References

- Abramovici, F.: Numerical application of a technique for recovering the spectrum of a time function. *Geophys. J. R. Astron. Soc.* **32**, 65–78, 1973
- Aki, K.: Attenuation of shear waves in the lithosphere for frequencies from 0.05 to 25 Hz. *Phys. Earth Planet. Inter.* **21**, 50–60, 1980
- Anderson, D.L., Given, J.W.: Absorption band Q model for the earth. *J. Geophys. Res.* **87**, 3893–3904, 1982
- Anderson, D.L., Hart, R.S.: Attenuation models of the earth. *Phys. Earth Planet. Inter.* **16**, 289–306, 1978
- Anderson, D.L., Minster, J.B.: The frequency dependence of Q in the earth and implications for mantle rheology and Chandler wobble. *Geophys. J. R. Astron. Soc.* **58**, 431–440, 1979
- Berckhemer, H., Auer, F., Drisler, J.: High-temperature anelasticity and elasticity of mantle peridotite. *Phys. Earth Planet. Inter.* **20**, 48–59, 1979
- Berckhemer, H., Kampfmann, W., Aulbach, E., Schmeling, H.: Shear modulus and Q of forsterite and dunite near partial melting from forced-oscillation experiments. *Phys. Earth Planet. Inter.* **29**, 30–41, 1982
- Burdick, L.J.: t^* for S waves with a continental raypath. *Bull. Seismol. Soc. Am.* **68**, 1013–1030, 1978
- Clements, J.: Intrinsic Q and its frequency dependence. *Phys. Earth Planet. Inter.* **27**, 286–299, 1982
- Der, Z.A., McElfresh, T.W., O'Donnell, A.: An investigation of the regional variations and frequency dependence of anelastic attenuation in the mantle under the United States in the 0.5–4 Hz band. *Geophys. J. R. Astron. Soc.* **69**, 67–99, 1982
- Frasier, C.W., Chowdhury, D.K.: Effect of scattering on PcP/P amplitude ratios at Lasa from 40° to 84° distance. *J. Geophys. Res.* **79**, 5469–5477, 1974
- Futterman, W.I.: Dispersive body waves. *J. Geophys. Res.* **67**, 5279–5291, 1962
- Hanks, T.C.: The corner frequency shift, earthquake source models, and Q . *Bull. Seismol. Soc. Am.* **71**, 597–612, 1981
- Ito, K., Kennedy, G.C.: Melting and phase relations in a natural peridotite to 40 kilobars. *Am. J. Sci.* **265**, 519–538, 1967
- Kanamori, H.: Spectrum of short-period core phases in relation to the attenuation in the mantle. *J. Geophys. Res.* **72**, 2181–2186, 1967
- Kanamori, H., Anderson, D.L.: Importance of physical dispersion in surface wave and free oscillation problems: Review. *Rev. Geophys. Space Phys.* **15**, 105–112, 1977
- Karato, S.: Comment on the effect of pressure on the rate of dislocation recovery in olivine, by D. L. Kohlstedt et al. *J. Geophys. Res.* **86**, 9319, 1981
- Karato, S., Ogawa, M.: High pressure recovery of olivine: Implications for creep mechanism and creep activation volume. *Phys. Earth Planet. Inter.* **28**, 102–117, 1982
- Kohlstedt, D.L., Nichols, H.P.K., Hornack, P.: The effect of pressure on the rate of dislocation recovery in olivine. *J. Geophys. Res.* **85**, 3122–3130, 1980
- Kurita, T.: Spectral analysis of seismic waves, Part 1. Data windows for the analysis of transient waves. *Spec. Contrib. Geophys. Inst. Kyoto Univ.* **9**, 97–122, 1969
- Leblanc, G.S.J.: Truncated crustal transfer functions and fine crustal structure determination. *Bull. Seismol. Soc. Am.* **57**, 719–733, 1967
- Liu, H.-P., Anderson, D.L., Kanamori, H.: Velocity dispersion due to anelasticity; implications for seismology and mantle composition. *Geophys. J. R. Astron. Soc.* **47**, 41–58, 1976
- Lundquist, G.M., Cormier, V.F.: Constraints on the absorption band model of Q . *J. Geophys. Res.* **85**, 5244–5256, 1980
- Madariaga, R.: Dynamics of an expanding circular fault. *Bull. Seismol. Soc. Am.* **66**, 639–666, 1976
- Mitchell, B.J.: Frequency dependence of shear wave internal friction in the continental crust of eastern North America. *J. Geophys. Res.* **85**, 5212–5218, 1980
- 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
- Okada, H., Suzuki, S., Asano, S.: Anomalous underground structure in the Matsushiro earthquake swarm area as derived from a fan shooting technique. *Bull. Earthquake Res. Inst.* **48**, 811–833, 1970
- Roeker, S.W., Tucker, B., King, J., Hatzfeld, D.: Estimates of Q in Central Asia as a function of frequency and depth using the coda of locally recorded earthquakes. *Bull. Seismol. Soc. Am.* **72**, 129–150, 1982
- Ross, J.V., Ave'Lallemant, H.C., Carter, N.L.: Activation volume for creep in the upper mantle. *Science* **203**, 261–263, 1979
- Sacks, I.S.: Q_s of the lower mantle-A body wave determination. Carnegie Institution, Ann. Rep. Dir. Department of Terrestrial Magnetism, Year Book 79, 508–512, 1980
- Shimsoni, M., Ben-Menahem, A.: Computation of the divergence coefficient for seismic phases. *Geophys. J. R. Astron. Soc.* **21**, 285–294, 1970
- Singh, K., Fried, J., Aspel, R., Brune, J.: Spectral attenuation of SH -wave along the Imperial Fault and a preliminary model of Q in the region. *Bull. Seismol. Soc. Am.* **72**, 2003–2016, 1982
- Sipkin, S.A., Jordan, T.H.: Frequency dependence of Q_{ScS} . *Bull. Seismol. Soc. Am.* **69**, 1055–1079, 1979
- Touvenot, F.: Frequency dependence of the quality factor in the upper crust: A deep seismic sounding approach. *Geophys. J. R. Astron. Soc.* **73**, 427–447, 1983
- Tsujiura, M.: Frequency analysis of seismic waves (1). *Bull. Earthquake Res. Inst.* **44**, 873–891, 1966
- Ulug, A.: Frequenzabhängigkeit von Q seismischer Raumwellen im Erdmantel. *Ber. Inst. Meteorol. u. Geophys. Univ. Frankfurt*, **49**, 1983
- Wielandt, E., Streckeisen, G.: The leaf-spring seismometer: design and performance. *Bull. Seismol. Soc. Am.* **72**, 2349–2367, 1982

Received January 24, 1984, Revised June 6, 1984

Accepted June 8, 1984

The polarization of S waves in a heterogeneous isotropic Earth model

V.F. Cormier

Earth Resources Laboratory, Department of Earth, Atmospheric, and Planetary Sciences,
Massachusetts Institute of Technology, Cambridge, MA 02139, USA

Abstract. The polarization vector of an S wave propagating in a heterogeneous isotropic medium remains fixed with respect to a vector basis $(\hat{\mathbf{t}}, \hat{\mathbf{e}}_1, \hat{\mathbf{e}}_2)$, where $\hat{\mathbf{t}}$ is the unit tangent to the ray at a point on the ray and $\hat{\mathbf{t}}, \hat{\mathbf{e}}_1, \hat{\mathbf{e}}_2$ are mutually orthogonal. This vector frame forms the basis of the orthogonal co-ordinate frame, in which the ray is one co-ordinate line. It is not generally equivalent to the Frénet frame $(\hat{\mathbf{t}}, \hat{\mathbf{n}}, \hat{\mathbf{b}})$. At any point along a ray in a heterogeneous medium, $\hat{\mathbf{e}}_1$ and $\hat{\mathbf{e}}_2$ may have rotated about $\hat{\mathbf{t}}$ by an angle Θ with respect to the natural $SH(\hat{\mathbf{e}}_1)$ and $SV(\hat{\mathbf{e}}_2)$ directions at that point. In a geocentric, spherical co-ordinate system the derivative of Θ with respect to travel time τ along the ray in a medium with S velocity $v=v(r, \vartheta, \varphi)$ is

$$\frac{d\Theta}{d\tau} = \frac{\cot i}{r} \left[\sin \zeta \frac{\partial v}{\partial \vartheta} - \frac{\cos \zeta}{\sin \vartheta} \frac{\partial v}{\partial \varphi} \right],$$

where the incidence angle i and azimuthal angle ζ are defined at an instantaneous point on the ray, and where $v(r, \vartheta, \varphi)$ is a sufficiently continuous and slowly varying function for zeroth order ray theory to be applicable. At the source, one may set $\Theta=0$, $\hat{\mathbf{e}}_1=\hat{\mathbf{e}}_1$ and $\hat{\mathbf{e}}_2=\hat{\mathbf{e}}_2$. At any other point along the ray, the directions of $\hat{\mathbf{e}}_1, \hat{\mathbf{e}}_2$ can be determined from the value of Θ and the directions of $\hat{\mathbf{e}}_1, \hat{\mathbf{e}}_2$ at that point. $\hat{\mathbf{e}}_1, \hat{\mathbf{e}}_2$ are uniquely defined using the instantaneous tangent vector $\hat{\mathbf{t}}$ and radius vector $\hat{\mathbf{r}}$ to the point. The equations for ray trajectory and Θ are integrated through the three-dimensional velocity perturbation surrounding a descending slab. The results show that realistic three-dimensional structures will rarely generate more than 10° deviations in the orientation of the S polarization vector compared to that observed in a radially homogeneous Earth. This agrees with the general success in the use of S polarizations for focal mechanism solutions. It also suggests that larger deviations and the progressive complexity of the particle motion ellipse observed as time advances in the S waveform are primarily the consequences of phase interference and/or shear wave splitting due to anisotropy.

Key words: S -wave polarization – Ray tracing – Lateral heterogeneity

Introduction

The next generation of whole Earth models will likely include both lateral heterogeneity as well as general anisotropy. The effects of general anisotropy on the splitting and three-component particle motion of S waves are well known (e.g. Crampin, 1981), but the competing effects of lateral heterogeneity in an isotropic Earth model have generally been ignored.

Classical treatments start with Frénet's formulas for the path derivatives of three mutually orthogonal unit vectors $(\hat{\mathbf{t}}, \hat{\mathbf{n}}, \hat{\mathbf{b}})$, where $\hat{\mathbf{t}}, \hat{\mathbf{n}},$ and $\hat{\mathbf{b}}$ are respectively the tangent, normal and binormal at a point on a seismic ray. Frénet's formulas show that the S particle motion, as referred to the $(\hat{\mathbf{n}}, \hat{\mathbf{b}})$ directions, rotate around the ray at a rate equal to the local torsion T . Since this rate is equal but of opposite sign to the rate at which the $(\hat{\mathbf{n}}, \hat{\mathbf{b}})$ axes themselves rotate about the ray, it is assumed that the S polarization is not substantially changed through smoothly varying structures. A basic difficulty in discussing the rotation of the Frénet frame, however, is that the rotation only has good meaning if it is referred to some specified orientation. An example derivation and brief discussion of these results is given in Aki and Richards (1980), where they note that the assumption of essentially fixed SH/SV ratios for propagation through smoothly varying structures "has not thoroughly been investigated."

Červený and Hron (1980) have pointed out difficulties in deriving a ray series solution for S waves using the Frénet frame. Determination on $\hat{\mathbf{n}}$ and $\hat{\mathbf{b}}$ requires computation of the ray torsion T and curvature K . Determination of the amplitude coefficients (geometrical spreading) of S waves is difficult because the components of the displacement vector couple into both the $\hat{\mathbf{n}}$ and $\hat{\mathbf{b}}$ directions even in the zeroth order approximation of the asymptotic ray series. Pšenčík (1979) and Červený and Hron instead recommend seeking a ray series solution in a new ray centred co-ordinate system $(\hat{\mathbf{t}}, \hat{\mathbf{e}}_1, \hat{\mathbf{e}}_2)$. This system has several important advantages. The displacement vector of S waves remains fixed with respect to $(\hat{\mathbf{e}}_1, \hat{\mathbf{e}}_2)$ at all times. In a three-dimensionally varying medium, the orientation of the S displacement can be determined at any point on the ray by simply integrating a differential equation for the orientation of

the (\hat{e}_1, \hat{e}_2) frame. Although \hat{e}_1 and \hat{e}_2 can be expressed in terms of the Frénet frame, one does not actually need to calculate \hat{n} and \hat{b} or T and K . In the \hat{e}_1, \hat{e}_2 system the transport equations of both principal components of S waves are independent of one another to zero asymptotic order.

In this paper, the ray tracing equation for the polarization vector of S waves is derived using the \hat{e}_1 and \hat{e}_2 frame and a spherical co-ordinate system appropriate for a whole Earth model. An example calculation is given, integrating the S ray tracing equations through the three-dimensional velocity perturbation of a descending slab. This example is chosen because it represents a structure with possibly the largest deviation from radial symmetry in terms of scale length and percent velocity fluctuation likely to be encountered by a teleseismic ray. As such, it can be used to predict the maximum order of deviation of the S polarization vector from its orientation in a radially symmetric Earth.

Ray tracing equation for the polarization of S waves

Following Popov and Pšenčík (1976; 1978) and Červený and Hron (1980), one can introduce unit vectors \hat{e}_1, \hat{e}_2 and \hat{t} . \hat{t} is tangent to the ray. \hat{e}_1 and \hat{e}_2 can be chosen arbitrarily as long as \hat{t}, \hat{e}_1 and \hat{e}_2 are mutually orthogonal and form a right-handed system.

The vectors \hat{e}_1 and \hat{e}_2 are related to the basis vectors \hat{e}_1, \hat{e}_2 of the ray centred co-ordinate system by rotation through an angle Θ :

$$\begin{aligned}\hat{e}_1 &= \cos \Theta \hat{e}_1 - \sin \Theta \hat{e}_2, \\ \hat{e}_2 &= \sin \Theta \hat{e}_1 + \cos \Theta \hat{e}_2.\end{aligned}\quad (1)$$

The S polarization vector remains fixed with respect to the \hat{e}_1, \hat{e}_2 vectors. The angle Θ describes the rotation of the vector basis \hat{e}_1, \hat{e}_2 with respect to \hat{e}_1, \hat{e}_2 about the central ray as the ray propagates through the heterogeneous medium. Thus the S polarization at any point along the ray can be determined by knowing the initial direction of S polarization, the value of Θ at that point and the orientation of \hat{e}_1, \hat{e}_2 at that point.

The variation of Θ along a ray can be expressed by a differential equation in the co-ordinates commonly used in the tracing of rays through complete models of a spherical earth. These are the fixed geocentric co-ordinates (r, ϑ, φ) shown in Fig.1 and the incidence angle i and azimuthal angle ζ defined at a local point on the ray. i is the angle between the ray direction and the \hat{r} direction; ζ is the angle between $\hat{\vartheta}$ and the projection of the ray path into the $\vartheta\varphi$ plane. The vectors $\hat{t}, \hat{e}_1, \hat{e}_2$ can be expressed by components in the (r, ϑ, φ) system as

$$\begin{aligned}\hat{t} &= (\cos i, \sin i \cos \zeta, \sin i \sin \zeta), \\ \hat{e}_1 &= (0, -\sin \zeta, \cos \zeta), \\ \hat{e}_2 &= (\sin i, -\cos i \cos \zeta, -\cos i \sin \zeta).\end{aligned}\quad (2)$$

With this choice, the system $(\hat{t}, \hat{e}_1, \hat{e}_2)$ is mutually orthogonal, right-handed, and \hat{e}_1 is perpendicular to the vector \hat{r} . In a radially symmetric Earth, \hat{e}_1 is in the SH direction and \hat{e}_2 is in the SV direction. In a heterogeneous Earth, \hat{e}_1, \hat{e}_2 instantaneously define SH and SV directions along a ray with respect to the local tangent

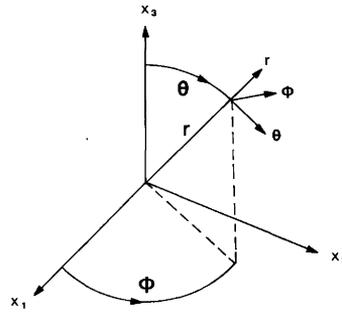


Fig. 1. Definition of the co-ordinate system (r, ϑ, φ) and $\hat{t}, \hat{\vartheta}, \hat{\varphi}$

to the ray and the radius vector \hat{r} . These directions are not, however, the SH and SV directions inferred from the projection of the ray into the plane defined by the source, receiver and centre of the Earth.

An equation for the derivative $\frac{d\Theta}{d\tau}$ with respect to travel time τ along a ray can now be obtained by comparing either of two equations for $\frac{d\hat{e}_1}{d\tau}$ or $\frac{d\hat{e}_2}{d\tau}$. The first such equation is obtained by differentiating Eq. (1) with respect to τ :

$$\frac{d\hat{e}_1}{d\tau} = \frac{d\hat{e}_1}{d\tau} \cos \Theta - \frac{d\hat{e}_2}{d\tau} \sin \Theta - (\hat{e}_1 \sin \Theta + \hat{e}_2 \cos \Theta) \frac{d\Theta}{d\tau}. \quad (3)$$

The second equation is a property of the ray centred co-ordinate system:

$$\frac{d\hat{e}_1}{d\tau} = (Vv \cdot \hat{e}_1) \hat{t}. \quad (4)$$

By equating the right hand side of Eq. (3) to that of Eq. (4), using the rules of differentiation given in Popov and Pšenčík (1976; 1978), the ray tracing equations for $\frac{di}{d\tau}$ and $\frac{d\zeta}{d\tau}$ and relations between spherical and Cartesian geocentric co-ordinates, the following result is obtained:

$$\frac{d\Theta}{d\tau} = \frac{\cot i}{r} \left[\sin \zeta \frac{\partial v}{\partial \vartheta} - \frac{\cos \zeta}{\sin \vartheta} \frac{\partial v}{\partial \varphi} \right]. \quad (5)$$

Equation (5) can now be integrated along with the other five equations (e.g. Julian and Gubbins, 1977; Aki and Richards, 1980) needed to determine the ray trajectory.

Example: the effects of a descending slab

As a test of the effects of laterally heterogeneous structure on the polarization of S waves, Eq. (5) was integrated for sources located in a three-dimensional velocity structure surrounding a descending slab. A 7% positive perturbation was added to the the S velocities of the 1-Hz isotropic PREM model (Dziewonski and Anderson, 1981). The perturbation was taken to be a Gaussian shaped zone centred on a dipping plane. An additional exponential factor controlled its decay with

depth. This simple perturbation has been successfully used to predict the order of the deviations of ray paths caused by slab structures (Toksöz et al., 1971). The parameters of the model can be easily modified to closely reproduce the velocity structure predicted by complete thermal models. Figure 2 shows the results of the calculation for the rotation angle Θ for a 540-km-deep earthquake occurring in a slab. The slab strike and dip was chosen to approximately model the Benioff zone dipping beneath the Kuril-Kamchatka arc. The initial take-off angle was held constant and the resulting epicentral distance varied between 85° to 110° for *SKS* waves. The results in Fig. 2 are for a 40-km-thick slab, which penetrates to the centre of the Earth. This experiment was designed to see how much of the *S* wave that leaves the source as *SH* can be rotated into *SV* by propagation through the three-dimensional structure surrounding the slab and then complete its path to the receiver as an *SKS* wave. The motivation for performing this experiment was to explain observations of *SKS* waves on both *SH* and *SV* components at azimuths for which focal mechanism solutions predict only small amounts of *SV* leaving the source (Cornier, 1984).

From Fig. 2 it is seen that this model rotates the initial orientation of the *S* particle motion vector by as much as ± 0.2 rad ($\pm 12^\circ$) for azimuths within 40° of the trend of the strike of the arc on the down-dip side. The rotation vanishes at azimuths perpendicular to the strike of the slab because the partial derivatives $\frac{\partial v}{\partial \vartheta}$ and $\frac{\partial v}{\partial \varphi}$ vanish in this ray direction. A shorter slab, which terminates at 1000 km depth, rotates the polarization angle by only about one-half as much, i.e. ± 0.1 rad ($\pm 6^\circ$).

The complete effects of a slab structure on the apparent *S* polarization must include the deviations it introduces in the ray path as well as the rotation of the *S* polarization vector. The initial *S* radiation that reaches the receiver leaves the source at a take-off angle and azimuth that differs from the angles calculated assuming a radially symmetric Earth. The polarization observed at a teleseismic station will be a sum of these two effects (Fig. 3).

Stated mathematically, consider the *S* wave radiated by a point double couple at initial angles ζ_0 , i_0 that reaches a particular receiver in a radially symmetric Earth. This receiver observes an *S* polarization vector described by an angle Φ in a plane perpendicular to the ray, where

$$\Phi = \tan^{-1} \frac{F^{SV}(\zeta_0, i_0)}{F^{SH}(\zeta_0, i_0)}, \quad (6a)$$

with F^{SV} and F^{SH} taken to be the far-field *SV* and *SH* radiation patterns functions of the double couple [e.g. p115 of Aki and Richards (1980)]. The actual *S* polarization observed by a receiver in a radially heterogeneous Earth will be radiated from the source at different initial angles ζ'_0 , i'_0 , and this initial *S* polarization will be rotated by an angle Θ during propagation to the receiver. The *S* polarization observed at the receiver is then given by

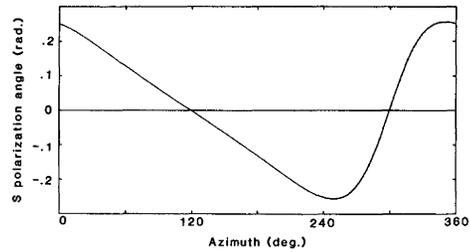


Fig. 2. The orientation of the *S* polarization and the vector basis \hat{e}_1 , \hat{e}_2 rotates about the ray as the ray traverses the three-dimensional velocity perturbation of a descending slab. The slab is assumed to have a +7% velocity contrast relative to a reference radially symmetric Earth (PREM). In the vicinity of the slab the velocity is assumed to be given by the

form $v_0(r) = v_0(r) \left\{ 1 + 0.07 \exp \left[- \left(\frac{\xi}{h'} \right)^2 - \frac{z}{d} \right] \right\}$, where ξ is the

distance from the axial plane of the slab. The example shown here is for a slab striking at azimuth 40° and dipping 68° to the NW. The initial take-off angle of a 540-km-deep source was held constant and is appropriate for *SKS* waves recorded at a distance of 75° in PREM. $h' = 40$ km and $d = \infty$. For $d = 1000$ km, the maximum deviation of the *S* polarization angle is about one-half as large

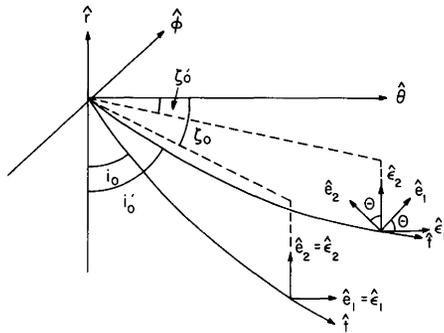


Fig. 3. The *S* polarization observed at a teleseismic receiver is affected in two ways by heterogeneous structure in the vicinity of the source. (1) Since the initial orientation of the *S* polarization vector depends on where the ray leaves the focal sphere, it must be calculated using angles i'_0 , ζ'_0 of the ray that has been perturbed by the heterogeneous structure. (2) The initial *S* polarization vector remains fixed with respect to the vector basis \hat{e}_1 , \hat{e}_2 , but propagation of the ray through the heterogeneous region rotates this vector basis by an angle Θ with respect to the instantaneous *SH* and *SV* directions \hat{e}_1 and \hat{e}_2

$$\Phi = \tan^{-1} \frac{-F^{SH}(\zeta'_0, i'_0) \sin \Theta + F^{SV}(\zeta'_0, i'_0) \cos \Theta}{F^{SH}(\zeta'_0, i'_0) \cos \Theta + F^{SV}(\zeta'_0, i'_0) \sin \Theta}. \quad (6b)$$

Test calculations with slab models show that the effect of path deviation generally does not cancel the effect of rotation of the particle motion vector. Integrated deviations of all angles through the heterogeneous source region are about the same order of magnitude. i_0 , ζ_0 and Θ all exhibit a maximum order of deviation of about 10° in the longest slab models. These deviations are too small to grossly bias measurements of the *S* polarization vector. Since the effects of lateral velocity heterogeneity on *S* polarization require either or both strong lateral velocity gradients or lateral velocity gradients maintained over a long portion

of the ray path, the slab tests can be used as an estimate of the maximum deviations in S polarization that may be practically observed on the Earth. Thus the small deviations found in the slab tests agree with the general success (Stauder, 1962) of the use of S polarizations in the determination of focal mechanism solutions.

Discussion and conclusions

Although the deviation of ζ_0 can be large (10° or more) in a heterogeneous source region, its effect on the apparent azimuth at the receiver is very small (less than several 0.1° s). This effect can be demonstrated by ray tracing and is shown schematically in Fig. 4. It also agrees with the behaviour of azimuth and slowness anomalies observed at large aperture arrays such as LASA and NORSAR (Berteussen, 1976). Thus if heterogeneity primarily exists in the source region and further back along the ray path, S waves can be accurately rotated into SH and SV components using the true great circle azimuth. The observed S polarization vector, however, may deviate as much as 10° from that predicted using the true focal mechanism because of the rotation of the S polarization vector.

Heterogeneity in the receiver region can affect both the apparent azimuth and the orientation of the S polarization vector. In this case, the SH and SV motion can be defined with respect to the local tangent to the ray. S motion can be resolved using the apparent rather than the great circle azimuth. The apparent azimuth may be determined from the three-dimensional orientation of the particle motion ellipse or from fitting a plane wave to the arrival times at a local array.

An important result of numerical experiments with slab structures is that although coherent, long scale length heterogeneities can cause deviations in the orientation of the S polarization vector of up to 10° , it is difficult to generate deviations much larger than this using any reasonable three-dimensional structure. It may generally be expected that deviations will be smaller than this estimated maximum because the estimate is based on zeroth order ray theory, which is strictly only valid at infinite frequency. At finite frequency, a body wave will tend to average the effects of heterogeneities over a wavelength. The small effects of lateral heterogeneity agree with the general success in the use of S wave polarizations for focal mechanism solutions. It does suggest that larger deviations in S polarizations are caused by phase interference and/or shear wave splitting due to anisotropy. These effects may be primarily responsible for the elliptical rather than linear particle motion observed in S waves and the progressive complexity in the particle motion ellipse commonly observed as time advances in the S waveform.

Acknowledgements. This research was supported by the National Science Foundation under grant EAR-82-18737.

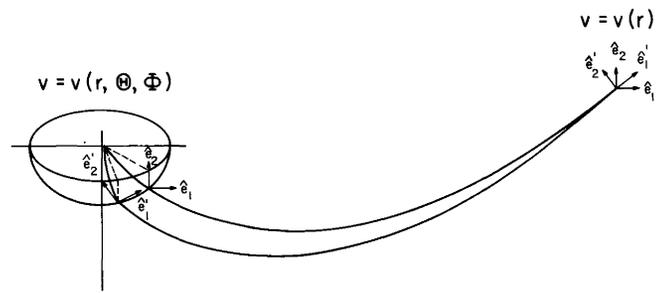


Fig. 4. Heterogeneous structure in the vicinity of the source primarily affects S polarization observed at a teleseismic receiver by rotating the vector basis \hat{e}_1, \hat{e}_2 . It does not have much effect on either the apparent azimuth or angle with respect to the vertical observed at the teleseismic station. Reciprocally, heterogeneous structure in the vicinity of the teleseismic receiver will affect the S polarization by rotating the vector basis \hat{e}_1, \hat{e}_2 and by generating deviations in both the apparent azimuth and the angle with respect to the vertical

References

- Aki, K., Richards, P.G.: Quantitative Seismology, pp 100-105, pp 724-725, W.H. Freeman and Co., San Francisco, 1980
- Berteussen, K.A.: The origin of slowness and azimuth anomalies at large arrays. *Bull. Seismol. Soc. Am.* **66**, 719-742, 1976
- Červený, V., Hron, F.: The ray series method and dynamic ray tracing system for three-dimensional inhomogeneous media. *Bull. Seism. Soc. Am.* **70**, 47-77, 1980
- Crampin, S.: A review of wave motion in anisotropic and cracked elastic-media. *Wave Motion* **3**, 343-391, 1981
- Cormier, V.F.: Some problems with S , SKS , and ScS observations and implications for the structure of the base of the mantle and the outer core. Submitted, *J. Geophys.*
- Dziewonski, A.M., Anderson, D.L.: Preliminary reference earth model. *Phys. Earth Planet. Int.* **25**, 297-356, 1981
- Julian, B.R., Gubbins, D.: Three-dimensional seismic ray tracing. *J. Geophys.* **43**, 95-113, 1977
- Popov, M.M., Pšenčík, I.: Ray amplitudes in inhomogeneous media with curved interfaces. *Geofysikalni Sbornik* **24**, 111-129, 1976
- Popov, M.M., Pšenčík, I.: Computation of ray amplitudes in inhomogeneous media with curved interfaces. *Studia Geophys. Geodaet.* **22**, 248-258, 1978
- Pšenčík, I.: Ray amplitudes of compressional, shear and converted body waves in three dimensional laterally inhomogeneous media with curved interfaces. *J. Geophys.* **45**, 381-390, 1979
- Stauder, W.: S -wave studies of earthquakes of the North Pacific, Pt. I: Kamchatka, *Bull. Seismol. Soc. Am.* **52**, 527-550, 1962
- Toksöz, M.N., Minear, J.W., Julian, B.R.: Temperature field and geophysical effects of a downgoing slab. *J. Geophys. Res.* **76**, 1113-1138, 1971

Received May 9, 1984; Revised version September 10, 1984
Accepted September 10, 1984

The relaxation of spherical and flat Maxwell Earth models and effects due to the presence of the lithosphere

Detlef Wolf

Department of Physics, University of Toronto, Toronto, Ontario, Canada, M5S 1A7

Abstract. The radial surface deflection of a spherical Earth model appropriate to time scales characteristic of glacial loading is compared with the associated half-space response. For that purpose the analytical solution for a Maxwell sphere surrounded by a thick elastic shell is derived under the assumption of incompressibility. The half-space approximation is deduced as a special case directly from the spherical solution. Comparison of the response spectra, for different thicknesses of the elastic shell, reveals only minor differences. In the spatial domain, the half-space approximation may, nevertheless, diverge significantly from the spherical solution. For a disk load radius $R=800$ km (Fennoscandia), the half-space approximation is adequate, whereas it is usually inappropriate if $R=1600$ km (Laurentia). The sensitivity of the response to the thickness of the elastic shell is also investigated. For $R=800$ km, the surface deflection in the central region below the load is fairly diagnostic of the shell thickness. If $R=1600$ km, the peripheral region is more sensitive to this parameter.

Key Words: Isostasy – Lithosphere – Maxwell continuum

Introduction

The isostatic adjustment of the Earth due to viscous flow in its mantle is a problem on which work has continued for nearly 50 years. The main results have been reviewed by Cathles (1975), Peltier (1982) and others. The effect of the lithosphere on the response characteristics has, however, been fully appreciated only during the past few years. Earlier work was inconclusive in this respect. This is exemplified by two interpretations of strandline data from the region of Pleistocene Lake Bonneville in the United States. Whereas Crittenden (1963) neglected the lithosphere and inferred a mantle viscosity of 10^{21} P from the data, Walcott (1970a) treated the mantle as inviscid and derived a value of 5×10^{22} Nm for the flexural rigidity of the lithosphere. Later, Nakiboglu and Lambeck (1982) re-interpreted the data, using an improved viscoelastic model incorporating both a lithosphere and a dynamic mantle, and suggested revised values for lithospheric thickness and mantle viscosity.

The modification of the response pattern of Maxwell Earth models by the presence of the lithosphere has been addressed by Peltier (1980) and by Wu and Peltier (1982). The analysis has recently been extended such that lithospheric thickness can be directly inferred from appropriate isostatic adjustment data (Peltier, 1984).

In the present investigation, the effect of the lithosphere on the response of Maxwell Earth models is analysed in greater detail. For this purpose straight-edged circular loads of radii $R=800$ km (Fennoscandia) or $R=1600$ km (Laurentia) are applied at times $t \geq 0$, and the relaxation of the surface is investigated. For loads comparable in scale to the larger disk, the sensitivity of the deflection in the peripheral region of the load to lithospheric thickness has already been demonstrated by Peltier (1984) in his analysis of relative-sea-level variations induced by the melting of the Laurentide ice sheet in Canada.

The smaller disk load may serve as a crude representation of the Fennoscandian ice sheet. On this load scale the significance of the lithosphere has been less clear. Cathles (1975, pp. 144–155) discussed the controversy and concluded that the effect of the lithosphere could be neglected when modelling the Fennoscandian uplift. His reasoning was, however, primarily based on the analysis of the wave-number spectrum of the load. Here we will study the effects of the lithosphere in the spatial domain. As will be shown, the shape of the deformed surface in the region formerly covered by the Fennoscandian ice load is quite sensitive to the presence of the lithosphere and may therefore be used to infer its thickness.

A second aspect of our analysis concerns the adequacy of flat-earth approximations when modelling the response due to loads of large diameter. Whereas half-space models are conventional for small loads such as lakes or volcanic islands, some uncertainty exists about the maximum load diameter tolerable in this approximation. The deglaciation-induced uplift of Fennoscandia, for example, has been modelled using both spherical and half-space theory (see Cathles, 1975, pp. 6, 173–180, for a review). The Laurentide ice load, on the other hand, has usually been regarded as too large for the application of flat-earth approximations (see Peltier, 1982).

All models to be discussed in the present study

apply to externally gravitating, incompressible and hydrostatically pre-stressed continua of uniform density. Thus, the gravitational effects disregarded in our models reduce to perturbations of the geopotential due to (a) the externally applied force field, viz. the "load", and (b) due to internal mass redistributions induced by this load.

Cathles (1975, pp. 72–83) has presented solutions for the surface deformation of elastic Earth models of increasing complexity and discussed the contributions of different gravitational effects. This also includes an assessment of the influence of the inviscid core on the observed response. From his study it can be inferred that the significance of the geometrical and gravitational effects of spherical Earth models is mainly confined to angular orders $n < 10$. As Cathles further shows, contributions (a) and (b) to the perturbation of the geopotential largely cancel each other, such that the net effect of self-gravitation is small. The influence of the Earth's core becomes perceptible at angular orders $n < 5$. The error committed when neglecting both self-gravitation and the core in the spherical solution amounts to 5–10 per cent of the surface deflection. For $n < 5$ this is smaller, by a factor of more than two, than the purely geometrical effect of sphericity.

The significance of self-gravitation for the relaxation of spherical Maxwell Earth models is not very well documented. As a crude guideline, strains in the elastic problem translate into strain rates in the associated viscoelastic problem. Therefore, changes in the deformation of an elastic sphere due to self-gravitation and the inviscid core are paralleled by corresponding changes in the relaxation times of a Maxwell sphere. As relaxation proceeds, the Maxwell continuum degenerates into an inviscid fluid, and the core loses its special role. Quantitative studies are, however, lacking so far. Recently, Dragoni et al. (1983) have derived analytical solutions for the deformation of a two-layer incompressible Maxwell sphere without a core in an attempt to examine whether earthquake-triggered movements of material can excite the Chandler wobble. The inclusion of self-gravitation into their analysis is reflected by the great complexity of their equations. If desired, appropriate modifications of their solution would permit a quantitative assessment of the effects of self-gravitation on the relaxation of spherical Maxwell Earth models.

In view of the limited objectives of the present study, the neglect of both self-gravitation and the inviscid core appears to be justified. In the following we will therefore derive the analytical solution describing the deformation of a two-layered, incompressible and pre-stressed Maxwell Earth model subject to an external gravity field and deduce the half-space limit directly from this solution. This approach will then allow us to obtain better constraints on the range of load scales amenable to flat-earth approximations.

Theory

Our derivation of the equations appropriate to the quasi-static deformation of a Maxwell sphere will be based on the correspondence principle. For geophysical applications, this method has been developed in papers

that began with Peltier (1974). Major results from this work are summarized in Peltier (1982). The main advantage of the correspondence principle is that it allows us to derive the viscoelastic solution from the solution for an associated elastic problem. For the present purpose it is sufficient to restrict our analysis to incompressible Earth models. Together with the assumption of external gravitation this significantly simplifies the general form of the momentum balance appropriate to the quasi-static deformation of a viscoelastic continuum (Peltier, 1974). Defining a total perturbation stress by

$$\sigma_{ij} = \sigma_{ij}^{(e)} - \rho g u_r \frac{r}{a_1} \delta_{ij}, \quad (1)$$

the momentum balance in Cartesian co-ordinates x_i is

$$\frac{\partial}{\partial x_j} \sigma_{ij} = 0. \quad (2)$$

In Eq. (1) $\sigma_{ij}^{(e)}$ denotes the elastic portion of the perturbation stress, and $\rho g u_r (r/a_1) \delta_{ij}$ accounts for stress advection in a hydrostatically pre-stressed sphere of radius a_1 and density ρ subject to gravitational acceleration g , with u_r the radial displacement component at the radial distance r (Love, 1911, pp. 89–93). The pre-stress term was first included in the viscoelastic formulation by Peltier (1974). Its importance was explicitly discussed by Wu and Peltier (1982), who noted that its consideration in the momentum balance is required in order that, for a Heaviside loading history, the correct solution be obtained in the inviscid limit ($t \rightarrow \infty$) of the viscoelastic continuum. Recently, the significance of pre-stress has been further analysed by Wolf (1984a).

Since incompressibility has been assumed, Hooke's law takes the form

$$\sigma_{ij}^{(e)} = -p^{(e)} \delta_{ij} + \mu \left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right), \quad (3)$$

with μ Lamé's second constant and u_i the displacement. $p^{(e)}$ denotes the elastic portion of the perturbation pressure p and

$$p^{(e)} = \lim_{\substack{\lambda \rightarrow \infty \\ \Delta \rightarrow 0}} (\lambda \Delta), \quad (4)$$

where λ is Lamé's first constant, and $\Delta = \partial u_i / \partial x_i$ denotes the dilatation (Love, 1927, pp. 255–257). Substituting for σ_{ij} in Eq. (2) yields

$$-\frac{\partial p}{\partial x_j} \delta_{ij} + \mu \frac{\partial^2 u_i}{\partial x_j^2} = 0, \quad (5a)$$

or, in vector notation,

$$-\nabla p - \mu \nabla \times \nabla \times \mathbf{u} = 0. \quad (5b)$$

Confining our analysis to axially symmetric problems, we obtain for the radial and zonal components of Eq. (5b), respectively

$$\frac{\partial}{\partial \theta} (2r \omega_\phi \sin \theta) + r^2 \mu^{-1} \sin \theta \frac{\partial p}{\partial r} = 0, \quad (6)$$

$$\frac{\partial}{\partial r}(2r\omega_\phi) - \mu^{-1} \frac{\partial p}{\partial \theta} = 0, \quad (7)$$

where $2\omega_\phi = (\nabla \times \mathbf{u})_\phi$ is implied. Here r , θ and ϕ denote the radial, zonal and azimuthal components of the coordinate vector, respectively, with $\theta=0$ on the axis of symmetry. For spherical co-ordinates, the equation of continuity $\nabla \cdot \mathbf{u} = 0$ becomes

$$\sin \theta \frac{\partial}{\partial r}(r^2 u_r) + r \frac{\partial}{\partial \theta}(u_\theta \sin \theta) = 0. \quad (8)$$

To solve the system of Eqs. (6)–(8) for u_r , u_θ and p , we seek solutions of the type

$$u_r = R_1(r) P_n(\cos \theta), \quad (9)$$

$$u_\theta = -R_2(r) \sin \theta P'_n(\cos \theta), \quad (10)$$

$$p = R_3(r) P_n(\cos \theta). \quad (11)$$

Here P_n is the Legendre polynomial of degree n , where $n=0, 1, \dots$, and a prime denotes differentiation with respect to the argument. Substitution in Eqs. (6)–(8) yields

$$n(n+1)(R_1 - rR'_1 - R_2) + \mu^{-1} r^2 R'_3 = 0, \quad (12)$$

$$R'_1 - 2R'_2 - rR''_2 + \mu^{-1} R_3 = 0, \quad (13)$$

$$rR'_1 + 2R_1 - n(n+1)R_2 = 0. \quad (14)$$

Equations (12)–(14) constitute a system of coupled ordinary differential equations for R_1, R_2 and R_3 . Eliminating R_2 and R_3 in Eqs. (12) and (13) yields

$$r^4 R_1^{(4)} + 8r^3 R_1^{(3)} + 2[6 - n(n+1)]r^2 R_1'' - 4n(n+1)rR_1' - n(n+1)[2 - n(n+1)]R_1 = 0. \quad (15)$$

The solution of Eq. (15) is

$$R_1 = A_n r^{n+1} + B_n r^{-n} + C_n r^{n-1} + D_n r^{-(n+2)}. \quad (16)$$

Substituting for R_1 and R'_1 in Eq. (14) we obtain

$$R_2 = \frac{1}{n(n+1)} [(n+3)A_n r^{n+1} - (n-2)B_n r^{-n} + (n+1)C_n r^{n-1} - nD_n r^{-(n+2)}]. \quad (17)$$

Solving Eqs. (12) and (13) for R_3 yields

$$r^2 R_3'' + 2rR_3' - n(n+1)R_3 = 0, \quad (18)$$

with the solution

$$R_3 = E_n r^n + F_n r^{-(n+1)}. \quad (19)$$

The integration constants E_n and F_n are not independent. Substituting in Eq. (12) and equating coefficients we obtain

$$nE_n = 2\mu(2n+3)A_n, \quad (20)$$

$$(n+1)F_n = 2\mu(2n-1)B_n. \quad (21)$$

For the elastic shell, where $\mu = \mu_1$, the solutions for displacement components and pressure are therefore of the type

$$u_r(r, \theta) = [A_n r^{n+1} + B_n r^{-n} + C_n r^{n-1} + D_n r^{-(n+2)}] P_n(\cos \theta), \quad (22)$$

$$u_\theta(r, \theta) = -\frac{1}{n(n+1)} [(n+3)A_n r^{n+1} - (n-2)B_n r^{-n} + (n+1)C_n r^{n-1} - nD_n r^{-(n+2)}] \sin \theta P'_n(\cos \theta), \quad (23)$$

$$p(r, \theta) = 2\mu_1 \left[\frac{2n+3}{n} A_n r^n + \frac{2n-1}{n+1} B_n r^{-(n+1)} \right] P_n(\cos \theta). \quad (24)$$

If we observe that

$$\sigma_{rr} = -p + 2\mu_1 \frac{\partial u_r}{\partial r}, \quad (25)$$

$$r\sigma_{r\theta} = \mu_1 \left(\frac{\partial u_r}{\partial \theta} - u_\theta + r \frac{\partial u_\theta}{\partial r} \right), \quad (26)$$

the solutions for the stress components can be shown to be of the type

$$\sigma_{rr}(r, \theta) = 2\mu_1 \left[\frac{n^2 - n - 3}{n} A_n r^n - \frac{n^2 + 3n - 1}{n+1} B_n r^{-(n+1)} + (n-1)C_n r^{n-2} - (n+2)D_n r^{-(n+3)} \right] P_n(\cos \theta), \quad (27)$$

$$\sigma_{r\theta}(r, \theta) = -\frac{2\mu_1}{n(n+1)} [n(n+2)A_n r^n + (n^2 - 1)B_n r^{-(n+1)} + (n^2 - 1)C_n r^{n-2} + n(n+2)D_n r^{-(n+3)}] \sin \theta P'_n(\cos \theta). \quad (28)$$

The integration constants A_n, B_n, C_n and D_n can be determined from the boundary conditions. If the elastic sphere is subject to a load pressure P_n , we have for the elastic perturbation stress at the surface $r=a_1$

$$\sigma_{rr}^{(e)}(a_1, \theta) = -P_n(\cos \theta), \quad (29)$$

$$\sigma_{r\theta}^{(e)}(a_1, \theta) = 0. \quad (30)$$

From the general solution for the deformation of a uniform elastic sphere of rigidity μ_2 , we obtain at the interface $r=a_2$ between mantle and shell (see Appendix A),

$$2\mu_2(n-1)(2n^2 + 4n + 3)u_r(a_2, \theta) = n a_2 \left[(2n+1)\sigma_{rr}(a_2, \theta) - 3(n+1)\frac{P_n}{\sin \theta P'_n} \sigma_{r\theta}(a_2, \theta) \right], \quad (31)$$

$$2\mu_2(n-1)(2n^2 + 4n + 3)u_\theta(a_2, \theta) = -a_2 \left[3\frac{\sin \theta P'_n}{P_n} \sigma_{rr}(a_2, \theta) - (2n^2 + n + 3)\sigma_{r\theta}(a_2, \theta) \right]. \quad (32)$$

Substituting in Eqs. (29)–(32) for $u_r, u_\theta, \sigma_{rr}$ and $\sigma_{r\theta}$ from Eqs. (22), (23), (27) and (28) at $r=a_2$, a system of four equations is obtained, which allows the determination of the four integration constants for the specified boundary conditions. After considerable algebraic manipulation we obtain

$$A_n = -\frac{n(n+1)(n^2-1)}{2\mu_1 D} [k_1 - k_2(a_2/a_1)^{-2} + k_3(a_2/a_1)^{-(2n+1)}], \quad (33)$$

$$B_n = \frac{n^2(n+1)(n+2)}{2\mu_1 D} [k_4 - k_2(a_2/a_1)^{-2} + k_5(a_2/a_1)^{2n+1}], \quad (34)$$

$$C_n = -\frac{n^2(n+1)(n+2)}{2\mu_1 D} [k_4 - k_6(a_2/a_1)^2 - k_3(a_2/a_1)^{-(2n+1)}], \quad (35)$$

$$D_n = \frac{n(n+1)(n^2-1)}{2\mu_1 D} [k_1 - k_6(a_2/a_1)^2 - k_5(a_2/a_1)^{2n+1}]. \quad (36)$$

Here the integration constants have been transformed according to $a_1^n A_n \rightarrow A_n$, $a_1^{-(n+1)} B_n \rightarrow B_n$, $a_1^{n-2} C_n \rightarrow C_n$, $a_1^{-(n+3)} D_n \rightarrow D_n$. k_1, \dots, k_6 and D are complicated expressions in terms of μ_1, μ_2, ρ and n (see Appendix B).

In the following we will be concerned with the radial surface displacement $u_r(a_1, \theta)$. Substituting for the integration constants in Eq. (22) and putting $r = a_1$ yields

$$u_r = T_n^{(e)} P_n, \quad (37)$$

where

$$T_n^{(e)} = \frac{n(n+1)(2n+1)a_1}{2\mu_1 D} [k_6(a_2/a_1)^2 - k_2(a_2/a_1)^{-2} + k_5(a_2/a_1)^{2n+1} + k_3(a_2/a_1)^{-(2n+1)}]. \quad (38)$$

Several special solutions may be derived from Eq. (37); for example, the solution for the elastic shell surrounding an inviscid mantle, which is obtained by taking $\mu_2 = 0$. A similar solution was previously discussed by Slichter and Caputo (1960). Solutions for the uniform elastic sphere result from $\mu_1 = \mu_2, a_2 = 0$ or $a_1 = a_2$.

In the present context it is of particular interest to compare Eq. (37) with the associated flat-earth solution for the same physical model. The two-layer elastic half-space can be treated as a separate boundary-value problem. More insight is gained, however, if Eq. (37) is reduced to the flat-earth solution. This approach will be employed in what follows.

If $h = a_1 - a_2$, we simply require that h/a_1 is a small quantity. Also n must be large enough that $k = n/a_1$ remains finite. If we observe that $a_2/a_1 = 1 - h/a_1$, put $n\theta = nr'/a_1 = kr'$ and substitute the asymptotic approximations for the constants in $T_n^{(e)}$, we obtain, after some algebra, the flat-earth approximation of Eq. (37) as

$$u_r(a_1) = -w(0) = -T^{(e)}(k) J_0(kr'), \quad (39)$$

where

$$T^{(e)}(k) = \left\{ \rho g - 2\mu_1 k \frac{(M^2 - 1)(kh)^2 + [M \cosh(kh) + \sinh(kh)]^2}{(M^2 - 1)kh - [M \sinh(kh) + \cosh(kh)] [M \cosh(kh) + \sinh(kh)]} \right\}^{-1}, \quad (40)$$

with $M = \mu_2/\mu_1$. Here we have used an asymptotic approximation for Legendre polynomials in terms of Bessel functions, viz. (e.g. Watson, 1944, p. 155)

$$\lim_{\substack{n \rightarrow \infty \\ \theta \rightarrow 0}} P_n(\cos \theta) = J_0(n\theta), \quad (41)$$

where J_0 is the Bessel function of the first kind and zeroth order. r' denotes the radial co-ordinate in the cylindrical system $z = a_1 - r$, $r' = \theta a_1$, and w is the vertical displacement in this system. For $M = 0$, the lower half-space becomes inviscid. If, on the other hand, $\mu_1 = \mu_2$, Eq. (39) is reduced to

$$w(0) = \frac{1}{\rho g + 2\mu_1 k} J_0(kr'). \quad (42)$$

This is the solution for the uniform (and incompressible) elastic half-space. If pre-stress is neglected, Eq. (42) reduces to $w(0) = J_0/(2\mu_1 k)$ (e.g. Sneddon, 1951, pp. 468-486; Jeffreys, 1976, pp. 265-267).

In the following, we will be concerned with circular load distributions q in the form of straight-edged

(spherical) caps or (plane) disks. Thus, for the sphere,

$$q(\theta) = \begin{cases} 0, & 0 \leq \theta < \beta \\ 1/2, & \theta = \beta \\ 1, & \beta < \theta \leq \pi \end{cases}, \quad (43)$$

and for the half-space

$$q(r') = \begin{cases} 0, & 0 \leq r' < R \\ 1/2, & r' = R \\ 1, & R < r' < \infty \end{cases}. \quad (44)$$

Here $\beta = R/a_1$ is the angular radius of the cap and R the radius of the disk. $q(\theta)$ can be expanded into a Legendre series

$$q(\theta) = \sum_{n=0}^{\infty} q_n P_n(\cos \theta), \quad (45)$$

where (e.g. Lebedev, 1972, pp. 68-69)

$$q_n = \begin{cases} (1 - \cos \beta)/2, & n=0 \\ [P_{n-1}(\cos \beta) - P_{n+1}(\cos \beta)]/2, & n=1, 2, \dots \end{cases}. \quad (46)$$

Similarly $q(r')$ can be written as the inverse zeroth-order Hankel transform, i.e.

$$q(r') = \int_0^{\infty} \hat{q}(k) k J_0(kr') dk, \quad (47)$$

where (e.g. Sneddon, 1951, p. 528)

$$\hat{q}(k) = \frac{R}{k} J_1(kR). \quad (48)$$

The generalization of our solutions for Maxwell continua is straightforward. Here we will only remark that the correspondence principle allows us to identify the solution of any elastic problem with the Laplace transform of the solution of an associated Maxwell problem, if we substitute

$$\lambda(s) = (\lambda s + \tau^{-1} K)/(s + \tau^{-1}), \quad (49)$$

$$\mu(s) = \mu s/(s + \tau^{-1}), \quad (50)$$

for λ and μ , respectively. In Eqs. (49) and (50) $\tau = \eta/\mu$, where η is the dynamic viscosity, $K = \lambda + 2/3\mu$ denotes the bulk modulus, and s is the Laplace transform variable of the time t .

Since the pronounced temperature dependence of viscosity renders most of the Earth's lithosphere elastic on a time scale of the order of 10 ka, $\eta_1 \rightarrow \infty$ and therefore $\mu_1(s) = \mu_1$, and the transformation only applies to μ_2 .

Before the inverse Laplace transform can be implemented, some algebraic operations must be applied to the elastic solution. These are outlined in Appendix C.

Numerical Results

In order to illustrate the effects of sphericity, we begin with a comparison of the response spectra between the

spherical solution and those obtained in the half-space approximation. In the following, superscript e will refer to the amplitude of the elastic or instantaneous portion of the response, whereas superscript v will denote the viscous or time-dependent part of the response. We have, from Appendix C, for the viscoelastic transfer function $T^{(ve)}(t)$ of the continuum subject to a Heaviside loading history $H(t)$

$$T^{(ve)}(t) = T^{(e)} - T^{(v,1)}[\exp(-s^{(1)}t) - 1] - T^{(v,2)}[\exp(-s^{(2)}t) - 1], \quad (51)$$

which, for $t \rightarrow \infty$, becomes

$$T^{(i)} = T^{(e)} + T^{(v,1)} + T^{(v,2)}, \quad (52)$$

where superscript i denotes the inviscid limit at $t \rightarrow \infty$. The response is therefore completely characterized by the quantities $T^{(e)}$, $T^{(v,1)}$, $T^{(v,2)}$, $s^{(1)}$, $s^{(2)}$. As mentioned in Appendix C, Eq. 51 applies both for the sphere and for the half-space.

For our numerical analysis we choose three different Earth models. Model A is uniform, i.e. without elastic shell. Since the Earth is known to have a lithosphere, the uniform model may appear inappropriate. Its main purpose in the present discussion is to illuminate the modification of the response by the presence of this structural feature. In Model B the lithosphere is 100 km thick. This is a standard value for continental regions based on much previous geophysical research. A similar value has, for example, been inferred by McConnell (1968) in his analysis of glacio-isostatic rebound in Scandinavia. Model C is characterized by an enhanced lithospheric thickness of 200 km. This reflects recent estimates proposed by Peltier (1984) from his interpretation of deglaciation-induced relative-sea-level rise and polar-wander information.

In all three models the density is taken to be $\rho = 3320 \text{ kg m}^{-3}$. This is considerably below the average value appropriate to the whole Earth but fairly characteristic of the material at a depth of 33 km (Bullen, 1963, pp. 232–235). Combined with $g = 9.81 \text{ m s}^{-2}$, this reduced value insures the correct surface deflection in the inviscid limit, which, for low angular orders, is essentially independent of the elastic parameters of the lithosphere and largely determined by the near-surface density contrast. For the shear modulus of the viscoelastic mantle the value $\mu_2 = 1.45 \times 10^{11} \text{ Nm}^{-2}$ is adopted. This is the mean value for the whole Earth (e.g. Wu and Peltier, 1982, p. 442). For the lithosphere we choose $\mu_1 = 0.67 \times 10^{11} \text{ Nm}^{-2}$, which is typical of a depth of 100 km (Bullen, 1963, pp. 232–235). In all models the dynamic viscosity is $\eta_2 = 10^{22} \text{ P}$, which appears to be fairly characteristic of both the upper and the lower mantle (Cathles, 1975; Peltier, 1982).

In discussing Figs. 1–3, we will postpone the analysis of the accuracy of the half-space approximation momentarily and focus on the properties of the spherical solution.

Figure 1 shows the transfer functions for the uniform Earth model as a function of the angular order n . The quasi-linear decrease of the elastic transfer function with a slope of -1 on the double-logarithmic plot (Fig. 1a) is familiar. If the effects of gravity and sphericity are neglected, the theoretical solution is (e.g. Jef-

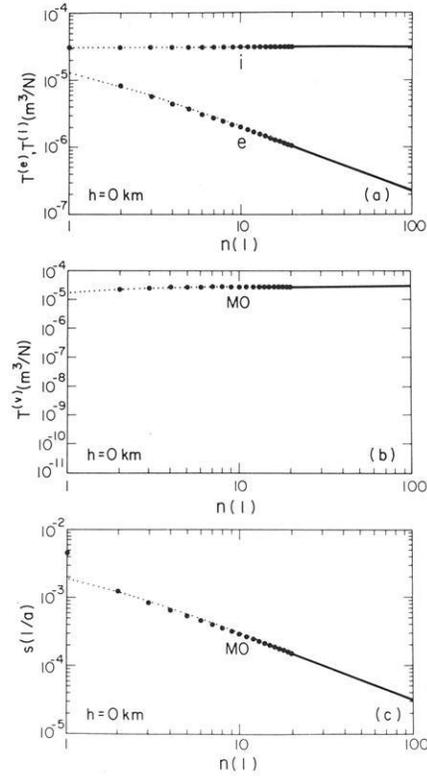


Fig. 1 a, b and c. **a** Elastic transfer functions $T^{(e)}$ and inviscid transfer functions $T^{(i)}$, **b** viscous transfer functions $T^{(v)}$ and **c** inverse relaxation times s as functions of angular order n for Model A. The *solid dots* refer to the spherical solution, the *dotted lines* illustrate the appropriate half-space approximations

frey, 1976, pp. 265–267)

$$T^{(e)}(k) = 1/(2\mu_1 k). \quad (53)$$

Since Model A is uniform, the time dependence is characterized by a single relaxation mode. As Fig. 1b shows, the viscous transfer function $T^{(v)}$ increases with increasing n . This simply compensates for the reduced elastic response at large angular orders such that the inviscid transfer function $T^{(i)}$ be independent of n . The inverse relaxation time $1/s$ increases almost linearly with angular order. For a viscous half-space proportionality holds rigorously such that (Haskell, 1935)

$$1/s = 2\eta_2 k/(\rho g). \quad (54)$$

Figure 2 displays the response pattern of Model B. Since the shear modulus of the lithosphere is smaller than the mantle value, $T^{(e)}$ is enhanced at larger angular orders. The time dependence of Model B, however, differs fundamentally from that of the uniform model and is characterized by two relaxation modes. Wu and Peltier (1982) calculated the radial distribution of the shear energy for each mode and differentiated between a fundamental mantle mode M0 and a shallower fundamental lithospheric mode L0. Their terminology is adopted here. As Fig. 2b shows, the contribution of L0 is usually insignificant except around $n = 7$, where M0 and L0 are of comparable, even though reduced, amplitude. The pronounced decrease of the inviscid transfer function $T^{(i)}$ for $n > 30$ is a consequence of the attenuating effect of the lithosphere

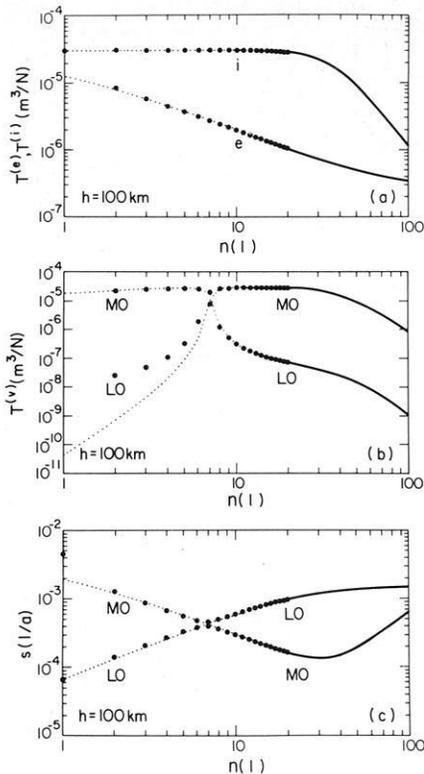


Fig. 2. Same as Fig. 1, but for Model B

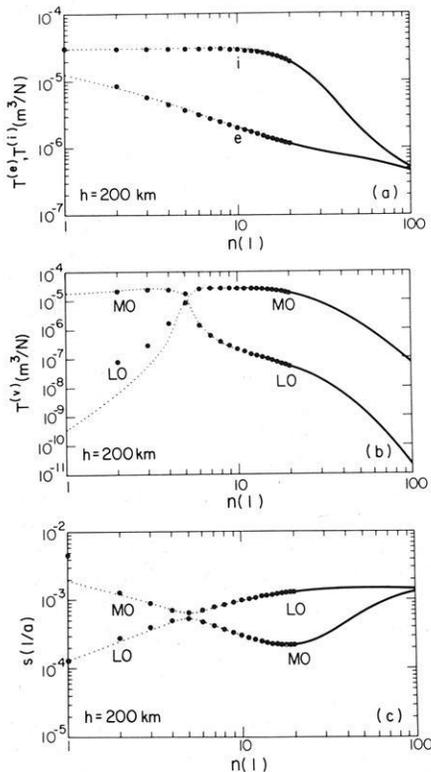


Fig. 3. Same as Fig. 1, but for Model C

at higher angular orders. If $2\pi a_1/n < h$, the lithosphere is nearly "opaque", and the response is essentially elastic. Compared to Model A, the relaxation time of the dominating mantle mode M0 is diminished at higher angular orders. This effect of the lithosphere is already

known from Newtonian Earth models (e.g. McConnell, 1968).

Figure 3 summarizes the response of Model C. The general character resembles that of Model B. Since the thickness of the lithosphere is increased, the viscous contribution to the response is reduced and can be neglected for $n > 100$ (Fig. 3a and b). The inverse relaxation times $s^{(1)}$ and $s^{(2)}$, which can be shown to become identical when $n \rightarrow \infty$, have almost merged at $n = 100$.

The most interesting feature of Figs. 1-3 is the close match between spherical and half-space transfer functions down to very low angular orders. This does not, however, apply to $n=0, 1$. At $n=0$, incompressible spherical models appear rigid and

$$T_0^{(e)} = 0. \quad (55)$$

For the half-space models, however, we find from Eq. (40) at $k=0$

$$T^{(e)}(0) = 1/(\rho g), \quad (56)$$

which is identical to the inviscid limit $T^{(i)}(0)$ for uniform loading. For the spherical model this is obtained from Eq. (38) for $n=1$, where

$$T_1^{(e)} = 1/(\rho g). \quad (57)$$

Apart from that, differences are generally confined to the lithospheric mode L0 of the viscous transfer function, whose amplitude is negligible at most angular orders.

The significance of the differences mentioned above for the response to physical loads is governed by their spectral representation. If high angular orders dominate the load spectrum, the low end of the response spectrum is not sampled efficiently and the half-space approximation is expected to be adequate. For extended loads, however, the lower angular orders become increasingly important. The Laurentide ice sheet, for example, had its energy concentrated near $n=5$ (as compared to $n=15$ for the Fennoscandian ice sheet). An additional effect is that at low orders the Legendre series Eq. (45) will differ from its approximation Eq. (47). This is a direct consequence of the identity for Legendre functions and Bessel functions expressed by Eq. (41). If n is finite and $\theta > 0$, Eq. (41) holds only approximately.

In the following figures, the radial surface displacement $u_r(a_1)$ due to circular loads, as calculated from the spherical solution Eq. (37), is plotted and compared with the vertical surface displacement $-w(0)$ according to the half-space approximation Eq. (39). We thus compare, for $R=800$ km and $R=1600$ km,

$$u_r(a_1, \theta) = - \sum_{n=0}^{\infty} T_n^{(ve)} q_n P_n(\cos \theta), \quad (58)$$

with

$$-w(r', 0) = - \int_0^{\infty} T^{(ve)}(k) \hat{q}(k) k J_0(k r') dk, \quad (59)$$

where $T_n^{(ve)}$ and $T^{(ve)}(k)$ are given by Eq. (51) and q_n and $\hat{q}(k)$ denote the appropriate load spectra according to Eqs. (46) and (48), respectively. In each case the thickness h_0 of the load is 3 km, while its density ρ_0 is 1000 kg m^{-3} .

Figures 4 and 5 show the deflection for the uniform

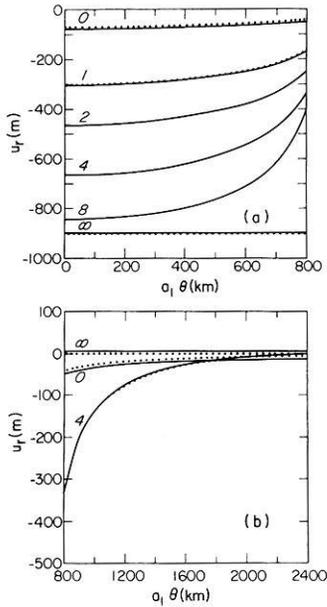


Fig. 4 a and b. Radial surface displacement u_r , as a function of distance $a_1 \theta$ from the load axis for **a** the central region and **b** the peripheral region, and for several times (in units of ka) after the emplacement of the load. The *solid lines* refer to the spherical solution, the *dotted lines* illustrate the appropriate half-space approximation. Results are for Model A and $R = 800$ km

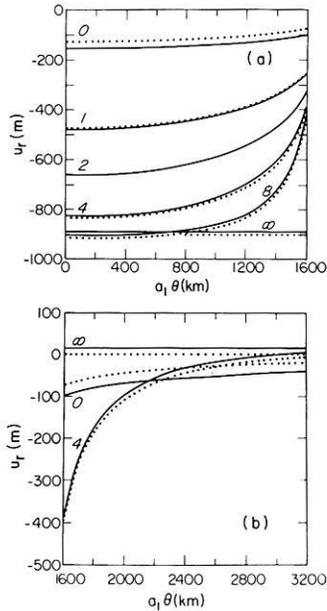


Fig. 5. Same as Fig. 4, but for Model A with $R = 1,600$ km

Model A. If $R = 800$ km, the half-space approximation is excellent. This also holds for $R = 1600$ km except for times directly following emplacement of the load. With increasing time, local compensation is approached. In that limit, the radius of the viscoelastic sphere is slightly increased (Figs. 4b and 5b). This reflects the conservation of volume for the incompressible model considered.

Figure 6 illustrates the response of Model B for $R = 800$ km. For this combination of lithospheric thickness and load radius, the half-space approximation is adequate. If $R = 1600$ km, the region peripheral to the

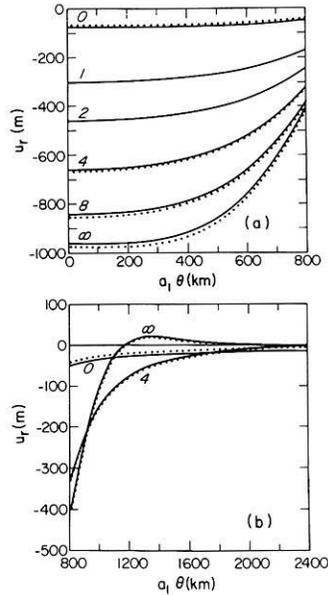


Fig. 6. Same as Fig. 4, but for Model B with $R = 800$ km

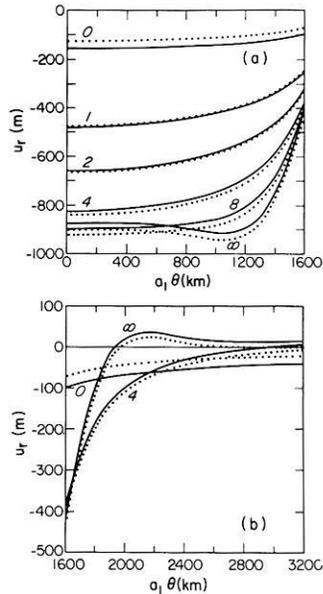


Fig. 7. Same as Fig. 4, but for Model B with $R = 1,600$ km

load requires special attention (Fig. 7b). Here displacements are reduced by about one order of magnitude compared to the central region below the load. This, however, results in relatively larger differences between the spherical and half-space solutions in the peripheral region. The maximum of the peripheral bulge, for example, is underestimated by about 40 per cent on the basis of half-space theory.

For Model C, which is characterized by a 200-km-thick lithosphere, the results are comparable to those of Model B. If $R = 800$ km (Fig. 8a and b), half-space solutions are again adequate. For $R = 1600$ km (Fig. 9a and b), they become marginal in the central region and inadequate in the peripheral region. As before, the maximum of the forebulge is underestimated by approximately 40 per cent.

A different aspect of our theoretical results is the sensitivity of the response to lithospheric thickness.

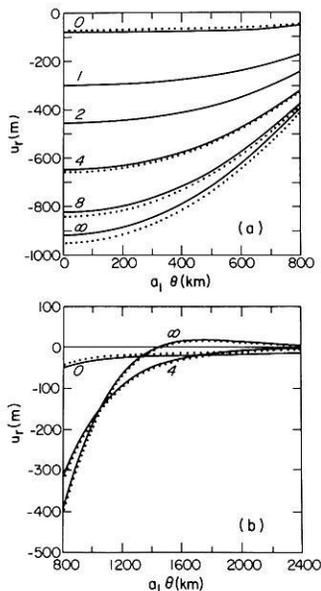


Fig. 8. Same as Fig. 4, but for Model C with $R = 800$ km

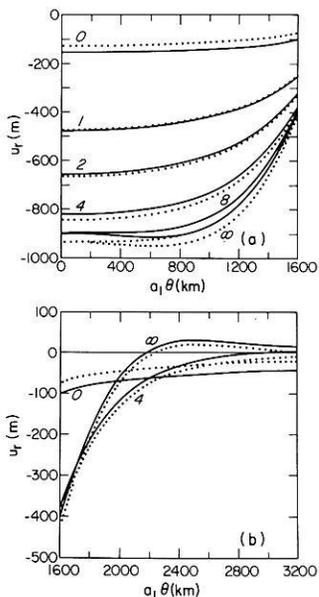


Fig. 9. Same as Fig. 4, but for Model C with $R = 1,600$ km

Since the lithosphere becomes increasingly “transparent” with decreasing angular order of the excitation, we expect that, for $R = 1600$ km, the central region below the load ($a_1 \theta < R$) reflects this insensitivity to a certain extent. As a comparison of Figs. 5a, 7a and 9a shows, the surface deflections are quite similar for distances $a_1 \theta < 800$ km. At larger distances from the axis, and particularly at larger times, the three models differ, however, and Models B and C display an edge effect. This overshoot is already familiar from elementary thin-plate solutions for square-edged loads (e.g. Brochie and Silvester, 1969) and has also been discussed for realistic Earth models by Wu and Peltier (1982). These authors also show that the edge effect is partially related to the load shape and vanishes if more realistic loads of parabolic or similar cross-sections are employed. As can be seen from Figs. 5b, 7b and 9b, the effect of the lithosphere becomes even more pronounced

in the periphery of the load ($a_1 \theta > R$), where the location, amplitude and shape of the bulge are closely controlled by lithospheric thickness.

If $R = 800$ km, the situation is different. Near the load axis the presence of the lithosphere causes the deflection in the inviscid limit to exceed that appropriate to local compensation (Figs. 4a, 6a and 8a). By comparing Figs. 6a and 7a we realize that this behaviour is best explained by the superposition of two edge effects. As for $R = 1600$ km, the excess displacement will, however, decrease for loads of smooth cross-section. The shape of the deflection curves is nevertheless distinct and reflects the effect of the lithosphere. The surface deflection in the peripheral region, on the other hand, seems less suitable for inferring lithospheric thickness. This is mainly related to the fact that, for $R = 800$ km, the volume displaced by the load is small and the bulge therefore not pronounced (Figs. 4b, 6b and 8b).

Discussion and Conclusions

Prior to completion of the work discussed in this paper, the adoption of either spherical or half-space models in the interpretation of isostatic adjustment data has been based more on intuition than on rigorous analysis. For the Fennoscandian ice sheet, for example, the flat-earth approximation has been considered adequate by most authors (e.g. Haskell, 1935; McConnell, 1968). The Laurentide glaciation in Canada, on the other hand, has almost exclusively been analysed in terms of spherical Earth models (see Peltier, 1982, for a summary). A notable exception is Walcott's (1970b) interpretation. This author used a two-dimensional load representation and a flat-earth model to infer the thickness of the lithosphere from tilt data of proglacial Lake Algonquin.

Our previous analysis has confirmed the adequacy of half-space approximations when modelling deformations induced by loads comparable in scale to the Fennoscandian ice sheet. The main advantage of analytical half-space solutions is clearly that they yield much more tractable algebraic expressions compared to those required for the spherical solution. Of some consequence may also be the fact that the numerical implementation of the inverse Hankel transform, based on Simpson's rule, converges faster than the Legendre series of the associated spherical problem.

We have further shown that, for load scales comparable to that of the Laurentide ice sheet, the neglect of sphericity usually yields incorrect results. This holds particularly for the deflection in the periphery of the load. Here deformations are also sensitive to the parameters characterizing the lithosphere and may be used to infer its thickness (Peltier, 1984). This, however, renders the use of spherical models mandatory in such interpretations.

In the past, Walcott's (1970b) inference of lithospheric thickness on the basis of tilt data from Pleistocene Lake Algonquin near the margin of the Laurentide ice sheet was repeatedly questioned (e.g. Nakiboglu and Lambeck, 1982). The main objection in this criticism has been his assumption of an inviscid mantle. Here we will briefly comment on the geometry of his

model, which is characterized by a two-dimensional load resting on a flat Earth.

Using simple half-space models, it may be shown that, for load scales comparable to the Laurentide ice sheet, circular disk loads may be replaced by two-dimensional loads when calculating the deflection near the load margin (Wolf, 1984b). This, however, does not interfere with the question whether sphericity can be neglected in this case. As a cursory inspection of Walcott's theoretical curves shows, the peripheral bulge is quite small, both in amplitude and lateral extension, in his flat-earth solution. Walcott interpreted it is a short-wavelength "static bulge" superimposed on a "dynamic bulge" of larger amplitude and extension, and associated the tilt data with the former.

Quite obviously, Walcott's differentiation between two such bulges is physically not justified. From the present analysis it is further clear that the theoretical bulge increases in amplitude and lateral extent, if the appropriate spherical solution is employed (Fig. 7b). The value of approximately 100 km for the thickness of the lithosphere inferred by Walcott must therefore be viewed with some caution. Since this value is also smaller, by a factor of two, than recent estimates of lithospheric thickness for the same region (Peltier, 1984), the Lake Algonquin tilt data require re-interpretation.

A second result of our analysis has been the demonstration of the overall sensitivity of the response to lithospheric thickness for load scales comparable to the Fennoscandian ice sheet. As discussed in the previous section, this sensitivity is a simple consequence of a shift in the load spectrum to higher angular orders, at which the lithosphere appears increasingly "opaque". This is at variance with Cathles' (1975, p. 153) conclusions, who believed that the strength of the lithosphere was not sufficient to affect the uplift in the central regions of Scandinavia. McConnell (1968), on the other hand, argued that a lithosphere was required in order that the observed decrease in the relaxation time spectrum at shorter wavelengths could be explained, and inferred a value of 120 km for its thickness. McConnell's whole analysis may, however, be criticized, since it is based on the incorrect assumption that the area deformed by the glacial load essentially coincides with the area covered by the load. Considering the limitations of the past interpretations and in view of the fundamental importance of the lithosphere for several geodynamic phenomena, the Fennoscandian uplift data therefore deserve renewed attention.

Appendix A

If we are interested in the solution for a uniform elastic sphere of radius a_2 and shear modulus μ_2 , the general solution of the coupled system of Eqs. (12)–(14) must yield finite field quantities at $r=0$. Then $B_n=D_n=0$ for $n=0, 1, \dots$, and the general solutions Eqs. (22)–(24), (27) and (28) simplify accordingly.

The constants A_n and C_n are determined from the boundary conditions at $r=a_2$. For generality, we allow for arbitrary pre-specified surface tractions and solve Eqs. (27) and (28) at $r=a_2$ for A_n and C_n . We obtain, since $B_n=D_n=0$,

$$2\mu_2 a_2^n (n-1)(2n^2+4n+3) A_n = -n(n^2-1) \left[\frac{\sigma_{rr}(a_2, \theta)}{P_n} + \frac{n\sigma_{r\theta}(a_2, \theta)}{\sin \theta P_n'} \right], \quad (60)$$

$$2\mu_2 a_2^{n-2} (n-1)(2n^2+4n+3) C_n = n \left[n(n+2) \frac{\sigma_{rr}(a_2, \theta)}{P_n} + \frac{(n+1)(n^2-n-3)\sigma_{r\theta}(a_2, \theta)}{\sin \theta P_n'} \right]. \quad (61)$$

Substituting for A_n and C_n in Eqs. (22) and (23) yields

$$2\mu_2 (n-1)(2n^2+4n+3) u_r(r, \theta) = \left(\frac{r}{a_2} \right)^{n-1} n a_2 \left\{ \left[n(n+2) - (n^2-1) \left(\frac{r}{a_2} \right)^2 \right] \sigma_{rr}(a_2, \theta) + \left[(n+1)(n^2-n-3) - n(n^2-1) \left(\frac{r}{a_2} \right)^2 \right] \frac{P_n}{\sin \theta P_n'} \sigma_{r\theta}(a_2, \theta) \right\}, \quad (62)$$

$$2\mu_2 (n-1)(2n^2+4n+3) u_\theta(r, \theta) = - \left(\frac{r}{a_2} \right)^{n-1} a_2 \left\{ \left[n(n+2) - (n+3)(n-1) \left(\frac{r}{a_2} \right)^2 \right] \cdot \frac{\sin \theta P_n'}{P_n} \sigma_{rr}(a_2, \theta) + \left[(n+1)(n^2-n-3) - n(n+3)(n-1) \left(\frac{r}{a_2} \right)^2 \right] \sigma_{r\theta}(a_2, \theta) \right\}. \quad (63)$$

If we put $r=a_2$, the required relation between the field quantities at a_2 is obtained.

Appendix B

For the constants in Eqs. (33)–(36) we obtain

$$D = [k'_1 k_4 + k_1 k'_4 - k'_2 k_6 (a_2/a_1)^2 - k_2 k'_6 (a_2/a_1)^{-2} - k'_3 k_5 (a_2/a_1)^{2n+1} - k_3 k'_5 (a_2/a_1)^{-(2n+1)}], \quad (64)$$

where

$$k_1 = n(n+2)(2n-1)(M-1) \cdot [(2n^2+4n+3)M+2n(n+2)], \quad (65a)$$

$$k_2 = (2n+1)(n^2-1)(M-1) \cdot [(2n^2+4n+3)M+2n(n+2)], \quad (65b)$$

$$k_3 = [2(n^2-1)M+2n^2+1] \cdot [(2n^2+4n+3)M+2n(n+2)], \quad (65c)$$

$$n(n+2)k_4 = (2n+3)(n^2-1)^2(M-1) \cdot [(2n^2+4n+3)M+2n(n+2)], \quad (65d)$$

$$k_5 = -2(n^2-1)(2n^2+4n+3)(M-1)^2, \quad (65e)$$

$$k_6 = [(2n+1)M - (n-2)](n^2-1)(2n^2+4n+3)M + 2n^2(n+2)(n^2+3n-1)(M-1) - [3M+2(n-1)](n+1)^2(n^2-n-3) - [(2n+1)(2n^2+n+3)+3(n+3)(n^2-1)](n+1)M, \quad (65f)$$

$$k'_1 = n^2(n+2)^2(2n-1), \quad (66a)$$

$$k'_2 = n(n+1)(2n+1)[(n-1)(n+2)+F_1], \quad (66b)$$

$$k'_3 = -n[(n+2)(2n^2+1) - (n+1)(2n+1)F_1], \quad (66c)$$

$$k'_4 = (n^2 - 1)^2(2n + 3), \quad (66 \text{ d})$$

$$k'_5 = (n + 1)[(n - 1)(2n^2 + 4n + 3) + n(2n + 1)F_1], \quad (66 \text{ e})$$

$$k'_6 = (n + 1)(n^2 - 1)(n^2 - n - 3) + n^2(n + 2) \cdot (n^2 + 3n - 1) - n(n + 1)(2n + 1)F_1. \quad (66 \text{ f})$$

Here we have used $M = \mu_2/\mu_1$ and $F_1 = \rho g a_1/(2\mu_1)$.

Appendix C

In the following, $T^{(e)}$ may be interpreted either as the elastic transfer function $T_n^{(e)}$ according to the spherical solution Eq. (38) or the half-space approximation $T^{(e)}(k)$ according to Eq. (40). In either case, $T^{(e)}$ can be written in the form

$$T^{(e)} = \frac{e_0 + e_1 M + e_2 M^2}{d_0 + d_1 M + d_2 M^2}, \quad (67)$$

where $M = \mu_2/\mu_1$. The coefficients are complicated expressions of the model parameters. With $M \rightarrow M(s) = \mu_2(s)/\mu_1$, where $\mu_2(s)$ is given by Eq. (50), Eq. (67) can also be interpreted as the Laplace transform

$$\tilde{T}^{(ve)}(s) = \frac{E_0 + E_1 s + E_2 s^2}{D_0 + D_1 s + D_2 s^2} \quad (68)$$

of the transfer function $T^{(ve)}(t)$ appropriate to the associated viscoelastic solution for an impulsive surface load $q\delta(t)$ (see Peltier, 1974, for details). Equation (68) is written in terms of powers of the Laplace transform variable s . The upper-case coefficients of Eq. (68) are easily expressed in terms of the lower-case coefficients of Eq. (67) and $M(s)$. The transform $\tilde{T}^{(ve)}(s)$ of the viscoelastic transfer function may be split into an elastic portion $T^{(e)}$ and a viscous portion $\tilde{V}(s)$, i.e.

$$\tilde{T}^{(ve)}(s) = T^{(e)} + \tilde{V}(s), \quad (69)$$

where

$$T^{(e)} = \lim_{s \rightarrow \infty} \tilde{T}^{(ve)}(s). \quad (70)$$

$T^{(e)}$ is given by Eq. (38) or Eq. (40) and interpreted as the Laplace transform of the instantaneous elastic response of the viscoelastic system to the impulsive load applied at $t=0$. The transform $\tilde{V}(s)$ of the viscous portion can be cast into the form

$$\tilde{V}(s) = T^{(v,1)} \frac{s^{(1)}}{s + s^{(1)}} + T^{(v,2)} \frac{s^{(2)}}{s + s^{(2)}}, \quad (71)$$

where $T^{(v,1)}$, $T^{(v,2)}$ and $s^{(1)}$, $s^{(2)}$ are given in terms of the coefficients in Eq. (68). Implementing the inverse Laplace transform yields

$$T^{(ve)}(t) = T^{(e)} \delta(t) + T^{(v,1)} s^{(1)} \exp(-s^{(1)} t) + T^{(v,2)} s^{(2)} \exp(-s^{(2)} t). \quad (72)$$

This is the system's impulse response. The response to a Heaviside loading history follows from convolving it with the impulse response. From Eq. (72) we obtain for $t \geq 0$

$$T^{(ve)}(t) = T^{(e)} - T^{(v,1)} [\exp(-s^{(1)} t) - 1] - T^{(v,2)} [\exp(-s^{(2)} t) - 1]. \quad (73)$$

Acknowledgements. This research was financially supported by a Natural Sciences and Engineering Research Council of Canada Postgraduate Scholarship. Special thanks are due to Richard Peltier for stimulating discussions and to Khader Khan for drafting the figures.

References

- Brothie, J.F., Silvester, R.: On crustal flexure. *J. Geophys. Res.* **74**, 6240-6252, 1969
- Bullen, K.E.: An introduction to the theory of seismology, 3rd edn. Cambridge: Cambridge University Press 1963
- Cathles, L.M.: The viscosity of the Earth's mantle. Princeton: Princeton University Press 1975
- Crittenden, M.D. jr.: Effective viscosity of the Earth derived from isostatic loading of Pleistocene Lake Bonneville. *J. Geophys. Res.* **68**, 5517-5530, 1963
- Dragoni, M., Yuen, D.A., Boschi, E.: Global postseismic deformation in a stratified viscoelastic Earth: effects on Chandler wobble excitation. *J. Geophys. Res.* **88**, 2240-2250, 1983
- Haskell, N.A.: The motion of a viscous fluid under a surface load. *Physics* **6**, 265-269, 1935
- Jeffreys, H.: The Earth, 6th edn. New York: Cambridge University Press 1976
- Lebedev, N.N.: Special functions and their applications. New York: Dover 1972
- Love, A.E.H.: Some problems of geodynamics. Cambridge: Cambridge University Press 1911
- Love, A.E.H.: A treatise on the mathematical theory of elasticity, 4th edn. Cambridge: Cambridge University Press 1927
- McConnell, R.K.: 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 response to surface loading with application to Lake Bonneville. *Geophys. J. R. Astron. Soc.* **70**, 577-620, 1982
- Peltier, W.R.: The impulse response of a Maxwell Earth. *Rev. Geophys. Space Phys.* **12**, 649-669, 1974
- Peltier, W.R.: Ice sheets, oceans, and the Earth's shape. In: Earth rheology, isostasy, and eustasy, N.-A. Moerner, ed.: pp. 45-63. New York: Wiley 1980
- Peltier, W.R.: Dynamics of the ice age Earth. *Adv. Geophys.* **24**, 1-146, 1982
- Peltier, W.R.: The thickness of the continental lithosphere. *J. Geophys. Res.* 1984 (in press)
- Slichter, L.B., Caputo, M.: Deformation of an Earth model by surface pressures. *J. Geophys. Res.* **65**, 4151-4156, 1960
- Sneddon, I.A.: Fourier transforms. New York: McGraw-Hill 1951
- Walcott, R.I.: Flexural rigidity, thickness, and viscosity of the lithosphere. *J. Geophys. Res.* **75**, 3941-3954, 1970 a
- Walcott, R.I.: Isostatic response to loading of the crust in Canada. *Can. J. Earth Sci.* **7**, 716-727, 1970 b
- Watson, G.N.: A treatise on the theory of Bessel functions, 2nd edn. Cambridge: Cambridge University Press 1944
- Wolf, D.: Thick plate flexure re-examined. *Geophys. J. R. Astron. Soc.* 1984 a (in press)
- Wolf, D.: On the relation between two-dimensional and axisymmetric loads in plate flexure problems. *J. Geophys.* **54**, 232-235, 1984
- Wu, P., Peltier, W.R.: Viscous gravitational relaxation. *Geophys. J. R. Astron. Soc.* **70**, 435-485, 1982

Received February 1, 1984; Revised version July 10, 1984

Accepted July 11, 1984

Plasma-gas interactions in planetary atmospheres and their relevance for the terrestrial hydrogen budget

Hans Uwe Nass and Hans Jörg Fahr

Institut für Astrophysik und Extraterrestrische Forschung der Universität Bonn, Auf dem Hügel 71, D-5300 Bonn 1, Federal Republic of Germany

Abstract. The relevance of charge exchange processes for the evaporation of neutral gases out of planetary gravitational fields has been known for quite a long time. For the planetary escape, however, collisions of the escaping constituent with both neutral and ionized atmospheric species operate in building up the escape flux. For the earth, hydrogen collisions with O-atoms and O^+ - and H^+ -ions lead to about comparable contributions to the hydrogen escape at heights below 800 km, whereas at larger heights only charge exchange collisions need to be considered. In the present work, the lower region is therefore described by particle and energy flux conservation concepts, whereas in the upper region free-flight kinetic concepts are used. We give solutions for the height profiles of the hydrogen density, temperature and escape flux. The inclusion of charge exchange collisions increases the scale height, and consequently a smaller decrease of the hydrogen density is given. The hydrogen temperature, however, is reduced by including this effect. The charge exchange processes contribute about 70% to the total hydrogen escape which is of the order of a few 10^7 atoms/cm² s at the plasmopause.

Key words: Plasma-gas interactions – Charge exchange collisions – Escape

Introductory remarks

It has been known for quite a long time that thermal loss mechanisms alone cannot account for the total hydrogen escape from the earth's atmosphere. Rather, it has been suspected that non-thermal, i.e. ion-induced, processes due to charge exchange with ionospheric O^+ - and H^+ -ions, also contribute to hydrogen escape. The extraordinary importance of such charge exchange processes also for non-terrestrial exospheres has just recently been stressed in papers by Cloutier et al. (1978), Eviatar et al. (1979), Ip and Axford (1980), Goertz (1980), Brown and Schneider (1981), Goertz and Ip (1982) and Ip (1982). An ion pick-up from Io's ionosphere via charge exchange collisions may contribute to its atmospheric losses. The underlying idea in all these papers is that due to charge exchange processes, newly created neutrals, most probably H- or O-atoms, are injected into the "escape branch" of the velocity distri-

bution function at a rate in excess of the counteracting thermalization rate. The population density in this branch can thus be enhanced compared to the thermal one, resulting in enhanced escape rates. In this context it may also be mentioned that charge exchange interaction of precipitating terrestrial keV O^+ -ions with atmospheric oxygen under geomagnetically disturbed conditions represents a source of escaping O-atoms (Torr et al., 1974, Torr and Torr, 1979). The insufficiency of the thermal escape concept to describe the terrestrial hydrogen geocorona has been pointed out by Bertaux (1975), Vidal-Madjar (1978) and Vidal-Madjar and Thomas (1978). There it was shown that the Lyman-Alpha inferred variation of the exobasic hydrogen density with thermospheric temperature is incompatible with the assumption of a Jeans-type escape of hydrogen (Jeans, 1925). Thus a charge-exchange-induced contribution to the terrestrial hydrogen escape has been proposed to solve this problem. Both charge exchange contributions and rarefied-gas-dynamical deviations from the Jeans escape, as discussed in Gross (1974), Fahr and Weidner (1977) and Fahr and Nass (1978), tend to weaken the temperature dependence of the net escape rate. Nevertheless, a quantitative value for the fractional escape contribution due to charge exchange processes has not yet been given in a rigorous treatment. In order to give as realistic an answer as possible, several different approaches towards this problem have been attempted, most of these being concerned with the specific problem of the terrestrial hydrogen escape where charge exchange collisions of H-atoms with protonospheric O^+ - and H^+ -ions play the dominant role (Maher and Tinsley, 1977; Chamberlain, 1977; Hodges et al., 1981; Shizgal and Lindenfeld, 1982). An extensive review of this work was recently given by Fahr and Shizgal (1983). In most of the earlier works, with the exception of Shizgal and Lindenfeld (1982), charge exchange escape contributions had been considered exclusively. However, as the latter authors could show, the resulting escape rates are substantially reduced if, in addition to charge exchange collisions, elastic collisions of H-atoms with O-atoms are also included. To arrive at a more conclusive answer concerning the fractional contributions by thermal and "non-thermal" processes we have reinvestigated this problem with special emphasis on the coupling of both processes.

Theoretical concept for the terrestrial case

In order to calculate the hydrogen loss we subdivide the upper atmosphere into two zones: (1) a transition region between 300 and 800 km where hydrogen atoms suffer collisions with both the neutral component of the background gas, i.e. with oxygen atoms, and with the ionized component, i.e. oxygen ions and protons; and (2) the region between 800 km and the plasmopause where nearly no elastic collisions and only a few charge exchange collisions between hydrogen atoms and fast protonospheric protons occur. Due to the presence of elastic collisions there exists a strong coupling in the lower region between the non-escaping and the escaping branch of the hydrogen velocity distribution function. To overcome the mathematical difficulties connected with this coupling we use the concepts of mass flow and energy flow continuity separately in this region, for these two categories of hydrogen particles. The build-up of the hydrogen escape in this region will be described by a parametrized quasi-Maxwellian representation of the hydrogen velocity distribution f_H with density, temperature and degree of anisotropy as unknown parameters. The determination of these parameters as functions of height is reached by integration of the continuity equations.

Due to the ineffectiveness of elastic collisions in the upper region in producing escaping particles we can use a simpler concept here and consider only local charge-exchange-induced production of escaping particles multiplied by their escape probabilities to space as a representation of the change of the escape flux with height. Deviating, however, from the practice of Shizgal and Lindenfeld (1982), we also consider the spherical geometry of the problem and thus take into account the effect of charge-exchange-induced escaping particles first moving to lower levels but with perigees greater than 800 km and therefore potentially contributing to the escaping branch of f_H . First we want to discuss the physics in the transition region and start with the description of the hydrogen velocity distribution function f_H .

Distribution function

To represent this distribution function we use here the following expression, given by Fahr and Nass (1978):

$$\begin{aligned} f_H(r, v') &= f_{ne}^\pm + f_e^+ + f_e^- \\ &= \frac{1}{G_1 + G_2 + G_3} \cdot \left\{ M(r, v') h(v_{esc} - v') \right. \\ &\quad + M(r, v') h(v' - v_{esc}) h\left(\frac{\pi}{2} - \theta\right) \\ &\quad \left. + \alpha(r) M(r, v') h(v' - v_{esc}) h\left(\theta - \frac{\pi}{2}\right) \right\} \quad (1) \end{aligned}$$

where $v' = |\tilde{v} - \tilde{v}_{Diff}|$ with \tilde{v}_{Diff} being that particular diffusion velocity at the 300 km level ensuring flux continuity of the H-atoms over the transition region. This diffusion velocity is small in comparison with thermal velocities, and thus we can identify v' with v . h is the step function, $M(r, v)$ is the Maxwellian distribution function according to a locally defined hydrogen tem-

perature $T_H(r)$, the subscripts "ne" and "e" stand for "non-escaping" and "escaping", and superscripts +, - for upward and downward distributions, respectively. The quantities G_1 , G_2 and G_3 are normalization factors, given explicitly by:

$$G_1 = \int_0^{v_{esc}} \int_0^{\pi/2} \int_0^{2\pi} f_{ne}^\pm d^3v = \text{erf}(\bar{v}_H) - \frac{2}{\sqrt{\pi}} \exp(-\bar{v}_H^2), \quad (2)$$

$$G_2 = \int_{v_{esc}}^{\infty} \int_0^{\pi/2} \int_0^{2\pi} f_e^+ d^3v = \frac{1}{2}(1 - G_1), \quad (3)$$

$$G_3 = \int_{v_{esc}}^{\infty} \int_{\pi/2}^{\pi} \int_0^{2\pi} f_e^- d^3v = \alpha \cdot G_2, \quad (4)$$

where

$\text{erf}(x)$ = error function,

$$v_{esc} = \sqrt{\frac{2\gamma M}{r}},$$

γ = gravitational constant,

M = mass of Earth,

r = distance from centre of Earth,

$$\bar{v}_H = \sqrt{\frac{\gamma M m_H}{k T_H r}},$$

m_H = mass of hydrogen atom,

k = Boltzmann constant.

The factor $\alpha(r)$ occurring in the distribution function (1) takes into account the fact that the distribution function f_e^\pm for escaping particles is increasingly anisotropic with r . This results from the continuous decrease of collision rates between hydrogen atoms and the background gas in higher atmospheric regions responsible for the production of the downward-flying particles with an energy greater than the escape energy. Obviously, below a certain height r_c , $\alpha(r)$ has to be equal to one, meaning that an isotropic Maxwellian prevails. Going upwards from this level α will decrease and at large heights r the value α will asymptotically tend to zero, meaning that there will be no downward-flying particles with escape energy. Hence one expects a specific functional relation between α and the collision probability which we formulate in the following manner

$$\alpha(r) = 1 - (1 - \beta) w_e(r, \infty), \quad (5)$$

where $w_e(r, \infty)$ is the angle averaged probability that a particle with an averaged escape velocity representative of the whole escape category can leave the spherical atmosphere without suffering any collisions above r :

$$w_e(r, \infty) = \int_0^{\pi/2} \exp\left\{-\sum_i q_i \bar{N}_{i,e}(r, \theta)\right\} \sin \theta d\theta. \quad (6)$$

Downward-flying particles with an energy greater than the escape energy are produced not only by collisions between upward-flying escaping particles and the background gas but also by charge-exchange collisions between fast protons and hydrogen or oxygen atoms. This means that under the additional effect of these collisions $\alpha(r)$ will not decrease towards zero as rapidly as in the case of elastic collisions alone. To account for this we introduce the quantity β in Eq. (5). This quan-

tity β has to be zero if there are no ionized hydrogen atoms in upper atmospheric regions. On the other hand, we can calculate β from the requirement that there must be a continuous transition between the transition region (300–800 km) and the higher atmospheric regions (800 km – plasmopause).

The quantity $\bar{N}_{i,e}(r, \theta)$ is the mean column density of the i 'th background gas component seen by an averaged escaping hydrogen atom at r flying upwards with an angle θ against the radius vector, i.e.:

$$\bar{N}_{i,e}(r, \theta) = \int_r^\infty n_i(r') \frac{\bar{v}_{i,e}^{rel}}{\bar{v}_e \cos \theta'} dr' \quad (7)$$

Here \bar{v}_e is the mean velocity of all escaping hydrogen atoms at r' .

$\bar{v}_{i,e}^{rel}$ is the mean relative velocity of a hydrogen atom moving with velocity \bar{v}_e with respect to the i 'th background gas component of temperature $T_i(r')$, particle mass m_i and Maxwellian velocity distribution $M[T_i(r')]$:

$$\bar{v}_{i,e}^{rel} = \sqrt{\frac{2kT_i}{\pi m_i}} \exp\left(-\frac{m_i \bar{v}_e^2}{2kT_i}\right) + \bar{v}_e \cdot \left(1 + \frac{kT_i}{m_i \bar{v}_e^2}\right) \operatorname{erf}\left(\sqrt{\frac{m_i \bar{v}_e^2}{2kT_i}}\right) \quad (8)$$

According to the conservation of angular momentum, the angle θ' of Eq. (7) is given by

$$\cos \theta' = \sqrt{1 - \left(\frac{r \bar{v}_e}{r' \bar{v}_e'}\right)^2 \sin^2 \theta} \quad (9)$$

By introducing the averaged escape velocity in Eq. (6) we give only an approximation to the escape probability. In principle we have to integrate this probability over the whole velocity distribution, but it can be shown that our approximation differs only slightly from the rigorous expression.

Energy flux continuity in the transition region

Because of the disparate masses of the collision partners, the energy coupling due to mutual elastic collisions between the light escaping gas constituent (hydrogen) and the relatively heavy neutral background gas (oxygen) is small. Thus, due to productions of escaping atoms, the energy that is locally extracted from the escaping gas constituent cannot be locally replaced by elastic collisions with oxygen. Therefore, there is no local balance between energy losses and gains. This unbalanced situation consequently leads to a height gradient of the temperature of the escaping constituent (T_H) in the transition region, whereas under geomagnetic quiet conditions the oxygen temperature T_O can be expected to remain at its asymptotic thermospheric temperature value T_0 . This is due to its high thermal conductivity and the non-existence of energy losses by escape. As was shown by Torr et al. (1974), precipitating O^+ -ions at higher latitudes under disturbed conditions can represent an energy source at exobasic levels. However, this has not been considered here since its influence on the averaged global escape is estimated

to be of minor importance. The hydrogen temperature T_H , however, may drop below T_0 in order to establish a local thermal energy source for the escaping constituent.

In a previous paper, Fahr and Nass (1978) also examined an additional energy loss for hydrogen embedded in oxygen. As is known, inelastic collisions between hydrogen and oxygen can give rise to the excitation of the 3P_1 state of oxygen followed immediately by an emission of a 63μ photon. This process in principle should be able to cool down hydrogen even further because oxygen, due to its high thermal conductivity, can be taken as an infinite heat reservoir with temperature T_0 . A more accurate examination of this specific process by Durrance and Thomas (1979), who took into account the characteristics of the optically thick radiation field of the 63μ emission line, shows, however, that if some simplifying assumptions with regard to the relevant inelastic cross-sections can be made, this specific temperature reduction can be neglected. Therefore, the energy situation in the transition region can be described by the following heat conduction equation:

$$\begin{aligned} \frac{d}{dr} \left(\kappa(r) \frac{d}{dr} T_H(r) \right) &= n_H n_O q_{H,O}^{el} \bar{v}_{H,O}^{rel} \varepsilon^{\frac{3}{2}} k [T_O - T_H] \\ &+ n_P n_H q_{P,H}^{ex} \bar{v}_{P,H}^{rel} \frac{3}{2} k [T_1 - T_H] \\ &+ n_P n_O q_{P,O}^{ex} \bar{v}_{P,O}^{rel} \frac{3}{2} k [T_1 - T_H] \\ &- n_H n_{O^+} q_{H,O^+}^{ex} \bar{v}_{H,O^+}^{rel} \frac{3}{2} k [T_1 - T_H] \\ &- \frac{d}{dr} \phi_e^{eff}(r). \end{aligned} \quad (10)$$

The terms on the right-hand side stand for the following processes:

- 1) energy coupling with oxygen,
- 2) energy coupling with protons,
- 3) energy gain due to collisions between protons and oxygen,
- 4) energy loss due to collisions between hydrogen and oxygen ions,
- 5) energy loss due to local escape contributions,

where

$$\begin{aligned} \kappa(r) &= \bar{\kappa} n_H \frac{\sqrt{T_H(r)}}{n_O q_{H,O}^{el} + n_P q_{P,H}^{ex} + n_{O^+} q_{H,O^+}^{ex}}, \\ \bar{\kappa} &= 1.23 \times 10^{-12} \text{ [g cm}^3 \text{ grad}^{-3/2} \text{ s}^{-3}\text{]}, \end{aligned}$$

$n_i(r)$ = density of the i 'th background gas component,

$$\varepsilon = 2m_H m_O / (m_H + m_O)^2,$$

T_0 = oxygen temperature,

T_1 = ion temperature (taken to be common for O^+ - and H^+ -ions).

The cross-section for elastic collisions between H- and O-atoms can be adopted with (Liwshitz and Singer, 1966):

$$q_{H,O}^{el} = 3 \times 10^{-15} \text{ cm}^2, \quad (11)$$

whereas the charge-exchange cross-sections are slightly velocity dependent and given by (Storm, 1970; Rapp, 1963)

$$q_{\text{P,H}}^{\text{ex}} = (1.64 \times 10^{-7} - 1.6 \times 10^{-8} \log(v_{\text{P,H}}^{\text{rel}}))^2 \quad (12)$$

with

$$10^5 \leq v_{\text{P,H}}^{\text{rel}} \leq 5 \times 10^7 \text{ cm/s}$$

and

$$q_{\text{P,O}}^{\text{ex}} = \frac{1}{3} (1.64 \times 10^{-7} - 1.6 \times 10^{-8} \log(v_{\text{P,O}}^{\text{rel}}))^2 \quad (13)$$

with

$$3 \times 10^5 \leq v_{\text{P,O}}^{\text{rel}} \leq 5 \times 10^7 \text{ cm/s.}$$

We use identical cross-sections for $\text{H}^+ - \text{O}$ collisions and for $\text{H} - \text{O}^+$ collisions, since in thermal equilibrium the ratio of the reaction rate $[\text{HO}^+]/[\text{H}^+\text{O}]$ is indicated to be 9:8 (Chamberlain, 1956).

The required mean relative velocities can be calculated by the following formula:

$$\begin{aligned} \bar{v}_{i,j}^{\text{rel}} = & \int_0^\infty \int_0^\pi \int_0^{2\pi} f_i d^3 v_i \\ & \cdot \int_0^\infty \int_0^\pi \int_0^{2\pi} f_j \sqrt{v_i^2 + v_j^2 - 2v_i v_j \cos \theta_j} d^3 v_j. \end{aligned} \quad (14)$$

If the velocity distribution functions of the background gas components are Maxwellians, which is a very reasonable assumption, then the integrations can be carried out analytically:

$$\begin{aligned} \bar{v}_{i,j}^{\text{rel}} = & \sqrt{\frac{8k}{\pi} \left(\frac{T_i}{m_i} + \frac{T_j}{m_j} \right)} \quad (i, j \neq \text{H}), \\ \bar{v}_{\text{H},j}^{\text{rel}} = & \frac{1 + \alpha}{2(G_1 + G_2 + G_3)} \sqrt{\frac{8k}{\pi} \left(\frac{T_{\text{H}}}{m_{\text{H}}} + \frac{T_j}{m_j} \right)} \\ & + \frac{1 - \alpha}{2(G_1 + G_2 + G_3)} I_9(\text{H}, j). \end{aligned} \quad (15)$$

(The factor $I_9(\text{H}, j)$ is explained in the Appendix.) The local energy flux associated with the escape flux at some height level r is given by

$$\phi_e^{\text{eff}}(r) = n_{\text{H}}(r) J_e^{\text{E}}(r) w_e(r, \infty) \quad (16)$$

with

$$\begin{aligned} J_e^{\text{E}} = & \int_{v_{\text{esc}}}^\infty \int_0^{\pi/2} \int_0^{2\pi} \left(\frac{1}{2} m_{\text{H}} v^2 \right) v \cos \theta f_e^+ d^3 v \\ = & \frac{1}{G_1 + G_2 + G_3} \frac{m_{\text{H}}}{2\sqrt{\pi}} \left(\frac{2k T_{\text{H}}}{m_{\text{H}}} \right)^{3/2} \\ & \cdot \exp(-\bar{v}_{\text{H}}^2) \left\{ \bar{v}_{\text{H}}^4 + \frac{1}{2} \bar{v}_{\text{H}}^2 + \frac{1}{2} \right\}. \end{aligned} \quad (17)$$

Gain and loss processes

The particle flux continuity equation can be written in the following form:

$$\text{div } J = \text{div } J_{ne}^+ + \text{div } J_{ne}^- + \text{div } J_e^+ + \text{div } J_e^-, \quad (18)$$

where the subscripts "ne", "e" stand for "non-escaping" and "escaping" particles and superscripts +, - for upward- and downward-directed flux, respectively.

This means that the total flux is built up by all particles crossing the height level r in any direction. Each contribution to the total flux J is derived by considering the relevant gain and loss processes responsible for the individual fluxes, i.e.

$$\begin{aligned} \text{div } J_{ne}^+ = & P_{ne}^{\text{N},+} + P_{ne}^{\text{I},+} + L_G, \\ \text{div } J_e^+ = & P_e^{\text{N},+} + P_e^{\text{I},+}, \end{aligned} \quad (19)$$

where the indices N, I stand for gain or loss processes due to collisions of hydrogen with the neutral and ionized background gas components, respectively. The term L_G is a gravitationally caused loss term and will be explained later.

The flux of hydrogen atoms at each height level r is composed of upward- or downward-flying particles starting at lower or higher levels r' and reaching r without suffering any collisions with the background gas. Thus contributions from all these levels r' have to be integrated, yielding

$$J_{ne}^+(r) = \frac{\int_{r_c}^r \left(\frac{r'}{r} \right)^2 n_{\text{H}}(r') v_{ne}^{\text{eff},+}(r', r) w_{ne}(r', r) dr'}{\int_{r_c}^r w_{ne}(r', r) dr'}. \quad (20)$$

An analogous equation describes the upward-directed flux of escaping particles. Here $w_{ne}(r', r)$ is the probability that the particles will fly without collisions from r' to r and is given by an expression similar to Eq. (6).

The quantity $v_{ne}^{\text{eff},+}(r', r)$ is the mean radial effusion velocity of non-escaping hydrogen atoms ascending from r' to r . We have to take into account that the flux of non-escaping particles at a height level r is built up only of those particles starting at a lower level r' and in fact arriving at the upper level r . This means that at r' these particles must have a velocity greater than a certain minimum velocity. Therefore, $v_{ne}^{\text{eff},+}(r', r)$ is given by:

$$\begin{aligned} v_{ne}^{\text{eff},+}(r', r) = & \int_{v_0}^{v_{\text{min}}} \int_0^{\theta_{\text{max}}} \int_0^{2\pi} f_{ne}^\pm v \cos \theta d^3 v \\ & + \int_{v_{\text{min}}}^{v_{\text{esc}}} \int_0^{\pi/2} \int_0^{2\pi} f_{ne}^\pm v \cos \theta d^3 v. \end{aligned} \quad (21)$$

Here v_0 is that particular minimum velocity belonging to an angle $\theta=0$ between the velocity vector of the particles and the radius vector. On the other hand, if the angle θ lies between 0 and θ_{max} , where

$$\theta_{\text{max}} = \arcsin \left(\frac{r}{r'v} \sqrt{v^2 - 2\gamma M \left(\frac{1}{r'} - \frac{1}{r} \right)} \right), \quad (22)$$

then these particles must have a minimum velocity which is between v_0 and v_{min} .

By introducing the following two quantities

$$\bar{v}_{\text{min}} = \sqrt{\frac{m_{\text{H}} v_{\text{min}}^2}{2k T_{\text{H}}}} \quad (23)$$

and

$$\bar{v}_0 = \sqrt{\frac{m_{\text{H}} v_0^2}{2k T_{\text{H}}}}, \quad (24)$$

the mean radial effusion velocity of non-escaping particles can be given analytically by

$$v_{ne}^{eff,+}(r', r) = \frac{0.5}{G_1 + G_2 + G_3} \sqrt{\frac{2kT_H}{\pi m_H}} \left(\frac{r}{r'}\right)^2 \cdot \{e^{-\bar{v}_0^2} - (1 + \bar{v}_{min}^2 - \bar{v}_0^2) e^{-\bar{v}_{min}^2}\} + \frac{0.5}{G_1 + G_2 + G_3} \sqrt{\frac{2kT_H}{\pi m_H}} \cdot \{(1 + \bar{v}_{min}^2) e^{-\bar{v}_{min}^2} - (1 + \bar{v}_H^2) e^{-\bar{v}_H^2}\}. \quad (25)$$

An expression for the mean radial effusion velocity of escaping particles is obtained more easily because particles having a velocity greater than the escape velocity at r' can always reach the upper level r , i.e.

$$v_e^{eff,+}(r') = \int_{v_{esc}}^{\infty} \int_0^{\pi/2} \int_0^{2\pi} f_e^+ v \cos \theta d^3 v = \frac{0.5}{G_1 + G_2 + G_3} \sqrt{\frac{2kT_H}{\pi m_H}} (1 + \bar{v}_H^2) e^{-\bar{v}_H^2}. \quad (26)$$

The quantity $v_e^{eff,-}(r', r)$ which enters the equation for the downward-directed flux of escaping particles J_e^- is given by

$$v_e^{eff,-}(r', r) = \int_{v_{esc}}^{\infty} \int_{\pi - \theta_g}^{\pi} \int_0^{2\pi} f_e^- v \cos \theta d^3 v = \alpha(r) \frac{0.5}{G_1 + G_2 + G_3} \sqrt{\frac{2kT_H}{\pi m_H}} \cdot \left\{ \left(\frac{r}{r'}\right)^2 (1 + \bar{v}_H^2) + \bar{v}_0^2 \right\} e^{-\bar{v}_H^2}, \quad (27)$$

where

$$v_0 = \sqrt{2\gamma M \left(\frac{1}{r} - \frac{1}{r'}\right)}. \quad (28)$$

Since a particle starting at an upper level r' must reach the lower level r in order to contribute to the flux considered there, its velocity vector can only attain angles with the radius vector within a certain range. The limiting angle θ_g is found by angular momentum considerations.

The flux considered is then given by

$$J_e^-(r) = \frac{\int_r^{r_G} \left(\frac{r'}{r}\right)^2 n_H(r') v_e^{eff,-}(r', r) w_e(r', r) dr'}{\int_r^{r_G} w_e(r', r) dr'}. \quad (29)$$

Here r_G is the limiting height of the transition region, i.e. $r_G = 800$ km.

To obtain an equation for the downward-directed flux of non-escaping particles we have to find an expression for the quantity $v_{ne}^{eff,-}$. But this is not trivial because this quantity is built up in two parts, namely (1) all particles produced at heights r' above r and flying downward; and (2) all particles starting at heights below r , reaching their apogee above r and again flying downward. To overcome these difficulties we reconsider the hydrogen velocity distribution f_H given in Eq. (1). Since this distribution is taken to be purely

isotropic for the non-escaping branch of the velocity space, the upward-directed flux of non-escaping particles must be consistently balanced by a corresponding downward-directed flux, i.e.

$$J_{ne}^-(r) = J_{ne}^+(r). \quad (30)$$

Let us now consider the gain and loss processes and start with those processes due to charge-exchange collisions. These collisions can be essentially characterized in the following simple way: the neutral and the ionized particle exchange an electron without changing their flight paths in doing so; this means there is no momentum transfer in such a collision [see for instance Mapleton (1972)]. Therefore, the production of upward-flying escaping hydrogen atoms resulting from collisions between protons and oxygen atoms is simply given by

$$P^{l,+} = n_O n_P \int_{v_{esc}}^{\infty} \int_0^{\pi/2} \int_0^{2\pi} q_{P,O}^{ex}(v_{rel}) f_P d^3 v_P \int_0^{\infty} \int_0^{\pi} \int_0^{2\pi} f_O v_{rel} d^3 v_O = n_O n_P q_{P,O}^{ex}(\bar{v}_{rel}) \int_{v_{esc}}^{\infty} \int_0^{\pi/2} \int_0^{2\pi} f_P d^3 v_P \int_0^{\infty} \int_0^{\pi} \int_0^{2\pi} f_O v_{rel} d^3 v_O, \quad (31)$$

where \bar{v}_{rel} is the mean relative velocity between the two particles and can be calculated by evaluating the six-fold integral of the second part in Eq. (31)

$$\bar{v}_{rel} = \frac{1}{2} \cdot \{I_7(P, O) - I_9(P, O)\}. \quad (32)$$

(The integrals I_1 through I_9 are explained in the Appendix.) Adding all the gain and loss processes in the same way for both upward-flying non-escaping and escaping hydrogen atoms, we come to the following two expressions, giving net production rates of H-atoms of both categories.

$$P_{ne}^{l,+} = n_H n_P \cdot \{q_{P,H}^{ex} \cdot \frac{1}{2} \cdot [I_9(P, H) - I_6(P, H)] + q_{P,H}^{ex} \cdot I_3(P, H) + \alpha \cdot q_{P,H}^{ex} \cdot I_4(P, H)\} - n_H n_P \cdot \{q_{P,H}^{ex} \cdot \frac{1}{2} \cdot I_9(H, P)\} + n_O n_P \cdot \{q_{P,O}^{ex} \cdot \frac{1}{2} \cdot I_9(P, O)\} + n_H \cdot n_{O^+} \cdot \{q_{H,O^+}^{ex} \cdot \frac{1}{2} \cdot I_9(H, O^+)\}. \quad (33)$$

$$P_e^{l,+} = n_H n_P \cdot \{q_{P,H}^{ex} \cdot \frac{1}{2} \cdot I_6(H, P) + \alpha \cdot q_{P,H}^{ex} I_2(P, H) + q_{P,H}^{ex} \cdot I_1(P, H)\} - n_H n_P \cdot \{q_{P,H}^{ex} \cdot \frac{1}{2} \cdot [I_7(H, P) - I_9(H, P)]\} + n_O n_P \cdot \{q_{P,O}^{ex} \cdot \frac{1}{2} \cdot [I_7(P, O) - I_9(P, O)]\} - n_H n_{O^+} \cdot \{q_{H,O^+}^{ex} \cdot \frac{1}{2} \cdot [I_7(H, O^+) - I_9(H, O^+)]\}. \quad (34)$$

The quantities $P_{ne}^{l,-}$ and $P_e^{l,-}$ are similarly found and given by:

$$P_{ne}^{l,-} = P_{ne}^{l,+} \quad (35)$$

$$P_e^{l,-} = n_H n_P \cdot \{q_{P,H}^{ex} \cdot \frac{1}{2} \cdot I_6(H, P) + \alpha \cdot q_{P,H}^{ex} \cdot I_2(P, H) + q_{P,H}^{ex} \cdot I_1(P, H)\} - n_H n_P \cdot \{\alpha \cdot q_{P,H}^{ex} \cdot \frac{1}{2} \cdot [I_7(H, P) - I_9(H, P)]\} + n_O n_P \cdot \{q_{P,O}^{ex} \cdot \frac{1}{2} \cdot [I_7(P, O) - I_9(P, O)]\} - n_H n_{O^+} \cdot \{\alpha \cdot q_{H,O^+}^{ex} \cdot \frac{1}{2} \cdot [I_7(H, O^+) - I_9(H, O^+)]\}. \quad (36)$$

The expression Eq. (35) holds since the velocity distribution functions entering the calculations are isotropic Maxwellians, whereas $P_e^{l,-}$ is not equal to $P_e^{l,+}$ because of the anisotropy of the distribution functions.

The gain processes due to elastic collisions between hydrogen and oxygen can be formulated in the following way:

$$P_e^{N,\pm} = \frac{0.5}{t_{rel}} (n'_{H,e} - n_{H,e}). \quad (37)$$

Here $n'_{H,e}$ is the equilibrium density of escaping hydrogen atoms if the hydrogen temperature were equal to the thermospheric temperature T_O , and t_{rel} is the relaxation time necessary for hydrogen to attain the oxygen temperature if there were no escape. This time was defined by Fahr and Nass (1978) as

$$t_{rel} = \frac{1}{q_{H,O}^{el} n_O \varepsilon \sqrt{\frac{8kT_H}{\pi m_H}}}. \quad (38)$$

The actual density of escaping hydrogen atoms is denoted by $n_{H,e}$. The quantities $n_{H,e}$ and $n'_{H,e}$ are thus given by

$$n_{H,e} = n_H \frac{G_2 + G_3}{G_1 + G_2 + G_3}, \quad (39)$$

$$\begin{aligned} n'_{H,e} &= n_H \int_{v_{esc}}^{\infty} \int_0^{\pi} \int_0^{2\pi} M(v, T_O) d^3v \\ &= n_H \cdot \left[(1 - \text{erf}(\bar{v}_H)) + \frac{2}{\sqrt{\pi}} \bar{v}_H e^{-\bar{v}_H^2} \right], \end{aligned} \quad (40)$$

$$\bar{v}_H = \sqrt{\frac{\gamma M m_H}{k T_O r}}. \quad (41)$$

The factor 0.5 in Eq. (37) indicates that we only deal with the production into the upper/lower half of the velocity space. A corresponding expression for non-escaping H-atoms is obtained in analogy to Eq. (37)

$$P_{ne}^{N,\pm} = \frac{0.5}{t_{rel}} (n'_{H,ne} - n_{H,ne}). \quad (42)$$

Finally, we return to the term L_G of Eq. (19). As already mentioned, this term is a gravitationally caused loss term taking into account the fact that not all non-escaping particles leaving a lower level r' will finally reach r . This means L_G is simply the gradient of the upward-directed flux of non-escaping particles weighted by the probability that these particles do not suffer any collision between their origin and the upper level r , i.e.

$$L_G = \frac{\int_{r_c}^r \frac{d}{dr} \left\{ \left(\frac{r'}{r} \right)^2 n_H(r') v_{ne}^{eff,+}(r', r) \right\} w_{ne}(r', r) dr'}{\int_{r_c}^r w_{ne}(r', r) dr'}. \quad (43)$$

It should be mentioned here that we have a term similar to L_G in the continuity equation for downward-flying escaping particles. This term, however, is mainly due to the spherical geometry of the problem and accounts for those particles with perigees $r_p > r$.

Simplified concept for great heights

Above the 800-km level, collisions between hydrogen and the background gas components become less and less important. This is reflected by the fact that the value of $\alpha(r)$, which is a certain measure of the collision probability above r , is less than 0.2 (see next section). Therefore, the height levels are energetically decoupled from each other. In order to obtain an expression for the escape flux one only has to sum up the local gains into the escape regime. Due to the rare collisions above 800 km, we can also assume that the velocity distribution function for non-escaping hydrogen atoms will stay isotropic. Taking advantage of the fact that above 800 km only collisions between hydrogen and protonospheric protons need be considered, the fluxes are given by

$$\begin{aligned} J_e^+(r) &= \left(\frac{r_G}{r} \right)^2 J_e^+(r_G) w_e(r_G, r) \\ &+ \int_{r_G}^r \left(\frac{r'}{r} \right)^2 n_{H,ne} n_P \\ &\cdot \left\{ q_{P,H}^{ex} \cdot \frac{1}{2} \cdot I_6(H, P) \right\} w_e(r', r) dr' \\ &+ \int_{r_G}^r \left(\frac{r'}{r} \right)^2 n_{H,ne} n_P \\ &\cdot \left\{ q_{P,H}^{ex} \frac{1}{2} \sqrt{1 - \left(\frac{\bar{v}_e(r_G) r_G}{\bar{v}_e r'} \right)^2} I_6(H, P) \right\} \tilde{w}_e(r_p, r) dr', \end{aligned} \quad (44)$$

$$\begin{aligned} J_e^-(r) &= \int_r^{r_{pp}} \left(\frac{r'}{r} \right)^2 n_{H,ne} n_P \\ &\cdot \left\{ q_{P,H}^{ex} \frac{1}{2} \left[1 - \sqrt{1 - \left(\frac{\bar{v}_e(r) r}{\bar{v}_e r'} \right)^2} \right] I_6(H, P) \right\} w_e(r', r) dr', \end{aligned} \quad (45)$$

where r_{pp} is the height of the plasmopause. The subscript "ne" indicates that we deal only with the non-escaping part of the total hydrogen density, whereas collisions of escaping hydrogen atoms with protons are considered as ineffective since they probably reproduce an escaping H-atom. The third term of Eq. (44) takes into account downward-flying hydrogen particles produced at r' with a perigee $r_p \geq r_G$, which contribute to the escape flux at r (see Fig. 1). The terms given in curly brackets represent the product between charge-exchange cross-section and mean relative velocity. Since the term $\tilde{w}_e(r_p, r)$ differs slightly from the aforementioned collision probabilities, it is given here explicitly

$$\begin{aligned} \tilde{w}_e(r_p, r) &= \int_{\pi/2}^{\pi-\theta_g} \left\{ \exp \left[-2 \int_{r_p}^{r'} q_{P,H}^{ex} n_P \bar{v}_e^{el} dr'' / \right. \right. \\ &\quad \left. \left. \left(\bar{v}_e' \sqrt{1 - \left(\frac{r_p v_p}{r'' \bar{v}_e'} \right)^2} \right) \right] \right. \\ &\quad \cdot \exp \left[- \int_{r'}^r q_{P,H}^{ex} n_P \bar{v}_e^{el} dr'' / \right. \\ &\quad \left. \left. \left(\bar{v}_e' \sqrt{1 - \left(\frac{r' \bar{v}_e}{r'' \bar{v}_e'} \right)^2 \sin^2 \theta} \right) \right] \right\} \sin \theta d\theta, \end{aligned} \quad (46)$$

where \bar{v}_e is the mean velocity of escaping hydrogen atoms at r' but refers to the ion temperature, and \bar{v}_e' is

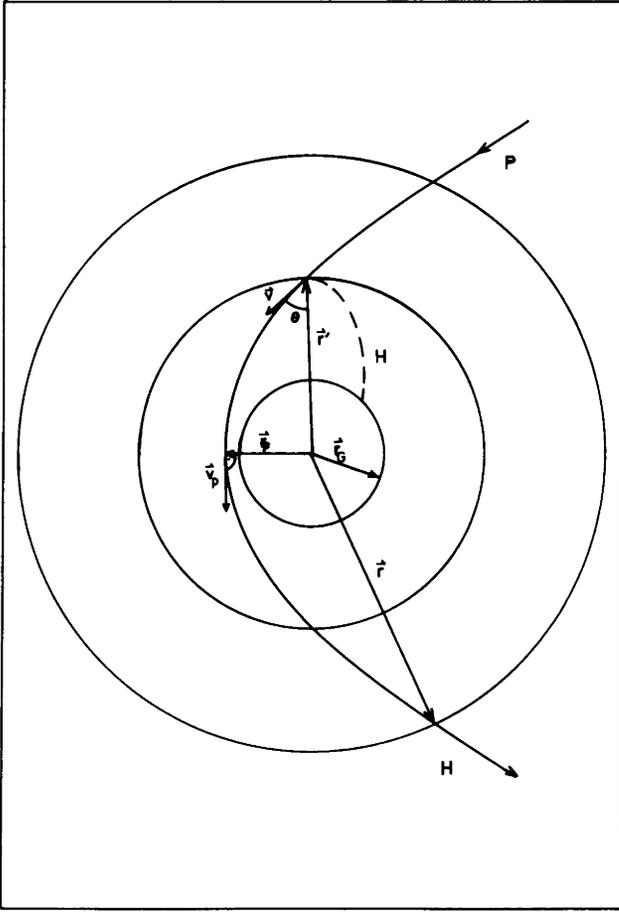


Fig. 1: A downward-flying hydrogen particle produced at r' with perigee $r_p \geq r_G$ has a certain probability of contributing to the escape flux at r

the velocity at r' . r_p and v_p are the height of the perigee and the velocity at this height, respectively (see Fig. 1).

Results for the terrestrial case

Input parameters and numerical procedure

In order to calculate the density and temperature distribution of hydrogen we need height profiles for the ion densities and the ion temperature as well as for the density profile of atomic oxygen. The latter can be obtained from a barometric law with temperature T_O modified by a factor first given by Chamberlain (1963), taking into account the reduction of the populated velocity space volume with increasing geocentric distance. Thus one obtains

$$n_O(r) = n_O^B(r) \cdot \left\{ 1 - \sqrt{1 - \left(\frac{r_c}{r}\right)^2} \exp\left(-\frac{\gamma m_O M}{k T_O r} \frac{r_c}{r_c + r}\right) \right\}, \quad (47)$$

where

$$n_O^B(r) = n_O(r_c) \exp\left[\frac{\gamma m_O M}{k T} \cdot \left(\frac{1}{r} - \frac{1}{r_c}\right)\right]. \quad (48)$$

The density value at the lower boundary of the transition region $n_O(r_c)$ is taken from the CIRA model (Jacchia, 1972). The ion data are obtained from the

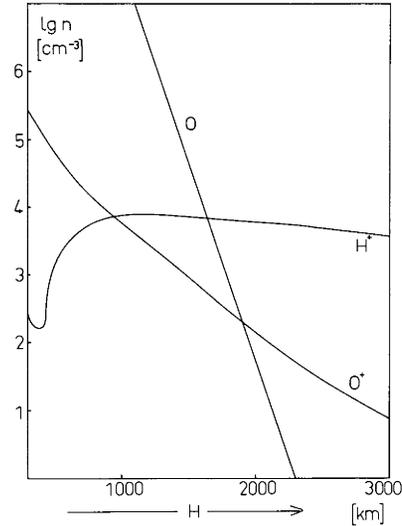


Fig. 2: The density profiles of oxygen O, oxygen ions O^+ and protons H^+ in a 1,000 K exosphere

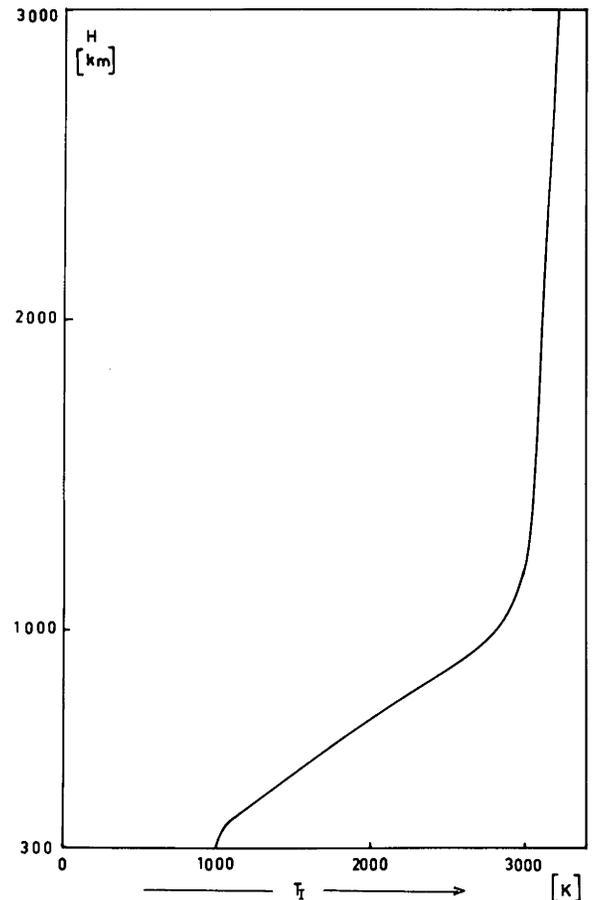


Fig. 3: Ion temperature distribution for an exospheric temperature of 1,000 K

“International Reference Ionosphere” (Rawer et al., 1978) valid up to 1,000 km. For greater heights we use a data representation in accordance with Raitt et al. (1975). Thus we obtain the density profiles and ion temperature distribution for an exospheric temperature of 1,000 K, as shown in Figs. 2 and 3.

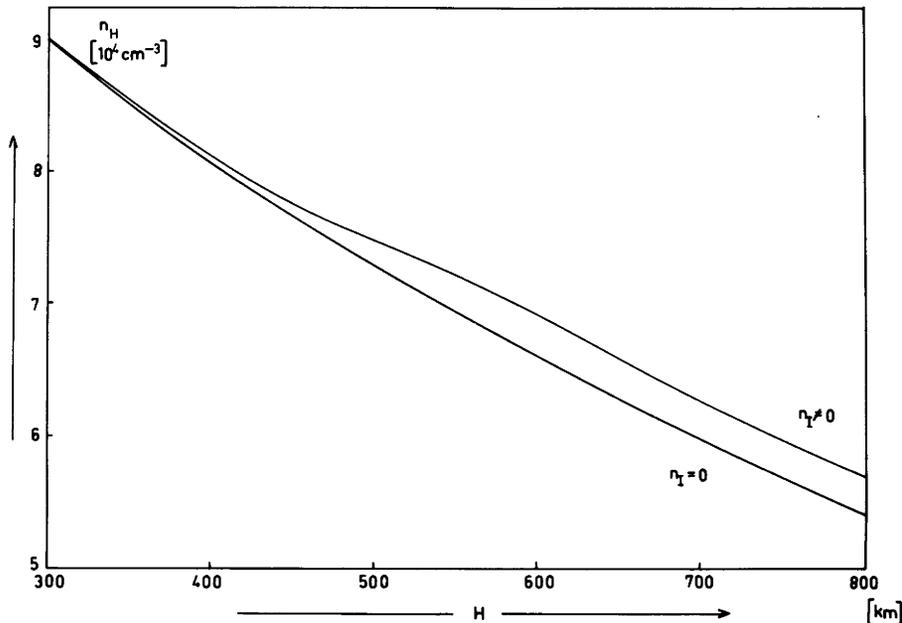


Fig. 4: The hydrogen density distribution. The *upper curve* ($n_1 \neq 0$) is obtained by taking into account, in addition to elastic collisions, collisions between neutral and ionized particles. The *lower curve* is obtained by neglecting the latter collisions ($n_1 = 0$).

To obtain the hydrogen temperature at each height level r we have to solve the heat conduction equation, Eq. (10). We solved this differential equation by using an extrapolation method (Bulirsch and Stoer, 1966), starting with the following boundary conditions:

$$T_H(r_c) = 1,000 \text{ K},$$

$$\left. \frac{dT_H}{dr} \right|_{r_c} = 0$$

and then integrating upwards step by step. The hydrogen density values entering these calculations can be obtained from the particle continuity equation, Eq. (18). In order to keep the integration error small, we choose a very narrow step width $\Delta r = r' - r$ of only 2 km. This small stepping also justifies our further expression, Eq. (25), for the mean radial effusion velocity of non-escaping particles. In this expression we obviously neglect all satellite particles with perigees between r and $r' = r - \Delta r$. But, of course, the number of these particles becomes negligible if we take this narrow step width. The solution of Eq. (18) at r requires a knowledge of downward-directed fluxes of escaping and non-escaping particles involving hydrogen densities and temperatures above r which are not yet available at this stage of calculation. In order to circumvent this difficulty, we start in a so-called zeroth iteration with a reduced particle continuity equation considering only upward-flying H-atoms:

$$\begin{aligned} \text{div } J &= \text{div } J_{ne}^+ + \text{div } J_e^+ \\ &= P_e^I + P_{ne}^I + P_e^N + P_{ne}^N + L_G. \end{aligned} \quad (49)$$

In this way we obtain a first-order solution for the hydrogen density and temperature distribution at each height level r up to the limiting height r_G of the transition layer. In a next iteration step – now beginning at r_G – we calculate n_H and T_H down to the height r_c , but now taking into account the whole particle continuity equation, Eq. (18). This iteration process is continued

until

$$\max_{r_c \leq r \leq r_G} \left| \frac{n_H^{i-1}(r) - n_H^i(r)}{n_H^i(r)} \right| = \text{const.}, \quad (50)$$

where $n_H^i(r)$ is the hydrogen density at height level r after the i 'th iteration. Fortunately, we need only 4–5 iterations to fulfill the condition (50).

Density and temperature distribution

The hydrogen density values calculated with the above-mentioned method are shown in Fig. 4. The upper curve is obtained by taking into account, in addition to elastic collisions, collisions between neutral and ionized particles; the lower curve is obtained by neglecting the latter collisions ($n_1 = 0$). The enhanced density values of the upper curve express the fact that collisions between protons and neutral oxygen atoms can produce neutral hydrogen atoms, whereas the counteracting process consisting of collisions between hydrogen atoms and oxygen ions is not very effective.

In order to illustrate the iteration process, the maximum relative differences of the calculated densities as a function of the iteration step are given in Table 1.

The relatively large differences after the second iteration step depends on the fact that, for the first time in both the upward- and downward-directed fluxes, hy-

Table 1

i	$\left \frac{n_H^{i-1} - n_H^i}{n_H^i} \right _{\max}$	
	$n_1 = 0$	$n_1 \neq 0$
0	—	—
1	1.2×10^{-3}	2.11×10^{-3}
2	1.3×10^{-2}	1.50×10^{-2}
3	3.0×10^{-4}	1.50×10^{-3}
4	-5.0×10^{-4}	-1.00×10^{-3}
5	3.0×10^{-4}	1.50×10^{-3}

drogen densities derived with the help of the whole particle continuity equation, Eq. (18), were used.

In Fig. 5 we show the resulting hydrogen temperature profiles, and, as expected, there is a temperature decrease with height. Due to the additional temperature coupling between hydrogen and the hot, ionized background gas components, this temperature decrease is less pronounced compared to that decrease caused by elastic collisions between hydrogen and oxygen atoms only.

The temperature reduction for the latter case ($n_I = 0$) obtained at the boundary of the transition region, i.e. $\Delta T_H(r_G) = 17$ K, can be compared with a value given previously by Fahr and Weidner (1977). They assume a linear temperature decrease over a certain height range Δr , so that the energy flux due to escape is balanced by a heat flux originating at lower heights.

$$\phi_e^{eff} = \kappa(r) \frac{T_O - T_H}{\Delta r}.$$

They obtained an estimate of the dimension Δr over which this temperature reduction takes place by calculating the average number of collisions needed to transfer the thermal energy difference $k \cdot \Delta T_H$ between oxygen and hydrogen to the hydrogen atoms. By using this simplified energy continuity equation they obtained a value of $\Delta T_H = 25$ K. Comparing this value to the one presented here, it can be deduced that their simple model overestimates the temperature reduction by a factor of 1.5, but that the expected effect was qualitatively well described.

In Fig. 6 we show the height dependence of the factor α needed for the escape branch of the hydrogen velocity distribution f_H . As mentioned above, a quantity β has to be calculated from the requirement of a continuous transition between the transition region and the higher atmospheric regions if charge-exchange collisions are considered ($n_I \neq 0$). This was done by comparing the downward-directed flux J_e^- calculated by Eq. (29) with the corresponding flux evaluated by Eq. (45). The two quantities become equal if

$$\beta = 0.15.$$

A value of two earth radii was taken for the height of the plasmopause in these calculations.

Above the 800 km level, $\alpha(r)$ can be calculated from the ratio J_e^-/J_e^+ . In Fig. 6 these values are also shown. As expected, $\alpha(r)$ decreases with height monotonically.

Escape flux

The total escape flux, i.e.

$$J_e^{tot}(r) = J_e^+(r) - J_e^-(r), \quad (51)$$

is shown in Figs. 7 and 8. The fluxes are calculated with the help of Eqs. (29), (44) and (45). The dashed-dotted line shows the result if collisions between neutral and ionized particles are also taken into account ($n_I \neq 0$), while the solid line gives the result only for elastic collisions taken into consideration ($n_I = 0$). It can be seen that only above the 1,000 km level does the production of escaping hydrogen atoms by charge-ex-

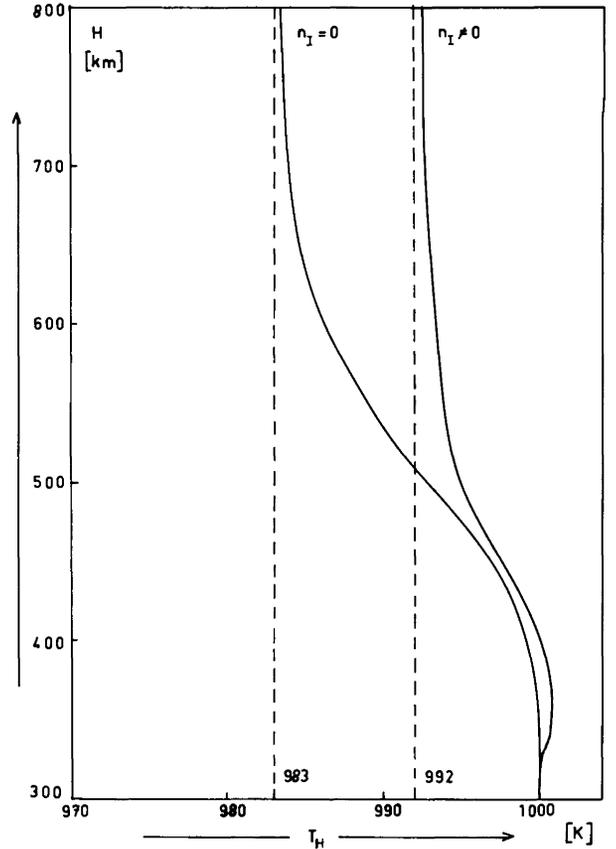


Fig. 5: The height profiles of the exospheric hydrogen temperature are shown. The labelling of the profiles has the same meaning as in Fig. 4

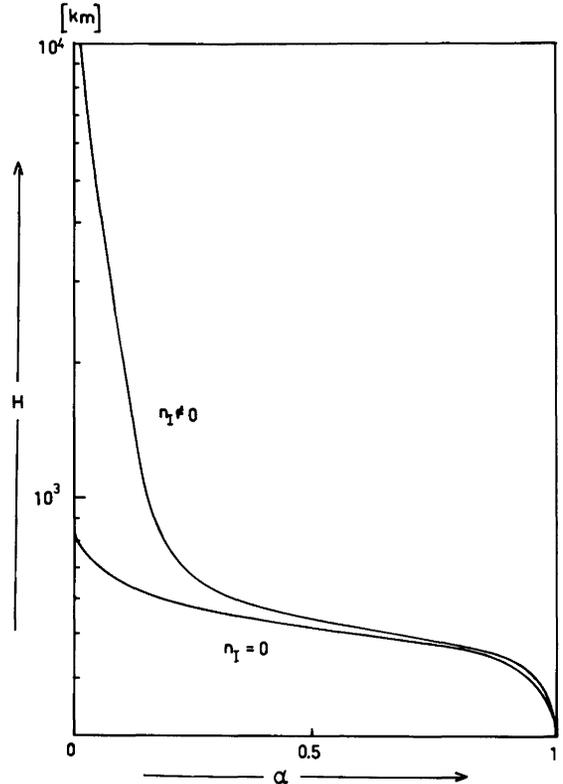


Fig. 6: Height profiles of the factor α needed for the escape branch of the hydrogen velocity distribution f_H [see Eq. (1)]. The labelling of the profiles has the same meaning as in Fig. 4

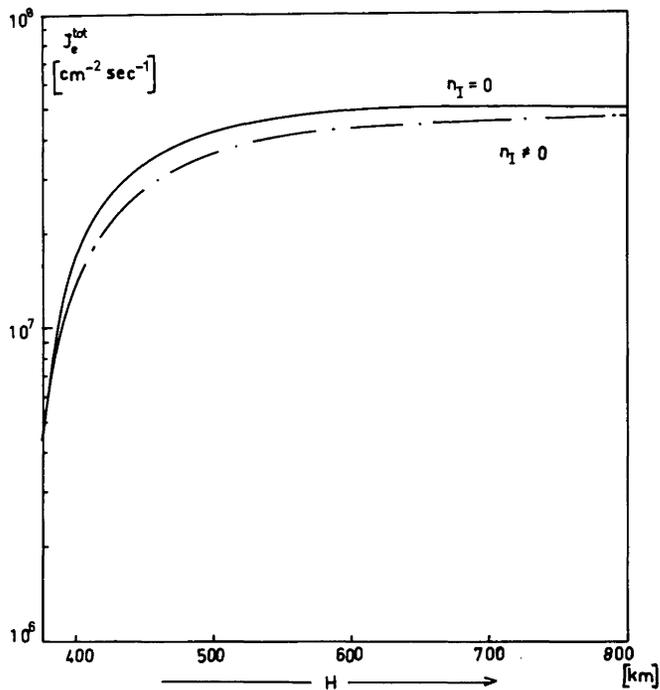


Fig. 7: The height dependence of the total escape flux is shown. The *dashed-dotted line* shows the result if collisions between neutral and ionized particles are taken into account ($n_I \neq 0$), while the *full line* gives the result only for elastic collisions taken into consideration ($n_I = 0$)

change processes lead to an enhancement of the total escape flux to values well above the pure thermal escape flux. Below this height the situation is somewhat reversed since charge-exchange collisions enhance both the upward-directed and the downward-directed escape flux. This means that the enhanced background gas density, due to the presence of the ions, partly impedes the escape flux of hydrogen atoms.

In Fig. 9 we show the height dependence of the upward-directed escape flux J_e^+ for both cases. Ob-

viously, the resulting flux given by the dashed-dotted line ($n_I \neq 0$) everywhere exceeds the flux given by the full line ($n_I = 0$). This is due to the fact that collisions between neutral and ionized particles lead to an enhancement of the local hydrogen density, especially the partial density in the escape branch, and therefore an enhanced escape flux J_e^+ results. At heights where no further production of escaping particles takes place, the decrease of J_e^+ must be proportional to $1/r^2$. This fact is indicated by the dashed line in Fig. 9.

An interesting result is found by comparing the pure thermal escape flux J_e^t with the total flux J_e^{tot} at the plasmopause

$$\frac{J_e^{tot}}{J_e^t} = 2.4. \quad (52)$$

This means that the ion-enhanced terrestrial escape flux is of the same order as the pure thermal one.

If the value of the total escape flux at the plasmopause, i.e. $1.7 \times 10^7 \text{ cm}^{-2} \text{ s}^{-1}$, were extrapolated downwards to a height of 500 km, one would arrive at

$$J_e^{tot}(500 \text{ km}) = 1.3 \times 10^8 \text{ cm}^{-2} \text{ s}^{-1}. \quad (53)$$

This value can be compared to $1.5 \times 10^8 \text{ cm}^{-2} \text{ s}^{-1}$ given by Hunten and Strobel (1974) and Liu and Donahue (1974a; b; c), who calculated the upward diffusive hydrogen flux at the turbopause level produced by dissociation processes of H-bearing molecular constituents at lower heights. As is obvious, both values are in good agreement, showing that our theory of planetary escape is able to account for the continuity of the hydrogen flux.

Critical review of the results

In this section we want to compare our results concerning ion-induced flux contributions to the terrestrial H-escape with results of earlier publications in this field.

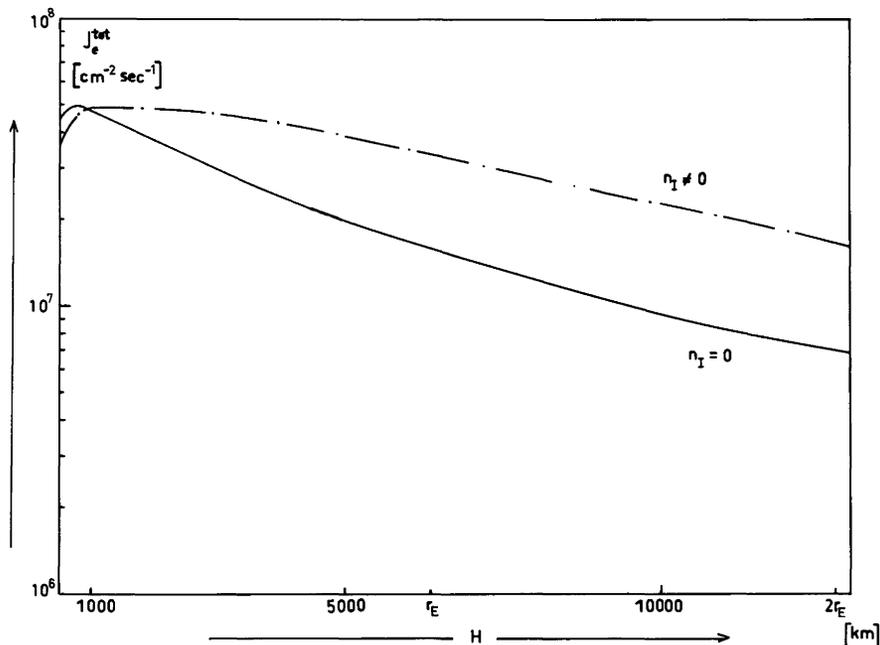


Fig. 8: As Fig. 7, but the height dependence is shown up to the plasmopause

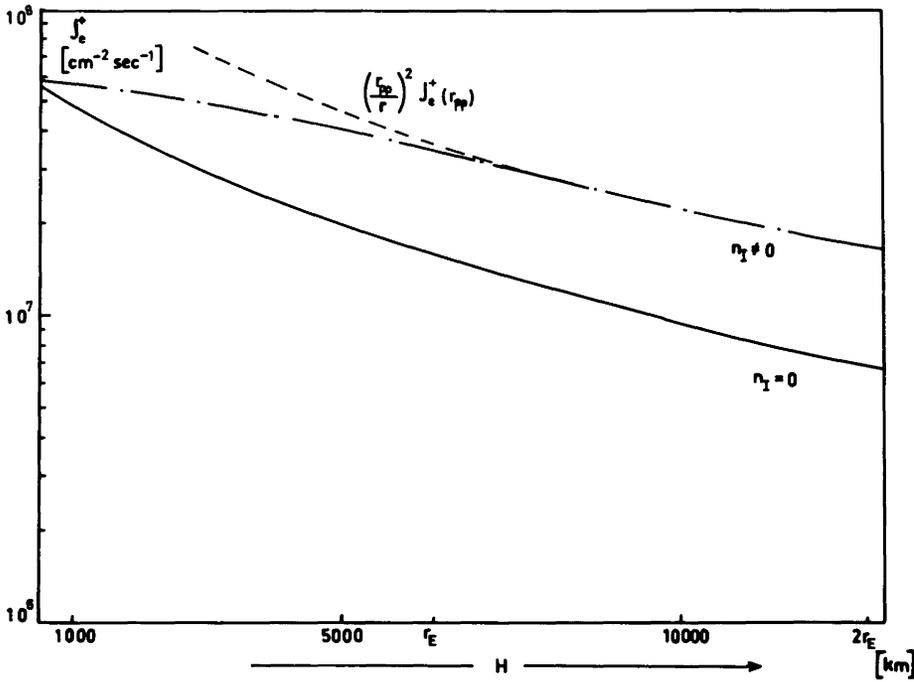


Fig. 9: The height dependence of the upward-directed escape flux is shown. The labelling of the profiles has the same meaning as in Fig. 7. The *dashed curve* is obtained if the value of the upward-directed escape flux were extrapolated downwards according to an inverse square law

As in all earlier contributions to this subject, it was demonstrated here that, due to the “charge-exchange contact” between neutral hydrogen and plasmaspheric ions ($T_1 \approx 2-4T_H$), the local H-density in the escape branch of f_H and hence the escape fluxes J_e are enhanced. However, as it turns out from our calculations the resulting enhancement is less pronounced than was claimed in earlier publications (Tinsley, 1973; Maher and Tinsley, 1977; Chamberlain, 1977; Maher, 1980; Maher and Tinsley, 1978; Hodges et al., 1981; Shizgal and Lindenfeld, 1982). A detailed analysis for this was given in a recent review by Fahr and Shizgal (1983).

Here we only want to mention the work of Shizgal and Lindenfeld (1982) which is the most suitable for a comparison with our results, with regard to ion-induced escape contributions to the terrestrial hydrogen escape. The latter authors use their former concept of local escape contributions (Lindenfeld and Shizgal, 1979) but expanded by inclusion of local charge-exchange collisions. Different height layers are considered as individually contributing to the total escape according to the local production rate of escaping H-atoms and the probability of an escape from this level to infinity is taken into account. After simplification of the collision process by the assumption of a constant charge-exchange cross-section and of Maxwell-Boltzmann distribution functions for the collision partner, the production term can be evaluated in a semi-analytic form.

The advantage of this approach clearly is its mathematical simplicity and clarity. This makes it very appropriate in studying the principal reaction of ion-induced escape fluxes to variations of the ion temperature T_1 and of the neutral temperature T_0 . However, it has to be kept in mind that this approach treats the ion-induced escape as de-coupled from the thermal one, i.e. no consistent treatment of local thermalization processes is included.

The main cause for their overestimation of the ion-induced escape fluxes is the assumption of constant ion

temperature T_1 and ion density n_1 over the entire region above the critical height r_c up to the plasmopause. In particular, the very large escape flux enhancements that result for low temperatures T_0 and high values of T_1/T_0 should be drastically reduced if a variation of T_1 with height, obeying $T_1 = T_0$ at r_c , were considered. In view of this, enhancement factors (of 1–3 over thermal escape) seem to be indicated by the results of Shizgal and Lindenfeld (1982), which are entirely compatible with our results.

In our approach, we have carefully taken into account the local H-particle productions into the escape category due to both elastic and inelastic collisions using a consistently determined distribution function f_H for the calculations of collisional productions and losses, i.e. the depopulation of the escape branch has also been taken consistently into account. The thermalization or relaxation of the local function f_H towards the local background distribution function was also taken into account using energy and particle flow continuities in the transition layer above r_c . The quantity that is not consistently determined in this approach is the ion distribution function taken to be a Maxwellian with the local ion temperature $T_1 = T_1(r)$. This is probably why even our results are still on the high side.

Appendix

In order to give expressions for the mean relative velocities, the following four integrals must be evaluated:

$$I_1(1, 2) = \int_{v_{esc}}^{\infty} \int_0^{\pi/2} f_1 d^3 v_1 \int_{v_{esc}}^{\infty} \int_0^{\pi/2} f_2 v_{rel} d^3 v_2, \quad (54)$$

$$I_2(1, 2) = \int_{v_{esc}}^{\infty} \int_0^{\pi/2} f_1 d^3 v_1 \int_{v_{esc}}^{\infty} \int_{\pi/2}^{\pi} f_2 v_{rel} d^3 v_2, \quad (55)$$

$$I_3(1, 2) = \int_0^{v_{esc}} \int_0^{\pi/2} f_1 d^3 v_1 \int_{v_{esc}}^{\infty} \int_0^{\pi/2} f_2 v_{rel} d^3 v_2, \quad (56)$$

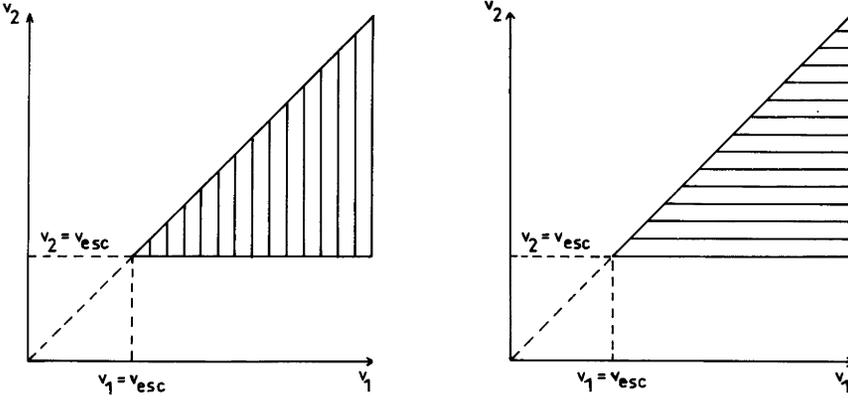


Fig. 10: The region of integration is shown for the double integral mentioned in the text

$$I_4(1, 2) = \int_0^{v_{esc}} \int_0^{\pi/2} f_1 d^3 v_1 \int_{v_{esc}}^{\infty} \int_0^{\pi/2} f_2 v_{rel} d^3 v_2, \quad (57)$$

where

$$d^3 v_i = v_i^2 \sin \theta_i d\theta_i d\psi_i dv_i,$$

$$f_i = \left(\frac{m_i}{2\pi k T_i} \right)^{3/2} \exp\left(-\frac{m_i v_i^2}{2k T_i}\right) \quad (i = 1, 2),$$

$$v_{rel} = \sqrt{v_1^2 + v_2^2 - 2v_1 v_2 \cos \theta_2}.$$

The integration over the angles can be carried out immediately by taking into account that

$$\int v_{rel} \sin \theta_2 d\theta_2 = \int \frac{v_{rel}^2}{v_1 \cdot v_2} dv_{rel}.$$

This yields for the integral (54):

$$\begin{aligned} & \int_{v_{esc}}^{\infty} \int_0^{\pi/2} f_1 d^3 v_1 \int_{v_{esc}}^{\infty} \int_0^{\pi/2} f_2 v_{rel} d^3 v_2 \\ &= 4\pi^2 \left(\frac{m_1}{2\pi k T_1} \right)^{3/2} \left(\frac{m_2}{2\pi k T_2} \right)^{3/2} \int_{v_{esc}}^{\infty} v_1^2 \exp\left(-\frac{m_1 v_1^2}{2k T_1}\right) dv_1 \\ & \cdot \left\{ \int_{v_{esc}}^{v_1} \left[\frac{-2(v_1 - v_2)^3}{6v_1 v_2} + \frac{2}{6v_1 v_2} (v_1^2 + v_2^2)^{3/2} \right] v_2^2 \right. \\ & \cdot \exp\left(-\frac{m_2 v_2^2}{2k T_2}\right) dv_2 \\ & \left. + \int_{v_1}^{\infty} \left[\frac{-2(v_2 - v_1)^3}{6v_1 v_2} + \frac{2}{6v_1 v_2} (v_1^2 + v_2^2)^{3/2} \right] v_2^2 \right. \\ & \cdot \exp\left(-\frac{m_2 v_2^2}{2k T_2}\right) dv_2 \left. \right\}. \end{aligned}$$

The first double integral can be rewritten after consideration of the region of integration. Here we first have to integrate over a vertical strip from $v_2 = v_{esc}$ to $v_2 = v_1$ and then sum up all these strips (see Fig. 10). The same region of integration is covered by integrating first over a horizontal strip from $v_1 = v_2$ to $v_1 = \infty$ and then sum up all these strips, i.e.

$$\int_{v_{esc}}^{\infty} \int_{v_{esc}}^{v_1} \dots dv_2 dv_1 = \int_{v_{esc}}^{\infty} \int_{v_{esc}}^{\infty} \dots dv_1 dv_2.$$

If we now introduce new variables of integration ($v_1 \rightarrow v_2$, $v_2 \rightarrow v_1$) in the above-mentioned integral, we

come to the following expression for integral (54)

$$\begin{aligned} & \int_{v_{esc}}^{\infty} \int_0^{\pi/2} f_1 d^3 v_1 \int_{v_{esc}}^{\infty} \int_0^{\pi/2} f_2 v_{rel} d^3 v_2 \\ &= 4\pi^2 \left(\frac{m_1}{2\pi k T_1} \right)^{3/2} \left(\frac{m_2}{2\pi k T_2} \right)^{3/2} \\ & \cdot \left\{ \int_{v_{esc}}^{\infty} v_1^2 \exp\left(-\frac{m_2 v_1^2}{2k T_2}\right) dv_1 \int_{v_1}^{\infty} \left[\frac{-2(v_2 - v_1)^3}{6v_1 v_2} \right. \right. \\ & \left. \left. + \frac{2}{6v_1 v_2} (v_1^2 + v_2^2)^{3/2} \right] v_2^2 \exp\left(-\frac{m_1 v_2^2}{2k T_1}\right) dv_2 \right. \\ & \left. + \int_{v_{esc}}^{\infty} v_1^2 \exp\left(-\frac{m_1 v_1^2}{2k T_1}\right) dv_1 \int_{v_1}^{\infty} \left[\frac{-2(v_2 - v_1)^3}{6v_1 v_2} \right. \right. \\ & \left. \left. + \frac{2}{6v_1 v_2} (v_1^2 + v_2^2)^{3/2} \right] v_2^2 \exp\left(-\frac{m_2 v_2^2}{2k T_2}\right) dv_2 \right\}. \end{aligned}$$

As can be seen by inspection of this expression, now only one double integral has to be evaluated. The result for the second integral is then gained from the result of the first one by changing m_1 into m_2 and T_1 into T_2 . The integration of the given double integral is straightforward and has not been given in detail. Furthermore, it can be shown that:

$$I_5(1, 2) = \int_{v_{esc}}^{\infty} \int_0^{\pi} f_1 d^3 v_1 \int_{v_{esc}}^{\infty} \int_0^{\pi} f_2 v_{rel} d^3 v_2 = 2 \cdot \{I_1(1, 2) + I_2(1, 2)\} \quad (58)$$

and

$$I_6(1, 2) = \int_0^{v_{esc}} \int_0^{\pi} f_1 d^3 v_1 \int_{v_{esc}}^{\infty} \int_0^{\pi} f_2 v_{rel} d^3 v_2 = 2 \cdot \{I_3(1, 2) + I_4(1, 2)\}. \quad (59)$$

In addition, Eq. (58) yields the well-known result ($v_{esc} \rightarrow 0$)

$$I_7(1, 2) = \int_0^{\infty} \int_0^{\pi} f_1 d^3 v_1 \int_0^{\infty} \int_0^{\pi} f_2 v_{rel} d^3 v_2 = \sqrt{\frac{8k}{\pi} \left(\frac{T_1}{m_1} + \frac{T_2}{m_2} \right)} \quad (60)$$

and with the help of Eqs. (58), (59) and (60) we yield

$$I_8(1, 2) = \int_0^{v_{esc}} \int_0^\pi f_1 d^3 v_1 \int_0^{v_{esc}} \int_0^\pi f_2 v_{rel} d^3 v_2$$

$$= \{I_7(1, 2) - (I_5(1, 2) + I_6(1, 2))\} - I_6(2, 1) \quad (61)$$

and

$$I_9(1, 2) = \int_0^{v_{esc}} \int_0^\pi f_1 d^3 v_1 \int_0^\infty \int_0^\pi f_2 v_{rel} d^3 v_2$$

$$= I_8(1, 2) + I_6(1, 2). \quad (62)$$

References

- Bertaux, J.L.: Observed variations of the exospheric hydrogen density with the exospheric temperature. *J. Geophys. Res.* **80**, 639, 1975
- Brown, R.A., Schneider, N.M.: Sodium remote from Io. *Icarus* **48**, 519, 1981
- Bulirsch, R., Stoer, J.: Numerical treatment of ordinary differential equations by extrapolation methods. *Numer. Math.* **8**, 1, 1966
- Chamberlain, J.W.: Excitation in nebulae: charge transfer and the Cassiopeia radio source. *Ap. J.* **124**, 390, 1956
- Chamberlain, J.W.: Planetary coronae and atmospheric evaporation. *Planet. Space Sci.* **11**, 901, 1963
- Chamberlain, J.W.: Charge exchange in a planetary corona: Its effect on the distribution and escape of hydrogen. *J. Geophys. Res.* **82**, 1, 1977
- Cloutier, P.A., Daniell, J.R., R.E., Dessler, A.J., Hill, T.W.: A cometary ionosphere model for Io. *Astrophys. Space Sci.* **55**, 93, 1978
- Durrance, S.T., Thomas, G.F.: The 63 μm radiation field in the earth's thermosphere and its influence on the atomic hydrogen temperature. *Planet. Space Sci.* **27**, 795, 1979
- Eviatar, A., Siscoe, G.L., Mekler, Y.: Temperature anisotropy of the jovian sulfur nebula. *Icarus* **39**, 450, 1979
- Fahr, H.J., Nass, H.U.: Concerning the structure of the transition layer between the terrestrial thermosphere and exosphere. *Ann. Geophys.* **34**, 219, 1978
- Fahr, H.J., Shizgal, B.: Modern exospheric theories and their observational relevance. *Rev. Geophys. Space Sci.* **21**, 75, 1983
- Fahr, H.J., Weidner, B.: Gas evaporation from collision determined planetary exospheres. *Mon. Not. R. Astr. Soc.* **180**, 593, 1977
- Gross, P.G.: A lower limit to Jeans' escape rate. *Mon. Not. R. Astron. Soc.* **167**, 215, 1974
- Goertz, C.K.: Io's interaction with the plasma torus. *J. Geophys. Res.* **85**, 2949, 1980
- Goertz, C.K., Ip, W.-H.: On the structure of the Io torus. *Planet. Space Sci.* **30**, 855, 1982
- Hodges, R.R., Rhorbough, R.P., Tinsley, B.A.: The effect of the charge exchange source on the velocity and temperature distributions and their anisotropies in Earth's exosphere. *J. Geophys. Res.* **86**, 6917, 1981
- Hunten, D.M., Strobel, D.F.: Production and escape of terrestrial hydrogen. *J. Atmos. Sci.* **31**, 305, 1974
- Ip, W.-H., Axford, W.I.: A weak interaction model for Io and the jovian magnetosphere. *Nature* **283**, 180, 1980
- Ip, W.-H.: On charge exchange and knock-on processes in the exosphere of Io. *Astrophys. J.* **262**, 780, 1982
- Jacchia, L.G.: *Cospar international reference atmosphere*. Akademie Verlag, Berlin 1972
- Jeans, J.H.: *An introduction to the kinetic theory of gases*. Cambridge: University Press 1925
- Lindenfeld, M.J., Shizgal, B.: Non-Maxwellian effects associated with the thermal escape of a planetary atmosphere. *Planet. Space Sci.* **27**, 739, 1979
- Liu, S.C., Donahue, T.M.: The aeronomy of hydrogen in the atmosphere of earth. *J. Atmos. Sci.* **31**, 1118, 1974a
- Liu, S.C., Donahue, T.M.: Mesospheric hydrogen related to exospheric escape mechanisms. *J. Atmos. Sci.* **31**, 1466, 1974b
- Liu, S.C., Donahue, T.M.: Realistic model of hydrogen constituents in the lower atmosphere and escape flux from the upper atmosphere. *J. Atmos. Sci.* **31**, 2238, 1974c
- Liwshitz, M., Singer, S.F.: Thermal escape of neutral hydrogen and its distribution. *Planet. Space Sci.* **14**, 541, 1966
- Maher, L.J.: The latitudinal variation of the charge exchange induced atomic hydrogen escape flux. *J. Geophys. Res.* **85**, 4621, 1980
- Maher, L.J., Tinsley, B.A.: Atomic hydrogen escape rate due to charge exchange with hot plasmaspheric ions. *J. Geophys. Res.* **82**, 689, 1977
- Maher, L.J., Tinsley, B.A.: The diurnal and solar cycle variation of the charge exchange induced hydrogen escape flux. *Planet. Space Sci.* **26**, 855, 1978
- Mapleton, R.A.: *Theory of charge exchange*. New York: Wiley-Interscience 1972
- Raitt, W.J., Schunk, R.W., Banks, P.M.: A comparison of the temperature and density structure in high and low speed thermal proton flows. *Planet. Space Sci.* **23**, 1103, 1975
- Rawer, K., Ramakrishnan, S., Bilitza, D.: *International reference ionosphere*. International Union of Radio Science (URSI), Brüssel, 1978
- Rapp, D.: Accidentally resonant asymmetric charge exchange in the protonosphere. *J. Geophys. Res.* **68**, 1773, 1963
- Shizgal, B., Lindenfeld, J.M.: A simple kinetic theory calculation of terrestrial atomic hydrogen escape fluxes induced by charge exchange collisions. *J. Geophys. Res.* **87**, 853, 1982
- Storm, D.A.: *The calculation of cross sections in various ion-atom collisions*. Ph. D. dissertation. Polytech. Inst. of Brooklyn, 1970
- Tinsley, B.A.: The diurnal variation of atomic hydrogen. *Planet. Space Sci.* **21**, 686, 1973
- Torr, M.R., Walker, J.C.G., Torr, D.G.: Escape of fast oxygen from the atmosphere during geomagnetic storms. *J. Geophys. Res.* **79**, 5267, 1974
- Torr, M.R., Torr, D.G.: Energetic oxygen: A direct coupling mechanism between the magnetosphere and thermosphere. *Geophys. Res. Lett.* **6**, 700, 1979
- Vidal-Madjar, A.: The Earth hydrogen exobase near a solar minimum. *Geophys. Res. Lett.* **5**, 29, 1978
- Vidal-Madjar, A., Thomas, G.E.: The terrestrial hydrogen problem. *Planet. Space Sci.* **26**, 863, 1978

Received: July 20, 1984; Accepted September 12, 1984

Increased ionospheric absorption connected with Pc 1 pulsations after geomagnetic storms

F. Märcz

Geodetic and Geophysical Research Institute of the Hungarian Academy of Sciences, H-9401 Sopron POB 5, Hungary

Abstract. Short-period magnetic pulsations (Pc 1 and IPDP) recorded during IMS at high latitudes in Finland have been studied around selected geomagnetic storms. As at lower latitudes, Pc 1 pulsations at five Finnish stations are generally missing during the main phase of the storm. However, they frequently occur some days later. Both statistical analyses (by the superposed epoch technique) and the review of individual events have shown that the after-effect in mid-latitude ionospheric absorption is generally significant following storms associated with clear Pc 1 after-effects at each high-latitude station. Using several parameters as indicators, the sequence of processes of importance to the after-effect in ionospheric absorption can be traced in the individual cases presented. The electron precipitation into the lower ionosphere (leading to the enhancement of ionospheric absorption) is usually attributed to resonant pitch angle scattering. The efficiency of this loss process can certainly depend on the occurrence of plasmaspheric whistler mode turbulence which is generally enhanced in the post-storm period. It is also suggested that periodic VLF emissions are involved in the processes generating Pc 1 pulsations. Ground-based Pc 1 observations can help to identify the probable appearance of VLF and/or ELF waves in the plasmasphere.

Key words: Geomagnetic activity – Short-period pulsations – Plasmaspheric waves – Electron precipitation – Lower ionosphere – Absorption after-effect

Introduction

Occurrences of short-period magnetic pulsations (Pc 1 and IPDP) recorded during IMS years at five high-latitude stations (Kevo, Sodankylä, Oulu, Jyväskylä and Nurmijärvi) in Finland have been published in the form of quick-look tables (Pikkarainen et al. 1982). A recent study (Märcz et al. 1984) has reviewed the latitude dependence of these types of pulsations and their association with certain magnetic and ionospheric parameters. Previously, a clear relation of the after-effect in ionospheric absorption to pearl-type (Pc 1) pulsations was shown on the basis of middle latitude data (Märcz and Verö 1977). The above-mentioned quick-look tables made possible the investigation of the dependence of high-latitude Pc 1 and IPDP events on geomagnetic disturbances. The main purpose of this paper is to confirm the connection of the after-effect in mid-latitude

ionospheric absorption with Pc 1 pulsations on the basis of pulsation observations along a meridional chain of high-latitude stations. Additionally, it will be shown that the post-storm increase of Pc 1 activity can be regarded as an indicator of the plasmaspheric turbulence and wave-particle interaction generally enhanced during the recovery phase of the magnetospheric storm. This is a signature obtained from ground-based records and might be useful, especially in the case of missing satellite observations. Results derived from statistical analyses and individual case studies are presented.

Data and methods of investigation

The data published in the quick-look tables contain Pc 1 and IPDP events between October 1976 and December 1979 for each station as a function of time (UT). The two types of pulsations have not been separated from each other in the tables. Nevertheless, Pc 1s occur mostly during the morning-day hours and IPDPs during the afternoon-evening hours. In order to yield an appropriate activity measure for both types, the individual occurrences were summarized for each day separately from 0000 to 1200 UT (a.m.) and from 1200 to 2400 UT (p.m.). For a.m. (Pc 1) pulsations a further separation was carried out between 0000 and 0600 UT and between 0600 and 1200 UT respectively, allowing independent analyses with night-time and day-time data.

On the basis of the daily sums of three-hourly Kp -indices (ΣKp), days of geomagnetic disturbances were selected and used as key days for superposed epoch analyses. The two combined criteria applied for the selection are as follows: on key days ΣKp must be ≥ 30 and in addition, on two out of the three days preceding the key day the actual ΣKp value must not exceed the long-term average determined for the corresponding year. These criteria resulted in the selection of disturbances reaching distinct peak values which were also well separated from each other. Altogether, 46 key days were found for the superposed epoch analyses.

The duration of a.m. and p.m. pulsations has been studied in two independent analyses covering the interval between the third day preceding and the tenth day following the key day. Furthermore, ionospheric absorption has been analysed by separating also the cases of storms with clear Pc 1 events and those without them. The ionospheric night absorption data determined at mid-latitude in Kühlungsborn (GDR) by the A3 method (at 245 kHz) are used (HHI

Geophys. Data, 1976–1979). In the detailed study of an individual event the Průhonice absorption data (at 185 kHz) are also presented for comparison (Ionospheric Data, Praha, 1978).

Results

The occurrence of a.m. (Pc 1) and p.m. (IPDP) pulsations on days around geomagnetic disturbances

The changes in the summarized durations around the 46 key days are shown in Fig. 1 for both a.m. (left) and p.m. (right) pulsations observed at five high-latitude stations. The features of the variations are rather different in the two sections of Fig. 1 hinting at individual responses to the geomagnetic disturbances. This confirms that two, really quite independent types of pulsations have been separated on the basis of their daily occurrence.

The a.m. events actually represent the Pc 1s; this is also indicated by their clear increase for several days following the geomagnetic disturbances (after-effect), as found earlier for the middle latitudes (März and Verö 1977). Another feature indicated by the study of März and Verö (1977), viz. Pc 1 activity is lowest at the time of peak geomagnetic activity (around the 0 day), seems to be common at high latitudes too.

The p.m. occurrences are presented on the right, of Fig. 1. There is only one characteristic change in dependence on increased geomagnetic activity. Short-period p.m. occurrences, accepted as IPDP events, are intensified by the geomagnetic disturbances at certain stations and are generally more frequent in the post-storm interval than before the storm.

Dividing the a.m. occurrences into two groups from 0000 to 0600 UT and from 0600 to 1200 UT, the night-time and day-time Pc 1 occurrences can be analysed separately. This has been done in Fig. 2 (left and right) where the variations are generally similar. However, certain peculiarities also appear. At each station, the night-time Pc 1 activity (left) is quite low on days close to the storm (0 day) and a rather broad occurrence maximum can be seen between the days +5 and +9. The latter proves the uniform appearance of the after-effect in night-time Pc 1s at high latitudes.

For day-time Pc 1s (Fig. 2, right), the decrease of the pre-storm activity level towards the storm-day is generally slight. However, it is quite important at the lowest latitude (Nurmijärvi). The occurrence maximum is expressed by a distinct peak on the fifth day following the key day, at all stations except Kevo. In spite of the exceptions mentioned, the after-effect is clearly present in the day-time Pc 1s too. Finally, a comparison of the two sections in Fig. 2 confirms that Pc 1 activity is generally higher during day-time hours (06–12) than at night (00–06), especially at the three stations situated at higher latitudes.

After-effect in ionospheric absorption associated with the appearance of Pc 1 events at high latitudes

Superposed epoch analyses. In Fig. 3 (top), mean values of ΣKp indicate the variation of geomagnetic activity around the 46 days selected as key days for the previous analyses. The horizontal dashed line represents the averaged activity determined for the period between October 1976 and De-

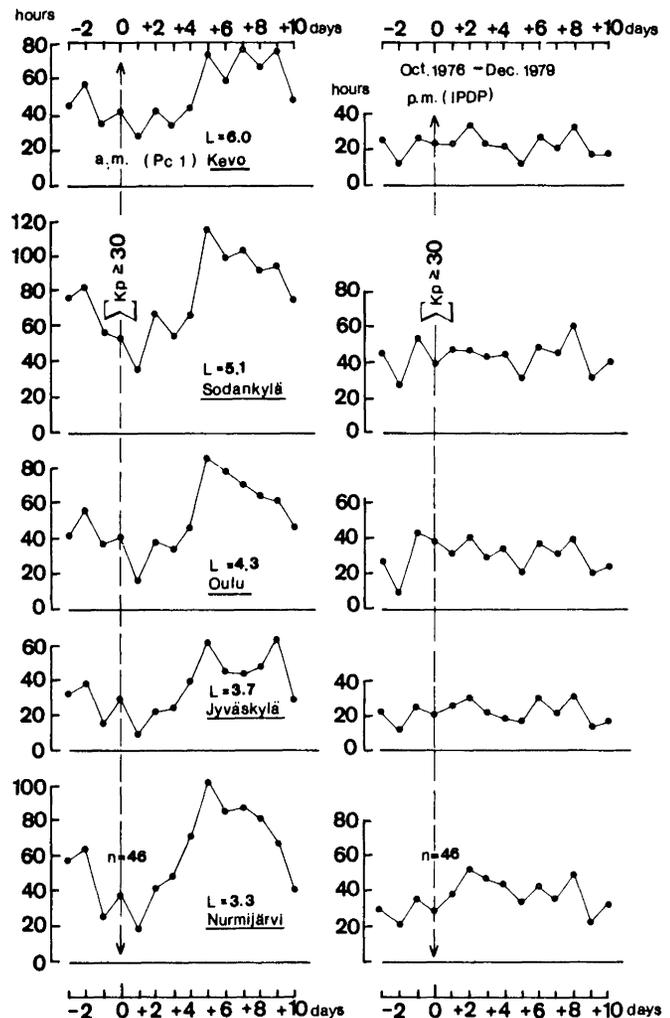


Fig. 1. Changes in the total duration of short-period magnetic pulsations observed at five stations in Finland around the 46 selected geomagnetic storms, between October 1976 and December 1979. *Left:* for a.m. pulsations accepted as Pc 1s. *Right:* for p.m. pulsations accepted as IPDPs

ember 1979. It is evident that geomagnetic activity approaches its average level soon after the disturbances. Consequently, the enhanced Pc 1 activity following the storm (Figs. 1 and 2) can be regarded as an after-effect.

Figure 3 (bottom) shows the mean departure of ionospheric night absorption (determined at 245 kHz at Kühlungsborn, GDR) from the corresponding monthly medians, around the selected 46 geomagnetic disturbances. The highly increased ionospheric absorption on 0 day can be identified as the primary storm effect. There is also a smaller, but rather prolonged absorption enhancement in the interval when the geomagnetic activity has returned to its normal level. This additional increase can be interpreted as an after-effect. On the basis of the error bars indicated in Fig. 3, the significance of the primary storm effect is apparent, but this can hardly be said of the after-effect.

It is known that not every geomagnetic disturbance is followed by enhanced ionospheric absorption (e.g. Lauter and Knuth, 1967). Nevertheless, it was shown that, in the case of a clear after-effect in Pc 1 pulsations, a post-storm enhancement is also most likely to occur in ionospheric absorption (März and Verö, 1977). This connection was

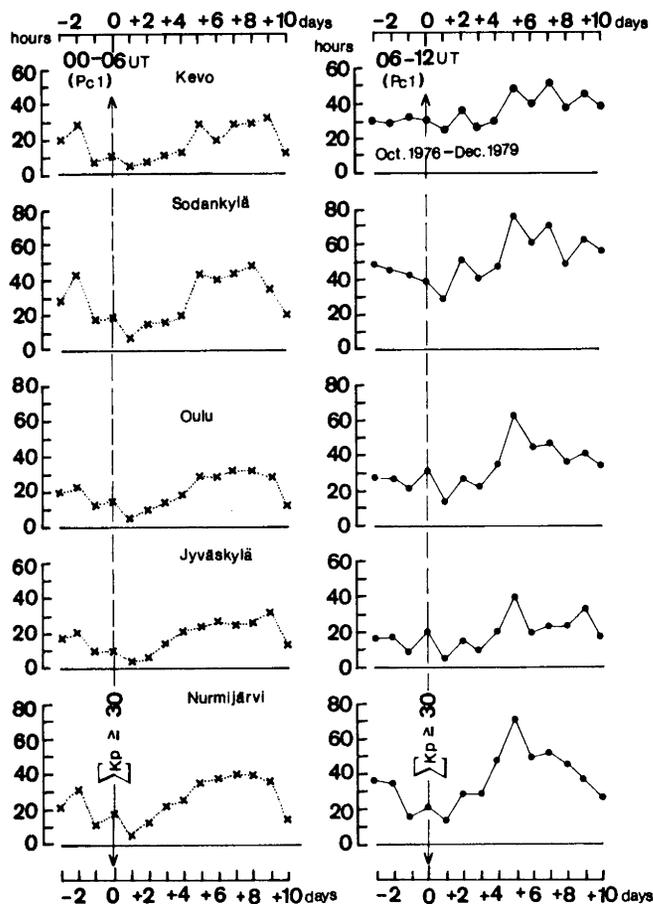


Fig. 2. Changes in the total duration of night-time Pc 1s (left) and day-time Pc 1s (right) around the same storms as in Fig. 1

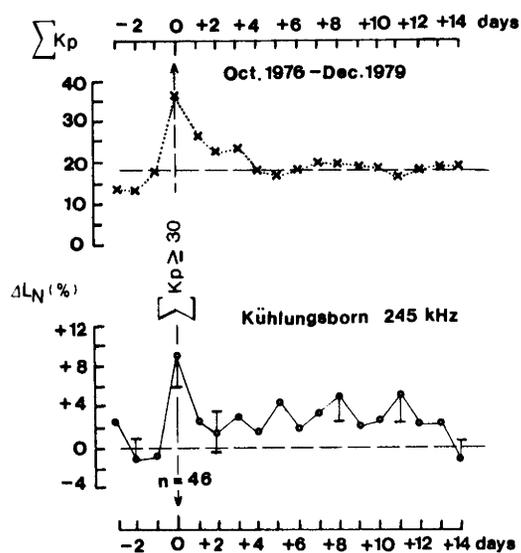


Fig. 3. Top: Variations in geomagnetic activity indicated by mean values of the daily K_p -sums around the 46 selected storms. (The horizontal dashed line indicates the average activity determined for the investigated interval.) Bottom: Mean departure of ionospheric night absorption from corresponding monthly medians (ΔL_N given in percent) around the same storms as in the top figure

revealed by means of mid-latitude pulsation observations (at Irkutsk and Nagycenk). Here, Pc 1 pulsation data from five high-latitude stations are used for proving the connection on an extended data base and for introducing some new aspects.

The occurrence of an after-effect in the pearl-type pulsations has been checked by determining a suitable ratio (R) of the Pc 1 durations following and preceding the geomagnetic storm. (In each individual case, the summarized daily durations of the interval between days +5 and +9 were divided by the summarized and appropriately weighted durations of the pre-storm interval between days -3 and -1.) Events with $R > 1$ can be accepted as after-effects. However, for the analysis of rather clearer effects an even stronger criterion ($R \geq 1.7$) was used. Two independent ratios were determined for each selected storm, one from the night-time (00-06 UT) and another from the day-time (06-12 UT) Pc 1 occurrences. Data from all the stations mentioned were treated in this way. Consequently, both the day-time and the night-time appearance (or lack) of the Pc 1 after-effect could be traced in a wide latitude range and, on this basis, different kinds of events might be distinguished.

Two special groups of events were chosen for testing the relationship between the high-latitude Pc 1 occurrences and the mid-latitude ionospheric absorption enhancements. One group consisted of events when the selected geomagnetic disturbances were followed by Pc 1 after-effects at each of the five stations. The other group of events served as a control. For these geomagnetic storms, the Pc 1 after-effects completely failed to appear at every high-latitude station. Considering both kinds of events, ionospheric night absorption determined at mid-latitude was analysed by the superposed epoch method and the results are presented in Fig. 4.

The night occurrences (00-06 UT) of the Pc 1s indicated clear after-effects for ten storms at all stations. (The mean value determined from the individual ratios of the five stations was higher than 1.7 in each case.) The corresponding increase in the ionospheric night absorption is quite significant (Fig. 4b). For the eight storms chosen without a Pc 1 after-effect at all five stations, no regular absorption variation appears (Fig. 4a). On the basis of day-time (06-12 UT) pulsations, again ten storms were selected when the Pc 1 after-effects occurred at all the stations. (The dates of the storm-days were only partly the same as those of the previous selection on the basis of night-time Pc 1s. However, the mean value of the ratios was at least 1.7 for these cases too.) The corresponding after-effect for ionospheric absorption shown in Fig. 4d is also significant, though its amplitude is somewhat lower than that in Fig. 4b. Storms without any after-effect in day-time Pc 1s (for all the stations) are rather rare and only four events of this kind were selected out of the 46 storms studied. In Fig. 4c the changes in ionospheric absorption for these four events are quite irregular, partly due to the small data set. For the evaluation of the results, error bars are given at main maxima and local minima approaching the average level (Fig. 4b and d). The error bars drawn at the horizontal dashed line (representing the average absorption level) have been determined from data of the pre-storm interval between days -3 and -1 (Fig. 4a-d).

For further analysis, all geomagnetic storms followed by after-effects in Pc 1 pulsations (i.e. the required criterion was fulfilled at each station) were collected into a common

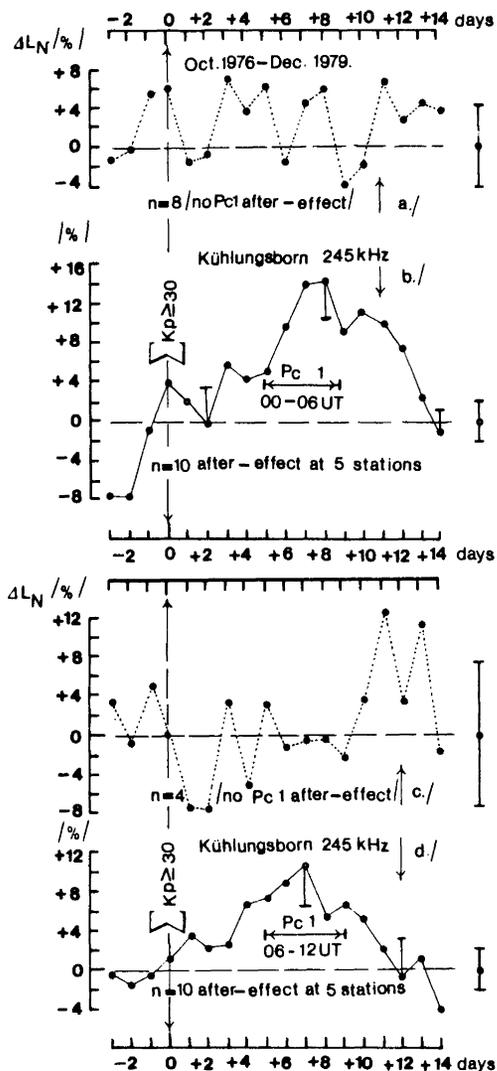


Fig. 4a-d. Mean departure of ionospheric night absorption from corresponding monthly medians for different storm events selected on the basis of Pc 1 pulsations: the after-effect in night-time Pc 1 pulsations a is missing b is clear and at each investigated station; the after-effect in day-time Pc 1 pulsations c is missing d is clear and at each investigated station

group, whether the after-effect occurred in the night-time (00–06 UT) or the day-time (06–12) pulsations. Altogether 16 events were selected from the original set (46). The changes in ionospheric night absorption around these storms are shown in Fig. 5b. The primary storm effect on the days 0 and +1 is followed by a long-lasting after-effect with a peak on day +7. The enhancements are rather significant, as indicated by the error bars. For comparison, the variations of ionospheric absorption for the total number (46) of geomagnetic storms (Fig. 5a) and for the rest of the events (30) without a clear Pc 1 after-effect (Fig. 5c) are also shown. In both cases, there is an important primary storm effect, but the after-effect is not significant. This indirectly confirms the importance of geomagnetic storms followed by strong Pc 1 activity with regard to the absorption after-effect.

Study of individual events. Figure 6 has been compiled to illustrate, by means of a case study, the following: the absorption after-effect on the basis of A3 measurements at

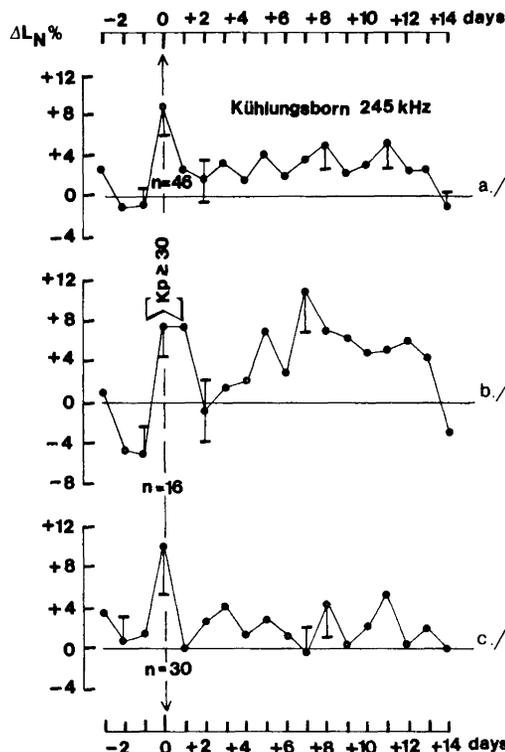


Fig. 5a-c. Mean departure of ionospheric night absorption from corresponding monthly medians for different storms: a for the 46 selected storms (as in Fig. 3, bottom) b for storms followed by clear after-effects in Pc 1 pulsations at all stations (without regard to the time of Pc 1 occurrences) c for the rest of the storms

two mid-latitude propagation paths (with different frequencies), the occurrences of Pc 1 pulsations at five high-latitude stations and the changes in geomagnetic activity indicated by ΣKp indices, as well as the equatorial Dst -variations. The Dst -index can be regarded as a suitable indicator of the hot plasma injections in the main phase and during the recovery phase of the magnetospheric storm (Lauter et al. 1977).

From a detailed analysis of the magnetospheric storm event in September/October 1978, it is clear that the highly increased geomagnetic activity around the 0 day (29 September) is not favourable for the appearance of Pc 1 pulsations. At the same time, the large positive departure ($>40\%$) in ionospheric absorption from the corresponding monthly median, determined for Kühlungsborn (245 kHz), indicates a prominent primary storm effect, but the latter is quite small at a lower latitude (Průhonice, 185 kHz). After one day the Dst -field recovers strongly from about -200 nT and, at the start of its slow recovery phase, a peak of enhanced absorption appears at Kühlungsborn as a signature of the after-effect. On 4 October, the recovery phase is interrupted by a minor decrease of Dst , thus a new particle injection into the plasmasphere could occur. This is associated with increased Pc 1 activity at each station which can be regarded as a signature of enhanced plasmaspheric turbulence including VLF and/or ELF wave activity. Presumably, these magnetospheric conditions were favourable for effective wave-particle interactions and particle precipitation. As a consequence, ionospheric absorption increased once again at Kühlungsborn and reached its peak value at Průhonice. Similar processes can also be traced

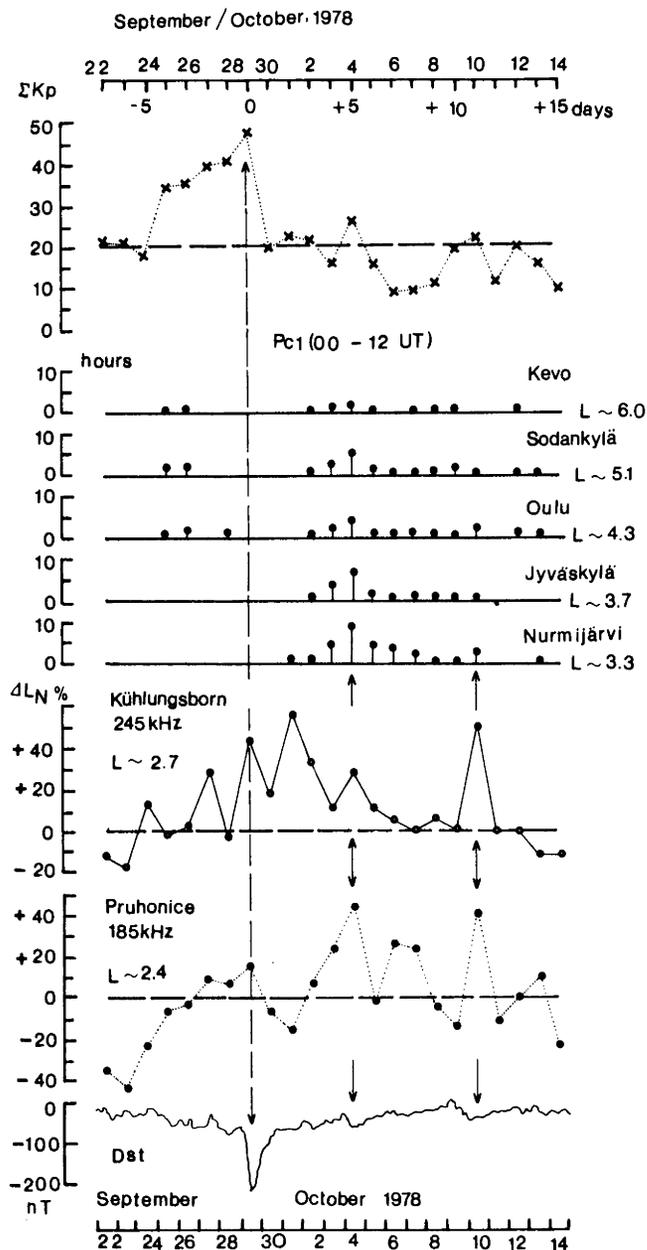


Fig. 6. Variations in several parameters during the September/October, 1978 event. (The parameters shown from top to bottom are the daily ΣKp -values, the Pc 1 durations at five high-latitude stations, the ionospheric night absorption at two mid-latitude stations and the Dst -index)

on 10 October, the enhancement in Pc 1 activity being quite moderate, however.

In Fig. 7, the January 1978 event shows, in its initial section, features essentially similar to those in Fig. 6, with the exception of a rather slight primary storm effect in ionospheric absorption. The minor re-enhancement of the Dst -field on 10 January should be associated with a particle injection and the simultaneous maximum of increased ionospheric absorption hints at the efficiency of the particle precipitation in spite of the somewhat declining Pc 1 activity. From 12 January the latter becomes reinforced while the Dst -field shows only a slight change, thus an enhanced plasmaspheric wave activity could be responsible for a new particle precipitation indicated by the repeated increase in

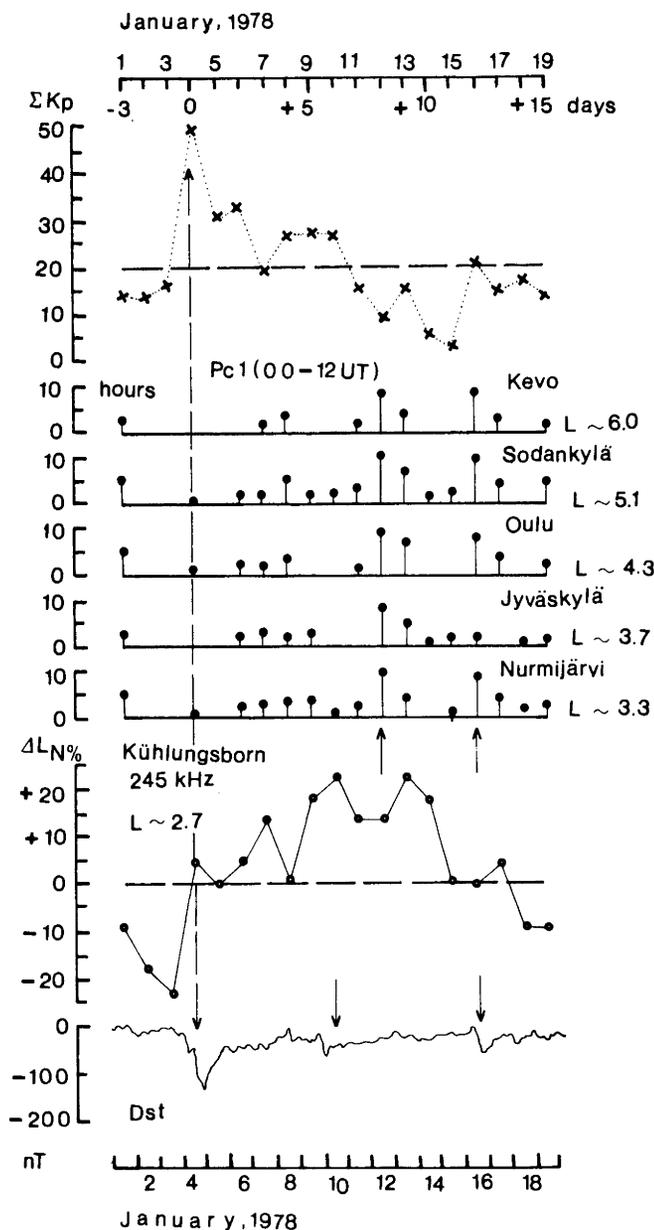


Fig. 7. As Fig. 6 but for the January 1978 event. (Ionospheric absorption is only shown for Kühlungsborn, as Pruhonice data are partly missing)

ionospheric absorption on 13 January. Finally, on 16 January both the Dst -field and the Pc 1 activity are strengthened, but the following increase in ionospheric absorption is of no importance. It can be argued that the slot region electrons have been depleted to a high degree by the previous loss processes. Consequently the new particle injection and the probable presence of plasmaspheric waves could only result in a slight electron precipitation. This might be unsatisfactory for influencing the radio wave absorption, but effective for Pc 1 generation. (Processes which are probably responsible for Pc 1 generation will be discussed in the next section.)

Discussion

The excessive ionospheric absorption of radio waves at middle latitudes well after the geomagnetic storm (i.e. the after-

effect) is due to increased electron density in the lower ionosphere. The increase in electron density (between about 95 and 75 km) results from enhanced ionization due to high energy electrons (>40 keV) precipitating from the outer radiation belt of the magnetosphere (Lauter and Knuth, 1967). During the main phase of the storm both protons (ions) and electrons are injected into the inner magnetosphere. The increased radial fluxes result in filling up the slot region and the plasmapause moves inwards to lower L values. It has been considered that the interaction between the cold plasma of the plasmasphere and the hot plasma of the plasma-sheet is an important process for particle precipitation (Lauter et al. 1977). From the start of the recovery phase the plasmapause moves outwards and the loss processes become effective. As the trapped electrons are lost to the atmosphere by resonant interaction with whistler mode turbulence (Spjeldvik and Thorne, 1975), the occurrence of VLF and/or ELF waves in the plasmasphere is another vital condition for particle precipitation. Some of these processes are thought to be involved in the generation of Pc 1 pulsations, thus an explanation for the connection of increased ionospheric absorption with Pc 1s should be sought for on this basis.

According to the initial ideas of several authors, Pc 1 pulsations result from bunches of trapped particles bouncing between conjugate points (e.g. Jacobs, 1970). Their generation might be due to ion cyclotron resonance and the outer plasmasphere seemed to be the most probable generation region (e.g. Lewis et al., 1977).

Bell (1976) presented a model of artificial pulsation generation which is based on a several-stage process involving pulsed VLF transmissions. According to the Bell (1976) mechanism, Pc 1 pulsations can be stimulated in the following way: from a ground- or satellite-based VLF transmitter, repetitive pulses are injected into the magnetosphere where VLF emissions are triggered and energetic electrons can be precipitated into the ionosphere due to wave-particle interaction; the precipitated flux of electrons modifies the conductivity of the lower ionosphere (D and E regions), inducing periodic changes in the current flow which in turn result in the generation of Pc 1 ULF waves. Consequently, the ULF magnetic field perturbations can be observed directly as pulsations on the ground.

Fraser-Smith and Helliwell (1980) suggested that these processes should also be effective for repetitive VLF activity occurring naturally in the magnetosphere. Recent investigations by Sato (1984) actually yielded experimental evidence for this mechanism. A close correlation between periodic VLF emissions ($T \sim 5.6$ s) and short-period magnetic pulsations was found at Syowa Station ($L \sim 6$) in Antarctica.

Conclusions

The connection of increased ionospheric absorption with Pc 1 pulsations, revealed for mid-latitudes by März and Verö (1977) and confirmed in the present study for high latitudes, is not inconsistent with the previously discussed findings. As mentioned, the increase of ionospheric absorption following certain geomagnetic storms is generally attributed to energetic electrons precipitating from the magnetosphere. If the electrons are precipitated by periodic VLF emissions with a repetition time corresponding to the

period range of Pc 1s (0.2–5 s), the generation of this type of pulsations can also be expected on the basis of the Bell (1976) mechanism and in accordance with the results of Sato (1984). The latter author (Sato, 1984) has shown that periodic VLF emissions could be observed simultaneously at ISIS 2 satellite level ($\sim 1,400$ km altitude) in the wide latitude range from $L \sim 3.5$ to $L \sim 14.0$ and on the ground at Syowa. Furthermore, at this station the magnetic pulsations also had the same periods as the repetition time of VLF emissions.

On this basis we have to suggest that both the post-storm increase in ionospheric absorption at mid-latitudes and the enhanced Pc 1 activity at high latitudes (which generally appear simultaneously, as shown for individual cases in Figs. 6 and 7) might originate from the same sources through processes including the electron precipitation due to VLF emissions in the magnetosphere. Actually, the present study has confirmed that ground-based Pc 1 records can be used as an appropriate tool for selecting geomagnetic storms which are probably followed by plasmaspheric turbulence (in the VLF or ELF band) associated with an effective particle precipitation, which finally results in the increase of radio wave absorption.

References

- Bell, T.F.: ULF wave generation through particle precipitation induced by VLF transmitters. *J. Geophys. Res.* **81**, 3316–3326, 1976
- Fraser-Smith, A.C., Helliwell, R.A.: Stimulation of Pc 1 geomagnetic pulsations by naturally occurring repetitive VLF activity. *Geophys. Res. Lett.* **7**, 851–853, 1980
- HHI Geophys. Data, Vol. 27–30, Berlin, 1976–1979
- Ionospheric Data, Observatories Průhonice and Panska Ves, Czechoslovak Acad. Sci. Geophys. Inst. Praha, 1978
- Jacobs, J.A.: Geomagnetic micropulsations. Berlin. Heidelberg, New York: Springer 1970
- Lauter, E.A., Knuth, R.: Precipitation of high energy particles into the upper atmosphere at medium latitudes after magnetic storms. *J. Atmos. Terr. Phys.* **29**, 411–417, 1967
- Lauter, E.A., Bremer, J., Grafe, A., Deters, I., Evers, K.: The post-storm ionisation enhancements in the mid-latitude D-region and related electron precipitation from the magnetosphere. HHI-STP-Report, No 9, Berlin, 1–86, 1977
- Lewis, P.B. Jr., Arnoldy, R.L., Cahill, L.J. Jr.: The relation of Pc 1 micropulsations measured at Siple, Antarctica, to the plasmapause. *J. Geophys. Res.* **82**, 3261–3271, 1977
- März, F., Verö, J.: Ionospheric absorption and Pc 1-type micropulsations following enhanced geomagnetic activity. *J. Atmos. Terr. Phys.* **39**, 295–302, 1977
- März, F., Verö, J., Benze, P.: Short-period pulsations observed in Finland during the IMS and certain associations with magnetic and ionospheric parameters. *Acta Geodaet. Geophys. Mont. Hung.* **19**, (3–4) 1984 (in press)
- Pikkariainen, T., Lukkari, L., Niskanen, J., Kangas, J.: Quick-look tables of Pc 1 and IPDP magnetic pulsation events recorded in Finland during IMS-years 1976–1979. Department of Physics, University of Oulu, Report No. 92. 1–78, 1982
- Sato, N.: Short-period magnetic pulsations associated with periodic VLF emissions ($T \sim 5.6$ s). *J. Geophys. Res.* **89**, 2781–2787, 1984
- Spjeldvik, W.N., Thorne, R.M.: The cause of storm after-effect in the middle latitude D-region. *J. Atmos. Terr. Phys.* **37**, 777–795, 1975

Received: May 2, 1984; Revised Version: August 25, 1984
Accepted: September 12, 1984

Some aspects of whistler duct lifetimes at low latitudes

Lalmani*

Max-Planck-Institut für Aeronomie, D-3411 Katlenburg-Lindau, Federal Republic of Germany

Abstract. Naturally occurring low-latitude whistlers provide a powerful ground-based technique for probing the inner magnetosphere. Whistler data recorded during a 5-h period at our low-latitude ground stations, Gulmarg ($L=1.28$), Nainital ($L=1.16$) and Varanasi ($L=1.11$), were used to determine the lifetime of whistler ducts. Duct lifetimes as short as 50 min were observed, a result which has an important implication for current theories of ducts at low latitudes.

Key words: Whistlers – Magnetosphere – Ducts – Ionosphere – Plasmopause – Dispersions – Causative sferics – Electric field.

Introduction

Over the last decade, whistlers have become a very important tool for probing the plasmaphere and beyond. Since the pioneering work of Storey (1953), the observation of whistlers has been continued over a wide range of high-to-low latitudes (Allcock, 1960; Helliwell, 1965; Iwai and Ohtsu, 1967; Somayajulu et al., 1972; Cerisier, 1973; Singh et al., 1977; Hayakawa and Tanaka, 1978; Carpenter, 1983). It is now well established that most ground-observed whistlers are guided between opposite hemispheres by magnetic field-aligned enhancements of ionization termed “whistler ducts”. The theory of trapping and propagation of whistlers in ducts has been discussed by several authors (e.g. Smith, 1961; Walker, 1972; Laird and Nunn, 1975). By using these ducted whistlers the dynamics and structure of the magnetospheric thermal plasma have been studied (Carpenter and Park, 1973; Y. Corcuff, 1975). However, Walker (1976) recently showed that there remain many unsolved problems in the whistler propagation itself, especially in the lower exosphere and ionosphere, such as the excitation of ducts and their lifetimes, leakage from them, and the transmission properties through the ionosphere.

Determination of duct lifetimes is important for evaluating theories of duct formation at low latitudes. On the basis of detailed sonographic study, Somayajulu and Tantry (1968) have found that it might take less than 1 h for ducts to form and that once the duct has been formed, it might stay alive for a few hours to a few days. Ducts at mid-latitudes have been observed to persist for 10–20 h (Park

and Carpenter, 1970) in the vicinity of the plasmopause. Recently duct lifetimes at mid-latitudes as short as 30 min have been reported by making simultaneous observations from adjacent whistler stations (Hansen et al., 1983). The lifetime of ducts deduced by the low-latitude workers seems to be widely distributed from 30 min to a few days. Therefore, further study is needed of the lifetime of the ducts at low latitudes.

As a result of the increasing importance and versatility of the whistler technique it has become clear that improved understanding of some aspects of whistler propagation is necessary at low latitudes. The purpose of this paper is to determine the lifetime of whistler ducts by making continued whistler data recorded at our low-latitude ground stations, Gulmarg ($L=1.28$; $24^{\circ}26'N$, $147^{\circ}09'E$) and Nainital ($L=1.16$; $19^{\circ}02'N$, $149^{\circ}45'E$) from 0020 to 0500 I.S.T. (Indian Standard Time) on 25 March 1971, and at Varanasi ($L=1.11$; $15^{\circ}06'N$, $159^{\circ}33'E$) from 0020 to 0300 IST on 19 March, 1977. Briefly, our approach is as follows. Suppose we consider an ideal situation in which only one duct is contributing to the whistler occurrence rate at a station and also suppose that this duct takes a finite time to grow to its mature state and to decay, finally merging with the background ionization. We should then expect the whistler occurrence rate recorded on the ground to show the corresponding rise and fall with time. If, as suggested by Okuzawa et al. (1971), the growth and decay of ducts are cyclic, it follows that the whistler occurrence rate shows a kind of periodicity. This ideal case can, in favourable conditions, be extended to a realistic situation in which a few active ducts are simultaneously contributing to the occurrence rate. The presence of any periodicity in the combined occurrence rate can be detected by the standard technique of power spectrum analysis, as suggested by Madden (1964). We find that whistler-duct lifetimes can be as short as 50 min, a result which has important implications for current theories of ducts at low latitudes.

Observations

The whistlers recorded at Gulmarg, Nainital and Varanasi were always of very high quality, and the number of whistlers recorded during magnetic storms was always large enough to be of statistical significance. During the mentioned period the whistler data were characterized by a consistently good whistler intensity and rate with well-defined components.

* Permanent address: Department of Physics, Regional Engineering College, Srinagar, Kashmir, India

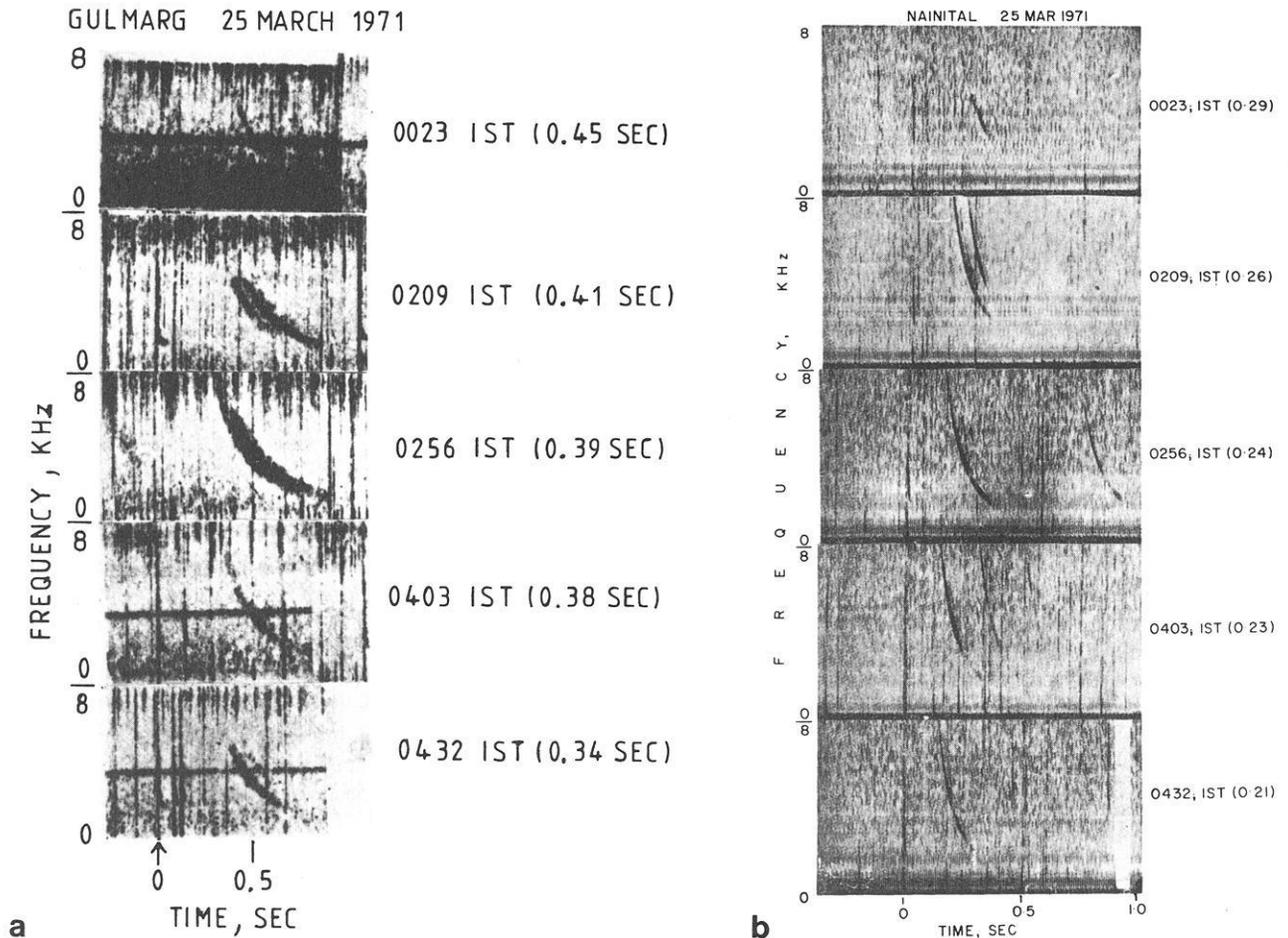


Fig. 1a + b. An example of spectrograms of whistlers recorded simultaneously at **a** Gulmarg and **b** Nainital on 25 March 1971, showing the variation of dispersion with time (figures shown by the side of time indicate the time delay at 5 kHz)

On 25 March 1971, whistlers in great numbers were observed at our field stations in Gulmarg and Nainital. The spurt in activity started around 0020 IST and lasted for about 5 h, ending finally at day break at 0520 IST. During this period the Kp index varied between 2 and 5. The mere occurrence rate itself was interesting and altogether several hundred whistlers were recorded and dispersion analyses made. The dispersion showed a remarkably smooth decrease within the observation period of 5 h.

In Fig. 1 we show a sequence on five sonograms corresponding to different times of occurrence recorded on 25 March 1971 at Gulmarg and Nainital. The first sonogram (0023 IST) shows a single trace, with energy limited to a frequency band of 3–5 kHz. This is a typical feature of low-latitude whistlers during magnetic disturbances. The first whistler to arrive, after the sudden commencement of the storm, has energies concentrated mainly around 4.5 kHz. This is probably a reflection of the concentration of energy in the mentioned frequency range at the source itself.

All the above sonograms have been arranged in such a way that causative sferics lie on a single vertical line. The figures shown in brackets by the side of time of occurrence indicate the delay at 5 kHz. It is at once evident from Fig. 1 that the delay time at 5 kHz decreases smoothly with time. This, of course, corresponds to a similar decrease in dispersion.

In Fig. 2 we show a sequence of five sonograms corre-

sponding to different times of occurrence during magnetic storms recorded at our low-latitude ground station Varanasi on 19 March 1977. Similar characteristics of smooth decrease in dispersion can also be seen in Fig. 2. The spurt in activities started around 0000 IST and lasted for about 5 h, and several hundred whistlers were recorded during this period.

The decrease in dispersions shown in Figs 1 and 2 (discussed later) indicate that some ducts formed within that portion of fields of view. Flux tube interchange has been proposed as a possible mechanism for duct formation, so it was of interest to determine whether the ducts are subject to any cross-L drift. Hence nose frequencies (f_n) of the non-nose whistlers recorded at our low-latitude stations were determined using the method of Ho and Bernard (1973) and plotted $f_n^{2/3}$ instead of f_n against IST, as shown in Fig. 3, so that the slope of the data points can be directly related to the convection electric field (Block and Carpenter, 1974; Park, 1976).

Results and discussions

In Fig. 4 we show the power spectrum analysis of the whistler occurrence rate at Gulmarg, Nainital and Varanasi on 25 March 1971 and 19 March 1977 to estimate duct lifetime. It is at once clear from Fig. 4 that the power spectra show peaks at certain selected frequencies. Thus, the time interval corresponding to the frequency separation between peaks

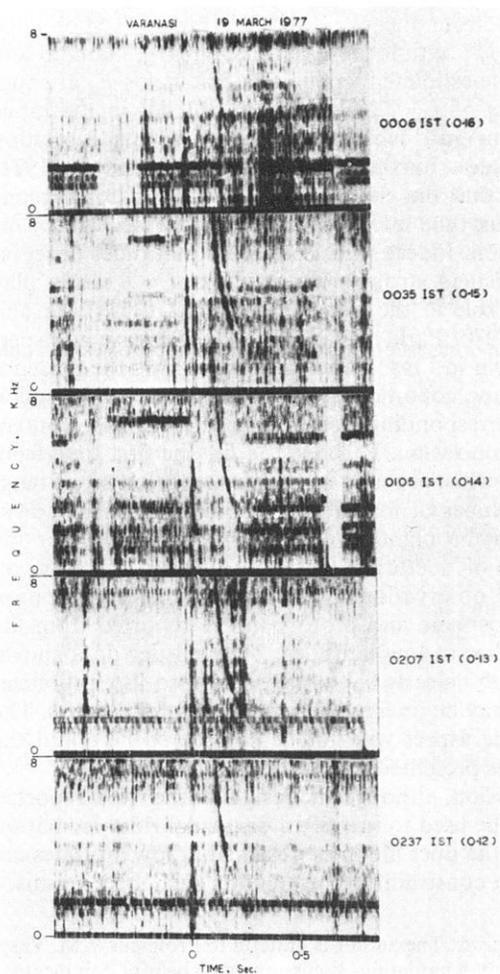


Fig. 2. An example of spectrograms of whistlers recorded at Varanasi on 19 March 1977, showing the variation of dispersion with time as in Fig. 1

is remarkably constant and is around 50 min. This result seems to indicate that some physical process with a periodicity of about 50 min is present in the whistler occurrence rate recorded at Gulmarg, Nainital and Varanasi. We wish to suggest that this periodicity signifies the continuous process of growth and decay of ducts as envisaged by Okuzawa et al. (1971). The spectrum corresponding to the data obtained at Nainital (Fig. 4b) and Varanasi (Fig. 4c) shows peaks at a frequency of about 2×10^{-3} rad/s and 1×10^{-3} rad/s and its harmonics respectively, having the constant time interval corresponding to the frequency separation between peaks. The power spectrum in Fig. 4a also shows peaks but the frequency separation between peaks is rather variable. In this case the average time interval corresponding to the frequency separation between peaks is about 1 h. Considering the fact that all the individual ducts which contribute to the whistler occurrence rate at a given station do not grow or decay in unison, we expect the frequency interval between peaks in the power spectrum to be variable. In this sense, the remarkable constancy of the frequency interval in the power spectrum shown in Fig. 4b and c should be regarded as fortuitous. It is also perhaps probable that at Nainital and Varanasi, which are

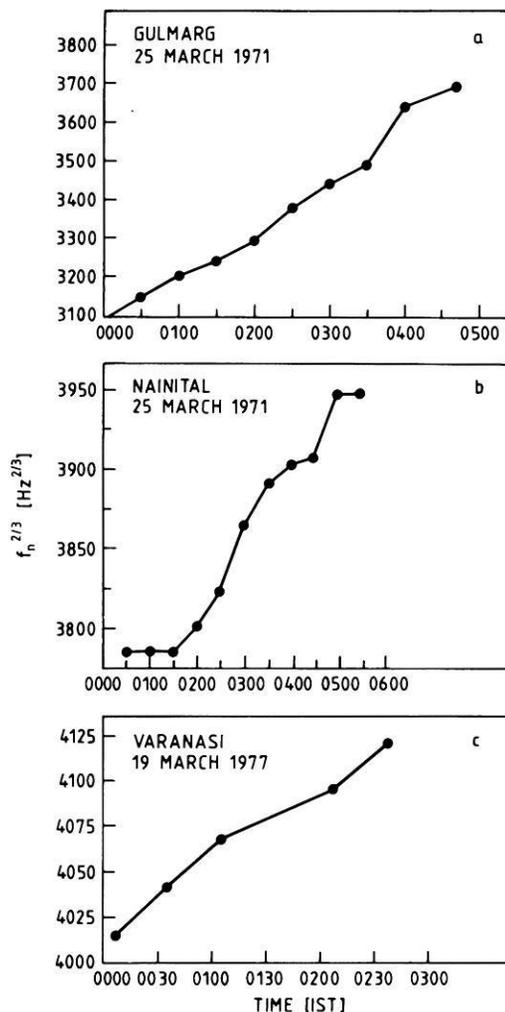


Fig. 3. Variation of whistler nose frequency (plotted on a linear scale in the two-thirds power of nose frequency), with local time observed at (a) Gulmarg, (b) Nainital and (c) Varanasi to estimate the westward component of electric field

at lower latitudes than Gulmarg, only very few ducts contribute to the occurrence rate. Such a conclusion is at least qualitatively in agreement with the fact that the "viewing area" of a ground station decreases with latitude (Carpenter, 1966). The period of about 50 min obtained from the data in Fig. 4 can be taken to represent the order of lifetime of ducts. The above value of about 50 min for the duct lifetime is comparable in order of magnitude with the estimate of other workers (Smith, 1960; Okuzawa et al., 1971; Park and Carpenter, 1970; Hansen et al., 1983). Especially Fig. 12 of Park and Carpenter (1970) seems to suggest that the average time for which the individual ducts may stay within the viewing area is around 2 h. Furthermore, the number of whistlers recorded at our low-latitude ground station is much less than the available sources (lightning flashes). This shows that the occurrence rate of whistlers at low latitudes is controlled more by the properties of the medium than by source effects.

Several duct-formation mechanisms have been proposed to date (Park and Helliwell, 1971; Cole, 1971; Walker, 1978; Thomson, 1978; Lester and Smith, 1980). An essential point common to several theories is that electric fields perturbing the magnetospheric plasma play a dominant role

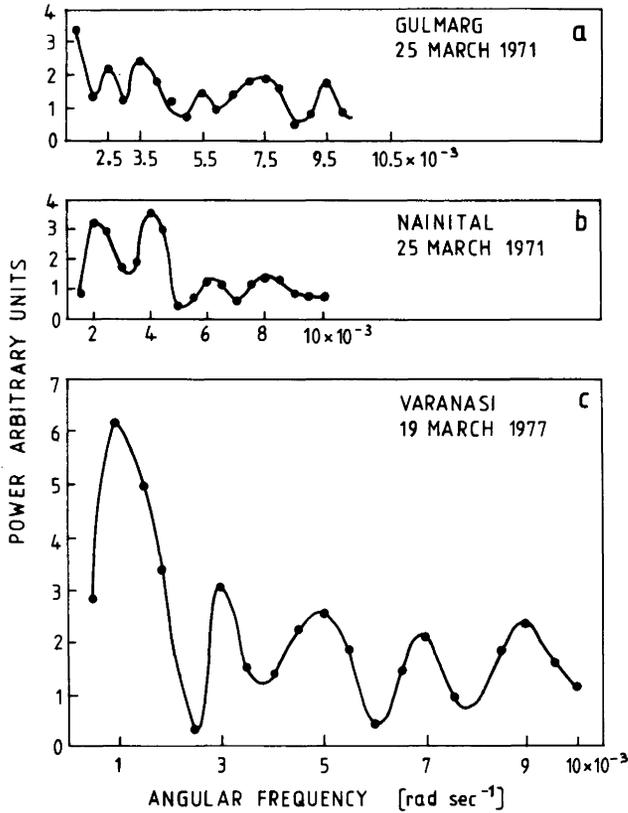


Fig. 4. Power spectra of the whistler occurrence rate observed at (a) Gulmarg, (b) Nainital and (c) Varanasi for whistler-duct lifetimes at low latitudes

in duct formation. Therefore, in the present paper an attempt was made to estimate the magnetospheric electric field during magnetic storm periods, using whistlers observed at Gulmarg, Nainital and Varanasi (shown in Figs. 1 and 2), from radial motions of discrete field-aligned whistler ducts, as indicated by changes in nose frequencies (Block and Carpenter, 1974). Park (1976) has shown that the observed changes in whistler nose frequency, f_n , were due to cross- L drift of ducts. The whistler nose frequency and the minimum equatorial gyrofrequency f_{Heq} , along the path of propagation are related (Block and Carpenter, 1974; Park, 1976) as

$$f_n \approx k f_{Heq} = k f_{H_0} (R_0/R)^3 \quad (1)$$

where $k = 0.38$ for diffusive equilibrium model of field-line distribution of ionization, and f_{Heq} and f_{H_0} are the equatorial gyrofrequencies at geocentric distances R and R_0 (earth's surface) respectively.

Specializing the hydromagnetic drift relation $V = E \times B / B^2$ to the magnetic equator, we obtain (in MKS units)

$$dR/dt = -(E_\omega/B_0) (R_0/R)^{-3} \quad (2)$$

where B_0 represents the geomagnetic field strength at the earth's surface and E_ω is the westward component of the magnetospheric electric field. From Eqs. (1) and (2) the convection electric field in the post mid-night sector, in a dipole model, in the equatorial plane is given (Block and Carpenter, 1974; Park, 1976) as

$$E_\omega = 2.07 \times 10^{-2} \frac{d(f_n^{2/3})}{dt} \text{ Vm}^{-1} \quad (3)$$

Thus, from Eq. (3) one can directly estimate the convection electric field from the slope of $f_n^{2/3}$. Figure 3 shows the variation of $f_n^{2/3}$ with local time for Gulmarg, Nainital and Varanasi. The estimated value of electric fields E_ω at equatorial heights of $L = 1.28, 1.16$ and 1.11 lies in the range of 0.1 to 0.7 mVm^{-1} which are sufficient for duct formation at low latitudes. Furthermore, Park and Helliwell (1971) have shown that the electric fields in the equatorial plane will cause flux tube interchange, a possible mechanism for duct formation. Indeed, most of the mid-latitudes observations of whistlers strongly suggest that $\vec{E} \times \vec{B}$ drifts play a dominant role in the transport of ionization (Park and Carpenter, 1970; Park, 1970, 1972). The decrease in dispersion, as shown in Figs. 1 and 2, clearly shows the presence of plasma-flow conditions because decrease in dispersion gives the corresponding decrease in the electron content of tubes of ionization. Park (1970), for the first time from mid-latitude whistler study of the electron content of magnetospheric tubes of ionization, reported the usual downward flux in the night-time across 1000 km level for the maintenance of nocturnal F-layer. However, there are no experimental observations of the interchange of ionization between ionosphere and protonosphere reported from the whistler studies at low latitudes. The whistler data shown in Figs. 1 and 2 clearly show the downward flux of ionization which may be interpreted in terms of $\vec{E} \times \vec{B}$ drifts. The details of this aspect will not be considered here and the results will be presented in a future report.

In conclusion, although the results of the work reported here cannot be used to identify a particular duct-formation mechanism, the duct lifetimes observed at low latitudes are an important constraint in formulating such mechanisms.

Acknowledgements. The author is grateful to Professor V.M. Vasyliunas for many stimulating discussions and helpful comments. I would also like to thank Professor O.N. Wakhlu and Dr. P.N. Khosa of REC, Srinagar, for their constant encouragement. Part of the work presented in this paper was done at the Department of Physics of REC Srinagar, India. A research fellowship granted by the Max Planck Society is gratefully acknowledged. Finally, we extend our thanks to Ms Deutsch for typing the manuscript.

References

- Allcock, G.M.: IGY whistler results, paper presented at 13th General Assembly, Union Radio Sci. Int. London, 1960
- Block, L.P., Carpenter, D.L.: Deviation of magnetospheric electric fields from whistler data in a dynamic geomagnetic field. *J. Geophys. Res.* **79**, 2783, 1974
- Carpenter, D.L.: Whistler studies of the plasmopause in the magnetosphere - I. Temporal variations in the position of knee and some evidence on plasma motions near the knee. *J. Geophys. Res.* **71**, 693, 1966
- Carpenter, D.L.: Some aspects of plasmopause probing by whistlers. *Radio Sci.* **18**, 917, 1983
- Carpenter, D.L., Park, C.G.: On what ionospheric workers should know about the plasmopause-plasmasphere. *Rev. Geophys. Space Phys.* **11**, 133, 1973
- Cerisier, J.C.: A theoretical and experimental study of non-ducted VLF waves after propagation through the magnetosphere. *J. Atmos. Terr. Phys.* **35**, 77, 1973
- Cole, K.D.: Formation of field-aligned irregularities in the magnetosphere. *J. Atmos. Terr. Phys.* **33**, 741, 1971
- Corcuff, Y.: Probing of the plasmopause by whistlers. *Ann. Geophys.* **26**, 363, 1975
- Hansen, H.J., Scourfield, M.W.J., Rash, J.P.S.: Whistler duct lifetimes. *J. Atmos. Terr. Phys.* **45**, 789, 1983

- Hayakawa, M., Tanaka, Y.: On the propagation of low-latitude whistlers. *Rev. Geophys. Space Phys.* **16**, 111, 1978
- Helliwell, R.A.: Whistlers and related ionospheric phenomena. Stanford Univ. Press, Stanford, Calif., 1965
- Ho, D., Bernard, L.C.: A fast method to determine the nose frequency and minimum group delay of a whistler when causative sferic is unknown. *J. Atmos. Terr. Phys.* **35**, 881, 1973
- Iwai, A., Ohtsu, J.: Long-term variation of whistler dispersion. *Proc. Res. Inst. Atmos. Nagoya Univ.* **14**, 51, 1967
- Laird, M.J., Nunn, D.: Full wave VLF modes in a cylindrically symmetric enhancement of plasma density. *Planet. Space Sci.* **23**, 1949, 1975
- Lester, M., Smith, A.J.: Whistler duct structure and formation. *Planet. Space Sci.* **28**, 645, 1980
- Madden, T.: Spectral, cross-spectral and bispectral analysis of low frequency electromagnetic data, in *Natural Electromagnetic phenomena below 30 kc/s* (Ed. Bleil, D.F.) p. 429, Plenum, New York, 1964
- Okuzawa, T., Yamanaka, K., Yoshino, T.: Characteristics of low-latitude whistler propagations associated with magnetic storms in March 1970. *Rep. Ionos. Space Res. Jap.* **25**, 17, 1971
- Park, C.G.: Whistler observations of the interchange of ionization between the ionosphere and the protonosphere. *J. Geophys. Res.* **75**, 4249, 1970
- Park, C.G.: Methods of determining electron concentrations in the magnetosphere from nose whistlers. *Radiosci. Lab. Stanford Univ. Tech. Rep. No. 3454-1*, 1972
- Park, C.G.: Substorm electric fields in the evening plasmasphere and their effects on the underlying F layer. *J. Geophys. Res.* **81**, 2283, 1976
- Park, C.G., Carpenter, D.L.: Whistler evidence of large scale electron-density irregularities in the plasmasphere. *J. Geophys. Res.* **75**, 3825, 1970
- Park, C.G., Helliwell, R.A.: The formation by electric fields of field-aligned irregularities in the magnetosphere. *Radio Sci.* **6**, 299, 1971
- Singh, R.N., Lalmani, Singh, R.P.: Whistlers recorded at Varanasi. *Nature Phys. Sci.* **266**, **40**, 1977
- Smith, R.L.: Guiding of whistlers in a homogeneous medium. *J. Res. Nat. Bur. Stand. Sect. D.* **64**, 505, 1960
- Smith, R.L.: Propagation characteristics of whistlers trapped in field-aligned columns of enhanced ionization. *J. Geophys. Res.* **66**, 3699, 1961
- Somayajulu, V.V., Tantry, B.A.P.: Effect of magnetic storms on duct formation for whistler propagation. *J. Geomagn. Geoelect.* **20**, 21, 1968
- Somayajulu, V.V., Rao, M., Tantry, B.A.P.: Whistlers at low latitude. *Indian J. Radio Space Phys.* **1**, 102, 1972
- Storey, L.R.O.: An investigation of whistling atmospherics. *Phil. Trans. Roy. Soc. London, Ser. A* **246**, 113, 1953
- Thomson, R.J.: The formation and lifetime of whistler ducts. *Planet. Space Sci.* **26**, 423, 1978
- Walker, A.D.M.: The propagation of very low frequency waves in ducts in the magnetosphere, 2. *Proc. R. Soc. London.* **A239**, 219, 1972
- Walker, A.D.M.: The theory of whistler propagation. *Rev. Geophys. Space Phys.* **14**, 629, 1976
- Walker, A.D.M.: Formation of whistler ducts. *Planet. Space Sci.* **26**, 375, 1978

Received February 17, 1984; Revised July 11, 1984

Accepted July 11, 1984

*Short communication***On a mixed quadratic invariant of the magnetic susceptibility tensor**

Vit Jelinek

Geofyzika Brno, Ječná 29a, 612 46 Brno, Czechoslovakia

Key words: Magnetic susceptibility anisotropy – Magnetic susceptibility tensor**Introduction**

Magnetic susceptibility of an anisotropic rock can be described by a second-order symmetric tensor. This tensor has several invariants of which commonly used is only the linear invariant closely associated with the mean susceptibility. The quadratic invariant does not have an explicit physical meaning.

In this paper, a particular mixed quadratic invariant of the susceptibility tensor is discussed which is numerically three times the variance of the principal susceptibilities. Its advantage consists in the fact that it can be determined using components of the susceptibility tensor in an arbitrary coordinate system; principal susceptibilities need not be known. This parameter can be used for characterizing the so called deviatoric component of susceptibility on one hand and statistical testing of the anisotropy of susceptibility on the other hand.

Magnetic susceptibility tensor

For a magnetically linear medium the relation between the intensity of the field \mathbf{H} and the induced magnetic polarization \mathbf{J} can be in a cartesian system of coordinates expressed by equation

$$\begin{bmatrix} J_1 \\ J_2 \\ J_3 \end{bmatrix} = \mu_0 \begin{bmatrix} k_{11} & k_{12} & k_{13} \\ k_{21} & k_{22} & k_{23} \\ k_{31} & k_{32} & k_{33} \end{bmatrix} \cdot \begin{bmatrix} H_1 \\ H_2 \\ H_3 \end{bmatrix} \quad (1)$$

or briefly

$$\mathbf{J} = \mu_0 \mathbf{kH}. \quad (2)$$

Constants k_{ij} are components of the symmetric tensor of magnetic susceptibility.

For the sake of brevity we shall hereafter, as a rule, leave out the word “magnetic”. Further, we shall not distinguish between the tensor and the matrix representing it.

The linear and the quadratic invariant

The susceptibility tensor may be interpreted geometrically by a quadric in the central position. The quadric is usually an ellipsoid, the so called susceptibility ellipsoid (Nagata 1961). Thus the invariants of quadrics are also invariants of the susceptibility tensor. For the invariants of quadrics see e.g. Rektorys (1969).

The commonly used linear invariant is given by the formula

$$I_1 = k_{11} + k_{22} + k_{33}. \quad (3)$$

It is simply related to the mean susceptibility:

$$\kappa = I_1/3. \quad (4)$$

Further, the quadratic invariant exists,

$$I_2 = \begin{vmatrix} k_{11} & k_{12} \\ k_{12} & k_{22} \end{vmatrix} + \begin{vmatrix} k_{22} & k_{23} \\ k_{23} & k_{33} \end{vmatrix} + \begin{vmatrix} k_{33} & k_{31} \\ k_{31} & k_{11} \end{vmatrix} \quad (5)$$

which has no simple physical meaning.

The mixed quadratic invariant

Let us form a mixed quadratic invariant from the above mentioned invariants,

$$I_M = \frac{2}{3} I_1^2 - 2I_2. \quad (6)$$

After some easy adaptations we get

$$I_M = k_{11}^2 + k_{22}^2 + k_{33}^2 - 3\kappa^2 + 2k_{12}^2 + 2k_{23}^2 + 2k_{31}^2. \quad (7)$$

If the susceptibility tensor is expressed in the system of principal directions, and thus the principal susceptibilities $\kappa_1, \kappa_2, \kappa_3$ are known, the Eq. (7) may be simplified to

$$I_M = \kappa_1^2 + \kappa_2^2 + \kappa_3^2 - 3\kappa^2 \quad (8)$$

or

$$I_M = (\kappa_1 - \kappa)^2 + (\kappa_2 - \kappa)^2 + (\kappa_3 - \kappa)^2. \quad (9)$$

From Eqs. (8) and (9) it follows that I_M is three times the variance of principal susceptibilities. It may be therefore assumed that it will be possible to assign to the invariant I_M certain practical sense. Before discussing this problem, we shall express the invariant I_M in still another way.

Representation of the mixed quadratic invariant by means of the deviatoric susceptibility tensor

Using the mean susceptibility, we can express the susceptibility tensor by a sum of two tensors

$$\mathbf{k} = \kappa \mathbf{1} + \mathbf{l}, \quad (10)$$

while $l_{11} + l_{22} + l_{33} = 0$. We shall say that the tensor $\kappa \mathbf{1}$ represents the isotropic component and the tensor \mathbf{l} the deviatoric component of the susceptibility tensor \mathbf{k} . The latter term is taken over from the theory of stress and strain, where the analogically defined quantity is called "the deviatoric stress" (Ramsay 1967).

It can be easily demonstrated that the mixed quadratic invariant of the tensor \mathbf{k} can be expressed by the components of tensor \mathbf{l} ,

$$I_M = l_{11}^2 + l_{22}^2 + l_{33}^2 + 2l_{12}^2 + 2l_{23}^2 + 2l_{31}^2 \quad (11)$$

and further,

$$I_M = \lambda_1^2 + \lambda_2^2 + \lambda_3^2, \quad (12)$$

where $\lambda_1, \lambda_2, \lambda_3$ are eigenvalues of the matrix representing the tensor \mathbf{l} .

Standard deviatoric susceptibility

It is often essential to express quantitatively the amount of the deviatoric susceptibility component. This is particularly important when measuring the susceptibility anisotropy using instruments which do not provide data on the isotropic component (torque magnetometer, induction anisometer with rotating sample).

For the measure of the amount of the deviatoric component we may take the quantity

$$\tilde{\kappa} = \sqrt{I_M/3}. \quad (13)$$

We shall term it the standard deviatoric susceptibility.

From Eqs. (12) and (13) it is evident that the standard deviatoric susceptibility really characterizes the quantity of the deviatoric component. It is obvious from Eqs. (9) and (13) that it equals the standard deviation of the principal susceptibilities.

The standard deviatoric susceptibility does not belong to the so called anisotropy factors for the very reason that it has the "dimension" of susceptibility. Moreover, it does not satisfy the definition given by Jelínek (1981).

Besides the quantitative description of the deviatoric component, the standard deviatoric susceptibility can be with advantage used for forming the normed tensor of the deviatoric component. The normed tensor is determined by the equation

$$\mathbf{L} = \frac{1}{\tilde{\lambda}} \mathbf{l}, \quad (14)$$

where $\tilde{\lambda} = \tilde{\kappa}$. The mean value of the tensor \mathbf{L} (defined similarly as the mean susceptibility) is zero, while the standard deviatoric value (defined in a similar manner as the standard deviatoric susceptibility) is unity which is very suitable for computation and numerical representation of results.

Test of anisotropy

When measuring the magnetic susceptibility by an A.C. bridge, the changes of the coil inductivity caused by the sample are evaluated (Girdler 1961; Fuller 1967; Jelínek 1973). The sample is inserted in the coil in a certain number of directions, e.g. 9, 15, 18 or more. The choice of the measuring directions is called the design of the experiment. The measured so called directional susceptibilities serve for computing the components of the susceptibility tensor, principal susceptibilities, and principal directions.

Certain differences can always be observed between the directional as well as principal susceptibilities. The purpose of the anisotropy test is to decide whether these differences are caused by the sample anisotropy or whether they are due to measuring errors only.

The test for anisotropy can be performed using the analysis of dispersion. Computations can be considerably simplified if the so called rotatable design of measuring directions (Hext 1963) is used. Hext suggested several rotatable designs, which, however, are not suitable for practical use. One design of 15 measuring directions, easily applicable, is described by Jelínek (1973).

Developing some ideas of Hext (1963), it is possible to construct relatively simply a test for the rotatable design. For the testing statistic it can be derived

$$F = \frac{2n}{75s^2} \hat{I}_M, \quad (15)$$

where n is the number of measuring directions, s^2 the estimate of dispersion in a single direction obtained from the sum of squares of residual errors, \hat{I}_M is the mixed quadratic invariant of the estimate of the susceptibility tensor.

If the sample is isotropic, the statistic F has F -distribution on 5 and $n-6$ degrees of freedom. If the statistic F exceeds the critical value on the chosen level of significance, the sample can be considered anisotropic.

Conclusion

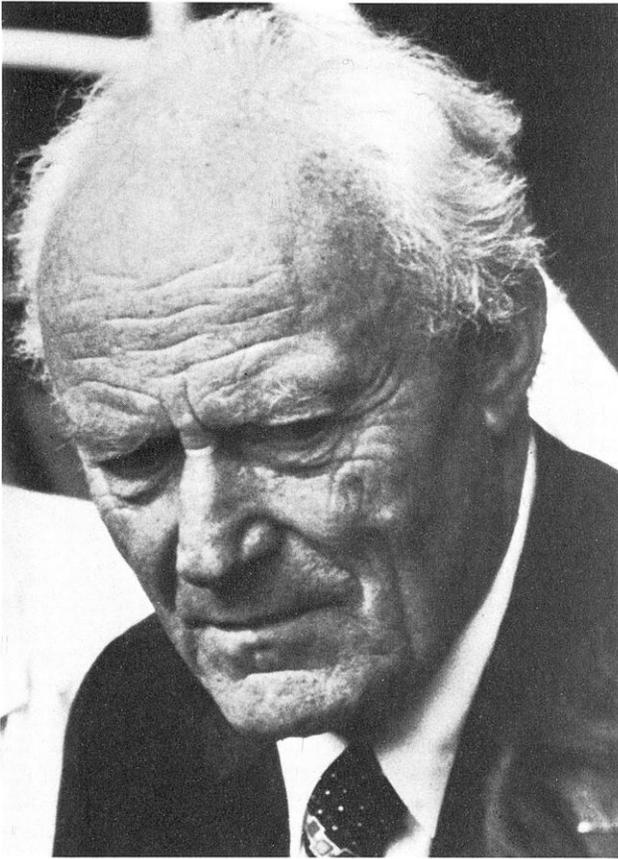
A mixed quadratic invariant of the tensor of magnetic susceptibility has been discussed. It equals three times the variance of the principal susceptibilities. From this invariant, the standard deviatoric susceptibility can be computed which characterizes the deviatoric component of susceptibility. The standard deviatoric susceptibility is of particular use when measur-

ing with an instrument which gives no information on the mean susceptibility (torque magnetometer, induction magnetometer with rotating sample). The mixed quadratic invariant also enables derivation of the test of anisotropy and simplifies its execution. It can be assumed that this invariant will simplify some calculations in studying the anisotropy of magnetic susceptibility of rocks.

References

- Fuller, M.D.: The A.C. bridge method. In: *Methods in palaeomagnetism*, D.W. Collinson, K.M. Creer, S.K. Runcorn, eds.: pp. 403–408. Amsterdam-London-New York: Elsevier Pub. Co. 1967
- Girdler, R.W.: The measurement and computation of anisotropy of magnetic susceptibility of rocks. *Geophys. J. RAS* **5**, 34–44, 1961
- Hext, G.R.: The estimation of second-order tensors, with related tests and designs. *Biometrika* **50**, 353–373, 1963
- Jelínek, V.: The precision A.C. bridge set for measuring magnetic susceptibility of rocks and its anisotropy. *Studia Geophys. Geodaet* **17**, 36–48, 1973
- Jelínek, V.: Characterization of the magnetic fabric of rocks. *Tectonophysics* **79**, T63–T67, 1981
- Nagata, T.: *Rock magnetism*. Tokyo: Maruzen 1961
- Ramsay, J.G.: *Folding and fracturing of rocks*. New York: McGraw-Hill 1967
- Rektorys, K.: *Survey of applicable mathematics*. London: Iliffe Books 1969

Received June 11, 1984; Accepted August 6, 1984

*In memoriam***Hans Closs (1907–1982)**

Am 2. Dezember 1982 verstarb plötzlich und unerwartet Prof. Dr. Hans Closs, Leitender Direktor und Professor i.R. in der Bundesanstalt für Bodenforschung und seit 1981 gewählter Präsident der Alfred-Wegener-Stiftung. Noch am Tage zuvor weilte er – wie uns schien – in gewohnter Frische unter uns und diskutierte mit uns seine aus der Fülle seines Wissens strömenden Gedanken zu zwei in Vorbereitung befindlichen geowissenschaftlichen Großprojekten. Mitten aus dieser vielseitigen Tätigkeit zum Wohle der Geowissenschaften in der Bundesrepublik Deutschland wurde er abberufen.

Hans Closs wurde am 1. Juli 1907 in Obersonthem im Kreis Schwäbisch-Hall geboren. Seine Ausbildung und sein beruflicher Werdegang hatten die folgenden Stationen:

Juli 1925 Reifeprüfung am Realgymnasium,
Schwäbisch-Hall

- ab WS 1925/26 Studium der Naturwissenschaften, insbesondere der Geologie, Mineralogie und Petrographie in Tübingen, Wien und Berlin
- 1933/34 Assistent am Institut für Mineralogie und Petrographie der Technischen Hochschule Berlin; Promotion bei Prof. W. Schmidt, Thema: „Quarzgefügestudien im östlichen Graubünden“
- ab 1.2.1935 Wissenschaftlicher Angestellter bei der Kommission zur Geophysikalischen Reichsaufnahme/Preußische Geologische Landesanstalt/Reichsstelle für Bodenforschung, Berlin
- 20.2.1941 Ernennung zum Bezirksgeologen bei der Reichsstelle für Bodenforschung
- ab 1948 Leiter der Abteilung Geophysik beim Amt für Bodenforschung (AfB), Celle/Hannover
- vom 1.12.1958 Leiter der Abteilung II „Laboratorien“ der Bundesanstalt für Bodenforschung (BfB), anfangs auch der Unterabteilung II/2 „Geophysikalische Laboratorien“
- bis 30.6.1972
- 1.7. bis Leiter der (neuen) Abteilung 3 „Geophysik“ der BfB
- 31.10.1972
- 1972–1976 Deutscher Vertreter in den Gremien des „Deep Sea Drilling Project“
- bis 1978 Mitherausgabe und Redaktion des Berichtsbandes über das Geodynamik-Projekt der Deutschen Forschungsgemeinschaft
- seit 1981 Präsident der Alfred-Wegener-Stiftung.

Drei Gesichtspunkte, nach denen einzelne Tätigkeiten im Wirken von Hans Closs eingeordnet werden können, sollen hier dargestellt werden.

Der Tektoniker, vor allem interessiert an den Alpen und dem Oberrheintalgraben: Der Beginn ist dokumentiert durch seine Promotionsarbeit, die ihn viele Monate in Graubünden in der Umgebung des Julierpasses verbringen ließ. Mehrere seiner gravimetrischen Studien bei der Geophysikalischen Reichsaufnahme liefern Beiträge zur Tektonik des Oberrheintalgrabens und zur Frage der Isotomie der Alpen. Er ist Mitautor der „Geotektonischen Karte von Nordwestdeutschland“ und deren umfangreichen Erläuterungen. Führend beteiligt er sich an den großen refraktionsseismischen Messungen, die in internationaler Kooperation zur

Erkundung des Untergrundes der Alpen durchgeführt werden. Seine Partner in der „Sous-Commission des Explosions Alpines“ waren vor allem Madame Labrouste und Carlo Morelli.

Lange Zeit war er Vizepräsident und Vertreter der Bundesrepublik bei der Commission International pour l'Exploration Scientifique de la Mer Mediterranée (CIESM) und Präsident des CIESM-Comité Géologique et Géophysique. Als Hauptkoordinator des Schwerpunktprogramms „Geodynamik des Mediterranen Raumes“ der Deutschen Forschungsgemeinschaft und Hauptschriftleiter des umfangreichen Abschlußbandes hat er wesentlich zum Gelingen des Programms beigetragen.

Viele Anregungen zu neuen Überlegungen über die Entstehung der Alpen und des Oberrheinalgrabens kamen aus den Forschungen des Deep Sea Drilling Project und seiner 1974 beginnenden internationalen Phase (IPOD), zu deren Entstehung Hans Closs sehr viel beigetragen hat. Davon zeugen eine ganze Reihe von Veröffentlichungen der letzten Jahre; umfangreiche Unterlagen zu einer Publikation über den Oberrheinalgraben fanden sich in seinem Nachlaß. Die komplette Liste seiner Veröffentlichungen findet sich im Heft 23 der Reihe E des Geologischen Jahrbuchs (Hannover 1982).

Der Anreger und Praktiker der Angewandten Geophysik:

Hans Closs begann seine Berufslaufbahn nach kurzer anderweitiger Tätigkeit bei der Kommission zur Geophysikalischen Reichsaufnahme, die das Gebiet des Deutschen Reiches systematisch mit den neuen Methoden der Geophysik im Hinblick auf Lagerstätten durchforstete. Die synoptische Auswertung der Ergebnisse für den nordwestdeutschen Raum führte zu der oben erwähnten Geotektonischen Karte. Nach 1945 beriet er viele Jahre die Erdölindustrie bei seismischen Untersuchungen.

Es lag nahe, erdöl- oder erdgasführende oder -höffige Strukturen auch in der Nordsee in Fortsetzung der vom Lande her bekannten Strukturen zu vermuten. 1958 begann Hans Closs mit seiner AfB-Gruppe gemeinsam mit dem Deutschen Hydrographischen Institut (DHI) und der Prakla GmbH mit dem Forschungsschiff Gauss des DHI die seismische Erkundung des Untergrundes der Nordsee, viele Jahre bevor die Industrie solche Untersuchungen aufnahm.

Schon Anfang der 50er Jahre hat Hans Closs auf Anwendungsmöglichkeiten der Ergebnisse mathematisch-physikalischer Forschungen zur Lösung verschiedenartigster praktischer geowissenschaftlicher Probleme hingewiesen. Er hat derartige Zielsetzungen beim Aufbau der Abteilung Geophysik des Amtes für Bodenforschung systematisch berücksichtigt mit der Einrichtung von Referaten für Geothermik, radioaktive Methoden einschließlich der Methoden zur physikalischen Altersbestimmung, Gesteinsphysik sowie Theorie und Mathematik. Dasselbe Thema hat er in seiner Präsidentenadresse 1962 bei der European Association of Exploration Geophysicists behandelt, woran man die Bedeutung erkennen kann, die er diesem Thema zugemessen hat.

Der wissenschaftliche Organisator und Förderer: Schon während seiner Tätigkeit bei der geophysikalischen Reichsaufnahme ist Hans Closs früh in Managementaufgaben mit hineingezogen worden und hineingewachsen. So war es praktisch ohne Diskussion, daß ihm nach dem Kriege die Leitung des Bereiches Geophysik des Amtes für Bodenfor-

schung übertragen wurde, als der Aufbau des Geologischen Dienstes für die drei westlichen Besatzungszonen erfolgte. Zusammen mit Franz Hallenbach hat er diesem Aufbau ganz wesentlich seinen Stempel aufgedrückt. Die Situation wiederholte sich noch einmal, als Ende 1958 die Bundesanstalt für Bodenforschung gegründet wurde und die Geophysik-Abteilung des Amtes für Bodenforschung im wesentlichen zum Niedersächsischen Landesamt für Bodenforschung kam. Speziell um den Aufbau der Seegeophysik in der Bundesrepublik und in der Bundesanstalt hat er sich dabei verdient gemacht. In seiner gesamten aktiven Zeit war er Mitglied der DFG-Senatskommission für Meeresforschung und viermal selbst Fahrtleiter auf dem Forschungsschiff Meteor. Von 1961 bis 1963 war Hans Closs 1. Vorsitzender der Deutschen Geophysikalischen Gesellschaft. Als Gründungsmitglied des Forschungskollegiums der Physik des Erdkörpers (FKPE) hat er zur guten Kooperation der geophysikalischen Institute und Abteilungen von Institutionen beigetragen. Noch heute beneiden viele Geologen die Geophysiker um dieses Gremium, das immer wieder zu fruchtbarem Gedankenaustausch zusammenfindet und neue interessante Projekte ins Leben ruft.

Im FKPE wurde auch die Idee eines seismologischen Zentralobservatoriums der Bundesrepublik angeregt. Eine amerikanische Station, die zur Nukleardetektion bei Gräfenberg in der Fränkischen Schweiz aufgebaut worden war, wurde der deutschen Wissenschaft angeboten. Hans Closs fand mit dem Bundesministerium für Wirtschaft einen Weg, einem solchen Observatorium die haushaltsmäßige Basis zu geben. Die Deutsche Forschungsgemeinschaft finanzierte die Investitionen zu einer der modernsten seismologischen Array-Stationen. Wieviel Überzeugungsarbeit steckte darin, alle beteiligten Stellen zu diesem Kooperationsprojekt zu bringen. Und solche Arbeit zu leisten, darin war Hans Closs ein Meister!

Das Bild des Wissenschaftlers und Wissenschaftsorganisations Hans Closs wäre unvollständig, wenn es nicht durch Ausführungen über seine Persönlichkeit ergänzt würde. Dies ist naturgemäß sehr schwierig und dürfte von Verschiedenen sehr unterschiedlich gesehen werden. Am hervorsteckendsten dürfte unbestritten seine Begeisterungsfähigkeit sein – die Fähigkeit, sich, vor allem aber auch seine Mitarbeiter, für eine Aufgabe zu begeistern und zu motivieren. Hinzufügen möchte ich seine Fähigkeit, Kritik nicht nur zu geben, sondern auch zu nehmen und zu berücksichtigen. Überraschend war für Außenstehende sein persönlicher bescheidener Lebensstil, den er auch nicht ablegte, als er zu Ehren kam.

In Anerkennung seiner Verdienste wurde er zum ersten Vorsitzenden der neu gegründeten Europäischen Geophysikalischen Gesellschaft gewählt. Viele andere Ehrungen wurden ihm zuteil. Ab 1965 war Hans Closs Mitglied der Deutschen Akademie der Naturforscher Leopoldina in Halle. 1970 verlieh ihm der Bundespräsident das Verdienstkreuz. 1971 ernannte ihn die Universität Bonn zum Honorarprofessor. Die Deutsche Geologische Gesellschaft verlieh ihm 1973 die Hans-Stille-Medaille. 1981 ernannte die Deutsche Geophysikalische Gesellschaft ihn zu ihrem Ehrenmitglied.

Er hat viele Freunde und Bewunderer zurückgelassen, die ihm lange über den Tod hinaus ein ehrendes Andenken bewahren werden.

Hans-Jürgen Dürbaum

Herbert Flathe (1919–1983)



Am 6. August 1983 starb Dr. Herbert Flathe, Direktor und Professor i.R. in der Bundesanstalt für Geowissenschaften und Rohstoffe, apl. Professor der Technischen Universität Braunschweig. Mit ihm ist ein hochangesehener Wissenschaftler und eine markante Persönlichkeit von uns gegangen. Herbert Flathe hat sich um die Entwicklung eines Spezialgebietes der Geophysik, der Geoelektrik, große Verdienste erworben.

Am 1. Juni 1919 ist er in Lengerich/Westfalen als einziger Sohn des Kaufmanns Martin Flathe geboren. Er besuchte dort die Volksschule, danach das Gymnasium in Osnabrück, wo er das vorgezogene Abitur mit „Auszeichnung“ bestand. Erst 17jährig konnte er 1937 mit dem Studium der Mathematik und Physik beginnen, zunächst im nahegelegenen Münster, dann in Hamburg.

Im Frühjahr 1939 wurde Flathe zum Reichsarbeitsdienst, bei Kriegsbeginn zur Wehrmacht eingezogen. Er war

in Frankreich, danach in Rußland an der Front und wurde dort verwundet. Zuletzt war er Ausbilder im Leutnant-Rang. Im Juli 1945 wurde er aus englischer Kriegsgefangenschaft entlassen. Er nahm im Sommersemester 1946 in Göttingen das Studium wieder auf und beendete es 1949 mit dem Staatsexamen in den Fächern Mathematik, Physik und Philosophie mit der Note „Auszeichnung“. Danach kehrte er nochmals nach Münster zurück, um 1951 bei Heinrich Behnke, der als Begründer einer Mathematik-Schule schon damals einen internationalen Ruf genoß, mit magna cum laude zu promovieren; Thema „Approximation analytischer Funktionen auf nichtgeschlossenen Riemannschen Flächen“. Ein Jahr zuvor heiratete er Irmgard Peersmann; aus dieser Ehe gingen drei Kinder hervor.

Mathematik und Philosophie spielten die vorherrschende Rolle im Studium von Herbert Flathe, und seine Fähigkeiten sprachen für einen Verbleib an der Hochschule. Der 32jährige entschied sich jedoch für eine Tätigkeit, bei der die praxisorientierte Anwendung der Physik dominierte. Hans Closs gewann den begabten Theoretiker für den Aufbau der Abteilung Geophysik des neuen Amtes für Bodenforschung in Hannover. Flathe war bereit, praktischer Feldgeophysiker zu werden, und brauchte nicht lange, die Umstellung zu meistern.

In den 50er Jahren hatte er bereits eine große Zahl von Objekten in allen Teilen des Bundesgebietes mit seinem Meßtrupp geoelektrisch untersucht; Ostfriesland, das Alpenvorland, das Oberrheintal und später West-Holstein waren bevorzugte Arbeitsgebiete. Vielfach führten seine Ergebnisse zur Standortfindung für ein Wasserwerk. Seine auf vielen hundert Messungen beruhende Festlegung der Süßwasser-/Salzwassergrenzlinie in der nordwestdeutschen Küstenregion hat im Prinzip heute noch Gültigkeit. Die Darstellung der Ergebnisse hatte stets eine besondere, den Autor Flathe anzeigende individuelle Note – sie hatte einen vorbildlichen Stil, war klar und verständlich auch in der Erklärung schwieriger physikalischer Vorgänge für den Nichtfachmann. Flathe orientierte sich besonders an den Wünschen und Forderungen der Geologen und suchte die Teamarbeit mit echter Partnerschaft.

Für die Verbesserung der Meßapparatur und der gewissenhaften Ausführungen der Feldarbeiten trug er – auch in den späteren Jahren noch – Sorge; schlechte Meßergebnisse, für die er einen untrüglichen Blick hatte, ließ er nicht gelten. Es nimmt nicht Wunder, daß er zu den Technikern, die seinen hohen Ansprüchen genügten, ein besonders herzliches Verhältnis hatte.

Als Hilfsmittel zur graphischen Auswertung von geoelektrischen Meßkurven war nach dem Kriege vom Reichsamt für Bodenforschung ein recht unvollständiger 3-Schichtmodell-Kurvenatlas gerettet, dessen Neubearbeitung und Erweiterung dringend erforderlich waren. Mitte der 50er Jahre gelang dem Mathematiker Flathe die Entdeckung der Möglichkeit, auf einfache Weise kompliziertere Schichtaufbaue für die geoelektrische Widerstandsmethode modellmäßig darzustellen. Eine exakte Überprüfung von Meßmaterial, das mehr als 3 Schichten birgt, war nun möglich. Für diese 1955 in Paris vor der European Association of Exploration Geophysicists vorgetragene Arbeit "A practical method of calculating geoelectrical model graphs for horizontally stratified media" erhielt er 1956 den Conrad-Schlumberger-Preis.

Zur Bewertung dieser Leistung muß man sich in die damalige „computerlose“ Zeit zurückversetzen. Um die Interpretationstechnik in der Geoelektrik voranzutreiben, bedurfte es Modellkurven für komplizierten Schichtenaufbau. Das Verfahren Flathe erlaubte die Berechnung derartiger Kurven mit Hilfe der damals ausschließlich zur Verfügung stehenden Tischrechenmaschinen. Flathe zeigte: Jede geoelektrische Modellkurve für den allgemeinen n-Schicht-Fall mit extrem leitfähigem bzw. nicht-leitendem Basement kann durch eine einfache Linearkombination von „Primelementkurven“ für drei Schichten zusammengesetzt werden. Die auftretenden Gewichte ergeben sich aus den Nullstellen gewisser Polynome, die in der Kernfunktion des „Stefanescu-Integrals“ enthalten sind. Die Anwendung dieser Technik führt zur Diskussion von Fällen mit mehr als 3 Schichten, die sich für die Interpretation von Meßkurven für hydrogeologische Problemstellungen als außerordentlich nützlich erwiesen. Einen Modellkurvenatlas für 5-Schichtfälle hat er hierzu selbst berechnet.

1958 wurde Flathe beim Amt für Bodenforschung zum „Landesgeologen“ auf Lebenszeit ernannt bei gleichzeitiger Übernahme in die Bundesanstalt für Bodenforschung. Er übernahm dort die Leitung des Referates „Allgemeines geophysikalisches Laboratorium“ und 1961 der neu eingerichteten Unterabteilung „Geophysikalisches Laboratorium“. 1966 wurde er zum Direktor und Professor an der Bundesanstalt ernannt. 1969 erhielt er zudem für seine außerordentlichen wissenschaftlichen Leistungen das Bundesverdienstkreuz am Bande.

Mehrmonatige Auslandseinsätze im Rahmen der technischen Zusammenarbeit bewiesen sein Können, die Widerstandsmethode als das Instrumentarium zur Grundwassererkundung einzusetzen. Zu nennen sind Einsätze in Paraguay (Chaco, 1959), Indonesien (Java, Madura, Bali, 1961), Jordanien (Jordantal, 1963/64), Singapore (1970), Griechenland (Sperchios Tal, 1972/73), Malaysia (Kelanton-Delta, 1974) und Indien (Mahanadi-Delta, 1976). Fast immer waren diese Einsätze verknüpft mit der Aufgabe, einheimische Wissenschaftler, Studenten und Techniker an die Geophysik heranzuführen oder über neueste Entwicklungen auf dem Gebiet der Geoelektrik zu berichten. Seine pädagogischen Fähigkeiten konnten sich dabei entfalten.

1963 nahm Flathe erstmalig einen Lehrauftrag für „Angewandte Geoelektrik“ an der Technischen Universität Braunschweig wahr, habilitierte sich 1965 im Fach „Angewandte Geophysik“ und wurde 1970 dort zum apl. Professor ernannt. Für viele Studenten, die durch Flathe's Ausbildung gegangen sind, wurde auf diese Weise der Weg zur Praxis eröffnet. Dank seiner guten Englisch-Kenntnisse

übernahm er gern Gastvorlesungen im Ausland, so an der Universität Karachi/Pakistan (1965), an der Universität Aarhus/Dänemark (1966), Universität Singapore (1970), Technische Universität Lulea/Schweden (1975), an der Benares Hindu-Universität/Indien (1976), an der School of Mines, Madrid/Spanien (1978) und Colorado School of Mines/USA (1978). Für mehrere internationale Trainingskurse wurde er von der UNESCO und anderen Organisationen als Lehrer angefordert (Kabul 1967, Delft 1969, Budapest 1969, Erice 1975, Clausthal 1976). Hieraus erwachsen viele bleibende Kontakte zu ausländischen Wissenschaftlern, mit denen er oft freundschaftlich verbunden war.

Mitte der 60er Jahre stellte Flathe sich eine Aufgabe, die ihn für ein Jahrzehnt wissenschaftlich fesseln sollte, nämlich die Förderung der Tiefengeoelektrik. Diese Arbeiten hatten zum Ziel, die geoelektrische Widerstandsmethode für große Tiefenaufschlüsse zu entwickeln und zu nutzen und die dabei gewonnenen Ergebnisse theoretisch zu deuten. Der Anstoß kam 1962 auf dem von der Deutschen Forschungsgemeinschaft veranstalteten Symposium „Erdmagnetische Tiefensondierung“. Man erwartete von der Widerstandsmethode Beiträge zur Krustenforschung in geringerer Tiefe, wo andere elektromagnetische Methoden damals keine Aussage zuließen. Höhepunkt der technischen Arbeiten des von ihm geleiteten Forschungsteams war 1967 die Tiefensondierung „Rheingraben“, bei der eine im Bau befindliche Hochspannungsleitung zwischen Lörrach und Karlsruhe auf einer Länge von 150 km für eine Schlumberger-Messung genutzt wurde. Für die Analyse der Meßergebnisse bediente er sich des Mailletschen Begriffs der Anisotropie, den er in den Mittelpunkt gestellt hat, um die Unterschiede von Bohrergebnissen einerseits und Ergebnissen einer Tiefensondierung andererseits zu deuten. Seine Arbeiten zu diesem Thema, insbesondere über die Deutung der 1970 ausgeführten Tiefensondierung an der Forschungsbohrung „Markdorf 2“, haben grundsätzliche Bedeutung gewonnen.

Anfang der 70er Jahre mußte Flathe eine nicht leichte Entscheidung fällen, die sein weiteres Berufsleben bestimmen sollte. Die von ihm geleitete Unterabteilung „Geophysikalische Laboratorien“ sollte in einem neuen Organisationsplan der Bundesanstalt als Abteilung „Geophysik“ ausgewiesen werden. Hans Closs ging 1972 in den Ruhestand und hinterließ eine auf mehr als 100 Mitarbeiter angewachsene Abteilung. Im Konflikt zwischen der Notwendigkeit, zunehmend administrative Aufgaben zu bewältigen, und der Möglichkeit, vorwiegend wissenschaftlich zu arbeiten, entschied er sich – wohl auch in Anbetracht seiner angegriffenen Gesundheit – dafür, in einem neu eingerichteten Referat „Geophysikalische Grundlagen“ seiner dominierenden Neigung nachzugehen.

Flathe nutzte die folgenden Jahre, um mit den neuen technischen Möglichkeiten, auf rasche Weise Modelle berechnen zu können, tiefer in die Interpretation und geologische Deutung der geoelektrischen Feldmessung einzudringen. Zum Inventar seines Dienstzimmers gehörte fortan eine Rechenanlage. Das von jüngeren Kollegen erarbeitete Verfahren zur automatischen Optimierung von Schichtmächtigkeiten und -widerständen aus den Meßdaten wandte er erstmalig auf Großobjekte an. Für ihn war es eine Freude, an der Lösung dieses „Umkehrproblems der Geoelektrik“ mitwirken zu können, für deren praxisgerechte Anwendung er seine Erfahrungen einbringen konnte.

Die insgesamt 36 Titel umfassenden Veröffentlichungen

sind im Festband E19 des Geologischen Jahrbuchs, der ihm zu seinem 60. Geburtstag gewidmet wurde, verzeichnet. Darunter sind vier (mit Koautoren verfaßte) größere Darstellungen: „Die geoelektrischen Verfahren der Angewandten Geophysik“ (in: „Lehrbuch der Angewandten Geologie“ von A. Bentz 1962), „Geoelektrik in der Wassererschließung“ (in: „Die Wassererschließung“ von H. Schneider 1972), „Geophysikalische Untersuchungen“ (in: „Abriß der Hydrogeologie“ von W. Richter und W. Lillich 1975) und „The smooth sounding graph“ (ein mit W. Leibold verfaßten Leitfaden 1976). Besonders dieses sind Werke, in denen man einen Wissenschaftler erkennt, der zu angewandter Forschung und Lehre gleichermaßen fähig ist.

Herbert Flathe zog sich in den letzten Jahren seines Lebens – gesundheitsbedingt – mehr und mehr aus der wissenschaftlichen Öffentlichkeit zurück. Wer ihn jedoch seit vie-

len Jahren kannte, sah in ihm nach wie vor den hochintelligenten Menschen, mit dem sich zu unterhalten vielfach großen Gewinn brachte. Viele seiner Freunde denken an faszinierende Diskussionen über Themen der Geschichte, Kunst, Theologie und Religionsphilosophie zurück. Seine tiefe Religiosität dokumentierte er mit selbstverständlicher Offenheit. Sein sehr bestimmt vorgetragenes, präzises Wissen war das tragende Element eines jeden Gespräches. Witz und geistreiche Einfälle würzten die Unterhaltung. Die vielen Eindrücke, die sein Berufsleben ihm schenkte, erlebte er intensiv und dankbar. Seinem Wunsch, Reisen in die Vergangenheit mit seiner Frau auszuführen und das Erlebte zu vertiefen, wurde schon ein halbes Jahr nach seiner vorgezogenen Pensionierung ein jähes Ende gesetzt.

Joachim Homilius

Book reviews

D.W. Collinson: Methods in Rock Magnetism and Palaeomagnetism, Techniques and Instrumentation Chapman and Hall, London New York, pp 503, 1983

During the last years palaeomagnetic investigations have become standard methods in many fields of geosciences. To fulfill the increasing pretension in the reliability of the results, more and more rockmagnetic experiments are made besides the routine palaeomagnetic measurements. Most instruments and methods which are used for the determination of rock magnetic properties were developed in research institutes and are not available from commercial companies. Constructional details and experiences often are not published or difficult to find. So many institutes spend time and manpower in constructional projects, a rather unnecessary and expensive way of obtaining own experiences. The most valuable aspect of the book in review is the compilation and preparation of all this information.

In my opinion this succeeded in a very clear manner. Some chapters deal with certain magnetic properties, as for instance magnetic susceptibility or NRM. After definition of a property, it is explained what it depends on, and which is the range of values to be expected from theory and from measurements already made. So the reader is enabled to choose from the following part of the chapter the measuring method satisfying his individual needs. As the precision and sensitivity, as well as constructional hints of different instruments are discussed in detail, it is possible to decide which one could be suitable and could be built with the existing facilities.

Other chapters contain a valuable amount of design criteria for standard devices in a palaeomagnetic laboratory, as controlled fields and field-free spaces, or even sampling of rocks. For a first judgement of the properties of such equipment there are given thumb rules and for comparison the specifications of many existing types are listed.

Finally, the book provides a brief introduction to fields as presentation and treatment of data, including extensive reference guides. Therefore, it is as well valuable for the experienced investigator as for the student starting with palaeomagnetism or rock magnetism. Unfortunately, the book is quite expensive so it will be found only in the libraries, and when asking for it, surely it will be already lent to someone else.

Harald Böhnel

H. Militzer und F. Weber: Angewandte Geophysik. Band 1. Gravimetrie und Magnetik. Springer Verlag, Wien/New York

Da seit dem Erscheinen von Haalcks Lehrbuch in den fünfziger Jahren keine deutschsprachige Gesamtdarstellung der angewandten Geophysik mehr erschienen ist, der wissenschaftliche Fortschritt auf diesem Gebiet seither aber sehr stürmisch verlaufen ist, ist die Idee von H. Militzer und F. Weber, ein neues Lehrbuch der angewandten Geophysik herauszubringen, nur zu begrüßen.

Dieses Lehrbuch soll 3 Bände enthalten, von denen der erste über Gravimetrie und Magnetik soeben erschienen ist. Band 2 soll den Titel „Geoelektrik-Geothermik-Radiometrie-Aerogeophysik“ und Band 3 den Titel „Seismik“ tragen. Diese Bände sollen im Abstand von je einem Jahr folgen.

Der Band „Gravimetrie und Magnetik“ beginnt mit einem Kapitel über die theoretischen Grundlagen der beiden Verfahren. Hier wird besonders deutlich, daß das Buch nicht für theoretische Puristen geschrieben ist, die gerade von diesem Kapitel sicherlich enttäuscht sein werden. Der eher anwendungsorientierte Leser, der sich mit den beiden Verfahren vertraut machen will, wird aber den Verzicht auf die bei mehr theoretisch orientierten Darstellungen übliche abstrakte und kompakte Schreibweise freudig begrüßen. Auch den Verzicht auf eine das wesentliche oft verstellende läng-

liche mathematische Beweisführung und deren Ersatz durch eine Vielzahl von Beispielen wird dieser Leser bestimmt nicht als unangenehm empfinden. Als gelungen muß auch der Versuch gewertet werden, die Gemeinsamkeit der theoretischen Grundlagen beider Verfahren herauszustellen.

Es folgen je ein Kapitel über angewandte Gravimetrie und angewandte Magnetik, in denen die geophysikalischen, geologischen und ingenieurtechnologischen Grundlagen am Anfang kurz umrissen sowie die Meßprinzipien kurz beschrieben werden. Anschließend werden klare Konzepte zur Durchführung der Messungen und der notwendigen Korrekturen und Reduktionen der gemessenen Daten beschrieben.

Während viele (einschließlich der englischsprachigen) Darstellungen der Gravimetrie und Magnetik damit fast abgeschlossen sind und der Interpretation der Daten nur wenige Seiten gewidmet werden, folgen bei H. Militzer und F. Weber noch ausführliche Kapitel über Paläo- und Archäomagnetik, Bearbeitung und Interpretation der gravimetrischen und magnetischen Meßergebnisse sowie über Beispiele komplexer gravimetrischer und magnetischer Untersuchungen, in denen anhand vieler Beispiele beschrieben wird, was man mit gravimetrischen und magnetischen Meßdaten alles „machen“ kann und wie man es „macht“.

Da der Inhalt des Buches sowohl sachlich als auch didaktisch sehr gut durchdacht ist, kann man Kritik nur an wenigen Details üben. So wäre es für den Anfänger zur Erklärung des Verlaufs des Geoids unter den Kontinenten bestimmt nützlich, das unter Geophysikern wohl allgemein bekannte Beispiel anzuführen, bei dem ein fiktives Kanalnetz die Kontinente durchzieht, in dem sich wie bei den Ozeanen die Wasseroberfläche entlang des Geoids einstellen würde. Dies könnte man dann noch durch eine kleine Zeichnung, bei der anhand eines Beispiels der Verlauf des Niveausphäroids, des Geoids und der Topographie verglichen wird, illustrieren. Ebenfalls hilfreich wäre sicherlich ein qualitatives Diagramm, bei dem die gegeneinander wirkenden Drehmomente über dem Ablenkungswinkel des Gravimetergestänges aufgetragen sind, um den Begriff der Astasierung zu verdeutlichen. Etwas zu sparsam ist für meinen Geschmack die Beschreibung des Absorptionszellenmagnetometers ausgefallen.

Da die Autoren das Prinzip „short writing makes long reading“ beherzigt haben und deshalb in ihren Darstellungen entsprechend ausführlich sind, ist das Buch leicht zu lesen, ohne daß dies mit einem Verlust an Niveau erkauft wird. Der Student, der autodidaktisches Arbeiten anhand von Büchern dem Vorlesungsbesuch vorzieht, bekommt durch dieses Buch ein umfassendes Wissen der Gravimetrie und Magnetik vermittelt, so daß jedem Studierenden der angewandten Geophysik trotz des nicht unbedingt auf die finanziellen Verhältnisse eines Studenten abgestimmten Preises (vielleicht könnte der Verlag die Herausgabe einer billigeren Paperback-Ausgabe in Erwägung ziehen) der Erwerb des Buches nur empfohlen werden kann. Da man auch einzelne Abschnitte lesen kann, ohne dauernd in vorderen Kapiteln die Bedeutung einer Größe oder eines Begriffes nachschlagen zu müssen, ist das Buch auch für denjenigen zu empfehlen, der mit der Materie bereits vertraut ist und gelegentlich sein Wissen zu dem einen oder anderem Punkt auffrischen will.

Detlev Doan

Berry, L.G., Brian Mason, R.V. Dietrich: Mineralogy, Concepts, Descriptions, Determinations. 2nd Ed., W.H. Freeman and Company, San Francisco 1983, pp 561

Diese von R.V. Dietrich neu bearbeitete Auflage der 1959 erstmalig erschienenen Mineralogy von L.G. Berry und Brian Mason ist vornehmlich als Lehrbuch für Studienanfänger der Mineralogie konzipiert. Es soll aber auch Geologen, Petrologen, Bodenkundlern und Festkörperphysikern als Nachschlagewerk dienen. Die Bespre-

chung in einer geophysikalischen Zeitschrift sollte darlegen, welchen Nutzen dieses Buch für einen Geophysiker haben könnte. Zunächst: für deutschsprachige Leser gibt es auf mineralogischem Gebiet – im Gegensatz zur geophysikalischen Literatur – sehr gute und moderne Lehrbücher. Hier wäre das Buch u. a. ergänzend zum Nachschlagen der englischen Terminologie nützlich. Doch kann dieses Buch auch einem Geophysiker nicht nur als Nachschlagewerk, sondern durchaus zur Aneignung mineralogischer Grundbegriffe dienen. Allerdings liegt der Schwerpunkt auf der Kristallographie und Kristallchemie einschließlich der ausführlichen Beschreibung von rd. 200 Mineralien, während ihre gesteinsbildenden Aspekte nur gestreift werden. Was z. B. den Stoffbestand der Erdkruste betrifft, wird hier von über 15 Jahre alten Vorstellungen der Krustenstruktur ausgegangen, die für die kontinentale Unterkruste noch eine durchgehende Basaltschicht annahmen. Auch werden nur Mittelwerte der gesamten Kruste ohne Aufgliederung in kontinentale und ozeanische Kruste gegeben, was aber leider verschwiegen wird. Der Text suggeriert vielmehr den Eindruck, als sei von der kontinentalen Kruste die Rede. Erst ein Datenvergleich mit den zitierten Originalarbeiten klärt darüber auf, daß sich die Ausführungen auf die Gesamtkruste beziehen.

Der kristallographische Teil geht vom Gitteraufbau aus und leitet in anschaulicher Weise die 14 Bravais-Gitter mit ihren Symmetrieeigenschaften ab. Darauf werden die Beziehungen zwischen Gitter und Kristallform entwickelt, die Nomenklaturen für Ebenen und Achsen eingeführt und ihre Messung und Darstellung in stereographischen Projektionen behandelt. Eine ausführliche Beschreibung der 32 Kristallklassen, der Gitterdefekte und der Zwillingbildungen schließt sich an. Das 3. und 4. Kapitel bringt die chemischen und physikalischen Eigenschaften der Minerale, das 5. Kapitel ihre Genese. Die Methoden zur Bestimmung von Natur und Gitterstruktur der Minerale werden im 6. Kapitel beschrieben. Das 7. Kapitel skizziert auf nur knapp 3 Seiten allgemeine Gesichtspunkte der Klassifikation und Namensgebung, während die 8 Kapitel des Teils II etwa 200 Mineralbeschreibungen nach den acht chemischen Klassen enthalten. Im Teil III finden sich Bestimmungstabellen, die nach den Merkmalen Glanz, Farbe und Härte geordnet sind. Jedem der ersten 7 Kapitel sind Listen weiterführender Literatur angefügt.

Das Buch ist gut ausgestattet. Die zumeist klare und anschauliche Darstellung wird durch eine Fülle guter Abbildungen und Photographien (leider nicht farbige) unterstützt.

K. Strobach

Falck-Ytter, Harald: Das Polarlicht, Nordlicht und Südlicht in mythischer, naturwissenschaftlicher und apokalyptischer Sicht. Mit Photographien von Torbjörn Lövgren. Verlag Freies Geistesleben, Stuttgart, 195 S., 1983

Ursprünglich war dieses hervorragend ausgestattete Buch, wie es in der Vorbemerkung heißt, vorwiegend als Bilddokumentation des Polarlichts gedacht. Doch auch in der nun vorliegenden Publikation sind die Bilder das Herausragende. Sie wurden von Torbjörn Lövgren, einem Mitarbeiter am Geophysikalischen Institut für Polarlichtbeobachtungen in Kiruna in Schweden aufgenommen und vermitteln als 21 ganzseitige Farbwiedergaben im Format 21 × 23 cm dem Leser einen faszinierenden Eindruck von einer den meisten Menschen in dieser Pracht unbekanntem Naturerscheinung. Zwei Farbtafeln enthalten eine Serie von 8 Aufnahmen im Abstand von 30 Sekunden, um einen Eindruck der zeitlichen Gestaltänderungen zu geben. Technische Daten zu den Photographien werden leider nicht genannt; die Belichtungszeiten dürften nach den mitphotographierten Sternspuren zwischen 10 und 20 Sekunden liegen.

Außer diesen hervorragenden Aufnahmen wurden noch 8 farbige Nordlicht-Lithographien des dänischen Malers Harald Moltke, die eindrucksvoll und von künstlerischem Wert sind, sowie ein farbiger Holzschnitt von Fridtjof Nansen aufgenommen. Einige Südlichtbilder als Stiche bzw. Zeichnungen erinnern daran, daß das Polarlicht beide Polkappen umfaßt. Sehr gute Photos der Sonnenkorona zur Zeit der Fleckenmaxima und Minima sowie eine Reihe von Graphiken, ferner historische Darstellungen von Nordlichtern tragen zur Veranschaulichung des Textes bei.

Ytter beginnt seine Ausführungen mit Schilderungen von Nordlichtabläufen, wobei auch das Problem von begleitenden Geräuschercheinungen angesprochen wird. Aber schon Alexander von Humboldt mußte feststellen: „Die Nordlichter scheinen weniger geräuschvoll geworden zu sein, nachdem ihr Auftreten genau registriert wird“. Die Schilderung des Polarlichts im mythischen Bewußtsein verschiedener Naturvölker, dann als Gegenstand der Natur- und Geistesgeschichte und endlich der modernen Naturwissenschaft schließt sich an. Dabei wird auch die Sonnenphysik ausführlich einbezogen. Das Anliegen Ytters geht allerdings über die Darstellung der rein physikalischen Inhalte hinaus. Er möchte die Naturtatsachen am Beispiel des Polarlichts als äußere Erscheinungen eines Geistig-Wesenhaften schildern, angesichts der fast durchweg materialistischen Auffassungen heutiger Naturwissenschaft ein berechtigtes Ziel. Ytter versucht dies dadurch, daß er Angaben aus den Schriften und Vorträgen Rudolf Steiners mit Inhalten der Polarlicht- und Sonnenforschung verknüpft. Die durchaus eigenen Vorstellungen, die er dabei entwickelt, sind jedoch für den Leser kaum nachvollziehbar, weil sie weder genügend logisch entwickelt noch erkenntnistheoretisch fundiert werden, insbesondere dort, wo Ytter bekannte Begriffe in einem abgewandelten Sinne gebraucht.

Das letzte Kapitel, das kommentierte Auszüge aus dem Epos „Nordlicht“ des Dichters Theodor Däubler enthält, entzieht sich eigentlich einer Besprechung in einer naturwissenschaftlichen Zeitschrift. Nur soviel: Däubler hat das Nordlicht vermutlich nie mit eigenen Augen gesehen, doch in 33000 Versen besungen. – Das Gute an Büchern ist, daß man sie nicht vollständig lesen muß.

K. Strobach

Lester C. King: Wandering Continents and Spreading Sea Floors on an Expanding Earth. John Wiley and Sons, Chichester, 232 pages, 1983

The author, a geologist from the university of Natal, South Africa, was involved mainly in studying the morphology of the earth, and now, as Emeritus, he “take(s) up the pen once more to re-scan the face of the globe” (see preface). Thus this book is rather a treatise than a textbook, covering the author’s experiences, beliefs and conclusions drawn throughout a scientific life.

The development of the earth is described from the beginning under the scope of plate tectonics. But in contrary to currently held beliefs the author explains drift as lateral motion caused by radial movements within the earth, and he promotes the theories of decreasing gravity and an expanding earth.

As intended, the geological history of the earth is scanned and highlighted with special topics and examples, which are often belonging to Africa, the author’s native place. The book is easy to read, although a basic knowledge in geology and geophysics is necessary to understand and to interpret the conclusions drawn by the author. The figures are drawn carefully, and a comprehensive list of references is given.

G. Jentzsch

Crustal structure beneath the Swabian Jura, SW Germany, from seismic refraction investigations*

D. Gajewski and C. Prodehl

Geophysical Institute, University of Karlsruhe, Hertzstr. 16, D-7500 Karlsruhe 21, Federal Republic of Germany

Abstract. As part of multidisciplinary investigations, combined refraction and reflection seismic experiments were carried out in the area of the geothermal anomaly of Urach, southwest Germany, in 1978 and 1979. Ten refraction and two reflection seismic profiles were recorded which reached maximum distances between 20 and 150 km.

This report presents a detailed velocity-depth model of crust and uppermost mantle, based on the available seismic refraction data. On all profiles a clear P_g phase with velocity 5.6 km/s is observed refracted within the uppermost crystalline basement. On some profiles the P_g phase is followed by a retrograde reflected or diving wave which has an apparent velocity of 6.1 km/s. The amplitudes of the corresponding prograde phase decrease considerably or disappear completely at shotpoint distances greater than 50 km. It is caused by a low-velocity zone with a mean velocity of 5.8 km/s extending from 8 to 19 km in depth and over more than 100 km in horizontal direction along the main profile parallel to the strike of the Swabian Jura. The lower boundary of the low-velocity zone is determined by a reflected phase which varies in quality depending on the inclination of the reflecting boundary. The western boundary of this low-velocity zone is located near the north-eastern edge of the Hohenzollerngraben. To the east it disappears west of the Nördlinger Ries. The structure of the crust southwest of the area containing the low-velocity zone is quite different from the structure towards the NE. The Moho is a first-order discontinuity dipping from 25 km depth in the SW to 26.5 km in the NE. Within the uppermost mantle high velocities (≥ 8.4 km/s) and strong velocity gradients are observed. A gravity investigation suggests high densities for the area of the low-velocity zone. This may be caused by fluid or gas phases or partially molten material, an explanation which is also supported by shear-wave observation.

Key words: Crustal structure – Swabian Jura – Urach geothermal anomaly – Seismic refraction – Low-velocity zone

Introduction

As part of a multidisciplinary investigation of the geothermal anomaly of Urach in southwest Germany, including

* Contribution No. 275 (55/SFB 108), Geophysical Institute, University of Karlsruhe

Offprint requests to: C. Prodehl

the 3334-m-deep drillhole Urach 3 (Haenel, 1982), detailed seismic reflection and refraction experiments were carried out in July 1978 and May 1979. Figure 1 shows the location of 10 seismic refraction lines varying in length from 20 to 150 km. The map of Fig. 15 shows the position of two detailed seismic reflection lines which were combined with special wide-angle observations. A description of the seismic refraction experiment and its data preparation is given by Jentsch (1980) and Jentsch et al. (1982). The reflection experiment and its interpretation are described by Bartelsen et al. (1982) and Meissner et al. (1982), the wide-angle survey of the same line by Trappe (1983).

The main purpose of the seismic refraction experiment was to determine depth contours, structure and velocity of the crystalline basement and to investigate whether the causes of the Urach geothermal anomaly might be located within the uppermost 5 km of the crust and whether the resolution of the velocity determination might be accurate enough (Jentsch et al., 1982; Prodehl et al., 1982). Accordingly, the shotpoints were arranged such that an optimal coverage, as well as several 20- to 30-km-long reverse lines, resulted in an average station spacing of 1–2 km.

Some of the profiles were extended to observation distances beyond 100 km, but beyond 30 km the station spacing was 4–5 km on average. In general, the quantity of data is sufficient to allow a quite detailed interpretation. However, it is evident that the arrangement of shotpoints and profiles may not be ideal for investigations of the whole crust because there is no properly reversed long profile and only a few profiles overlap for 30 km at most. As a consequence, the resulting model (see Fig. 6) is not unique, but there are other possible solutions which cannot be ruled out with the data available.

Geology

The geothermal anomaly of Urach is located southeast of Stuttgart at the northern margin of the Swabian Jura. The anomaly has an oval shape, its long axis being directed from WSW to ENE as shown in Fig. 1 by isolines of the geothermal gradient after Carlé (1974).

The whole area of the Urach geothermal anomaly and its surroundings is covered by Mesozoic sediments. The long axis of the anomaly runs parallel to the so-called “Albtrauf”, an erosional escarpment formed by the abrupt northern termination of the 300- to 500-m-thick Malm plate (Upper Jurassic limestones) which forms the Swabian Jura

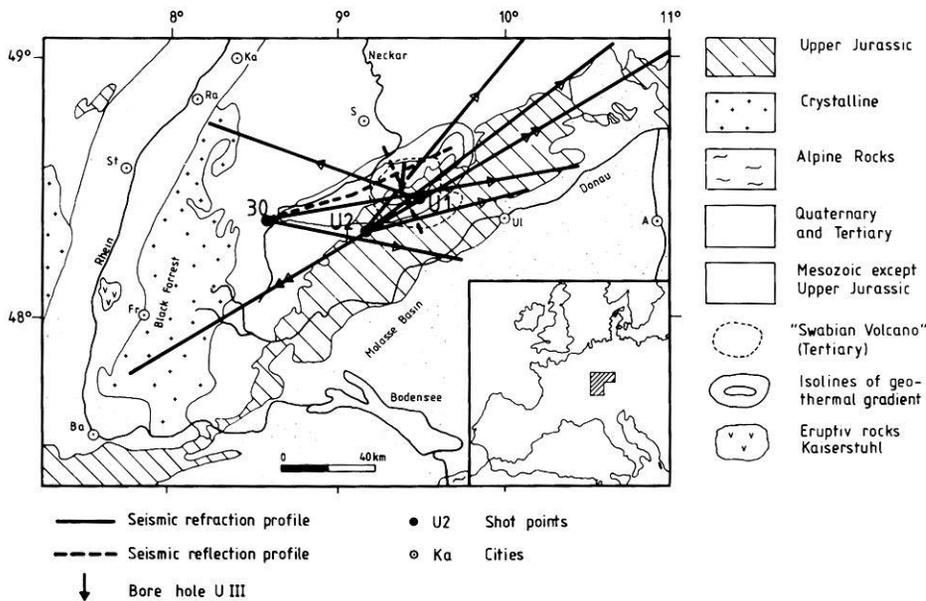


Fig. 1. Geological sketch map of southwest Germany showing the position of the seismic surveys in the area of the Urach geothermal anomaly. Cities: A Augsburg, Ba Basel, Fr Freiburg, Ka Karlsruhe, Ra Rastatt, S Stuttgart, St Strasbourg, Ul Ulm

mountains proper. This escarpment divides the area of investigation: the Swabian Jura in the south with maximum elevations of 850 m above sea level and the foreland to the north (Middle and Lower Jurassic and Triassic sediments) with maximum elevations of 450 m. More than 300 volcanic conduits (upper Miocene) of the so-called Swabian Volcano (Geyer and Gwinner, 1968; Mäussnest, 1974) have perforated the Mesozoic cover of this area. The basement consists mainly of granites and gneisses of the Variscan orogeny and has been found in some drillholes between 1 and 2 km depth below the surface, e.g. at 1.6 km depth in the Urach 3 drillhole mentioned above (Dietrich, 1982; Schädel, 1982).

The rise of the Black Forest and the Vosges in the west accompanied by the formation of the Rhinegraben and the subsidence of the Molasse Basin, i.e. the Alpine northern foreland, in the south are the most important tectonic features in the immediate neighbourhood of the South German Triangle within which the Urach geothermal anomaly is located.

One of the important fracture zones of the area of investigation is the (northwesterly striking) Hohenzollerngraben (Illies, 1982), which is also the site of major earthquake activity since the beginning of this century (Schneider, 1971; Turnovsky and Schneider, 1981) along a south-north trending line cutting through the Hohenzollerngraben near the crossing point of our main seismic refraction line, approximately 20 km WSW of shotpoint U2 (Fig. 1).

Correlation and interpretation of phases

As examples of the data recorded, the record sections of the four profiles U2-60, U1-60, U1-240 and U2-240 are shown in Figs. 2–5. U1 and U2 are shotpoints (Fig. 1), the number 60 or 240 is the azimuth of the corresponding profile. All records are normalized, i.e. the maximum amplitude of any record within the time window shown is scaled to a fixed width. With the exception of a few traces, the signal-to-noise ratio is generally sufficient. The travel-time curves shown in Figs. 2–5 are recalculated from the model shown

in Fig. 6. The phases are named in the same manner as that used by other authors (e.g. Giese et al., 1976; Prodehl, 1979).

For the distance range of the first 30 km from the shotpoints, where a very close station spacing was achieved, only selected traces have been plotted in the record sections of the long profiles as shown in Figs. 2–5. The detailed sections for this part are reproduced by Jentsch (1980). They were used for the correlation of phases concerning this distance range (see also Jentsch et al., 1982).

The arrivals a_0 belong to diving waves which travel in the sedimentary layers only. The apparent velocities vary from 3.8–4.2 km/s. With the data available it is not possible to deduce the fine structure of the sediments, but rather the existence of a gradual velocity increase with increasing depth was assumed. The phase a_0 is followed by a phase a'_0 in secondary arrivals which corresponds to a reflection from the crystalline basement underlying the sediments.

Beyond 8 km distance the first arrivals belong to the P_g phase, here named a_1 , which travels in the uppermost part of the crystalline basement. It gives velocities ≥ 5.6 km/s and can be correlated up to 30–40 km shotpoint distance. Jentsch et al. (1982) have used this phase for a time-term analysis in order to derive a depth contour map of the basement for the area of the Urach geothermal anomaly (Jentsch et al., 1982, Fig. 11).

At 30–40 km distance phase a_1 is overtaken by phase a_3 with an apparent velocity clearly above 6 km/s. The corresponding retrograde phase a_2 is less clearly visible in the record sections because it immediately follows phase a_1 in secondary arrivals and is, therefore, disturbed by interference with the first-arrival phase. This retrograde phase a_2 is produced by a depth range with strong positive velocity gradient in which the velocity increases from 5.7 to 6.05 km/s. The position of the travel-time curve system a_0 , a'_0 , a_1 , a_2 and a_3 varies from profile to profile indicating lateral variations within the upper crust (see also Jentsch et al., 1982).

Also, the amplitudes of phase a_3 characteristically vary through the area of investigation. On most profiles the energy of those phases decreases considerably between 50 and 70 km distance and may even vanish completely (see more

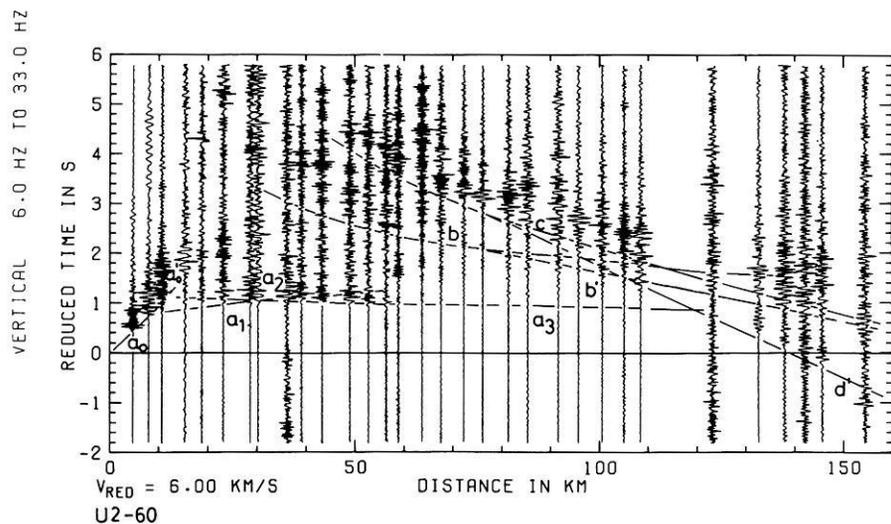


Fig. 2. Record section of the seismic refraction profile U2-60 observed from shotpoint U2 towards ENE (azimuth of 60°). a_0 , a_1 , a_3 , b' , d' refracted phases. a'_0 , a_2 , b , c reflected phases. The travel-time curves were calculated from the model of Fig. 6

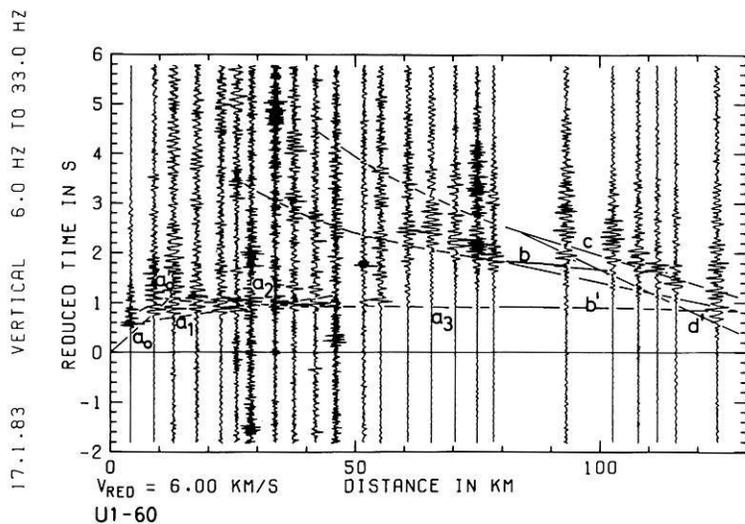


Fig. 3. Record section of the seismic refraction profile U1-60 observed from shotpoint U1 towards ENE (azimuth of 60°). For further explanations, see Fig. 2

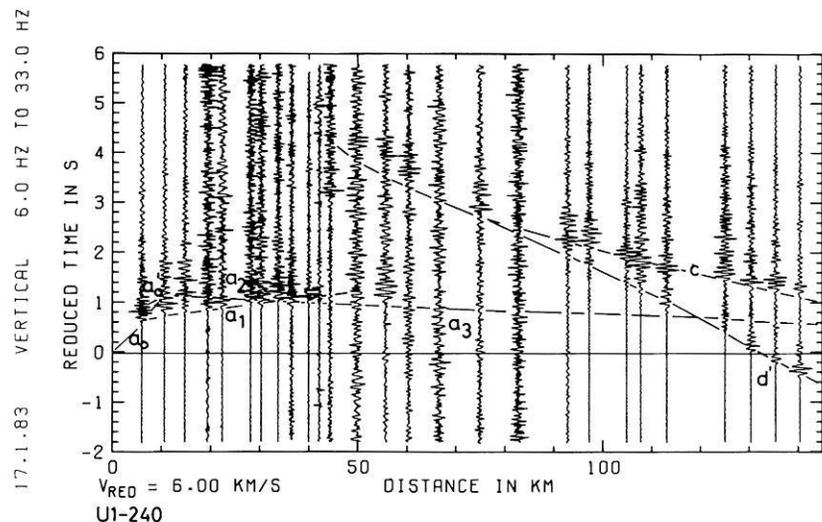


Fig. 4. Record section of the seismic refraction profile U1-240 observed from shotpoint U1 towards WSW (azimuth of 240°). For further explanations, see Fig. 2

detailed discussion below). It is explained by the existence of a pronounced low-velocity zone in the middle crust.

The lower boundary of this zone of velocity inversion in the middle crust is fixed by the retrograde travel-time branch of phase b . This phase varies from strong to weak

appearance in the record sections and is completely lost on profiles U1-240 and U2-240. On the other hand, phase a_3 can clearly be correlated to larger distances. Both observational features indicate that the low-velocity area terminates towards WSW. The variability of the amplitudes and

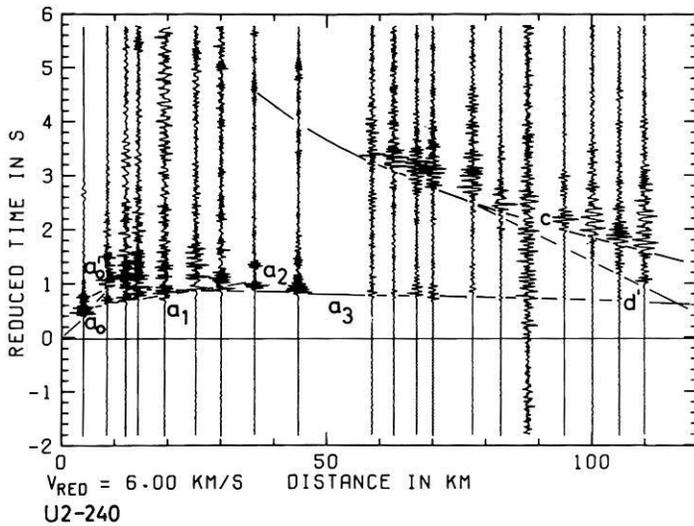


Fig. 5. Record section of the seismic refraction profile U2-240 observed from shotpoint U2 towards WSW (azimuth of 240°). For further explanations, see Fig. 2

of the travel time of phase *b*, where it exists, are explained by a strongly varying topography of the lower boundary of the low-velocity zone. For example, note the strong amplitudes of this phase on profile U1-60 in the distance range 50–80 km.

A prograde travel-time branch *b'* cannot safely be correlated. Either it does not exist or it is very weak and therefore hidden in the signal-generated noise of other phases. This leads to the conclusion that below the boundary determined by curve *b* the velocity gradient is only slightly positive or even negative.

On all record sections a phase *c*, interpreted as reflection from the crust-mantle boundary (commonly named $P^M P$, $P_M P$ or $P_m P$), can be clearly correlated over a wide distance range between 50 and 140 km. Only the profile U1-60 seems to be an exception. On the other hand, its apparent weakness could also be caused by the existence of the strongly developed phase *b* to which the amplitudes in this distance range are normalized. The critical distance of phase *c* is generally between 55 and 70 km shotpoint distance.

The records of Figs. 2–5 clearly indicate the existence of first arrivals which can be correlated as phase *d'* with apparent velocities of 8.0–8.2 km/s and which are interpreted as the P_n phase penetrating into the uppermost mantle. The strong increase of amplitudes of the P_n phase at distances > 120 km is probably caused by a positive velocity gradient within the underlying uppermost mantle. Or it may even be caused by the existence of an increasing velocity gradient, which may lead to the occurrence of a retrograde phase *d*₁ with apparent velocities of up to 8.4 km/s. This, however, is difficult to distinguish from the prograde phase *d'* (see also Ansoerge et al., 1979; Stangl, 1983).

All profiles of sufficient recording length with azimuths between 40° and 60° show these clear P_n amplitudes and high apparent velocities. Unfortunately, most profiles except the main line with azimuth 60° or 240° were quite short and consequently, e.g. on the profile 30–80 (see Fig. 1), a P_n phase could not be identified. Therefore, it is not possible to investigate the amplitude and velocity characteristics of the P_n phase in greater detail with respect to observational direction (Bamford, 1973; Fuchs, 1983).

The crustal model

For each profile a velocity-depth function was derived assuming that the velocity varies only with depth. For the profiles along the 60°-line from these velocity-depth functions, isolines of equal velocity were constructed leading to a first-approximation two-dimensional model. This served as a starting model for a ray-tracing procedure (Červený et al., 1977) varying the initial model until the theoretical travel times fitted the observed ones. In addition, the observed amplitudes were considered. As a result – but with the limitations of the first-order ray theory – kinematically and dynamically controlled two-dimensional velocity-depth distributions could be derived for the 60°-line (Fig. 6). More details of this procedure are described by Gajewski and Prodehl (1983). The velocity distribution is presented by lines of equal velocity. Boundary zones are indicated by a great density of velocity isolines, first-order discontinuities by superposition of several isolines.

The crystalline crust is covered by sediments in which the average velocity increases gradually with depth from 3.8 km/s near the surface to about 4.7 km/s at the basement, where it jumps discontinuously to about 5.5 km/s. Details of the structure of the crystalline basement and variations of its surface are described by Jentsch et al. (1982).

The most prominent feature of the crustal model shown in Fig. 6 is a low-velocity zone between 8 and 19 km depth which extends over 100 km in a horizontal direction and has a maximum thickness of 10 km. It terminates to the WSW near the Hohenzollerngraben and to the ENE near the Nördlinger Ries. The average velocity within this zone is 5.8 km/s.

The upper crust above this low-velocity zone is characterized by a strong velocity gradient. At about 3–5 km depth the velocity increases quite abruptly from 5.7 to 6.05 km/s.

At the lower boundary of the low-velocity zone the velocity increases discontinuously from 5.8 to 6.6 km/s. The lower crust, with an average velocity of 6.7 km/s, is between 8 and 11 km thick below and ENE of the low-velocity zone, but thins drastically west of the Hohenzollerngraben.

SWABIAN JURA

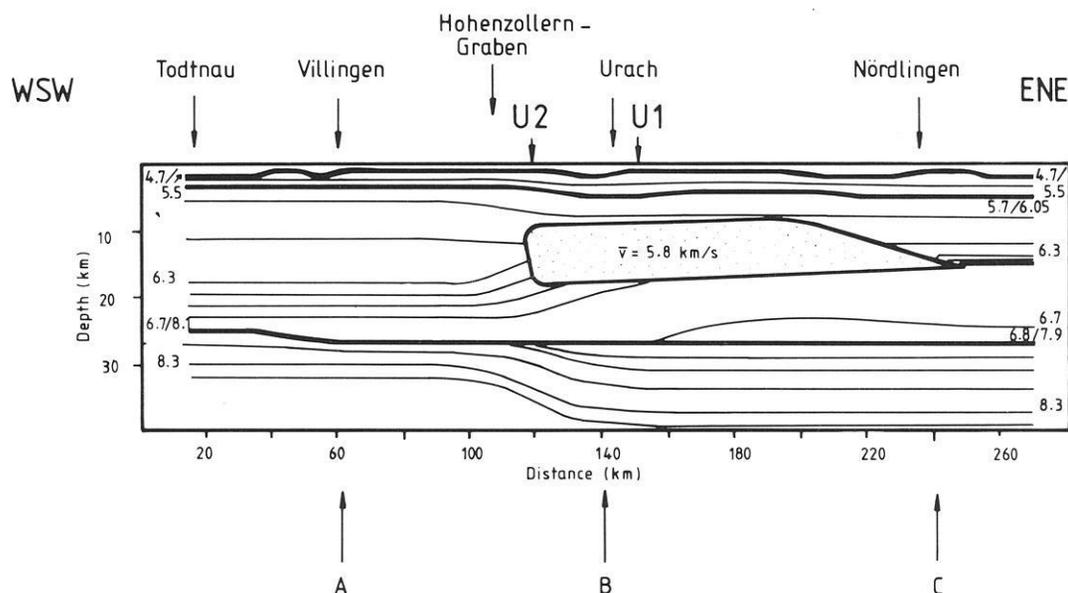


Fig. 6. Crustal model of the main profile observed through the Urach geothermal anomaly in a WSW-ENE direction. The contour interval of the lines of equal velocity is 0.1 km/s. Velocities < 5.6 km/s are not shown. Depth scale is exaggerated 2:1 vs distance scale. Dotted area: zone of reduced velocity. Arrows A, B, C refer to velocity-depth functions of the respective model in Fig. 14

The termination of the low-velocity zone to ENE and to WSW is not well constrained. As mentioned above, the shotpoint distribution is not ideal for investigations of the whole crust, in particular when such lateral variations exist. As a consequence, details of the position and the shape of the boundaries of the low-velocity zone in the middle crust towards ENE and WSW can be modelled in various ways without contradicting the observed data. The WSW boundary, especially, has been studied in more detail and Fig. 6 shows a possible model which seems to fit the observed travel-times and amplitudes best.

The 6.3–6.6 km/s isolines rise in the area of the Hohenzollerngraben and join there to form the bottom of the low-velocity zone.

The crust-mantle boundary is a first-order discontinuity. WSW, towards the Rhinegraben, it is located at 25 km depth. Towards ENE, it drops to 26.5 km under the Swabian Jura. In the same direction the velocity jump at the Moho changes from 6.7/8.1 to 6.7/7.9 km/s.

The model does not only show strong lateral variations with respect to the existence of the low-velocity zone, which is restricted to the Urach geothermal anomaly zone, but is also quite different WSW and ENE of this zone. The boundary between upper and lower crust towards the Black Forest is much more transitional.

Also, the uppermost mantle shows lateral variations of velocity reaching up to 8.4 km/s and a strong increase of velocity with depth (Ansorge et al., 1979; Stangl, 1983).

Amplitude characteristics

For the profiles U1-60, U2-60, U2-240 and U1-240 (Figs. 2–5) along the main line, ray theoretical synthetic seismo-

grams have been calculated based on the model in Fig. 6 (Figs. 7–10). Basically, the calculation of synthetic seismograms is more exact if the reflectivity method (Fuchs and Müller, 1971) is used which, in addition to amplitude studies, allows the study of frequency variations, e.g. such as the shift of the critical distance as a function of frequency. However, this method can, at present, only be applied for structures which change in vertical but not horizontal direction. For our problem, with vertically and laterally changing structures, only the ray theory could be applied using programs developed by Červený, Pšenčík and co-workers (e.g. Červený et al., 1977; Červený and Pšenčík, 1982).

Note that the distance axes of the observed record sections (Figs. 2–5) are numbered differently from those of the synthetic record sections (Figs. 7–10). The synthetic record sections use the same distances as are shown for the model in Fig. 6: shotpoint U2 corresponds to model-km 120, shotpoint U1 to model-km 150. Depending on the azimuth, distances increase (azimuth 60°, towards ENE) or decrease (azimuth 240°, towards WSW). All ray theoretical record sections are reduced with 6 km/s and show vertical components of ground displacement.

For the following reasons it did not seem reasonable to aim for an exact amplitude agreement of synthetically calculated and observed phases. Firstly, such a trial is prohibitive because of the limitations of the ray method. Secondly, the observed amplitudes are not only influenced by the deeper structure of the earth's crust, but are disturbed by many near-surface features as well as biased by instrumentally caused uncertainties. Rather, the aim was to explain the observed data qualitatively. This is best obtained by concentrating on the amplitude relations of various phases within the same seismogram.

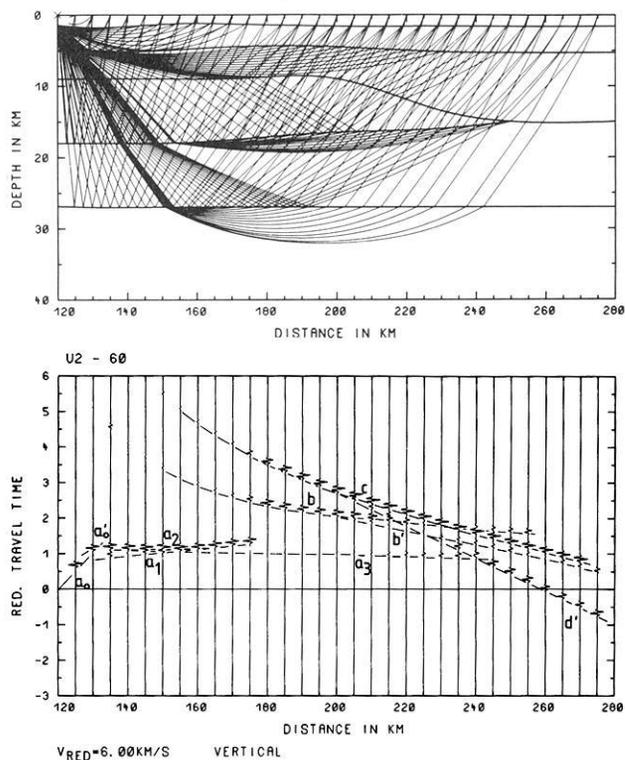


Fig. 7. Ray diagram for the model of Fig. 6 (*top*) and ray theoretical synthetic seismograms for shotpoint U2 (model-km 120) observed towards ENE (azimuth of 60°) with normalized amplitudes. The *heavy lines* in the ray diagram represent discontinuities of the model. Identification of phases corresponds to that of Figs. 2–5

The upper part of Figs. 7–10 contains ray diagrams for the corresponding profiles shown in Figs. 2–5. They show individual receiver positions along the profiles and the corresponding rays traversing the earth's crust. In the lower part of the figures the corresponding ray theoretical seismograms are shown. The synthetic seismograms are normalized in the same way as the observed data in Figs. 2–5.

The comparison of observed and computed seismograms (Figs. 2–5 and 7–10, respectively) shows pretty good agreement. All essential observed phases are also documented in the synthetic data and the amplitudes fit quite well too. If, on a seismogram, only one phase appears, the normalization procedure will amplify that phase, which in some cases may be irritating. Also, it may happen that the true amplitudes of the observed seismograms are larger than the ones on the adjacent traces. Then the synthetic section differs visually from the observed one even if more than one phase is seen on each seismogram. For example, this is the case for phase b which apparently increases in amplitude in the near-vertical reflection range (Fig. 7).

In the synthetic sections shown in Figs. 7–10, phases a_1 and a_2 cannot easily be distinguished because the two phases interfere with each other. In the observed seismograms, also, these phases cannot easily be separated because they follow each other with only a 0.2-s time difference.

For phase a_3 , in any case, this leads to sufficient agreement of synthetic and observed data. Looking at the ray diagrams (Figs. 7 and 8), one recognizes that the rays of these arrivals penetrate the depth range below 5 km but still above the low-velocity zone, an area with a relatively

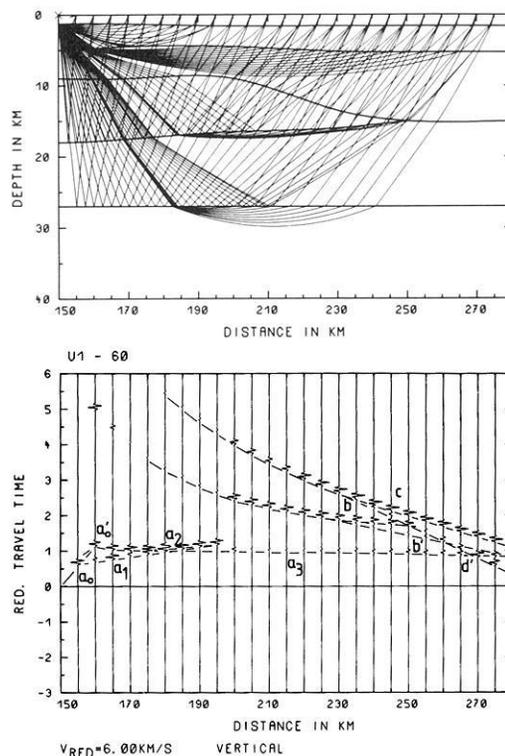


Fig. 8. Ray diagram for the model of Fig. 6 and ray theoretical synthetic seismograms for shotpoint U1 (model-km 150) observed towards ENE (azimuth of 60°), plotted with normalized amplitudes. The *heavy lines* in the ray diagram represent discontinuities of the model. Identification of phases corresponds to Figs. 2–5

small velocity gradient. Minor changes in the model parameters, e.g. shifting the upper surface of the low-velocity zone by only a few hundred metres, have significant consequences on this phase a_3 , because this considerably changes the distance at which some energy of this phase can still be observed. According to ray theory, a phase always disappears abruptly if caused by a low-velocity zone or a caustic region.

For profile U1-60 the agreement between observed (Fig. 3) and synthetic (Fig. 8) data is not so good, because the computed amplitude relations of phases b and c in the middle distance range do not fit the character of the observed records very well.

Ray theory always results in too small a value for the critical distance (caused by high-frequency approximation). For crustal investigations this shift of the critical distance towards smaller values is between 10 km and some tens of kilometres (Červený et al., 1977). As the frequencies of phases b and c are quite similar, however, this effect would shift the critical distance of both phases together so that the amplitude relation should not be altered by this effect.

So it can only be concluded that the reason for the disagreement between computed and observed record sections for profile U1-60 must be caused by the low-velocity zone rather than by the structure of the Moho. With the exception of this particular profile U1-60, all other profiles show clear $P^M P$ energy at distances greater than 50 km. In addition, the vertices of the rays of the reflection from the Moho on profile U2-60 are only 30 km from those on profile U1-60. The rays for both profiles penetrate through

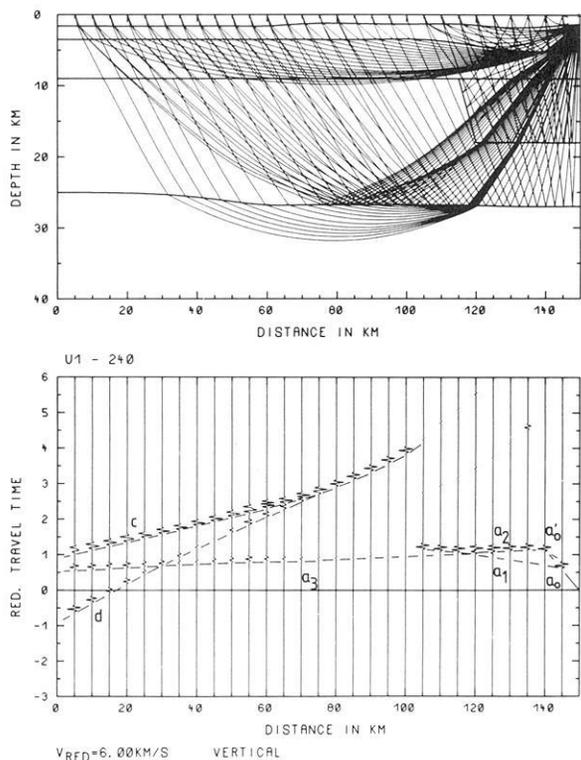


Fig. 9. Ray diagram for the model of Fig. 6 and ray theoretical synthetic seismograms for shotpoint U1 (model-km 150) observed towards WSW (azimuth of 240°), plotted with normalized amplitudes. For further explanations, see Fig. 8

the low-velocity zone, but only one profile, U2-60, shows “normal” amplitude characteristics for the $P^M P$ phase. It is unreasonable to assume that the Moho changes drastically within this 30-km horizontal distance, which would be necessary to shift the critical distance by more than 50 km.

The only reasonable solution, therefore, seems to assume strong lateral variations within the low-velocity zone itself. The assumed mean velocity of 5.8 km/s is an average value which certainly may be larger or smaller locally. Unfortunately, the seismic refraction data do not allow any definite conclusions with respect to this problem.

The synthetic data of profiles U1-240 and U2-240 again agree quite well with the observed data (Figs. 4 and 5 and Figs. 9 and 10.).

For both profiles, it is evident that in the observed and the computed record sections phase *b* is missing. On the other hand, the synthetic data of profile U1-240 show some energy resulting from reflections from the flanks of the low-velocity zone as can be seen in the ray diagram (Fig. 9, upper part). Such arrivals are also weakly seen in the observed data. The energy of that phase is quite low and, therefore, almost hidden by signal-generated noise in the data. The amplitude characteristics of phases *c* and *d'* ($P^M P$ and P_n) are well demonstrated by the synthetic data, only phase *a*₃ terminates too late in the computed sections.

In summary it can be concluded that, in general, good agreement could be reached between observed and com-

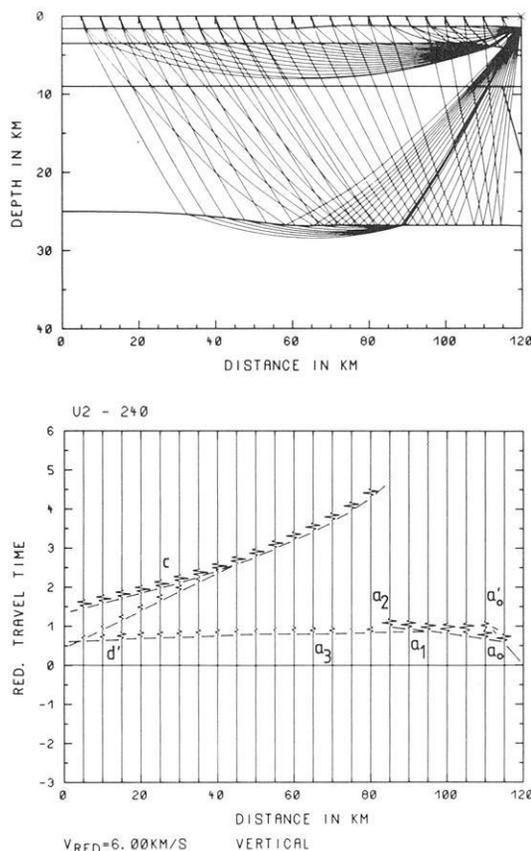


Fig. 10. Ray diagram for the model of Fig. 6 and ray theoretical synthetic seismograms for shotpoint U2 (model-km 120) observed towards WSW (azimuth of 240°) plotted with normalized amplitudes. For further explanations, see Fig. 8

puted amplitude data. Exceptions are caused mainly by influences of the low-velocity zone. Thus, it has to be assumed that this low-velocity zone possesses laterally varying fine structure which, however, cannot be resolved. Based on the amplitude information of the $P^M P$ and P_n phases, no indication of a stronger absorption within the low-velocity zone can be seen. Also phase *b*, the reflection from the bottom of this zone, does not allow any conclusions in that respect.

Comparison with results of the seismic reflection experiments

The model resulting from the interpretation of the near-vertical seismic reflection experiment (Bartelsen et al. 1982), located on a profile approximately 20 km NW of our 60° line (Figs. 1 and 15), is reproduced in Fig. 11. The complementary model based on wide-angle measurements along the same line of the near-vertical reflection observations was prepared by Trappe (1983) and is reproduced in Fig. 12. A direct comparison of the three models (Figs. 6, 11 and 12) is difficult because the model of the near-vertical reflection survey consists of layers with constant interval velocity, while the models of our seismic refraction survey and that from the wide-angle observations of the reflection line use velocity gradients, i.e. the velocity increases with increasing depth between velocity isolines. Nevertheless, the characteristic features can be compared.

All models show a low-velocity zone with a mean veloci-

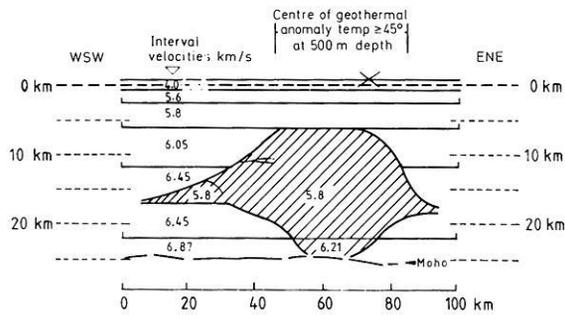


Fig. 11. Crustal model of Bartelsen et al. (1982, Fig. 13) of the Urach geothermal anomaly derived along the WSW-ENE seismic reflection line (Fig. 1) from near-vertical reflection observations

ty of 5.8 km/s. The extension in the horizontal direction differs. This may be caused by the difference in location (Fig. 15). The surface of this low-velocity zone of the “near-vertical” model (Fig. 11) is found at only 6 km depth, while that of the “wide-angle” model along the same line lies at about 9 km depth (Fig. 12) and thus corresponds to our result (Fig. 6) including the rise of this boundary towards ENE.

Bartelsen et al. (1982) concluded that this boundary could be interpreted as a transitional layer or as a lamellar structure with thin alternating low- and high-velocity layers, because they were able to prove the existence of horizons with increasing and decreasing impedance, at least in the WSW part of their line. On the other hand, no near-vertical reflections could be found from the top of the low-velocity zone in its central part.

So, the above-mentioned difference in depth may not be significant. In any case, it cannot be explained by different profile locations because the “near-vertical” model (Fig. 11) and the “wide-angle” model were derived for the same line. The agreement of the “wide-angle” model (Fig. 12) and our model (Fig. 6), for the upper crust including the low-velocity zone, is extremely good. This also holds for the structure and depth of the Moho. Significant deviations do exist for the crustal depth range immediately below the low-velocity zone. The velocities here are quite different. Only at the WSW and ENE “edges” of the models is the average velocity of all models approximately equal. With respect to the unfavourable shotpoint distribution of our experiment, the greatest uncertainties have to be faced here.

Figure 13 summarizes the results of all profiles in a fence

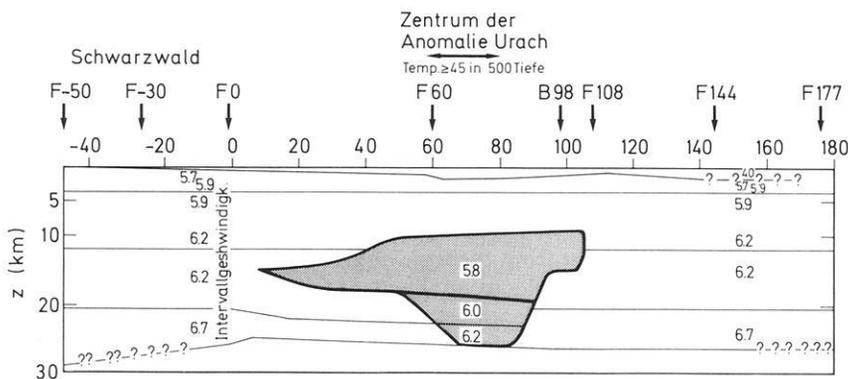


Fig. 12. Crustal model of Trappe (1983, Fig. 53) of the Urach geothermal anomaly derived along the WSW-ENE seismic reflection line (Fig. 1) from wide-angle reflection observations

diagram, including the seismic reflection line. As none of the profiles outside the 60°-line was reversed or overlapped, it is almost impossible to determine the three-dimensional extensions of the low-velocity zone.

A pronounced change in crustal structure is observed on the profiles pointing towards the Rhinegraben. In comparison to the crust east of the low-velocity zone, the lower crust in the west is thinner. The rise of the Moho towards the Rhinegraben may be connected with the general uplift of the graben and its shoulders (Edel et al., 1975; Giese, 1976), the distance between Urach and Rhinegraben being less than 100 km.

Comparison with crustal structure results of adjacent areas

For the adjacent areas a number of quite recent investigations exist (Deichmann and Ansorge, 1983; Stangl, 1983; Strössenreuther, 1982; Zucca, 1984) which complement the earlier publications (Aichele, 1976; Edel et al., 1975; Emter, 1971; Mueller et al., 1973; Prodehl et al., 1976).

Figure 14 shows velocity-depth functions of our two-dimensional model (Fig. 6) for three selected positions (Fig. 14a-c). The lower part of Fig. 14 shows some velocity-depth functions of other observations west (Fig. 14d) and east (Fig. 14e) of our area of investigation. The location of these surveys is shown in Fig. 15.

The models of Aichele (1976) and Strössenreuther (1982) show zones of velocity inversions with an average velocity of 5.8 km/s and around 10 km depth (Fig. 14e). Their low-velocity zones, however, are thinner than that in the centre of the Urach area (Fig. 14b).

Profile 09-240 is the continuation of our 60°-line into the Franconian Jura (Fig. 15). Contrary to the results of Aichele and Strössenreuther, Stangl (1983) does not find any velocity inversion from this profile. His model is quite similar to our velocity-depth function at model-km 240 (Fig. 14c and e). However, shotpoint 09 is more than 200 km from shotpoint U1 and, consequently, the turning points of corresponding rays are more than 100 km apart.

In conclusion, the area of the Swabian and Franconian Jura shows strong lateral variations of the crust. A discontinuity or transition zone at depths between 15 and 20 km in the middle crust seems to be typical for the South German Jura mountains (see Fig. 14e). It is not clear whether the low-velocity zone under the anomalous region of Urach is related to the one stated by Aichele (1976) and Strössenreuther (1982) for the Franconian Jura. Information is lacking on the crust under the area in between these two regions.

Urach Geothermal Anomaly

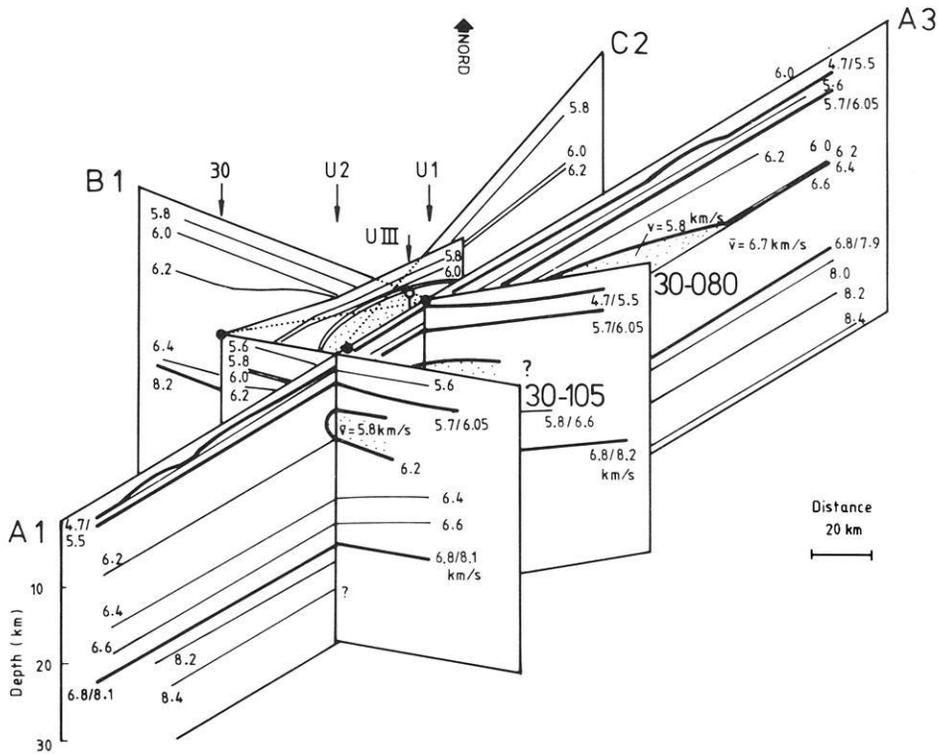


Fig. 13. Fence diagram showing the crustal structure in the area of the Urach geothermal anomaly based on seismic refraction and reflection measurements. Depth is exaggerated 2.3:1 vs horizontal distance. The contour interval of lines of equal velocity is 0.2 km/s. Velocities < 5.6 km/s are not shown. *Dotted area* indicates the position of a low-velocity zone

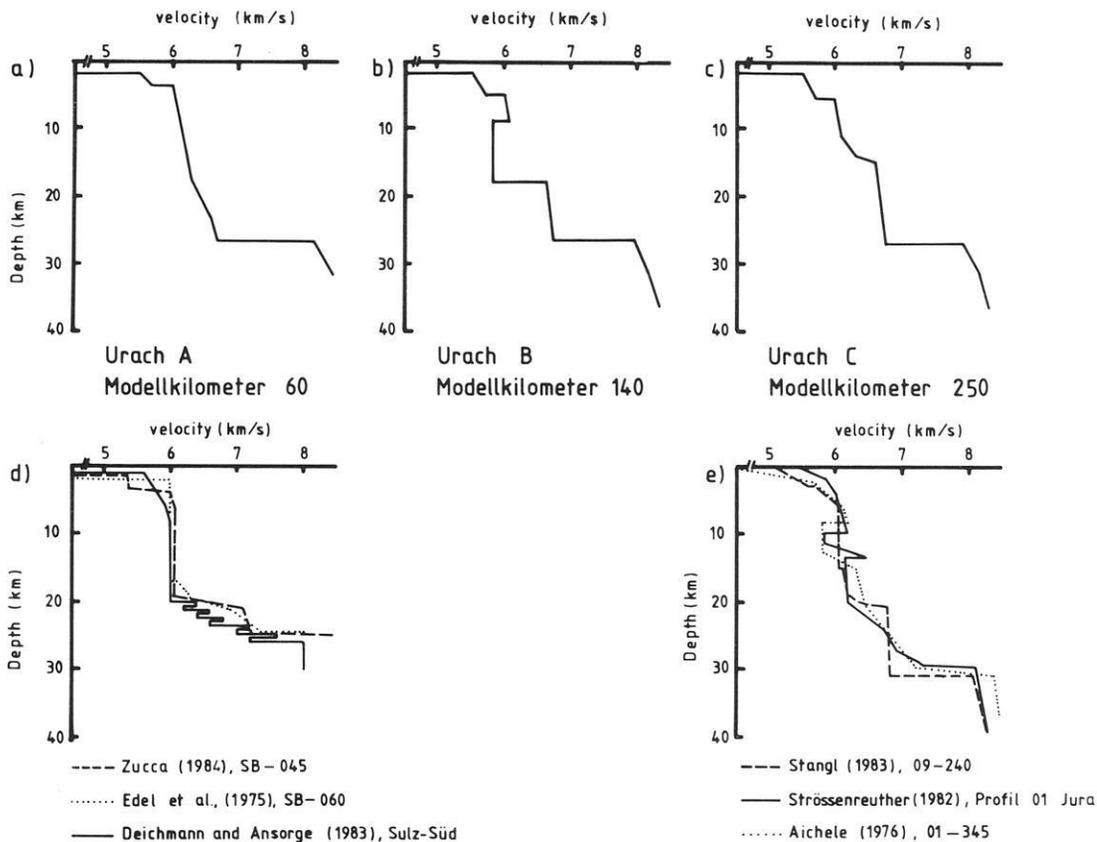


Fig. 14a-e. Velocity-depth functions representing various areas of the Urach geothermal anomaly (a-c) according to the model of Fig. 6 and adjacent areas to the west (d) and to the east (e)

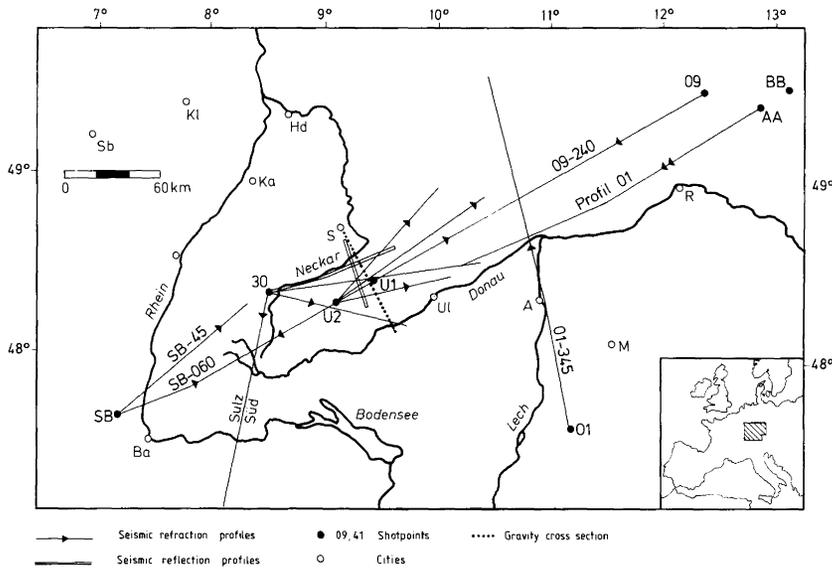


Fig. 15. Location map of seismic refraction and reflection profiles and a gravity cross-section in southwest Germany mentioned in Figs. 14 and 16 and in the text

Discussion and conclusions

The interpretation of a series of seismic refraction profiles in the area of the Urach geothermal anomaly has proved the existence of a low-velocity zone ($\bar{v}=5.8$ km/s) in the middle crust. How can such a zone be interpreted?

A gravimetric profile, oriented approximately perpendicular to our 60°-line and crossing it near shotpoint U1 (Figs. 15 and 16), shows a high-density body in the area of the geothermal anomaly (Makris et al., 1982), in apparent contrast to the existence of the low-velocity zone found by reflection and refraction seismics. The density contrast of that body with respect to the surrounding crustal material is up to 0.26 g/cm³, depending on the particular model of Makris et al. (1982). In this connection, the absence of near-vertical reflections from the surface of the low-velocity zone is of particular interest. According to the model of Bartelsen et al. (Fig. 11) it means that the velocity does not change here. On the other hand, both the "wide-angle" model of Trappe (Fig. 12) and our model (Fig. 6) do show *P*-wave velocities of up to 6.1–6.2 km/s above the low-velocity zone. This is not in contradiction to the seismic reflection result which gives interval velocities, i.e. an average velocity for the layer above the low-velocity zone.

A possible explanation for the fact that no near-vertical reflections are observed from the top of the low-velocity zone is that the impedance, i.e. the product of *P*-velocity and density, does not change. As the *P*-wave velocity is reduced from 6.2 to 5.8 km/s, the density must increase if the impedance is to remain constant. One can determine the density corresponding to 6.2 km/s after Nafe and Drake (see Talwani et al., 1959) and then compute the corresponding impedance. From that and the velocity of 5.8 km/s, a density of 3.0 g/cm³ results for the low-velocity body and a density contrast of $\Delta\rho=0.21$ g/cm³ with respect to the overburden. This contrast is within the range given by Makris et al. (1982) (Fig. 16).

Such an interpretation can explain both the decrease of seismic velocity and the increase of Bouguer gravity for the Urach geothermal anomaly. The combination of such high density and low *P*-wave velocity requires the existence of fluid or gas phases or partial melting of the rocks con-

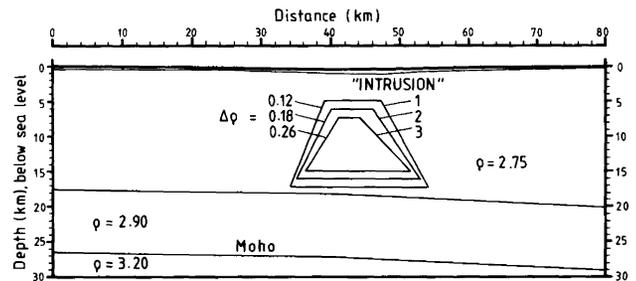


Fig. 16. Density model of Makris et al. (1982, Fig. 4) for a line crossing the Urach geothermal anomaly from NNW to SSE, perpendicular to the seismic lines of Figs. 2–12. For position of gravity cross-section, see Fig. 15

cerned (e.g. gabbro). This, however, leads to much higher temperatures in this depth range than has been assumed until now (e.g. Haenel and Zoth, 1982).

The assumption of the existence of partially molten material or fluid or gas phases is supported by an investigation of *S* waves. While clear wide-angle *P*-wave reflections from the lower boundary of the low-velocity zone can be observed on all profiles concerned (phase *b* in Figs. 2–4), *S*-wave reflections cannot be detected on any of the profiles as shown, for example, for profile U1-60 (Fig. 17). The record section is reduced with 3.46 km/s (corresponding to $V_p/V_s=\sqrt{3}$). The time scale is chosen so that the full range of observed *P* and of expected *S* phases can be detected. The full lines show the travel-time curves of the *P* waves and are the same as shown in Fig. 3. Dotted lines are travel-time curves of *S* waves calculated from the *P*-wave travel-times assuming a Poisson's ratio of 0.25, i.e. $V_p/V_s=\sqrt{3}$.

Figure 17 clearly demonstrates that there is no *S* wave recorded as a reflection from the bottom of the low-velocity zone. Also, the $S^M S$ phase, the *S*-wave reflection from the Moho, and the corresponding S_n wave, refracted in the uppermost mantle, cannot be seen in the record sections. The result is the same for the record sections of the horizontal components. While shotpoint U1 does not show any *S* energy from the middle and lower crust, a weak $S^M S$

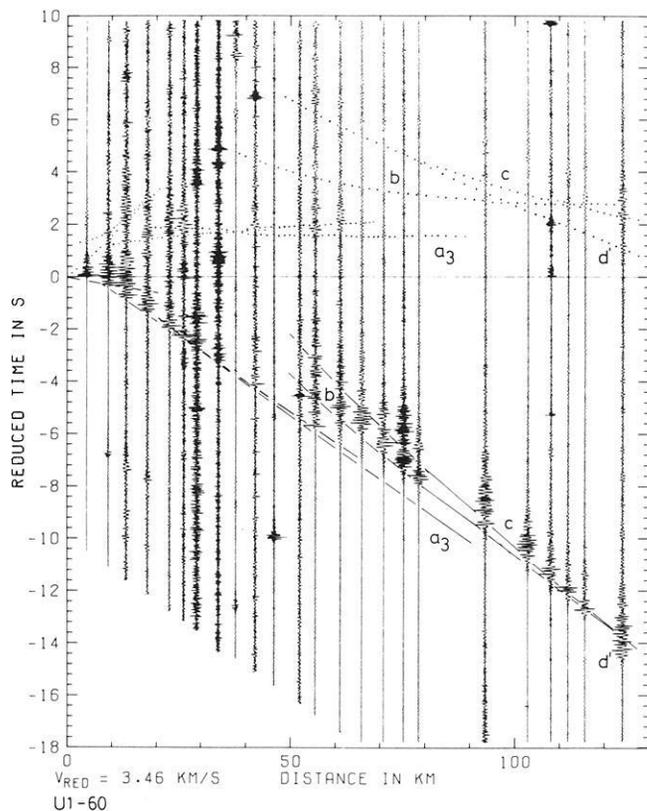


Fig. 17. Record section of the seismic refraction profile U1-60. Reduction velocity = 3.46 km/s. Solid lines are the travel-time curves calculated for P waves from the model of Fig. 6 and are the same as shown in Fig. 3. Dotted lines are travel-time curves for S waves calculated from the same model and assuming a Poisson's ratio of $P_p/V_s = \sqrt{3}$

phase is observed for shotpoint U2 in both directions. Its energy is considerably smaller than the corresponding $P^M P$ -phase energy in ENE direction, but stronger to the WSW. This is probably caused by the fact that, on this profile, the rays penetrate the low-velocity zone only on half of their total ray path.

As partial melts attenuate S waves stronger than P waves (Mavco, 1980), these observations may be a further indication for the existence of such partial melt.

If one explains the low-velocity zone with high density as magmatic intrusion, then this event should have happened about 1 million years ago according to a simple heat conduction model calculation (Lachenbruch et al., 1976; Chapman, personal communication), in order to preserve temperatures of about 700°C until present times. Such short time span, however, does not correspond with the occurrence of the known volcanic activity which happened about 15 million years ago in this area.

Considering the total crustal thickness, a rather thin crust exists under the Urach geothermal area which is similar to that found under the Rhinegraben proper, 60–80 km to the west. Taking into account the results of the profile Sulz-S (Deichmann and Ansorge, 1983) and the profiles extending from the Rhinegraben into the Black Forest (Edel et al., 1975), it seems that the crust remains shallow throughout the whole area between the Black Forest and the Swabian Jura, in contrast to the increase of crustal thickness west of the Rhinegraben. According to Aichele

(1976) and Strössenreuther (1982), the standard central European crustal thickness of about 30 km is present under the Franconian Jura, so that the transition must be located somewhere near the Nördlinger Ries.

In conclusion, the seismic investigations of the geothermal anomaly of Urach reveal a rather heterogeneous crustal structure. Beneath the centre of the anomaly a well constrained low-velocity zone is found whose extension coincides fairly well with the contours of the isolines of geothermal gradient. Though its exact western and eastern extension cannot be accurately fixed it can be stated that its western border coincides approximately with the position of the Hohenzollerngraben, a fracture zone which is cut by a N-S trending line of major anomalous earthquake activity.

The combination of gravity and seismic data may indicate the existence of partial melts of high density which may not be older than about 1 million years, in contrast to the known volcanic activity in this area which occurred about 15 million years ago.

The fact that the crust is almost as thin as that beneath the Rhinegraben proper may suggest that either the same anomalous mantle is present under the whole area reaching from the Rhinegraben to the Urach area or that two branches of anomalous mantle behaviour exist.

Acknowledgements. The investigation, as part of a multidisciplinary project under the responsibility of Dr. Ralph Haenel, Hannover, was jointly sponsored by the European Community (contract no. 071-76 EG 61A-23-5) and the Federal Ministry for Research and Technology (contract no. EG 4027A). The interpretation work was additionally supported by the German Research Association with the special research programme "Stress and stress release in the lithosphere" of the University of Karlsruhe (SFB 108).

We are grateful to K. Fuchs (Karlsruhe), D. Chapman (Clausthal, on leave from Salt Lake City), P. Giese (Berlin) and H. Wilhelm (Karlsruhe) for fruitful discussions. The Institute of Geophysics of the University of Kiel supplied original data and figures. W. Kaminski helped intensively during the digital data preparation on the Raytheon 500 computer of the Geophysical Institute of the University of Karlsruhe. Beate Aichroth and W. Friederich carried out many calculations which were performed on the Burroughs 7700 computer of the Computer Centre of the University of Karlsruhe.

References

- Aichele, H.: Interpretation refraktionsseismischer Messungen im Gebiet des Fränkisch-Schwäbischen Jura. Ph.D. Thesis, University of Stuttgart, 1976
- Ansorge, J., Bonjer, K.-P., Emter, D.: Structure of the uppermost mantle from long-range seismic observations in southern Germany and the Rhinegraben area. *Tectonophysics* **56**, 31–48, 1979
- Bamford, D.: Refraction data in western Germany – a time-term interpretation. *J. Geophys.* **39**, 907–927, 1973
- Bartelsen, H., Lueschen, E., Krey, Th., Meissner, R., Schmol, H., Walter, Ch.: The combined seismic reflection-refraction investigation of the Urach geothermal anomaly. In: *The Urach geothermal project*, Haenel, R. ed.: pp 231–245, Schweizerbart, Stuttgart, 1982
- Carlé, W.: Die Wärme-Anomalie der mittleren Schwäbischen Alb (Baden-Württemberg). In: *Approaches to Taphrogenesis*, Illies, H.J., Fuchs, K., eds: pp 207–212, Schweizerbart, Stuttgart, 1974
- Červený, Molotkov, I.A., Pšenčík, I.: Ray method in seismology, Univerzita, Karlova, Prag, 1977

- Červený, V., Pšenčík, I.: Seismic ray package, Fortran program. Prag, 1982
- Deichmann, N., Ansorge, J.: Evidence for lamination in the lower continental crust beneath the Black Forest (southwest Germany). *J. Geophys.* **52**, 109–118, 1983
- Dietrich, H.-G.: Geological results of the Urach 3 borehole and the correlation with other boreholes. In: *The Urach geothermal project*, Haenel, R., ed.: pp 49–58, Schweizerbart, Stuttgart, 1982
- Edel, J.B., Fuchs, K., Gelbke, C., Prodehl, C.: Deep structure of the southern Rhinegraben area from seismic refraction investigation. *J. Geophys.* **41**, 333–356, 1975
- Emter, D.: Ergebnisse seismischer Untersuchungen der Erdkruste und des oberen Erdmantels in Südwestdeutschland. Ph.D. Thesis, University of Stuttgart, 1971
- Fuchs, K.: Recently formed elastic anisotropy and petrological models for the continental subcrustal lithosphere in southern Germany. *Phys. Earth Planet. Inter.* **31**, 93–118, 1983
- Fuchs, K., Müller, G.: Computation of synthetic seismograms with the reflectivity method and comparison with observations. *Geophys. J. R. Astron. Soc.* **23**, 417–433, 1971
- Gajewski, D., Prodehl, C.: Zweidimensionales seismisches Modellieren mit der Strahlenmethode. *Berichtsband 1981–1983 des Sonderforschungsbereichs 108 der Universität Karlsruhe „Spannung und Spannungsumwandlung in der Lithosphäre“*, 209–227, 1983
- Geyer, O.F., Gwinner, M.P.: Einführung in die Geologie von Baden-Württemberg. Schweizerbart, Stuttgart, 1968
- Giese, P.: The basic features of crustal structure in relation to the main geologic units. In: *Explosion seismology in Central Europe*, Giese, P., Prodehl, C., Stein, A. eds.: pp 241–242, Springer, Berlin, Heidelberg, New York, 1976
- Giese, P., Prodehl, C., Stein, A.: *Explosion seismology in Central Europe*. Springer, Berlin, Heidelberg, New York, 1976
- Haenel, R. (ed.): *The Urach geothermal project (Swabian Alb, Germany)*. Schweizerbart, Stuttgart, 1982
- Haenel, R., Zoth, G.: Temperature measurements and determination of heat flow density. In: *The Urach geothermal project*, Haenel, R. (ed.): pp 81–88, Schweizerbart, Stuttgart, 1982
- Illies, J.H.: Der Hohenzollerngraben und Intraplatten-Seismizität infolge Vergitterung lamellärer Scherung mit einer Riftstruktur. *Oberrhein. geol. Abh.* **31**, 47–78, 1982
- Jentsch, M.: A compilation of data from the 1978–79 Urach, Baden-Württemberg, seismic-refraction experiment. Open-file report, 80-1. *Geophys. Inst. Karlsruhe*, 1980
- Jentsch, M., Bamford, D., Emter, D., Prodehl, C.: A seismic refraction investigation of the basement structure in the Urach geothermal anomaly, southern Germany. In: *The Urach geothermal project*, Haenel, R. ed.: pp 247–262, Schweizerbart, Stuttgart, 1982
- Lachenbruch, A.H., Sass, J.H., Munroe, R.J., Moses, T.H.: Geothermal setting and simple heat conduction models for the Long Valley caldera. *J. Geophys. Res.* **81**, 769–784, 1976
- Makris, J., Müller, K., Tödt, K.H.: Gravity measurements at the geothermal anomaly Urach. In: *The Urach geothermal project*, Haenel, R. ed.: pp 313–322, Schweizerbart, Stuttgart, 1982
- Mäussnest, O.: Die Eruptionenpunkte des Schwäbischen Vulkans. *Z. Dtsch. Geol. Ges.* **125**, 23–54, 277–352, 1974
- Mavco, G.M.: Velocity and attenuation in partially molten rocks. *J. Geophys. Res.* **85**, 5173–5189, 1980
- 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: *Geothermics and geothermal energy*, Čermák, V., Haenel, R., eds.: pp 285–292, Schweizerbart, Stuttgart, 1982
- Mueller, S., Peterschmitt, E., Emter, D., Ansorge, J.: Crustal structure of the Rhinegraben area. *Tectonophysics* **20**, 381–391, 1973
- Prodehl, C.: Crustal structure of the western United States. *U.S. Geol. Survey Prof. Paper* **1034**, 1979
- Prodehl, C., Ansorge, J., Edel, J.B., Emter, D., Fuchs, K., Mueller, S., Peterschmitt, E.: Explosion-seismology research in the central and southern Rhinegraben – a case history. In: *Explosion seismology in central Europe*, Giese, P., Prodehl, C., Stein, A., eds.: pp 313–328, Springer, Berlin, Heidelberg, New York, 1976
- Prodehl, C., Emter, D., Jentsch, M.: Seismic refraction studies of the geothermal area of Urach, southwest Germany. In: *Geothermics and geothermal energy*, Čermák, V., Haenel, R., eds.: pp 277–283, Schweizerbart, Stuttgart, 1982
- Schädel, K.: The geology of the heat anomaly of Urach. In: *The Urach geothermal project*, Haenel, R., ed.: pp 147–156, Schweizerbart, Stuttgart, 1982
- Schneider, G.: *Seismizität und Seismotektonik der Schwäbischen Alb*. Enke, Stuttgart, 1971
- Stangl, R.: Geschwindigkeits-Tiefen-Verteilungen von P-Wellen im oberen Mantel Süddeutschlands, die mit Laufzeit- und Amplituden-Beobachtungen auf zwei gegengeschossenen Langprofilen verträglich sind. *Diploma Thesis*, University of Karlsruhe, 1983
- Strössenreuther, U.: Die Struktur der Erdkruste am Südwestrand der Böhmisches Masse, abgeleitet aus refraktions-seismischen Messungen der Jahre 1970 und 1978/79. *Ph.D. Thesis*, University of München, 1982
- Talwani, M., Sutton, G.H., Worzel, J.C.: A crustal section across the Puerto Rico Trench. *J. Geophys. Res.* **64**, 1545–1555, 1959
- Trappe, H.: Eine Auswertung von Weitwinkelreflexionen auf dem Profil Urach und eine Korrelation zu den Ergebnissen der Steilwinkel-Reflexionsseismik. *Diploma Thesis*, University of Kiel, 1983
- Turnovsky, J., Schneider, G.: The seismotectonic character of the September 3, 1978, Swabian Jura earthquake series. *Tectonophysics* **83**, 151–162, 1981
- Zucca, J.J.: The crustal structure of the southern Rhinegraben from reinterpretation of seismic refraction data. *J. Geophys.* **55**, 13–22, 1984

Received December 19, 1983; Revised version October 15, 1984
Accepted October 16, 1984

Static deformations and gravity changes at the earth's surface due to atmospheric loading

W. Rabbel and J. Zschau

Institut für Geophysik, Christian Albrechts Universität, Olshausenstr. 40-60, D-2300 Kiel, Federal Republic of Germany

Abstract. Deformations and gravity changes at the Earth's surface due to regional and global air pressure variations are estimated for a radially stratified earth. The results are as follows:

- *Vertical displacements* of seasonal character have maximum amplitudes of ± 0.5 cm. (Anti-)Cyclones, however, can cause vertical displacements of up to ± 2.5 cm.
- *Horizontal displacements* have amplitudes less than ± 2.5 mm.
- *Horizontal principal strains* may have amplitudes up to 10^{-8} . They reduce to about $\pm 1.5 \cdot 10^{-9}$ for seasonal changes in the air pressure distribution.
- The *total gravity perturbation* consisting of the Newtonian attraction of air masses and of self-gravitation due to the elastic deformation may go up to ± 20 μ gal in the case of (anti-)cyclones, and ± 3 μ gal in the case of seasonal air pressure changes.
- The *total tilt* due to seasonal air pressure variations can be as high as ± 1.5 msec. For passing (anti-)cyclones this value may go up to ± 10 msec.

All the above values have to be modified in the direct vicinity of coastlines. The modification is only slight for the displacements and the secondary gravity effect, but it is important for the other components. There, the necessary modification may amount to several hundred percent depending on the type of deformation component and on the distance to the coastline.

Precise air pressure corrections of radial displacements and gravity changes cannot be achieved by using a single regression coefficient. Either the characteristic wavelengths of the pressure distribution have to be taken into account or the following two-coefficient correction equations have to be used:

$$\begin{aligned} \text{Radial displacement:} & \quad u = -0.90 \bar{p} - 0.35 (p - \bar{p}) \\ \text{Primary gravity:} & \quad g_p = 0.36 \bar{p} + 0.41 (p - \bar{p}) \\ \text{Secondary gravity:} & \quad g_s = -0.17 \bar{p} - 0.08 (p - \bar{p}) \\ \text{Total gravity:} & \quad g = g_p + g_s, \end{aligned}$$

with u = radial displacement in mm, g_p , g_s , g = primary, secondary and total gravity, respectively, in μ gal, p = local pressure variation in mbar, \bar{p} = average of the pressure variation in a surrounding area of 2,000 km (in mbar) and \bar{p} the same average, except for setting the pressure values equal to zero over ocean areas.

These corrections have been tested for seasonal air pressure variations and they have proved to be highly precise.

The average errors are less than 0.5 mm, 0.1 μ gal, 0.1 μ gal and 0.2 μ gal for the radial displacements, the primary, secondary and total gravity changes, respectively. The maximum errors are less than 1 mm in the case of the radial displacements, 0.3 μ gal and 0.2 μ gal for the primary and secondary gravity changes, respectively, and 0.4 μ gal for the total gravity changes. Due to a small, spatially constant error term these values apply strictly only to spatial differences of the above deformation components. The differences, however, can be taken between any two points on the Earth's surface.

Key words: Geodynamics – Atmospheric loading – Global deformation – Global positioning – Gravity variations

Introduction

Global deformations of the Earth involving significant gravity changes may be caused by either endogenic, tectonic forces or by exogenic influences such as Earth tides, ocean and atmospheric loading as well as snow coverage during winter. The classical approach of measuring these phenomena is the use of continuously recording gravimeters, tilt- and strainmeters distributed over the Earth's surface. However, this approach has one principle difficulty: because of instabilities of the instruments and their installations it can only provide reliable information on periodic deformations of periods not much longer than one day. Thus, its applicability is almost exclusively restricted to the Earth's tidal deformation. This situation has improved significantly with the development of the superconducting gravimeter (Prothero and Goodkind, 1972) and it has drastically changed with the application of geodetic space techniques such as Satellite Laser Ranging and Very Long Baseline Interferometry (Committee on Geodesy, National Research Council, 1981).

Using the latter technique for relative positioning, even the direct determination of continental drift rates of the order of centimetres per year appears to become possible in the very near future. Preliminary results have already given continental drift rates over the last 3 years which generally agree with those averaged over the geological past (Walter, 1984). Millimeter-scale accuracies are envisaged for similar systems to detect vertical surface motions associated with tectonic processes (Walter 1984). High-precision relative gravimetry, as well as absolute gravimetry, promises to substantially contribute to achieving this goal. These

methods as well as the space techniques do not suffer from the above-mentioned lack of stability. Thus, besides measuring displacements due to continental drift, they will also correctly record long-period displacements as, for instance, those caused by seasonal changes in the Earth's surface load. Not allowing for such externally induced deformations might, for instance, result in erroneous continental drift rates from relative positioning provided the displacements of external origin are of the same order of magnitude as those due to global plate motion.

The purpose of this paper is to give an idea of the magnitude of deformations and gravity changes caused by regional and global variations in the air pressure distribution. Displacements, horizontal principal strains, tilt and gravity effects are estimated quantitatively.

The next section deals with deformations of the Earth's surface due to regional air pressure variations which are quasi-periodic with a repetition time of several days. The effect of global seasonal deviations from mean atmospheric pressure is investigated in the following section and, finally, a simple but highly precise correction of radial displacements and gravity changes for air pressure variations is presented.

Deformations and gravity changes due to (anti-)cyclones

The loading functions

Cyclones and anticyclones are extreme air pressure lows and highs, respectively, showing deviations from mean atmospheric pressure (1,013 mbar) of at least some 10 mbar. The highest pressure observed for anticyclones is 1,080 mbar, which occurred in Siberia. The lowest value for non-tropical cyclones is about 925 mbar, while in the tropics the pressure may even fall below 900 mbar, (Baur 1948, Faust 1968). Both cyclones and anticyclones have spatial extensions between some hundred (tropical cyclones) and some thousand kilometres (continental anticyclones). Their duration is generally of the order of a few days and sometimes they can remain stable for weeks. Thus, inertia effects in the corresponding loading deformations can usually be neglected. An exception to this is the passage of cold or warm fronts, for which the characteristic times involved may be much shorter than 1 day (see Müller and Zürn 1983). However, such phenomena are not considered here.

In most cases the pressure distributions of (anti-)cyclones are geometrically simple. The isobars are near to being circular and, in a first approximation, the anomalous pressure may be described by the simple formula

$$p(r) = p_{max} * \exp(-r^2/r_0^2) \quad (1)$$

(Trubytsin and Makalkin 1976).

Here r is the distance from the centre of the (anti-)cyclone, p_{max} is the maximum pressure anomaly at the centre, and r_0 the distance at which the anomaly has dropped to p_{max}/e ($e = 2.718 \dots$).

For estimating the deformations and gravity changes at the Earth's surface due to the occurrence of (anti-)cyclones we have set r_0 to three characteristic values:

- $r_0 = 160$ km and 400 km, modelling "tropical cyclones" (steep pressure gradient, small spatial extension)
- $r_0 = 1,000$ km, modelling a "continental anticyclone" (small pressure gradient, large spatial extension).

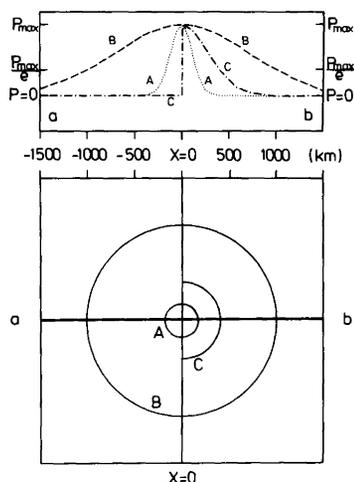


Fig. 1. Considered air pressure distributions. *Top:* Pressure amplitudes along profile a–b. *Bottom:* Two-dimensional distributions symbolized by (half-)circles with radius r_0 , where r_0 denotes the distance at which the pressure anomaly has dropped to $1/e$ of the maximum value. (A) Continuous distribution as given by Eq. (1), with $r_0 = 160$ km (steep gradient, small extension of the pressure anomaly). (B) Continuous distribution as given by Eq. (1), with $r_0 = 1000$ km (small gradient, large extension of the pressure anomaly). (C) Discontinuous distribution as given by Eq. (2), with $r_0 = 400$ km and a coastline at $x = 0$. Line a–b is the profile for which the deformation and gravity effects of Figs. 2–5 are calculated

All calculated deformations and gravity changes are normalized to $p_{max} = 1$ mbar so that results for any maximum pressure can be readily deduced.

Strictly speaking, simple loading functions as given by Eq. (1) can only be applied to anomalous air pressure on the continental surface far from any coastlines. On the ocean floor, passing cyclones cause a more complicated effective pressure distribution due to the reaction of the water masses. In general, this reaction is dynamical and is affected by water depth, geometry of the coastlines, velocity of the cyclone etc., in a highly complex way. Without any dynamical effects the ocean would react to air pressure changes like an inverse barometer and would compensate an air pressure low by raising the water level such that there is no pressure change on the ocean floor. These static conditions are roughly valid for monthly mean values (Thompson 1979) or for very slow-moving cyclones.

To account for the static reaction of the ocean we have additionally used the anomalous pressure distribution

$$p_H = H(x - x_0) * p(r), \quad (2)$$

where $p(r)$ is the same as in Eq. (1), $x = x_0 = \text{const.}$ defines a straight coastline and $H(x - x_0)$ is the Heaviside step function. Thus, there is zero pressure on the ocean side of the coastline and pressure distributions $p(r)$, as defined in Eq. (1), on the continental side.

Figure 1 shows the pressure distributions used in this section: two continuous distributions (Eq. 1) with $r_0 = A, B$ and one discontinuous distribution (Eq. 2) with $r_0 = C$ and $x_0 = 0$.

Method of calculation

The method of global loading calculations for a given Earth model and a given load distribution is now standard (see

Longman 1962; Farrell 1972; Zschau 1979a among others): loading love numbers are calculated first and from this the Earth's elastic response to a surface point load, the so-called Green's functions, are determined. The Green's functions are convolved with the load distribution in order to obtain the total displacements, strain, gravity and tilt effects for any desired point on the surface. The method takes account of self-gravitation and it usually involves a hydrostatic pre-stress term.

For the numerical evaluation of the convolution integrals we have used a grid system which is determined by meridians and parallels. All the Green's functions and loading distributions were taken as constant within one grid unit. The size of each grid unit was dependent on its distance to the surface point for which the loading calculations were carried out. However, it was always chosen such that a further refinement of the grid system did not change the results for any deformation component.

For the determination of the gravitational effects we have assumed the pressure variations to be due to a homogeneous perturbation of air density within a column of 8.4 km in height (standard atmosphere). The difference between the gravitational attraction of this upwards extending mass anomaly and the corresponding surface distribution of equal mass was taken into account by distance-dependent weighting functions.

Calculations have been carried out for a spherical and radially stratified Earth. We have used the Gutenberg-Bullen A Earth model as tabulated in Alterman et al. (1961). Test calculations with more up to date Earth models such as PREM (Dziewonski and Anderson 1981) did not give significantly different results. The introduction of anelasticity (see Zschau 1979c, d, 1980), however, may be important.

Estimates based on Q -model PREM and assuming $Q \sim \omega^\alpha$ (ω = circular frequency) with $\alpha = 0.15$ (see Smith and Dahlen 1981, Okubo 1982) gave an increase of the loading deformations by 10%–20% when seasonal air pressure variations were considered. But as the structure of anelasticity in the Earth's mantle and its frequency dependence is not sufficiently well known, we have not taken anelasticity into account.

Computational results

Figures 2–5 present the displacements, horizontal strain, gravity and tilt effects due to the surface pressure distributions A, B, C of Fig. 1. A, B, C are chosen to be 160 km, 1,000 km and 400 km, respectively. While the results in cases A and B are valid for any profile crossing the centre of the pressure anomaly, in case C they are valid along a profile perpendicular to the coastline (a–b perpendicular to $x=0$ in Fig. 1).

As mentioned above, all values are normalised to 1 mbar pressure at the (anti-)cyclone's centre. In all cases, the direction pointing upwards is taken to be positive. Thus, contrary to the common notation, gravity will be positive if the attracting force is directed upwards. This ensures that only one coordinate system is used. Note the logarithmic scale of the length coordinate.

(a) *Displacements (Fig. 2)*. In the case of continuous pressure distributions (A, B), the maximum displacements turn out to be in the range of $\mp(1-2.5)$ cm for the vertical com-

ponent and $\pm(1-2.5)$ mm for the horizontal component. This assumes a maximum pressure anomaly between ± 20 and ± 60 mbar. The vertical component takes its extreme values at the centre of the (anti-)cyclone, the horizontal component at its flanks.

The magnitude of the displacements is critically dependent on the spatial extension r_0 of the pressure distribution. For instance, the vertical displacements for the $r_0 = 1,000$ km pressure anomaly (B) are up to four times stronger than those of the $r_0 = 160$ km anomaly (A). From this it is clear that there cannot be any unique regression coefficient between local displacements and local air pressure changes which could be used to correct geodynamic measurements for air-pressure-induced surface displacements. The latter depend on the amplitude *as well as* on the spatial width of the pressure anomaly.

The introduction of a step function into the geometry of the load (C) modifies the displacements considerably; the symmetry with respect to the coastline ($x=0$) is lost. Furthermore, the order of magnitude of the vertical displacements is the same as in case A although the width of pressure distribution A is less than half of the width of pressure distribution C. With the horizontal displacements, the order of magnitude is the same as in case B, although the width of pressure distribution B is more than twice as high as that of C. Thus, vertical displacements are weakened in the vicinity of a coastline, whereas horizontal displacements are amplified.

(b) *Horizontal strains (Fig. 3)*. Figure 3 gives the horizontal principal strains along the profile a–b in Fig. 1. The maximum strain effects for the axial and azimuthal components corresponding to the continuous pressure distributions (A, B) are $\mp(4-12) \cdot 10^{-9}$ if, as before, the pressure anomaly is assumed to be $\pm(20-60)$ mbar at the centre of the (anti-)cyclone.

Compared to the displacements, the maximum amplitudes of the axial and azimuthal strains are not strongly determined by the width of the pressure distribution. The shape of the strain curves A, B is however, except for the sign, rather similar to that of the corresponding pressure curves. It suggests approximately $(-1.5) - (-2.0) \cdot 10^{-10}$ strain per mbar of local air pressure change below the centre of the pressure anomaly. This coefficient, however, is not applicable to the flanks of the anomaly and it is not well enough determined to be used for precise air pressure corrections of geodynamic measurements.

The assumption of discontinuous surface pressure (C) leads to extremely high near-coast axial strain values (more than $-4 \cdot 10^{-10}$ strain/mbar, with a change in sign at the location of the pressure discontinuity. The corresponding azimuthal strain component is described by a smooth curve similar to that for the vertical displacements.

(c) *Gravity effect (Fig. 4)*. Each high and low pressure area corresponds to a density anomaly in the air. It causes an anomaly in the Newtonian attraction which is commonly known as the primary gravitational effect. The secondary gravitational effect of air pressure changes is due to the elastic deformation of the solid Earth. It includes gravity changes due to the shift of the surface through the gravity field and due to the redistribution of mass in the Earth's interior commonly known as self-gravitation.

A look at Fig. 4 demonstrates that the primary gravity

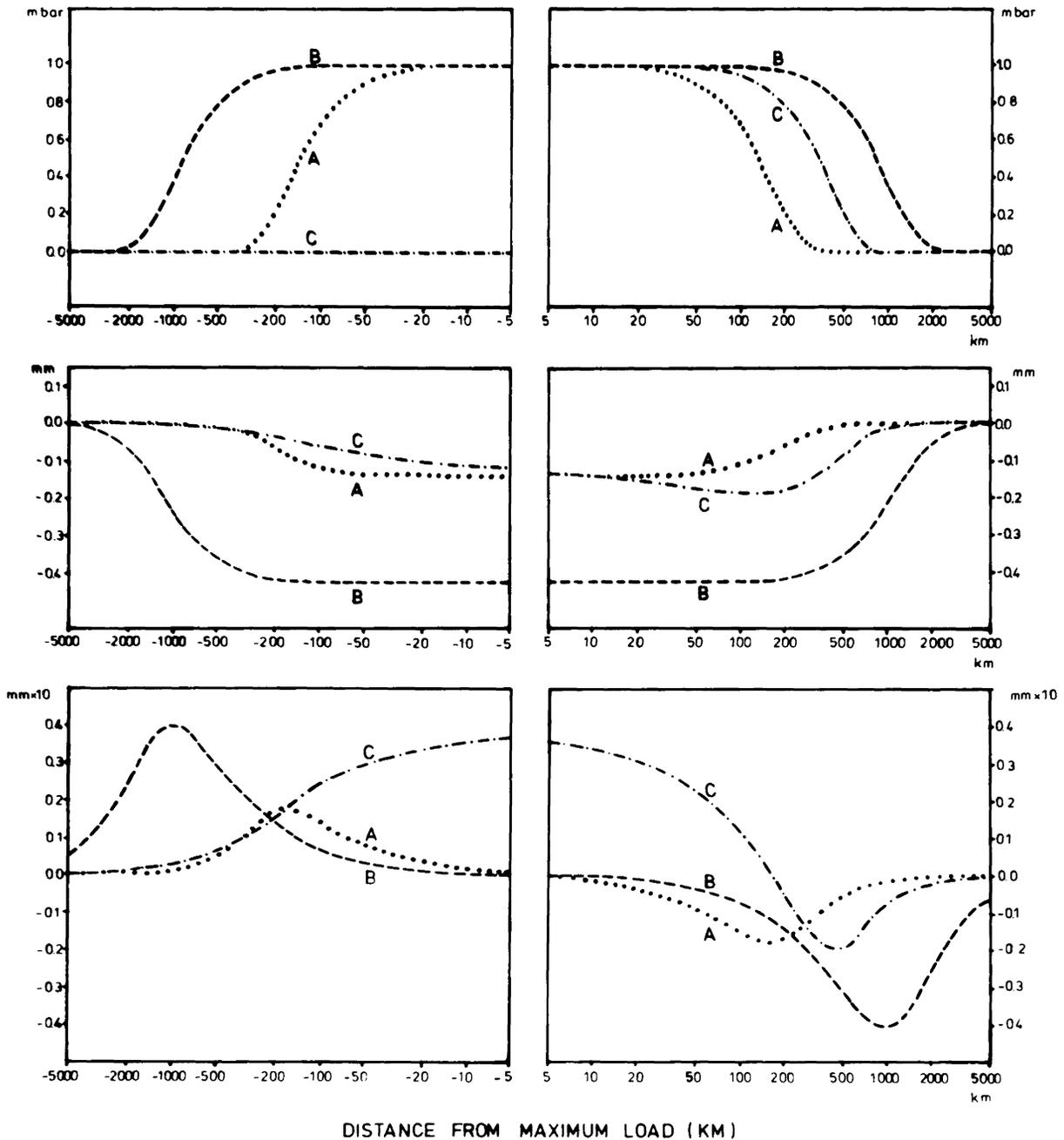


Fig. 2. *Top*: Axial cut through the anomalous air pressure distribution A, B, C of Fig. 1. *Middle*: Corresponding vertical displacement (positive upwards). *Bottom*: Corresponding horizontal displacements in the direction of profile a-b of Fig. 1

effect is dominant. It correlates remarkably well with the pressure variation. In the case of pressure distributions A and B, a good regression coefficient is $0.4 \mu\text{gal}/\text{mbar}$. For the secondary gravitational effect such a correlation is not seen as clearly. Thus, in case of the total gravitational effect the regression coefficient for distributions A and B may vary between 0.3 and $0.4 \mu\text{gal}/\text{mbar}$ resulting in a gravity change of around $\pm 20 \mu\text{gal}$ if, as before, a pressure change of $\pm 60 \text{ mbar}$ is assumed. The above regression coefficient is in good agreement with similar computations by Warburton and Goodkind (1977). It also agrees with the results obtained by Spratt (1982) who compared records of a superconducting gravimeter with local barometric pressure chan-

ges and came up with an averaged admittance of 0.29 – $0.34 \mu\text{gal}/\text{mbar}$.

All results are based upon the assumption that the density perturbation is homogeneous within a standard atmosphere of 8.4 km in height and that the perturbation can be estimated from the surface pressure by means of the hydrostatic approximation.

A strong increase of the regression coefficient between air pressure variations and the primary gravity effect seems to occur under air pressure anomalies in ocean areas. Here, the coefficient goes up to $0.8 \mu\text{gal}/\text{mbar}$ (curve C). This amplification is due to the additional effect of water level changes caused by the static compensation of air pressure

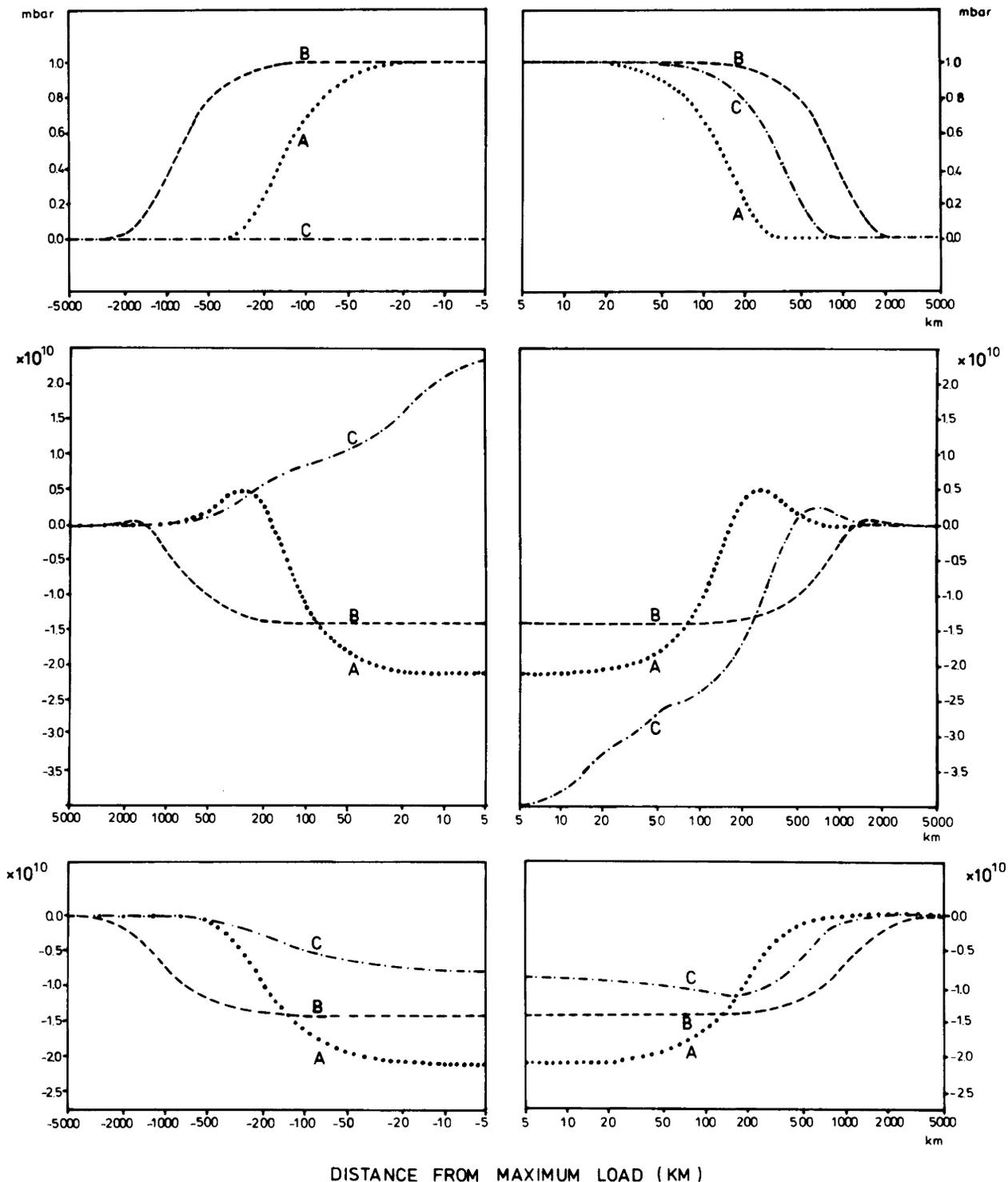
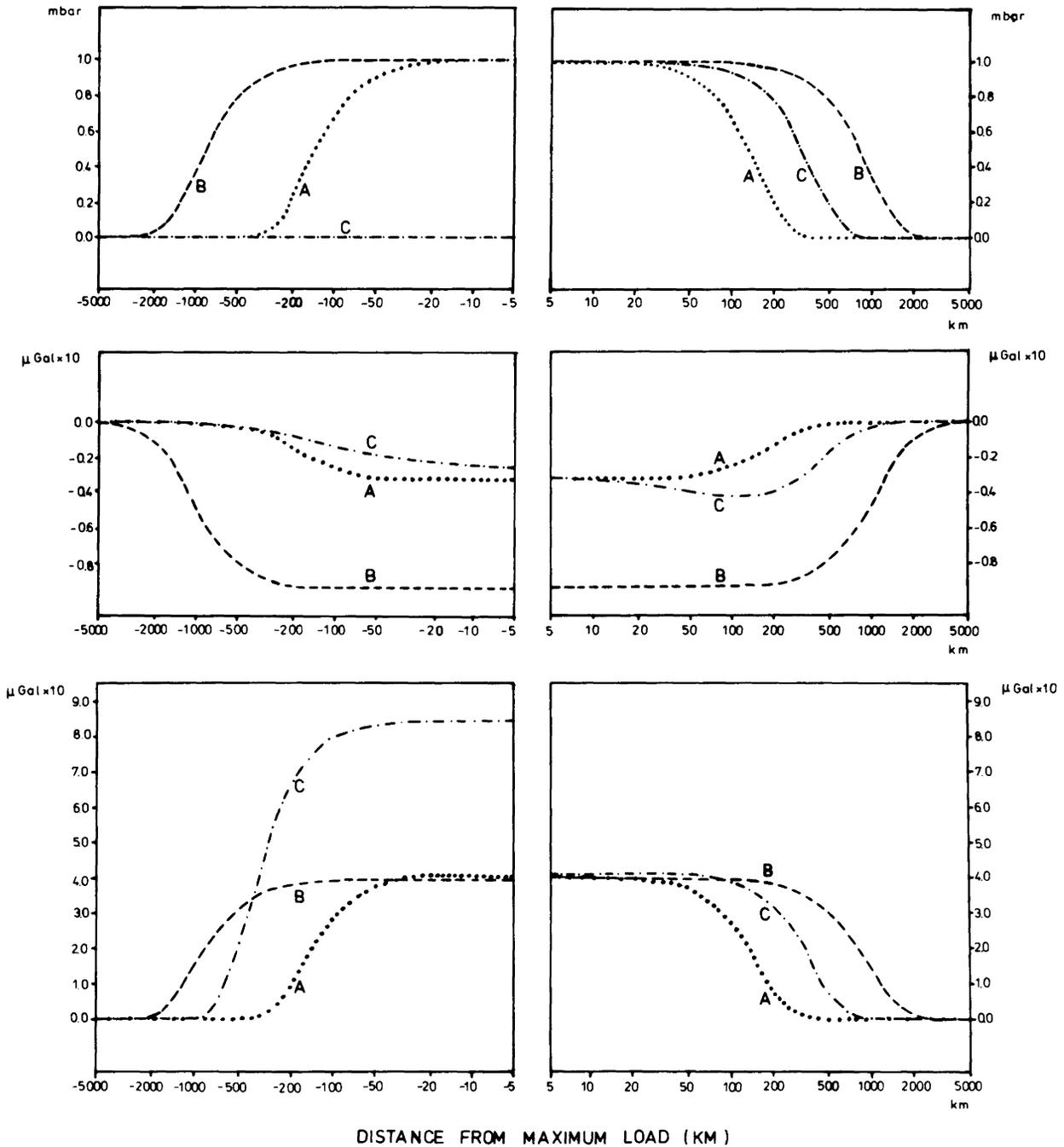


Fig. 3. *Top:* Axial cut through the anomalous air pressure distributions A, B, C of Fig. 1. *Middle:* Corresponding horizontal principal strain in the direction of profile a-b in Fig. 1 (axial strain). *Bottom:* Corresponding horizontal principal strain perpendicular to the direction of profile a-b in Fig. 1 (azimuthal strain). positive strain: increase of length (dilatation) negative strain: decrease of length (compression)

loading of the ocean. It will be measured by a gravimeter installed at mean sea level on a little island in the ocean whether the island is far from the coast or as near as 5 km to the next coastline. Only in the direct vicinity of the coastline is the regression coefficient lower and takes values between 0.8 and 0.4 $\mu\text{gal}/\text{mbar}$. This, however, neglects the height of the gravimeter station above mean sea level which,

at coastal stations, may significantly change the above coefficient.

Comparing Fig. 4 with Fig. 2 one can also deduce a regression coefficient between air-pressure-induced vertical displacements and secondary gravity changes. On average it is 0.46 $\text{cm}/\mu\text{gal}$ and is valid for both continuous (A, B) and discontinuous (C) pressure distributions. The corre-



DISTANCE FROM MAXIMUM LOAD (KM)

Fig. 4. *Top*: Axial cut through the anomalous air pressure distributions A, B, C of Fig. 1. *Middle*: Corresponding gravity effect caused by the elastic deformation (secondary effect). *Bottom*: Corresponding gravity effect caused by the Newtonian gravitation of the anomalous air mass (primary effect). Gravity is positive upwards!

sponding correlation is much better than that between the displacements and the air pressure variations. Unfortunately, such a good correlation is not obtained for the total gravity. Thus, the desired air pressure correction of global positionings with space techniques by means of gravity recordings would still require the difficult separation of the secondary gravity effect from the primary one.

(d) *Tilt* (Fig. 5). Similar to gravity, air-pressure-induced tilt consists of a primary effect, which is the deviation of the vertical due to the gravitational attraction of air, and a secondary effect which is the elastic deformation including

the tilt of the surface and a secondary deviation of the vertical due to self-gravitation of the deforming Earth.

Contrary to gravity, the secondary tilt component is larger than the primary one. In the case of continuous pressure distributions (A, B) the extreme values of tilt are found under the flanks of (anti-)cyclones as in the case of horizontal displacements. In contrast to the displacements, however, the magnitude of the maximum tilt is less dependent upon the width of the pressure anomaly than upon the pressure gradient. Taking ± 20 mbar and ± 60 mbar again for the average and extreme pressure anomalies, respectively, in the centre of (anti-)cyclones, one obtains a maximum

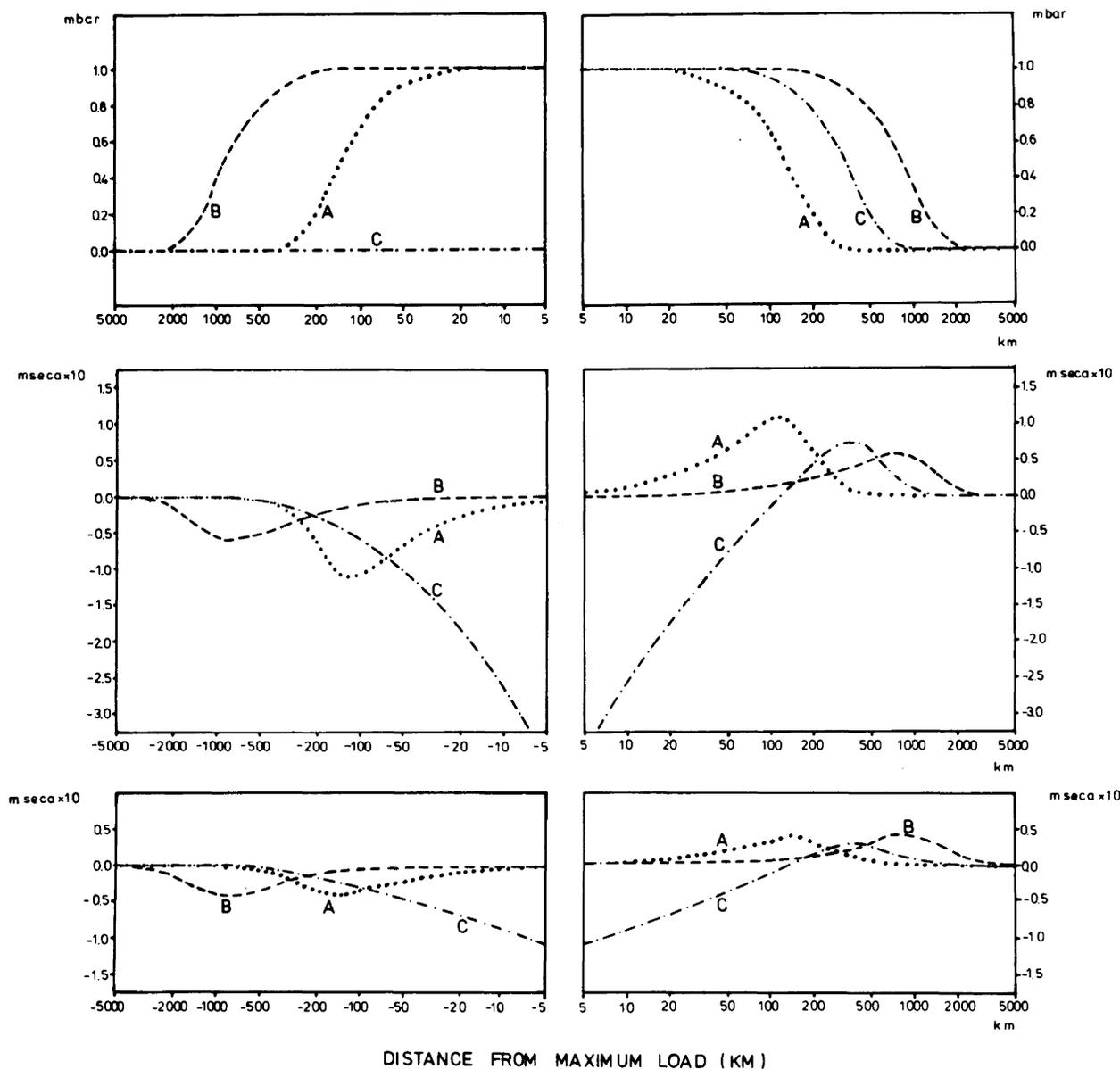


Fig. 5. *Top:* Axial cut through the anomalous air pressure distributions A, B, C of Fig. 1. *Middle:* Corresponding tilt caused by the elastic deformation (secondary effect). *Bottom:* Corresponding tilt (deflection of the vertical) caused by the Newtonian gravitation of the anomalous air mass (primary effect). All tilts are valid for the direction of profile a-b in Fig. 1 (positive, if the tip of a vertical pendulum moves into (-x)-direction)

tilt anomaly as high as $\pm(3-10)$ msec. For the discontinuous pressure distribution (C) this maximum increases by tens of msec. Thus, in the vicinity of coastlines air-pressure-induced tilt may be expected to be extremely high.

Deformations due to seasonal air pressure variations

While the last section dealt with the atmospherical loading at periods in the range of several days, in this section the effects of seasonal air pressure variations will be estimated. It is well known that monthly mean values of air pressure can be roughly characterized by a standing wave with extreme values in winter and summer and zero points in spring and autumn. Hence, one can get an idea of the maximum air-pressure-induced deformation and gravity effects by comparing the states of deformation between January and July.

The corresponding difference in the air pressure distribution is shown in Fig. 6. It is tabulated in Munk and Macdonald (1960), with gridpoints spaced every 10° in latitude and 20° in longitude.

The Munk and Macdonald pressure maps have been convolved with the Green's functions for the Gutenberg-Bullen A Earth model to yield the differences in the state of deformation and gravity between January and July. Again, the inverse barometer response has been assumed for oceanic areas. The results of the convolution are presented in Figs. 7-11, showing maximum seasonal variations of more than 1 cm for vertical displacements, 1 mm for horizontal displacements, 3×10^{-9} for horizontal strain, nearly 6 μgal for gravity and around 3 msec for tilt. In the case of horizontal displacements and tilt these extreme values are obtained for a region in the neighbourhood of Siberia and in the case of gravity the maximum value is

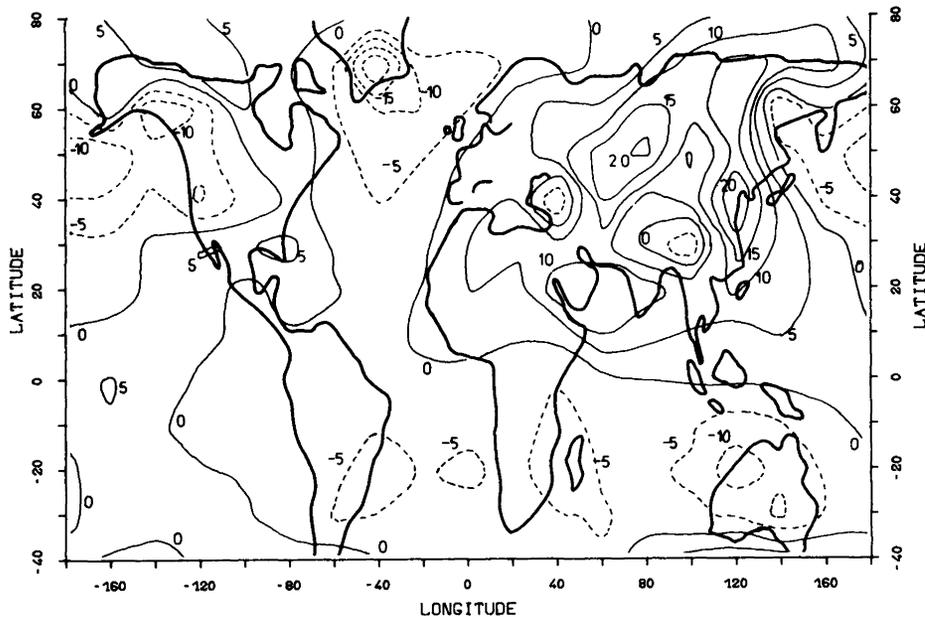


Fig. 6. Seasonal deviation of air pressure from atmospheric mean (mbar): "January minus July", according to Munk and Macdonald (1960)

obtained for Greenland. In all other cases they are valid for Siberia.

Air pressure corrections for radial displacement and gravity

Regarding the enormous precision and high stability of the superconducting gravimeter and the recent developments in precise point positioning with space methods, gravity changes and radial displacements seem to be the components of deformation in global geodynamics which require air pressure corrections most necessarily.

A simple procedure for correcting radial displacements is not yet available. Gravity records are usually corrected by using a regression coefficient around $0.3 \mu\text{gal}/\text{mbar}$. However, it was pointed out in the previous section that there is no one such regression coefficient, but that the coefficient is critically dependent on the wavelength of the loading. This is demonstrated in Fig. 12 which shows the seasonal gravity changes of Fig. 10 as a function of the local air pressure changes. Deviations from the best fitting regression line exceed the $1 \mu\text{gal}$ level. If one considers that the internal precision of the superconducting gravimeter lies around a tenth of a microgal, the shortcoming of the above correction will immediately become obvious.

The same applies to the possible correction of the radial displacements by means of one regression coefficient. It is seen from Fig. 13 that the deviations from the best fitting regression line may be as high as 4 mm. At present this correction error may not be serious for many purposes. But it will be too high if NASA can realize its plan to detect vertical surface motions with millimetre-scale accuracies (Walter 1984).

In order to improve these air pressure corrections we have prepared Tables 1–3. They give the air pressure regression coefficients for radial displacements and for primary and secondary gravity changes, respectively, as a function of the characteristic radius r_0 of the pressure anomaly and as a function of position with respect to the centre of the anomaly. The pressure distributions are assumed to obey Eq. (1). The positions with respect to the centre are determined by the relative pressure amplitudes p/p_{max} .

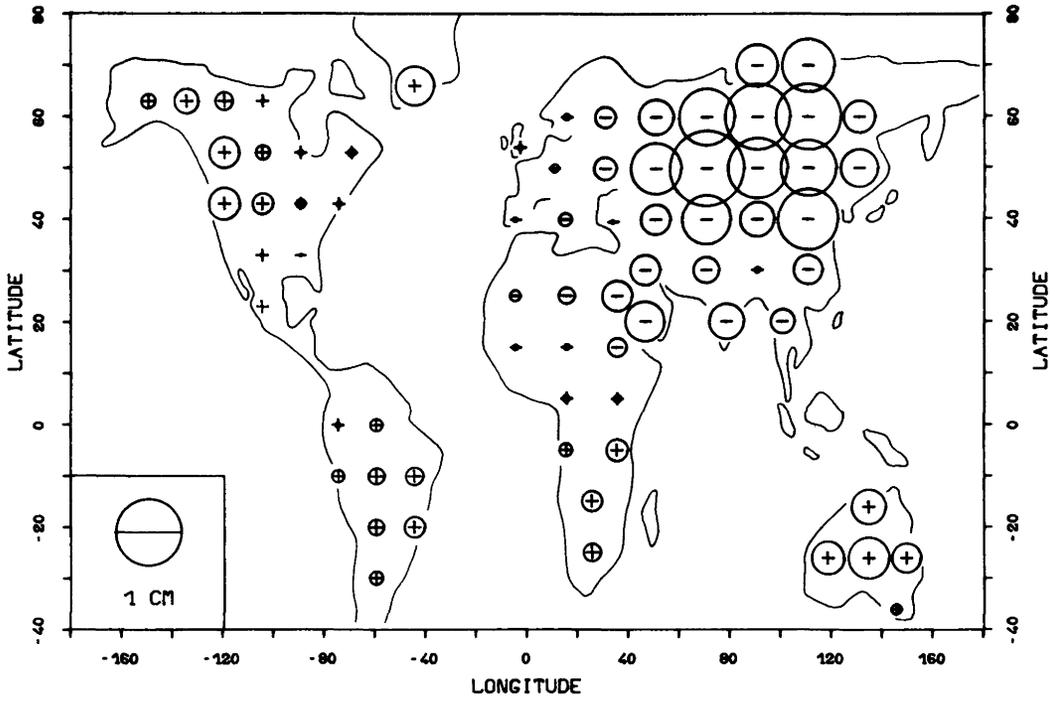
Figures 14–16 show these regression coefficients graphically. For the radial displacements the coefficient changes from approximately $-0.1 \text{ mm}/\text{mbar}$ at $r_0 = 160 \text{ km}$ to $-0.9 \text{ mm}/\text{mbar}$ at $r_0 = 5,500 \text{ km}$. For the primary gravity changes the coefficient lies between about $0.41 \mu\text{gal}/\text{mbar}$ and $0.31 \mu\text{gal}/\text{mbar}$ and for the secondary gravity changes the possible regression coefficients are between $-0.02 \mu\text{gal}/\text{mbar}$ and $-0.18 \mu\text{gal}/\text{mbar}$. It is obvious from these numbers that, for an accurate air pressure correction, an estimate of the characteristic wavelength of the pressure perturbation is absolutely necessary. With such an estimate the use of Tables 1–3 will give better results than the simple correction with only one regression coefficient.

It should be mentioned, however, that in general the line of regression between air pressure and a deformation component does not go through the origin but has the form

$$w = C_1 p + C_2 p_{\text{max}} \quad (3)$$

(see also Figs. 12 and 13), where w stands for any of the components u , g_p or g_s and where C_1 and C_2 are coefficients slightly dependent on r_0 and p/p_{max} . Only the coefficient C_1 is given in Tables 1–3 and shown in Figs. 14–16. The constant term $C_2 p_{\text{max}}$ will disappear, if not the deformation component itself, but its difference between two points at the Earth's surface is considered. These points may even belong to different cyclones because the effective constant term is always the sum of the constant terms of all cyclones on the Earth's surface. Therefore, it is the same for every point on the surface. From Figs. 12 and 13 it is seen that in practice the shifts of the regression lines are very small. Thus, the consideration of differences between the displacements or gravity variations at two different points is only necessary if the required precision of the air pressure correction is in the millimeter range and a fraction of a microgal.

The use of Tables 1–3 for the correction of air pressure effects on radial displacements and gravity requires an estimate of the characteristic wavelengths contained in a pressure distribution. Besides involving a certain amount of subjectivity, such an estimation may also involve somewhat tedious processes like two-dimensional filtering, smoothing



VERTICAL DISPLACEMENT

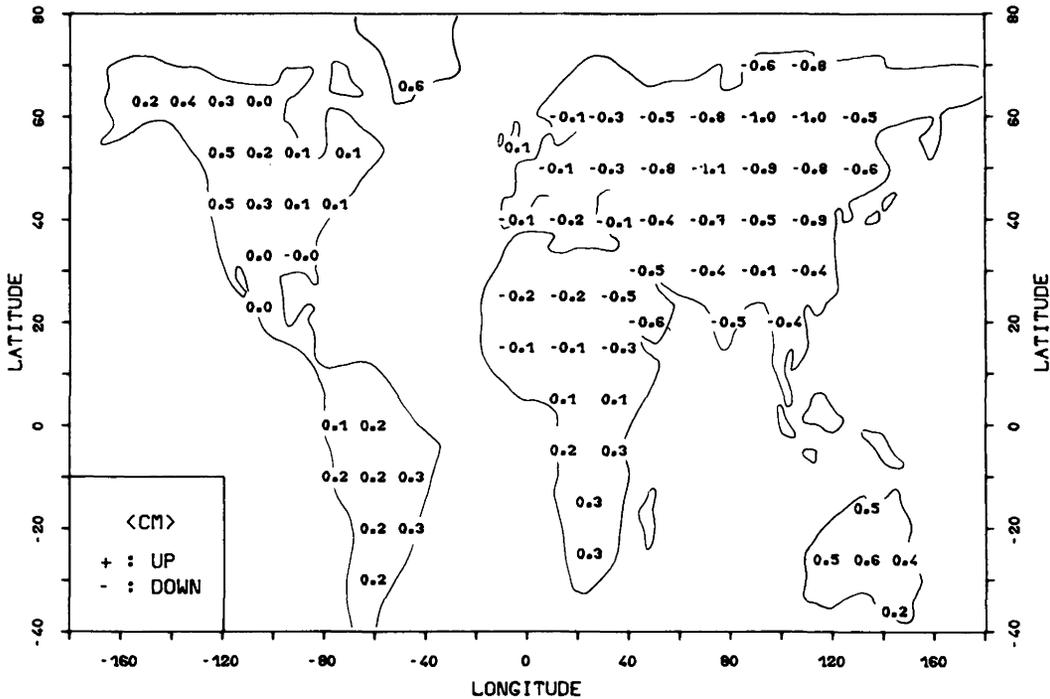


Fig. 7. Vertical displacements due to seasonal changes in the global air pressure distribution (Fig. 6): mean deviation “January minus July”

of the pressure distribution etc., which are not always easily done without a computer. We are therefore proposing a second and by far simpler method, which uses only two regression coefficients; one belonging to the long-wavelength loading and the other belonging to the short-wavelength loading. The proposed air pressure corrections are as follows:

Radial displacement: $u = -0.90 \bar{p} - 0.35 (p - \bar{p})$
 Primary gravity: $g_p = 0.36 \bar{p} + 0.41 (p - \bar{p})$ (4)
 Secondary gravity: $g_s = -0.17 \bar{p} - 0.08 (p - \bar{p})$
 Total gravity: $g = g_p + g_s.$

u will be in mm and the gravity values will be in μgal if the pressure values p, \bar{p}, \bar{p} are taken in mbar. p is the

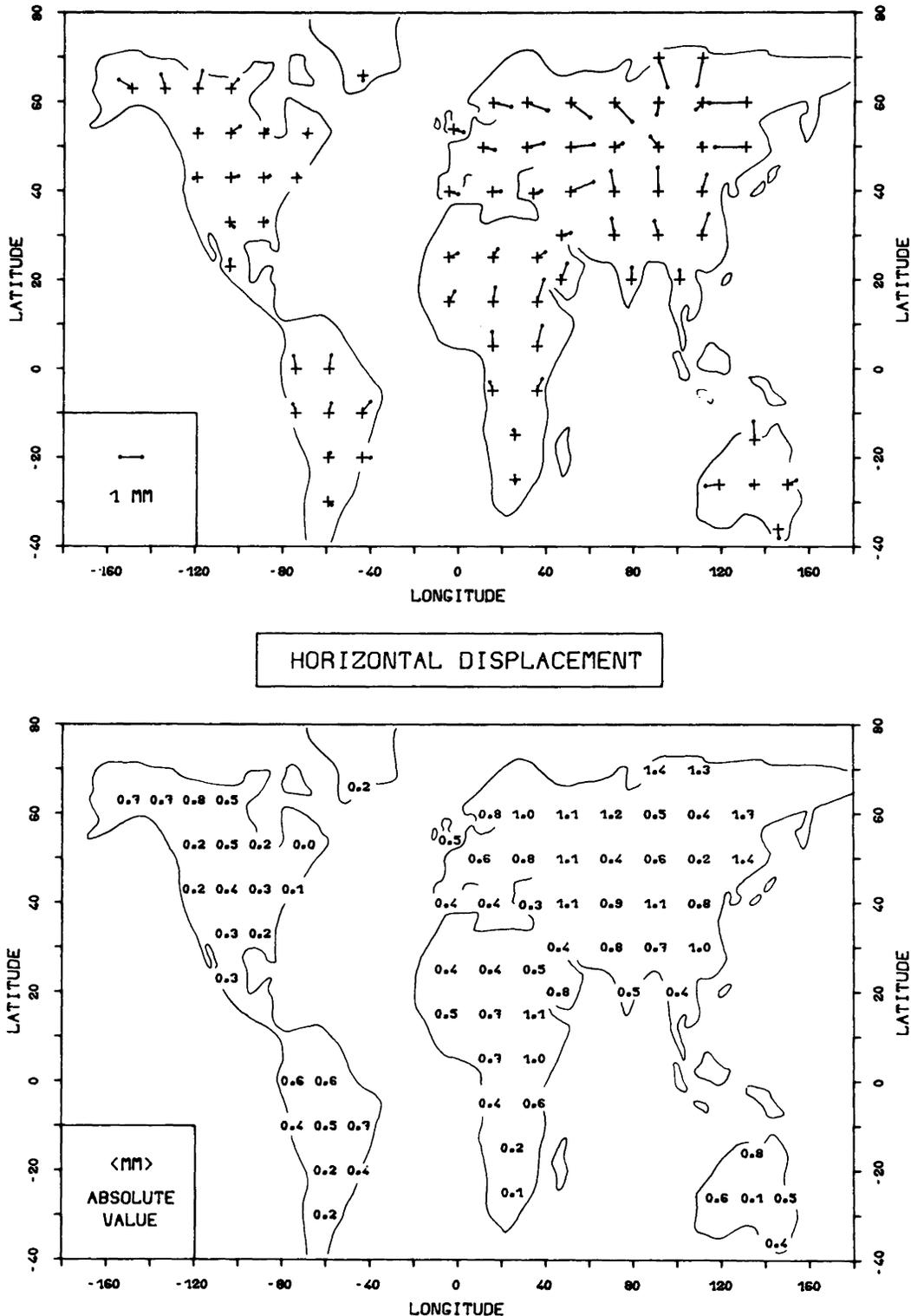


Fig. 8. Horizontal displacements due to seasonal changes in the global air pressure distribution (Fig. 6): mean deviation "January minus July"

measured pressure variation at the surface point under investigation. \bar{p} denotes the long-wavelength component of the pressure variation. It is obtained by averaging the pressure variations in a surrounding area of 2,000 km. \bar{p} is obtained in the same way, except that the pressure changes for the ocean areas are set to zero. This is necessary in order to account for the inverse barometer effect of the

oceans which influence the radial displacements and the secondary gravity changes.

We have applied the above empirical corrections to the seasonal variations of Figs. 6, 7 and 10. For averaging the pressure values we have used a $1,000 \times 1,000$ km² grid system consisting of 16 squares in each of which the average was estimated by inspection. The corrections determined

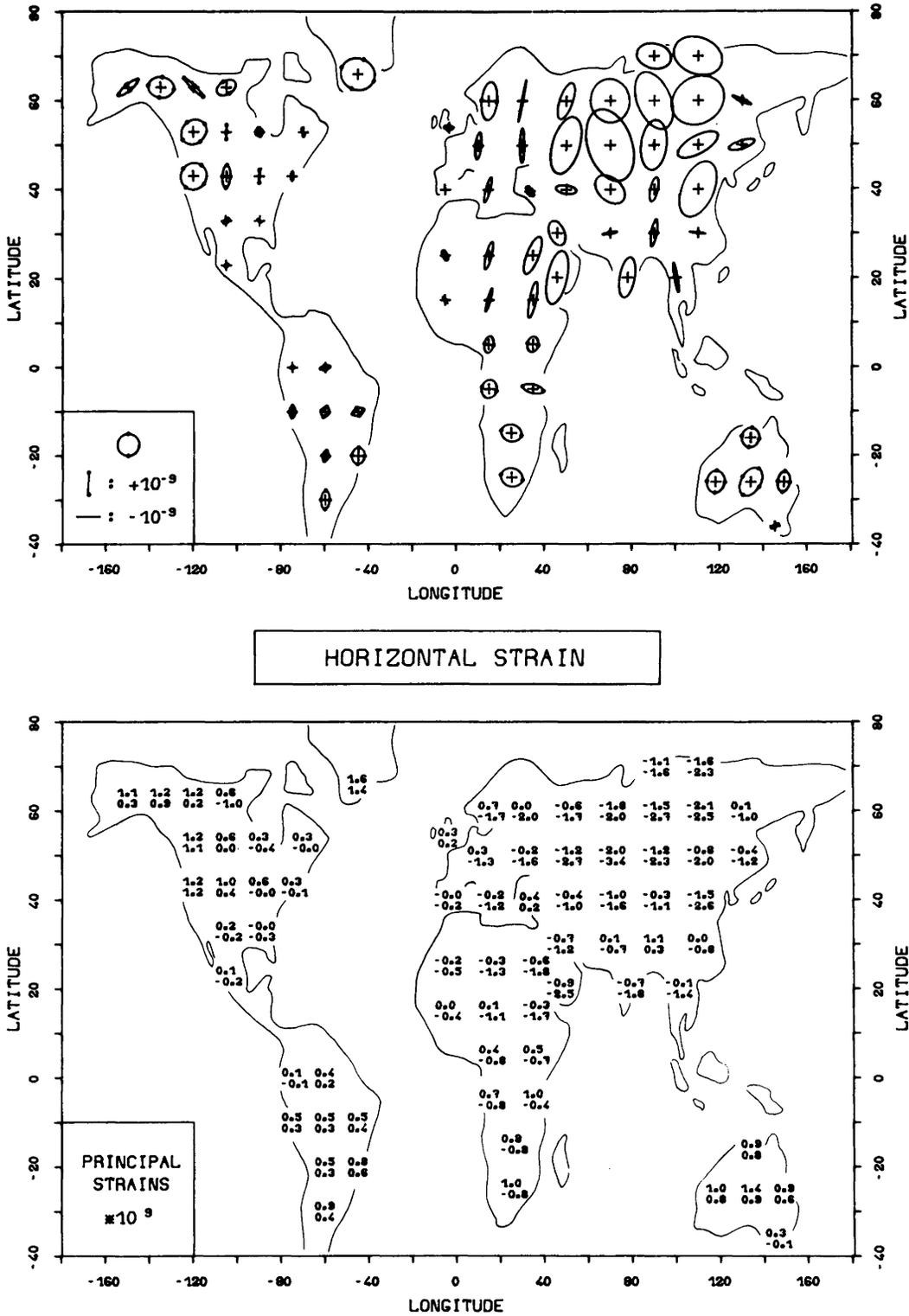


Fig. 9. Horizontal principal strains due to seasonal changes in the global air pressure distribution (Fig. 6): mean deviation “January minus July”. *Top*: The principal horizontal strain components coincide with the principal axes of the ellipses. The length of the principal axis is proportional to the principal strains. *Dots* at the end of a principal axis denote dilatation, otherwise compression is meant. *Bottom*: “+” dilatation, “-” compression

in this way are compared in Figs. 17–20 with the corrections as obtained from exact loading calculations by means of Green’s functions. Except for a small constant term there is excellent agreement between both. The mean differences are less than 0.5 mm, 0.1 μgal , 0.1 μgal and 0.2 μgal for

the radial displacements, the primary, secondary and total gravity changes, respectively. At maximum, the error in the estimation of the air-pressure-induced radial displacements is ± 1 mm. For the primary and secondary gravity changes the maximum estimation error is ± 0.3 μgal and

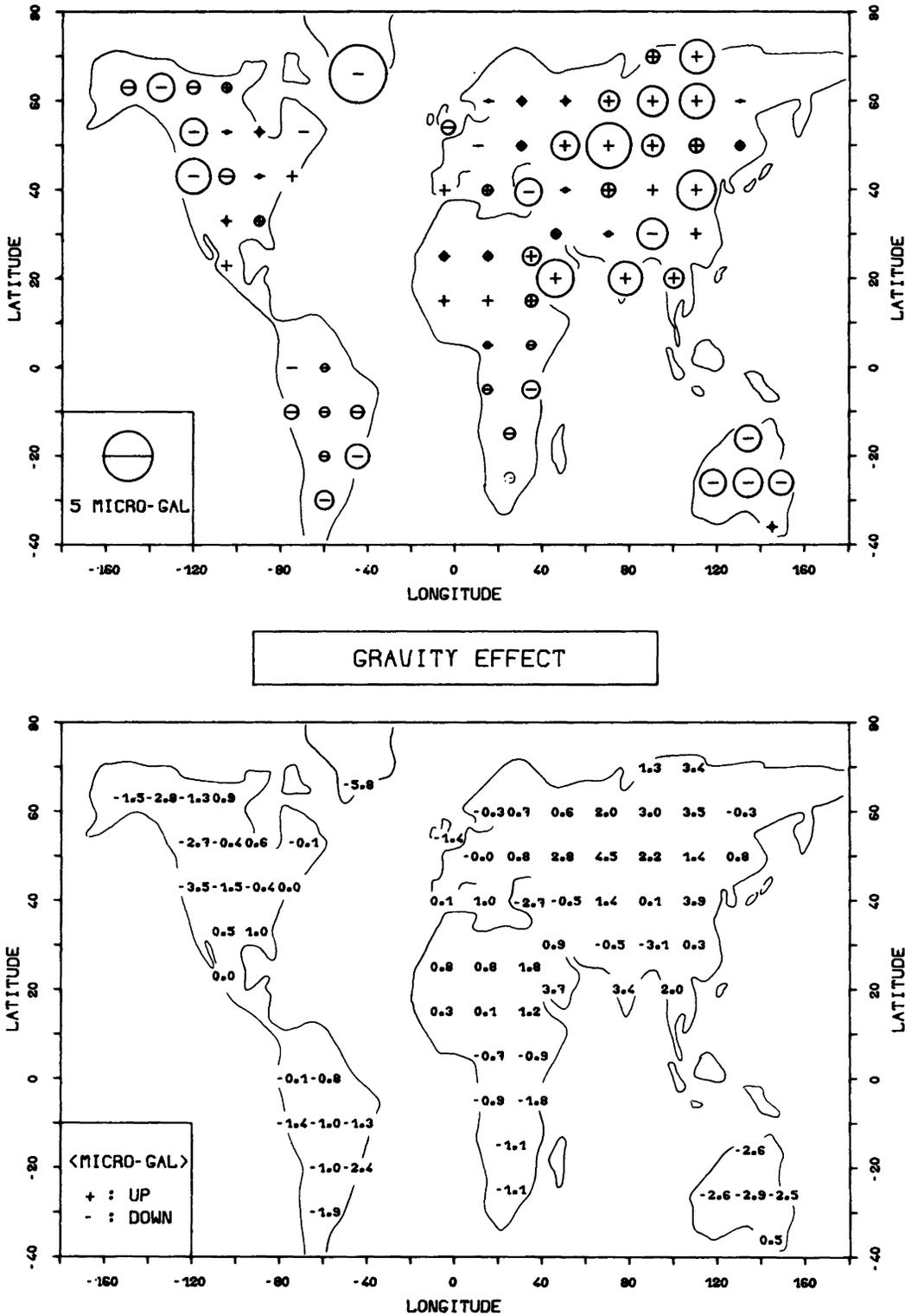


Fig. 10. Total gravity effect due to seasonal changes in the global air pressure distribution (Fig. 6): mean deviation "January minus July". Gravity is positive upwards!

$\pm 0.2 \mu\text{gal}$, respectively, and for the total gravity changes $\pm 0.4 \mu\text{gal}$. This does not take the constant term into account, which is due to distant pressure variations and which is the same for every point on the surface (see above). For the radial displacements it turns out to be practically zero (less than 0.1 mm), for the primary and secondary gravity changes its absolute value is less than $0.2 \mu\text{gal}$ and $0.1 \mu\text{gal}$,

respectively, and for the total gravity changes less than $0.3 \mu\text{gal}$. As explained above, the constant term will not be important if gravity or displacement differences between two points on the Earth's surface are considered. The estimation errors involved in Eq. (4) are, on average, four times smaller than those obtained when using the best fitting line of regression between gravity changes and displacements

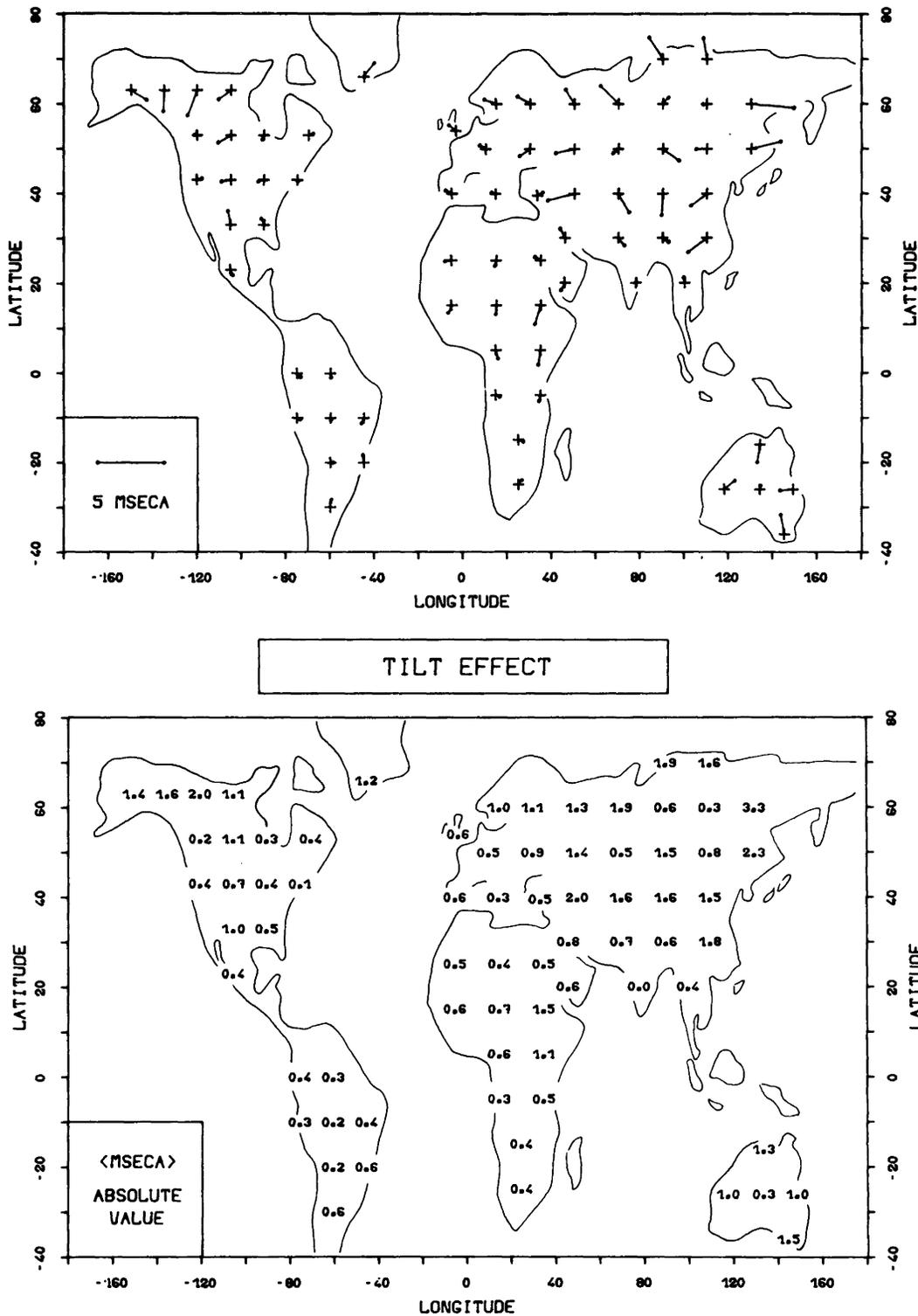


Fig. 11. Total tilt due to seasonal changes in the global air pressure distribution (Fig. 6): mean deviation "January minus July"

on the one hand and pressure variations on the other hand. Thus, the proposed correction method is not only simple but it is also highly precise and fulfils the present requirements of geodynamics.

Conclusions

The magnitude of air-pressure-induced deformation and gravity effects far from any coastline can reach more than

10% of the corresponding body tide effects. For tilt and strain the air pressure influence may even be of the same order of magnitude as the body tide. Maximum values are tabulated in Table 4.

These results have to be revised if the loading distribution is not continuous but contains a discontinuity due to the existence of a coastline. Vertical displacements, azimuthal strains and secondary gravity changes are slightly weakened near the coast, whereas all the other deformation

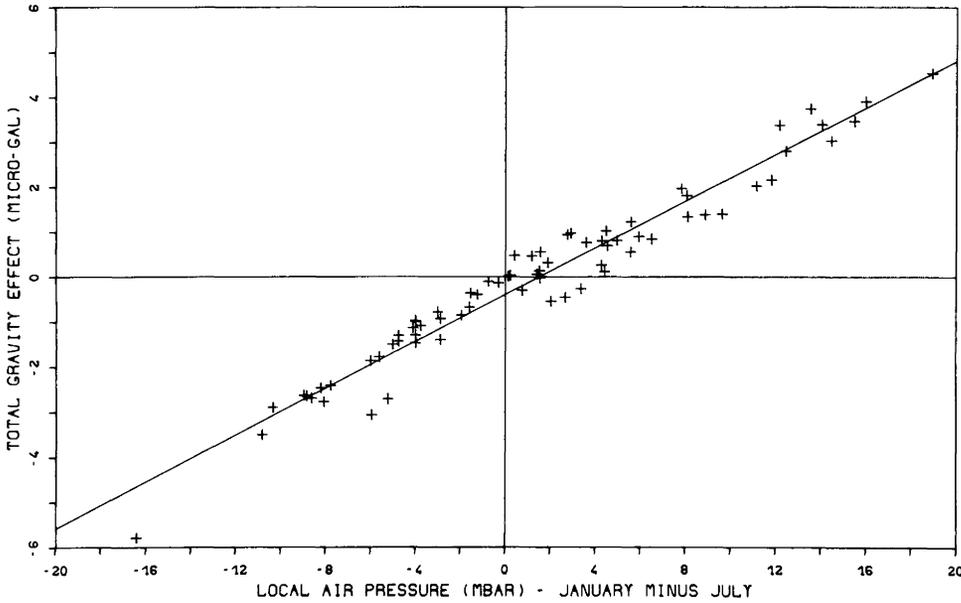


Fig. 12. Total gravity effect due to seasonal changes in the global air pressure distribution (Fig. 6) as a function of local air pressure. Deviations from the best fitting regression line exceed the $1 \mu\text{gal}$ level!

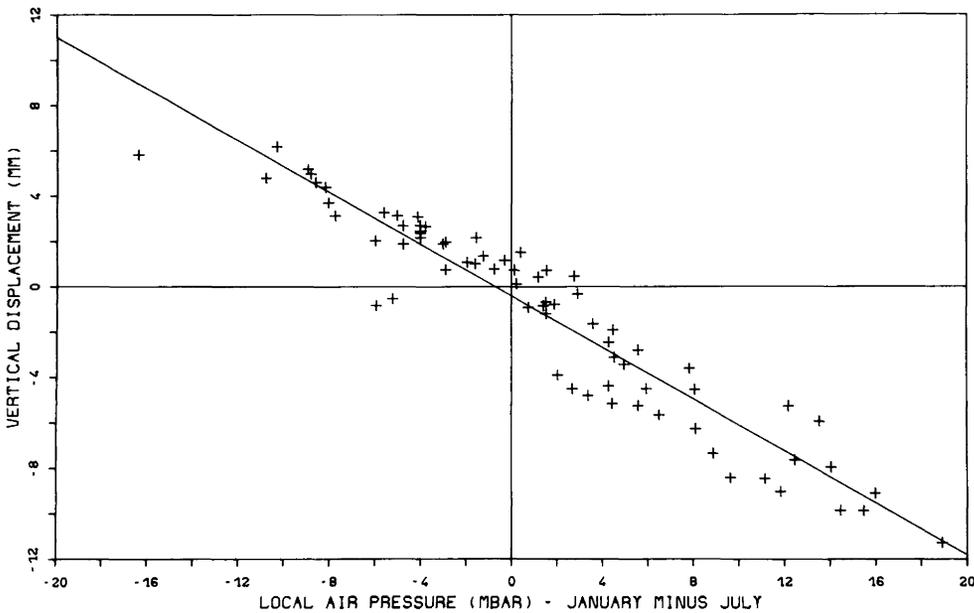


Fig. 13. Vertical displacement due to seasonal changes in the global air pressure distribution (Fig. 6) as a function of local air pressure. Deviations from the best fitting regression line may be as high as 4 mm!

components are amplified up to several hundred percent in the vicinity of coastlines, the amplification depending on the deformation components, geometry of the coastline, distance to the coastline and of course on the size and amplitude of the pressure anomaly itself. In addition, even the sign of a deformation component may change due to the existence of a coastline.

Further complications may arise from the dynamic reaction of the ocean to air pressure loading which was not taken into account here.

The vertical displacements in Table 4 coincide well with

those given by Trubytsin and Makalkin (1976) who carried out analytical calculations for cyclones above a homogeneous half-space. They are also in good agreement with the seasonal vertical movements given by Stolz and Larden (1979). In all cases the displacements are of centimetre-order. This corresponds to the precision which is presently approached for the measurements of regional and global base lengths using geodetic space techniques such as VLBI (Very Long Baseline Interferometry) and Laser Ranging. Even millimetre-scale accuracies are already envisaged (Walter 1984). A correction of such measurements for air-

Table 1. Regression coefficient $C_1 = du/dp$ (mm/mbar) for vertical displacement [see Eq. (3)]. p : local pressure, P_{max} : maximum pressure at the centre of a cyclone, r_o : radius of a cyclone [see Eq. (1)]

P/P_{max} (%)	20	35	50	65	80	95
r_o (km):						
160	-0.134	-0.114	-0.105	-0.099	-0.096	-0.092
300	-0.207	-0.182	-0.171	-0.156	-0.156	-0.148
450	-0.269	-0.236	-0.220	-0.214	-0.204	-0.198
600	-0.312	-0.280	-0.262	-0.251	-0.245	-0.238
800	-0.363	-0.329	-0.310	-0.296	-0.287	-0.283
1,000	-0.416	-0.371	-0.348	-0.335	-0.326	-0.319
1,300	-0.479	-0.425	-0.404	-0.385	-0.374	-0.368
1,600	-0.544	-0.481	-0.452	-0.435	-0.421	-0.412
2,000	-0.629	-0.557	-0.520	-0.495	-0.480	-0.469
2,500	-0.734	-0.646	-0.600	-0.572	-0.552	-0.537
3,000	-0.824	-0.726	-0.677	-0.644	-0.620	-0.606
3,500	-0.890	-0.794	-0.743	-0.710	-0.684	-0.668
4,000	-0.925	-0.846	-0.798	-0.766	-0.741	-0.724
4,500	-0.933	-0.878	-0.839	-0.809	-0.788	-0.773
5,000	-0.915	-0.890	-0.863	-0.842	-0.824	-0.813
5,500	-0.874	-0.885	-0.875	-0.862	-0.850	-0.843

Table 2. Regression coefficient $C_1 = dg_p/dp$ ($\mu\text{gal}/\text{mbar}$) for primary gravity [see Eq. (3)]. p : local pressure, p_{max} : maximum pressure at the centre of a cyclone, r_o : radius of a cyclone [see Eq. (1)]

p/p_{max} (%)	20	35	50	65	80	95
r_o (km)						
160	0.4084	0.4041	0.4014	0.3996	0.3985	0.3932
225	0.4147	0.4081	0.4057	0.4059	0.4080	0.4090
450	0.4148	0.4107	0.4098	0.4155	0.4135	0.4096
600	0.4097	0.4114	0.4098	0.4106	0.4141	0.4113
800	0.4045	0.4091	0.4106	0.4095	0.4099	0.4125
1,000	0.4008	0.4052	0.4071	0.4082	0.4088	0.4093
1,300	0.3981	0.3989	0.4037	0.4034	0.4041	0.4068
1,600	0.3904	0.3936	0.3970	0.4001	0.4014	0.4019
2,000	0.3796	0.3876	0.3916	0.3929	0.3954	0.3969
2,500	0.3679	0.3776	0.3828	0.3861	0.3888	0.3891
3,000	0.3574	0.3679	0.3749	0.3784	0.3809	0.3833
3,500	0.3466	0.3585	0.3663	0.3712	0.3738	0.3765
4,000	0.3344	0.3497	0.3581	0.3635	0.3670	0.3693
4,500	0.3237	0.3407	0.3504	0.3557	0.3600	0.3632
5,000	0.3144	0.3320	0.3425	0.3489	0.3532	0.3569
5,500	0.3055	0.3239	0.3353	0.3421	0.3469	0.3502

Table 3. Regression coefficient $C_1 = dg_s/dp$ ($\mu\text{gal}/\text{mbar}$) for secondary gravity [see Eq. (3)]. p : local pressure, p_{max} : maximum pressure at the centre of a cyclone, r_o : radius of a cyclone [see Eq. (1)]

p/p_{max} (%)	20	35	50	65	80	95
r_o (km)						
160	-0.0305	-0.0262	-0.0241	-0.0229	-0.0221	-0.0212
300	-0.0465	-0.0412	-0.0388	-0.0354	-0.0354	-0.0336
450	-0.0603	-0.0530	-0.0495	-0.0481	-0.0460	-0.0446
600	-0.0703	-0.0631	-0.0589	-0.0565	-0.0552	-0.0535
800	-0.0816	-0.0740	-0.0697	-0.0666	-0.0646	-0.0636
1,000	-0.0931	-0.0832	-0.0782	-0.0754	-0.0733	-0.0717
1,300	-0.1056	-0.0946	-0.0901	-0.0861	-0.0837	-0.0825
1,600	-0.1175	-0.1054	-0.0998	-0.0963	-0.0935	-0.0916
2,000	-0.1319	-0.1192	-0.1124	-0.1077	-0.1049	-0.1029
2,500	-0.1480	-0.1338	-0.1261	-0.1213	-0.1180	-0.1153
3,000	-0.1611	-0.1461	-0.1384	-0.1330	-0.1292	-0.1269
3,500	-0.1710	-0.1562	-0.1486	-0.1434	-0.1393	-0.1369
4,000	-0.1772	-0.1645	-0.1569	-0.1520	-0.1481	-0.1455
4,500	-0.1809	-0.1703	-0.1636	-0.1587	-0.1554	-0.1531
5,000	-0.1824	-0.1740	-0.1684	-0.1644	-0.1612	-0.1592
5,500	-0.1818	-0.1761	-0.1719	-0.1684	-0.1659	-0.1642

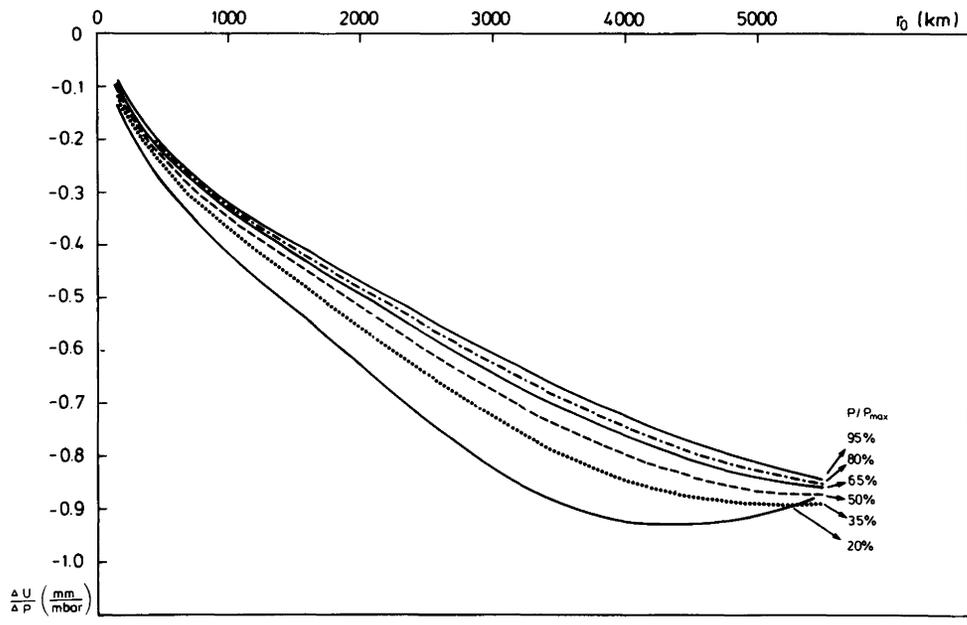


Fig. 14. Regression coefficient $C_1 = du/dp$ for vertical displacement [see Eq. (3)]. p : local pressure, p_{max} : maximum pressure at the center of a cyclone, r_0 : radius of a cyclone [see Eq. (1)]

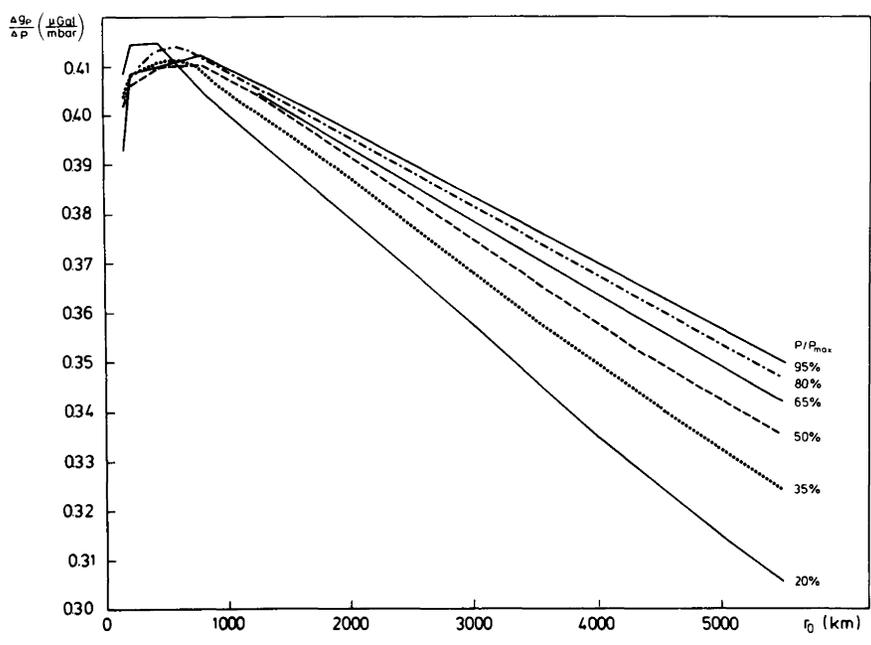


Fig. 15. Regression coefficient $C_1 = dg_p/dp$ for primary gravity [see Eq. (3)]. p : local pressure, p_{max} : maximum pressure at the centre of a cyclone, r_0 : radius of a cyclone [see Eq. (1)]

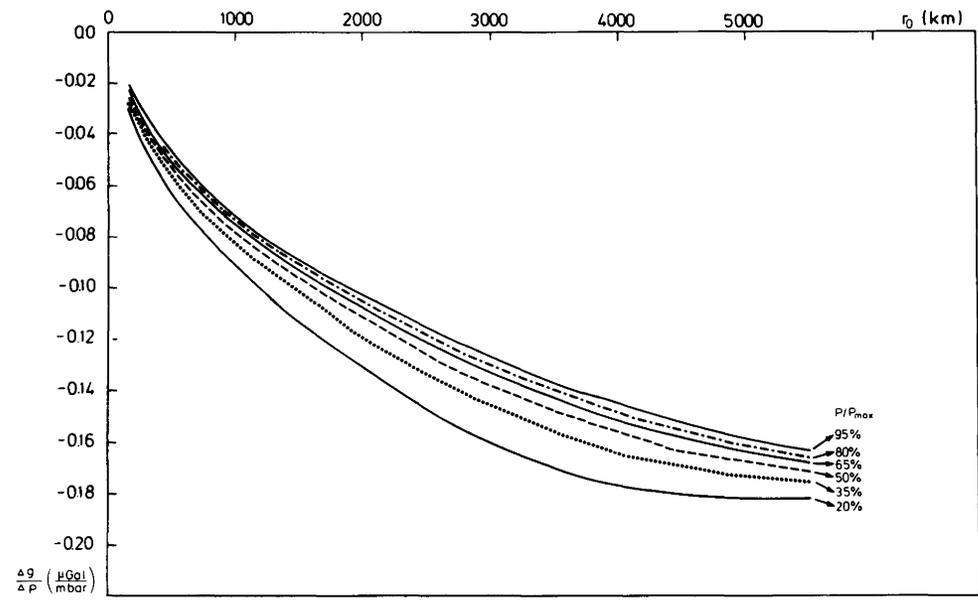


Fig. 16. Regression coefficient $C_1 = dg_s/dp$ for secondary gravity [see Eq. (3)]. p : local pressure, p_{max} : maximum pressure at the center of a cyclone, r_0 : radius of a cyclone [see Eq. (1)]

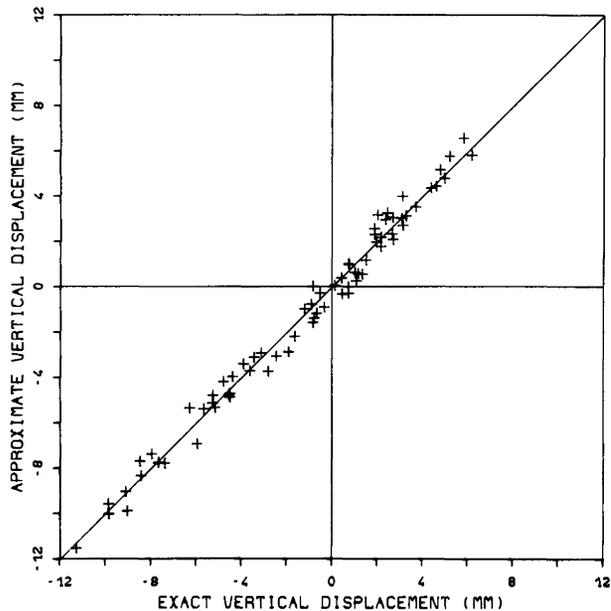


Fig. 17. Approximate vertical displacements according to Eq. (4) as a function of the corresponding values based on exact loading calculations. Results apply to seasonal changes in the global air pressure distribution shown in Fig. 6. Deviations from the exact calculations are less than 1 mm!

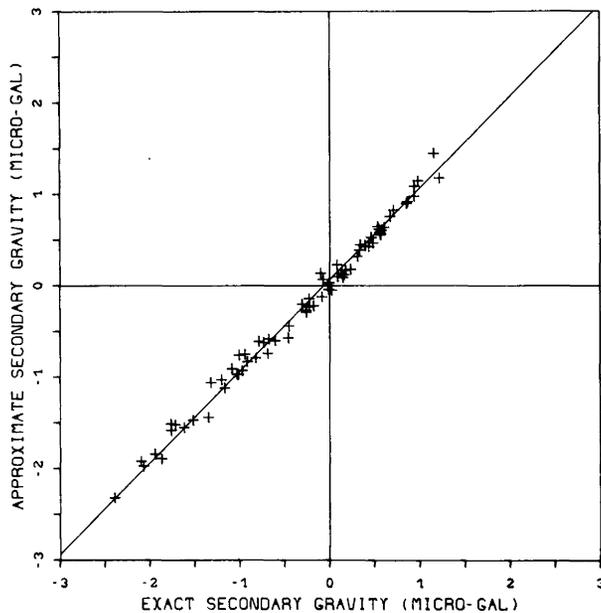


Fig. 19. Approximate secondary gravity effect according to Eq. (4) as a function of the corresponding values based on exact loading calculations. Results apply to seasonal changes in the global air pressure distribution shown in Fig. 6. Deviations from the exact calculations are less than 0.2 μgal !

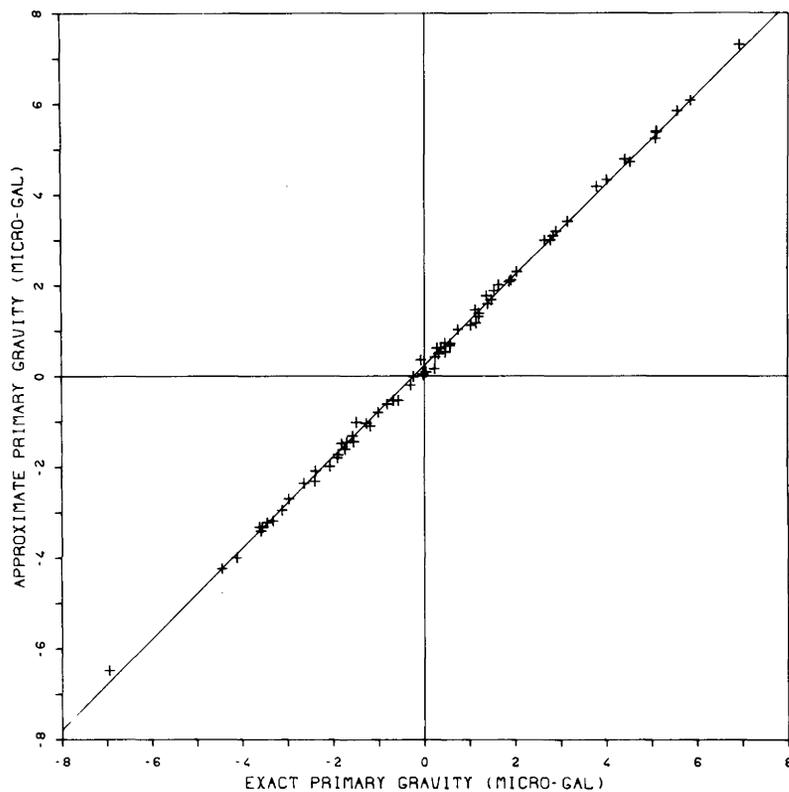


Fig. 18. Approximate primary gravity effect according to Eq. (4) as a function of the corresponding values based on exact loading calculations. Results apply to seasonal changes in the global air pressure distribution shown in Fig. 6. Deviations from the exact calculations are less than 0.3 μgal !

pressure-induced vertical displacements may therefore become desirable in the very near future. Air-pressure-induced horizontal displacements, however, seem to be negligible.

Although these results presented for vertical displacements seem to require an air pressure correction of regional and global base length measurements, such a correction

does not appear to be easy. Scherneck (1983) proposes to use gravimeter recordings in order to correct global baseline measurements for loading tide vertical displacements. He comes up with a regression coefficient of approximately 0.4 cm/ μgal for loading tides within a distance of 2,000 km. In the air pressure problem, such a regression coefficient

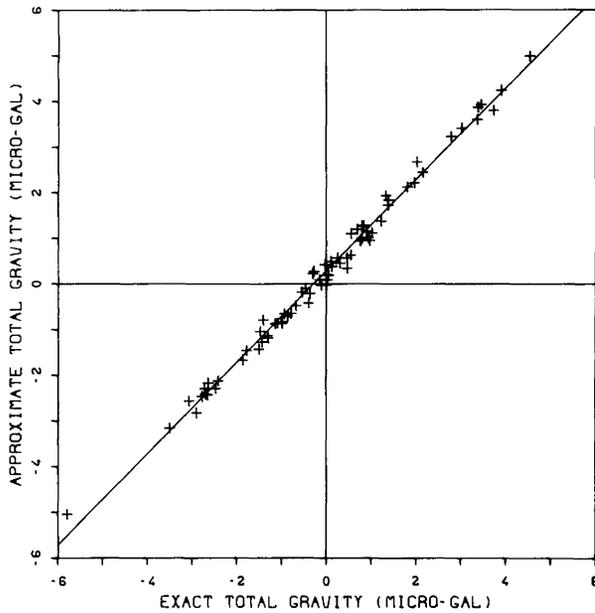


Fig. 20. Approximate total gravity effect according to Eq. (4) as a function of the corresponding values based on exact loading calculations. Results apply to seasonal changes in the global air pressure distribution shown in Fig. 6. Deviations from the exact calculations are less than $0.4 \mu\text{gal}$!

Table 4. The Earth's response to deviations of air pressure from mean atmospheric pressure. (Values are approximate maximum values. They may have to be modified near coastlines)

Period	Some days	One year
Pressure variation (mbar)	± 60	± 10
Vertical displacement (cm)	∓ 2.5	∓ 0.5
Horizontal displacement (mm)	± 2.5	± 0.5
Horizontal strain ($\cdot 10^9$)	∓ 10	∓ 1.5
Gravity effect (μgal) (positive upwards)	± 20	± 3
Tilt effect (msec)	± 10	± 1.5

can only be given reliably for the relation between air-pressure-induced displacements and the *secondary* gravity effect. The *primary* gravity effect which is the dominant one, does not correlate well enough with the displacements. A sufficient correction of base-length measurements using gravity recordings would therefore require the separation of the secondary gravity effect from the primary one. This, however, can probably not be done accurately enough at present.

The use of the wavelength dependent regression coefficients given in Tables 1–3 is more accurate. However, it involves the estimation of characteristic wavelengths and their amplitudes in a pressure distribution, which may be laborious in many cases. The use of Eq. (4) seems to be more adequate. It only involves two regression coefficients

and needs a weather-chart covering the surrounding area of 2,000 km around the surface point under investigation. The procedure of the air pressure correction is very simple and does not need any computer. Nevertheless, it is highly precise, allowing correction of seasonal radial displacements with an accuracy of at least 1 mm.

Among all continuous recording ground-based geodynamic instruments such as gravimeters, tiltmeters and strainmeters, probably the only one for which the long-term stability is good enough to reliably measure the small rates of tectonic movements is the superconducting gravimeter. Such movements may result in a few microgal gravity change within one year. Comparing this with the $\pm 20 \mu\text{gal}$ gravity perturbation within some days and $\pm 3 \mu\text{gal}$ within one year which may be introduced by air pressure variations, it is immediately clear that, besides air pressure corrections of radial displacements, such corrections are also necessary for gravity changes. Again the use of Eq. (4) seems to be most adequate for this purpose. It can provide air pressure corrections of seasonal gravity changes within $\pm 0.4 \mu\text{gal}$.

Acknowledgements. We are grateful to L. Bittner, P. Sadowiak and A. Wahlen for the preparation of the figures and to K. Helbig for typing the manuscript.

References

- Alterman, Z., Jarosch, H., Pekeris, C.L.: Propagation of Rayleigh waves in the Earth. *Geophys. J.R. Astron. Soc.* **4**, 219–241, 1961
- Baur, F.: Einführung in die Großwetterkunde. Wiesbaden: Dietrichsche Verlagsbuchhandlung 1948
- Committee on Geodesy, National Research Council: Geodetic monitoring of tectonic deformation – toward a strategy. Washington D.C.: National Academy Press 1981
- 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
- Faust H.: Der Aufbau der Erdatmosphäre. Braunschweig: Vieweg, 1968
- Longman, J.M.: A Green's function for determining the deformation of the Earth under surface mass loads. *J. Geophys. Res.* **67**, 845–850, 1962
- Müller, T., Zürn, W.: Observation of gravity changes during the passage of cold fronts. *J. Geophys.* **53**, 155–162, 1983
- Munk, W.H., Macdonald, G.J.F.: The rotation of the Earth. Cambridge University Press, 1960
- Okubo, S.: Theoretical and observed Q of the Chandler Wobble – Love number approach. *Geophys. J.R. Astron. Soc.* **71**, 647, 1982
- Prothero, W.A., Goodkind, J.M.: Earth-tide measurements with the superconducting gravimeter. *J. Geophys. Res.* **77**, 926–936, 1972
- Scherneck, H.G.: Crustal loading affecting VLBI sites. Department of Geodesy Report No. 20, University of Uppsala, 1983
- Smith, M.O., Dahlen, F.A.: The period and Q of the Chandler Wobble. *Geophys. J. Astron. Soc.* **64**, 223–281, 1981
- Spratt, R.J.: Modelling of the effect of atmospheric pressure variations on gravity. *Geophys. J.R. Astron. Soc.* **71**, 173–186, 1982
- Stolz, A., Larden, D.R.: Seasonal displacement and deformation of the Earth by the atmosphere. *J. Geophys. Res.* **84**, 6185–6194, 1979
- Thompson, K.R.: Regression models for monthly mean sea level. *Marine Geodesy* **2**, 269–290, 1979
- Trubysin, A.P., Makalkin, A.V.S.: Deformation of the Earth's Crust due to atmospherical cyclones. *Izv. Acad. Sci. USSR, Phys. Solid Earth* **12**, 343–344, 1976

- Walter, L.S.: Geodynamics. Proceedings of a workshop held at the Airlie House, Airly, Virginia, February 15–18, 1983; NASA Conference Publication 2325, 1984
- Warburton, J.R., Goodkind, J.M.: The influence of barometric pressure variations on gravity. *Geophys. J.R. Astron. Soc.* **48**, 281–292, 1977
- Zschau, J.: Auflastzeiten. Habilitation-thesis. University of Kiel, 1979a
- Zschau, J.: Phase shifts on tidal sea load deformations of the Earth's surface due to low viscosity layers in the interior. *Proceed. 8th Intern. Symp. Earth Tides.* 372–398, Bonn 1979c
- Zschau, J.: Tidal friction in the solid Earth: Loading tides versus body tides. *Proceed. 8th Intern. Symp. Earth Tides.* 62–94, Bonn 1979d
- Zschau, J.: The influence of the Earth's viscosity on deformations by marine tidal surface loads. In: *Earth Rheology, isostasy and eustasy.* N.-A. Mörner (ed): pp 161–167. John Wiley and Sons, 1980

Received January 31, 1984; revised version November 27, 1984

Accepted November 30, 1984

The normal modes of a uniform, compressible Maxwell half-space

Detlef Wolf

Department of Physics, University of Toronto, Toronto, Ontario, Canada, M5S 1A7

Abstract. The analytical solution for the load-induced deformation of a uniform, compressible, hydrostatically pre-stressed elastic half-space is derived. The solution is correct to first order in the quantity ε , which is inversely proportional to the wave number k of the deformation. Usually ε is very small compared with unity for Earth deformations on a scale amenable to the half-space approximation. Since pre-stress advection is included in the analysis, the correspondence principle allows us to solve the field equations governing the deformation of the associated Maxwell half-space. The viscoelastic solution shows that the relaxation of the Maxwell continuum is characterized by a fundamental mode and a rapidly decaying overtone of much smaller amplitude. In the incompressible limit the overtone is not excited. The significance of the results for the relaxation of the Earth's mantle is briefly discussed.

Key words: Compressibility – Maxwell continuum – Normal modes

Introduction

The response of the Earth's lithosphere or mantle to applied surface loads, such as volcanic islands, sedimentary basins or glacial loads, has frequently been modelled using Maxwell continua (e.g. Walcott, 1970; Beaumont, 1978; Lambeck and Nakiboglu, 1981; Nakiboglu and Lambeck, 1982). In these models the Earth's compressibility was, however, usually neglected. This simplification was probably motivated by the view that allowing for compressibility would be unlikely to change the results markedly.

In a series of papers starting in 1974, Peltier developed a general theory for the relaxation of self-gravitating, compressible Maxwell Earth models. The formalism was applied to infer the Earth's viscosity stratification from deglaciation-induced relative-sea-level variations (see Peltier, 1982, for a summary).

One of several interesting aspects of Peltier's investigations was the recognition of the complicated response pattern of "realistic" Maxwell Earth models whose elastic structure is taken from seismological Earth models. The relaxation of realistic models is characterized by a multitude of discrete and exponentially decaying

modes carrying distinct proportions of the total strain energy (Peltier, 1976; Wu and Peltier, 1982).

Each non-adiabatic density contrast, for example, is associated with a characteristic mode. This is also a feature of purely Newtonian viscous models (Parsons, 1972). Each viscosity contrast, such as the contrast near the base of the lithosphere, causes additional modes in the Maxwell model. This is a feature not paralleled in Newtonian viscous models, in which discontinuities in viscosity primarily modify the relaxation of the fundamental mantle mode associated with the density jump at the Earth's surface (McConnell, 1968). Some of the higher modes have relaxation times that are very short compared with the relaxation time of the fundamental mantle mode. Peltier (1976) therefore termed them transition modes. Usually these modes are only poorly excited, however.

Wu and Peltier (1982) also studied the effects due to compressibility on the relaxation of Maxwell continua and compared the response of compressible and incompressible Maxwell models in the Laplace-transform domain. They showed that, whereas the initial elastic response of the compressible model is characterized by significantly enhanced deformation, the final inviscid response is identical to that of the incompressible approximation.

In the following we will further examine the modifications introduced by compressibility. For this purpose we will study the relaxation of a particularly simple Maxwell Earth model. The analysis will be based on the formal solution for a uniform, compressible and pre-stressed elastic half-space. Application of the correspondence principle and normal-mode analysis will then allow us to show that the associated Maxwell half-space is characterized by the usual fundamental mode *and* an "overtone" of short relaxation time. The geophysical consequences of this will be briefly discussed by considering a characteristic numerical example.

Theory

Although the model of a uniform Maxwell half-space is elementary, it may serve as a first approximation when studying deformations of the Earth's mantle on a time-scale characteristic of deglaciation events. The solution of the equivalent elastic problem is published in several

textbooks (e.g. Jeffreys, 1976, pp. 265–267). If the elastic solution is to be used to solve the associated Maxwell problem, it must be modified, however, and gravitational restoring forces must be included.

This has recently been discussed for incompressible continua (Wolf, 1985a, b). In this approximation the governing equations can be re-formulated in terms of the total perturbation stress. If the continuum is compressible, this simple method fails and a more general approach is required. In the following, we will transform the field equations governing the deformation of a compressible, pre-stressed elastic half-space into a simultaneous first-order differential system. This system can be solved using standard matrix methods.

We confine ourselves to axisymmetric loading problems and use cylindrical co-ordinates r, ϕ, z . Then the stress-strain relations are

$$\sigma_{rz} = \mu \left(\frac{\partial w}{\partial r} + \frac{\partial u}{\partial z} \right), \quad (1)$$

$$\sigma_{zz} = (\lambda + 2\mu) \frac{\partial w}{\partial z} + \frac{\lambda}{r} \frac{\partial}{\partial r} (ru). \quad (2)$$

Taking the first-order Hankel transform of Eq. (1) and the zeroth-order Hankel transform of Eq. (2), with respect to the radial co-ordinate r , we obtain

$$\hat{u}'_1 - k \hat{w}_0 - \frac{1}{\mu} \hat{\sigma}'_{rz1} = 0, \quad (3)$$

$$\frac{\lambda k}{\lambda + 2\mu} \hat{u}_1 + \hat{w}'_0 - \frac{1}{\lambda + 2\mu} \hat{\sigma}'_{zz0} = 0. \quad (4)$$

Here u, w, σ_{rz} and σ_{zz} denote the radial and vertical displacement components and the appropriate elastic stress components, respectively. Parameters λ and μ are Lamé's first and second constants. A circumflex denotes Hankel transformation of zeroth or first order, as indicated by the subscript; symbol k denotes the Hankel transform variable or wave number. A prime is used to indicate differentiation with respect to the vertical co-ordinate z .

The two components of the equilibrium equation are

$$\frac{\partial \sigma_{rr}}{\partial r} + \frac{\partial \sigma_{rz}}{\partial z} + (\sigma_{rr} - \sigma_{\phi\phi})/r + \rho g \frac{\partial w}{\partial r} = 0, \quad (5)$$

$$\frac{\partial \sigma_{rz}}{\partial r} + \frac{\partial \sigma_{zz}}{\partial z} + \sigma_{rz}/r + \rho g \frac{\partial w}{\partial z} = 0. \quad (6)$$

The last term in Eqs. (5) and (6), respectively, accounts for stress advection in a hydrostatically pre-stressed elastic continuum of density ρ . The external gravity field g is assumed to act in the positive z -direction. Density changes due to the dilatation of the material have been neglected. Upon first-order Hankel transformation of Eq. (5) and zeroth-order Hankel transformation of Eq. (6) we obtain

$$-\frac{4(\lambda + \mu)\mu k^2}{\lambda + 2\mu} \hat{u}_1 - \rho g k \hat{w}_0 + \hat{\sigma}'_{rz1} - \frac{\lambda k}{\lambda + 2\mu} \hat{\sigma}'_{zz0} = 0, \quad (7)$$

$$\rho g \hat{w}'_0 + k \hat{\sigma}'_{rz1} + \hat{\sigma}'_{zz0} = 0, \quad (8)$$

where σ_{rr} and $\sigma_{\phi\phi}$ have been eliminated using the appropriate stress-strain relations.

Equations (3), (4), (7) and (8) may be written in matrix form. With $D = d/dz$ we obtain

$$\begin{bmatrix} D & -k & -\frac{1}{\mu} & 0 \\ \frac{\lambda k}{\lambda + 2\mu} & D & 0 & -\frac{1}{\lambda + 2\mu} \\ -\frac{4(\lambda + \mu)\mu k^2}{\lambda + 2\mu} & -\rho g k & D & -\frac{\lambda k}{\lambda + 2\mu} \\ 0 & \rho g D & k & D \end{bmatrix} \begin{bmatrix} \hat{u}_1 \\ \hat{w}_0 \\ \hat{\sigma}'_{rz1} \\ \hat{\sigma}'_{zz0} \end{bmatrix} = 0. \quad (9)$$

This is the generalization, for a pre-stressed continuum, of the first-order system derived by Farrell (1972). If we assume solutions of the type $\exp(mz)$, the attenuation constants m will be roots of the secular determinant

$$(\lambda + 2\mu)m^4 - 2(\lambda + 2\mu)k^2 m^2 + (\lambda + 2\mu)k^4 + \rho g m^3 - \rho g k^2 m = 0. \quad (10)$$

We find

$$m_{1,2} = \pm k, \quad (11)$$

$$m_{3,4} = -\frac{\lambda + \mu}{2(\lambda + 2\mu)} \varepsilon k \pm k \left[1 + \frac{(\lambda + \mu)^2 \varepsilon^2}{4(\lambda + 2\mu)^2} \right]^{1/2}, \quad (12)$$

where $\varepsilon = \rho g / [(\lambda + \mu)k]$ has been introduced. As discussed by Cathles (1975, pp. 38–39), ε will in general be very small compared with unity for deformations of the Earth whose scale is sufficiently small to be modelled by half-space approximations. Neglecting higher-order terms, we therefore have

$$m_{3,4} = \pm k \left[1 \mp \frac{\lambda + \mu}{2(\lambda + 2\mu)} \varepsilon \right], \quad (13)$$

which is correct to first order in ε .

If pre-stress is neglected, $g = 0$ and therefore $\varepsilon = 0$. If, on the other hand, the continuum is incompressible, $\lambda \rightarrow \infty$ and again $\varepsilon = 0$. In both cases the secular determinant, Eq. (10), has two double roots, and Eq. (9) represents a degenerate system. The combined effects of compressibility and pre-stress advection therefore remove the degeneracy of the system and cause a "gravitational splitting" of the attenuation constants, which has some similarity with the rotational splitting of the eigenfrequencies in the theory of the Earth's free oscillations. From Eq. (13) it is also evident that the attenuation of the elastic field quantities with depth is not solely determined by the lateral scale of the load but is also influenced by the material properties of the elastic continuum.

The eigenfunctions belonging to the four eigenvalues given by Eqs. (11) and (13) are calculated using matrix methods described by Frazer et al. (1938, pp. 61–70, 156–172). We obtain, correct to first order in ε ,

$$\begin{bmatrix} \hat{u}_1 \\ \hat{w}_0 \\ \hat{\sigma}'_{rz1} \\ \hat{\sigma}'_{zz0} \end{bmatrix} = \begin{bmatrix} \mp 1 - \varepsilon \\ 1 \\ -2\mu k \mp \mu k \varepsilon \\ \pm 2\mu k - \lambda k \varepsilon \end{bmatrix} A_{1,2} \exp(m_{1,2} z), \quad (14)$$

$$\begin{bmatrix} \hat{u}_1 \\ \hat{w}_0 \\ \hat{\sigma}_{rz1} \\ \hat{\sigma}_{zz0} \end{bmatrix} = \begin{bmatrix} \mp 1 - \varepsilon/2 \\ 1 \pm \frac{\mu}{2(\lambda + 2\mu)} \varepsilon \\ -2\mu k \mp \mu k \frac{\mu}{\lambda + 2\mu} \varepsilon \\ \pm 2\mu k - \lambda k \varepsilon \end{bmatrix} A_{3,4} \exp(m_{3,4} z). \quad (15)$$

The following analysis will be limited to a uniform half-space. To be consistent with the direction of the gravity field adopted in Eqs. (5) and (6), the continuum must occupy the region $z > 0$. If we impose the usual boundary conditions,

$$\hat{\sigma}_{rz1}(z=0) = 0, \quad (16a)$$

$$\hat{\sigma}_{zz0}(z=0) = -\hat{q}_0, \quad (16b)$$

$$\tilde{T}^{(ve)}(s) = \frac{3(\lambda + 2\mu)s^2 + 2(3\lambda + 4\mu)\tau^{-1}s + (3\lambda + 2\mu)\tau^{-2}}{3[2\mu k(\lambda + \mu) + \rho g \lambda]s^2 + 2[\mu k(3\lambda + 2\mu) + \rho g(3\lambda + \mu)]\tau^{-1}s + \rho g(3\lambda + 2\mu)\tau^{-2}}, \quad (24)$$

we obtain, from Eqs. (14) and (15),

$$A_2 = \frac{2(\lambda + 2\mu) - \mu \varepsilon}{(\lambda + \mu) \varepsilon} \frac{\hat{q}_0}{2\mu k + \lambda k \varepsilon}, \quad (17)$$

$$A_4 = \frac{-2(\lambda + \mu) + (\lambda + 2\mu) \varepsilon}{(\lambda + \mu) \varepsilon} \frac{\hat{q}_0}{2\mu k + \lambda k \varepsilon}. \quad (18)$$

The solution for the vertical surface displacement, the quantity of prime geophysical interest, then becomes

$$\hat{w}_0(z=0) = \hat{q}_0 \left[1 + (2 - \varepsilon) \frac{\mu}{2(\lambda + \mu)} \right] / (2\mu k + \lambda k \varepsilon). \quad (19)$$

If pre-stress is neglected, $\varepsilon = 0$ and

$$\hat{w}_0(z=0) = (\lambda + 2\mu) \hat{q}_0 / [2\mu k(\lambda + \mu)], \quad (20)$$

which is the familiar solution for the non-gravitating half-space. If compressibility is neglected, $\lambda \rightarrow \infty$ and

$$\hat{w}_0(z=0) = \hat{q}_0 / (2\mu k + \rho g). \quad (21)$$

This is identical to the solution discussed in Wolf (1985b), which was, however, derived directly from the incompressible field equations. Equation (21) is correct without restrictions on the wave number k . Since we will be applying our elastic solution to solve the corresponding Maxwell problem, it is illuminating to reduce Eq. (19) to its inviscid limit, $\mu = 0$, which is also the infinite-time limit for Maxwell continua (Wu and Peltier, 1982). Then

$$\hat{w}_0(z=0) = \hat{q}_0 / (\rho g), \quad (22)$$

which expresses local compensation of the load by buoyancy forces.

If $\varepsilon \ll 1$, Eq. (19) is simplified to

$$\hat{w}_0(z=0) = (\lambda + 2\mu) \hat{q}_0 / [2\mu k(\lambda + \mu) + \rho g \lambda]. \quad (23a)$$

This may also be written in terms of Poisson's ratio ν .

Since $2\nu = \lambda/(\lambda + \mu)$, we obtain

$$\hat{w}_0(z=0) = \hat{q}_0(1 - \nu) / (\mu k + \nu \rho g). \quad (23b)$$

This equation is slightly different from that proposed by Nakiboglu and Lambeck [1982, Eq. (24)] as the solution to the same problem which is considered here. Their solution was, however, derived from physically unreasonable boundary conditions (see Wolf, 1985) and assumed to be universally valid for any value of k .

According to the correspondence principle (Appendix A), Eq. (23a) can be interpreted as the Laplace transform $\tilde{w}(s)$ of the impulse response of the associated Maxwell continuum. If λ and μ are replaced by Eqs. (30) and (31) and if the Laplace transform $\tilde{T}^{(ve)}(s) = \tilde{w}(s)/\hat{q}$ of the viscoelastic transfer function is introduced, Eq. (23a) is thus transformed to

where the subscript has been dropped. Taking the inverse Laplace transform (Appendix B) yields

$$T^{(ve)}(t) = T^{(e)} \delta(t) + T^{(v,1)} s^{(1)} \exp(-s^{(1)} t) + T^{(v,2)} s^{(2)} \exp(-s^{(2)} t), \quad (25)$$

for the impulsive forcing $\hat{q}(k) \delta(t)$ or

$$T^{(ve)}(t) = T^{(e)} - T^{(v,1)} [\exp(-s^{(1)} t) - 1] - T^{(v,2)} [\exp(-s^{(2)} t) - 1], \quad (26)$$

for the Heaviside loading event $\hat{q}(k) H(t)$.

The explicit formulae for the viscous transfer functions (normal modes) $T^{(v,1)}$, $T^{(v,2)}$ and the associated inverse relaxation times $s^{(1)}$, $s^{(2)}$ are not very illuminating. In the incompressible limit, $\lambda \rightarrow \infty$, however, simple analytical expressions result for the inverse relaxation times, viz.

$$s^{(1)} = \tau^{-1}, \quad (27)$$

$$s^{(2)} = \rho g \tau^{-1} / (2\mu k + \rho g). \quad (28)$$

Numerical example

The portion of the relaxation carried by the fundamental mode in the uniform model was analysed previously to some extent (Nakiboglu and Lambeck, 1982; Wu and Peltier, 1982; Wolf, 1984). The existence of an overtone with a relaxation time close to the Maxwell time [and close to the relaxation times of the transition modes of realistic Earth models identified by Peltier (1976)] in this simple model was, however, appreciated previously and is therefore also discussed here.

The effects due to compressibility are displayed in Fig. 1, where $T^{(e)}$, $T^{(v,1)}$, $T^{(v,2)}$, $s^{(1)}$ and $s^{(2)}$ are plotted as functions of wave number. For definiteness we have taken $\lambda = 0.80 \times 10^{11} \text{ N m}^{-2}$, $\mu = 0.67 \times 10^{11} \text{ N m}^{-2}$ and $\rho = 3,380 \text{ kg m}^{-3}$. These values are fairly characteristic of the Earth at 100-km depth (Bullen, 1963, pp. 232-235). The viscosity is $\eta = 10^{21} \text{ Pa s}$, which appears to be typical of the upper mantle (e.g. Cathles, 1975). A cut-off angular order of $n = ka = 5$ has been chosen for the

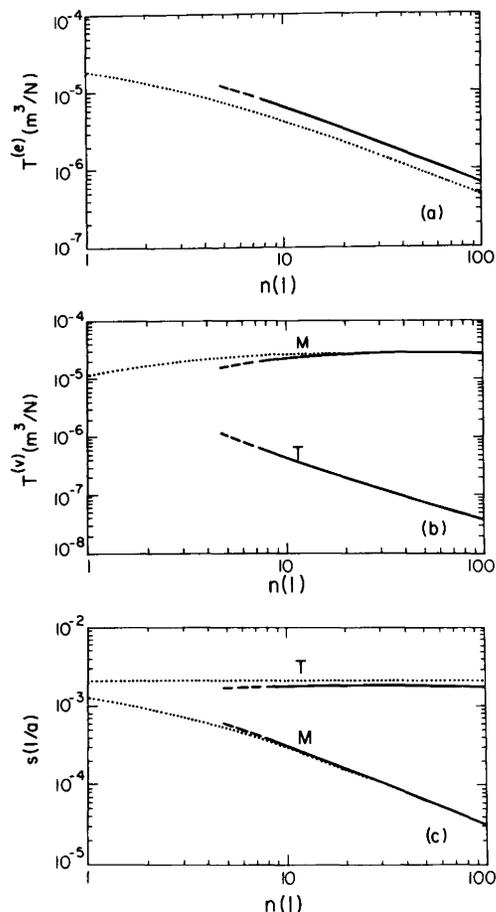


Fig. 1a–c. **a** Elastic transfer function $T^{(e)}$, **b** viscous transfer function $T^{(v)}$ and **c** inverse relaxation time s as function of angular order n according to compressible solution (solid) and incompressible approximation (dotted); symbols M and T denote relaxation mode

compressible model, where a is the Earth's radius. For this value of n we find $\varepsilon = 0.29$ and thus $\varepsilon^2 = 0.08$.

The elastic transfer function $T^{(e)}$ is illustrated in Fig. 1a. Inspection of the diagram shows that $\nu = 0.272$ (which corresponds to $\lambda = 0.80 \times 10^{11} \text{ Nm}^{-2}$) leads to an increase in $T^{(e)}$ by approximately 40% compared with $\nu = 0.5$ (which corresponds to $\lambda \rightarrow \infty$). This is what is required by Eq. (23b).

The relaxation of a particular deformation is governed by the spectral characteristics of the viscous response. Figure 1c shows that the neglect of compressibility causes an insignificant change in the relaxation times of either mode. This has already been noted for the fundamental mode (Wu and Peltier, 1982).

In Fig. 1b the amplitude spectra $T^{(v,1)}$ and $T^{(v,2)}$ are displayed. The fundamental mode of the compressible model is characterized by reduced amplitude compared with the incompressible approximation. Since the sum $T^{(e)} + T^{(v,1)} + T^{(v,2)}$ is independent of the value adopted for λ [see Eq. (22)], this is consistent with the fact that the compressible model's elastic response is enhanced (Fig. 1a).

The overtone is only excited if $\lambda < \infty$. It is therefore intimately related to the compressibility of the material. For $\lambda = 0.80 \times 10^{11} \text{ Nm}^{-2}$, it carries a small fraction of

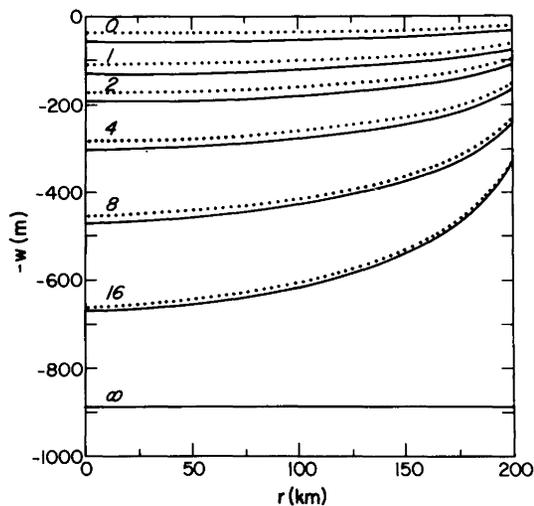


Fig. 2. Vertical surface displacement w as function of distance r from load axis for several times (in units of ka) after emplacement of load according to compressible solution (solid) and incompressible approximation (dotted)

the total viscous response. The relative strength of the overtone increases with decreasing angular order. This trend is expected to continue for $n < 5$, which is, however, below the range of validity of our approximation.

In order to assess the geophysical significance of these differences, we study the response in the spatial domain for a square-edged disk load of 200-km radius, 3-km thickness and a density of $1,000 \text{ kg m}^{-3}$, and for the Heaviside loading event $\hat{q}(k)H(t)$. In taking the inverse Hankel transform (Appendix C), the limits of integration have been fixed to $n = 5$ and $n = 1,000$. The truncation at the lower end primarily affects the accuracy of the response near the elastic limit. The behaviour at large times is more influenced by the cut-off at $n = 1,000$. The total truncation error is, however, small and amounts to a few per cent at the most.

In Fig. 2 the vertical surface deflection is shown. Since the distribution of the model parameters is uniform, a peripheral bulge does not develop as relaxation proceeds (Nakiboglu and Lambeck, 1982; Wolf, 1984). The discussion has therefore been confined to the central region below the load. From inspection of the figure it is evident that effects due to compressibility are noteworthy only during the initial period of relaxation following the emplacement of the load. At $t = 4$ ka the differences in deflection have already decreased to approximately 10%; at $t = 16$ ka the discrepancy is only a few per cent.

Conclusion

The geophysical significance of compressibility is clearly contingent upon whether the earlier or later phases of relaxation are sampled by the observations. In interpretations of glacio-isostatic rebound, for example, the lithosphere is usually regarded as elastic and the adjustments are mainly controlled by the viscosity of the Earth's mantle. Since relaxation times are characteristically between 1 and 10 ka (Wu and Peltier, 1982, Fig. 10) and since deglaciation was almost complete at 8 ka before present, the majority of the post-glacial-

adjustment data sample the intermediate or later phases of relaxation. During the early phases when effects due to compressibility are more noticeable, relaxation is also markedly influenced by the unloading event itself. The shape of the ice-sheet and the details of the deglaciation history are, however, not known with great accuracy. Consequently, the interpretation of isostatic-adjustment data from this time interval is subject to considerable uncertainties (e.g. Wolf, 1985c). We may therefore conclude that incompressible Maxwell continua are adequate representations of the Earth's mantle in most instances.

Appendix A

Correspondence principle

According to the correspondence principle (Biot, 1954; Peltier, 1974; Cathles, 1975, pp. 23–29), the solution to an elastostatic problem can be interpreted as the Laplace transform of the quasi-static response, to an impulsive load $q\delta(t)$, of the associated Maxwell continuum governed by Laplace-transformed stress-strain relations

$$\tilde{\sigma}_{ij} = \lambda(s) \tilde{\sigma}_{kk} \delta_{ij} + 2\mu(s) \tilde{\epsilon}_{ij}, \quad (29)$$

provided that

$$\mu(s) = \lambda s / (s + \tau^{-1}), \quad (30)$$

$$\lambda(s) = (\lambda s + K \tau^{-1}) / (s + \tau^{-1}). \quad (31)$$

Here the tilde denotes Laplace transformation with respect to time t ; s is the Laplace-transform variable. $K = \lambda + 2\mu/3$ denotes the bulk modulus and $\tau = \eta/\mu$ the Maxwell time, with η the dynamic viscosity. Since large values of s correspond to short-time-scale behaviour and vice versa, we realize, from inspection of Eqs. (29)–(31), that the instantaneous response, $t = 0$, is elastic. Constitutive relations appropriate to long times after the loading event are obtained by observing that $K(s) = \lambda(s) + 2\mu(s)/3 = K$, which is readily verified from Eqs. (30) and (31). Then Eq. (29) may be written in the alternative form

$$\tilde{\sigma}_{ij} = K \tilde{\epsilon}_{kk} \delta_{ij} - 2/3 \mu(s) \tilde{\epsilon}_{kk} \delta_{ij} + 2\mu(s) \tilde{\epsilon}_{ij}. \quad (32)$$

In the limit of $t \rightarrow \infty$ Eq. (30) vanishes, and Eq. (32) takes the form

$$\tilde{\sigma}_{ij} = K \tilde{\epsilon}_{kk} \delta_{ij} \quad (33a)$$

or, in the time domain,

$$\sigma_{ij} = K \epsilon_{kk} \delta_{ij}. \quad (33b)$$

These are the inviscid constitutive relations for a compressible continuum.

Appendix B

Normal modes

The Laplace transform $T^{(ve)}(s)$, of the viscoelastic transfer function describing the relaxation of the Maxwell continuum following an impulsive forcing $\hat{q}\delta(t)$, can be split into an elastic portion $T^{(e)}$ and a viscous portion $\tilde{V}(s)$. We may therefore write

$$\tilde{T}^{(ve)}(s) = T^{(e)} + \tilde{V}(s), \quad (34)$$

where

$$T^{(e)} = \lim_{s \rightarrow \infty} \tilde{T}^{(ve)}(s). \quad (35)$$

Here $T^{(e)}$ represents the Laplace transform of the instantaneous elastic response of the viscoelastic continuum to the

impulsive load. The transform $\tilde{V}(s)$ of the viscous portion can be cast into the form

$$\tilde{V}(s) = \frac{T^{(v,1)} s^{(1)}}{s + s^{(1)}} + \frac{T^{(v,2)} s^{(2)}}{s + s^{(2)}}, \quad (36)$$

where $T^{(v,1)}$, $T^{(v,2)}$ and $s^{(1)}$, $s^{(2)}$ are complicated functions of the model parameters. Taking the inverse Laplace transform yields

$$T^{(ve)}(t) = T^{(e)} \delta(t) + T^{(v,1)} s^{(1)} \exp(-s^{(1)} t) + T^{(v,2)} s^{(2)} \exp(-s^{(2)} t). \quad (37)$$

This is the system's impulse response. The response to a Heaviside loading event $\hat{q}H(t)$ follows from convolving it with the impulse response. From Eq. (37) we obtain, for $t \geq 0$,

$$T^{(ve)}(t) = T^{(e)} - T^{(v,1)} [\exp(-s^{(1)} t) - 1] - T^{(v,2)} [\exp(-s^{(2)} t) - 1]. \quad (38)$$

Appendix C

Inverse Hankel transform

We are concerned with square-edged, circular disk loads

$$q(r) = \begin{cases} 0, & 0 \leq r < R \\ 1, & R < r < \infty \end{cases}, \quad (39)$$

where R is the radius of the disk. The load distribution $q(r)$ can be written as the inverse zeroth-order Hankel transform

$$q(r) = \int_0^\infty \hat{q}(k) k J_0(kr) dk, \quad (40)$$

where (e.g. Sneddon, 1951, p. 528)

$$\hat{q}(k) = J_1(kR) R/k. \quad (41)$$

The response in the spatial domain is then obtained from

$$w(r, 0) = \int_0^\infty T^{(ve)}(k) \hat{q}(k) k J_0(kr) dk, \quad (42)$$

for the half-space, where $T^{(ve)}(k)$ denotes the viscoelastic transfer function.

Acknowledgements. This research was financially supported by a Natural Sciences and Engineering Research Council of Canada Postgraduate Scholarship. Discussions with Richard Peltier have been very helpful.

References

- Beaumont, C.: The evolution of sedimentary basins on a viscoelastic lithosphere: theory and examples. *Geophys. J. R. Astron. Soc.* **55**, 471–497, 1978
- Biot, M.A.: Theory of stress-strain relations in anisotropic viscoelasticity and relaxation phenomena. *J. Appl. Phys.* **25**, 1385–1391, 1954
- Bullen, K.E.: An introduction to the theory of seismology, 3rd edn. Cambridge: Cambridge University Press 1963
- Cathles, L.M.: The viscosity of the Earth's mantle. Princeton: Princeton University Press 1975
- Farrell, W.E.: Deformation of the Earth by surface loads. *Rev. Geophys. Space Phys.* **10**, 761–797, 1972
- Frazer, R.A., Duncan, W.J., Collar, A.R.: Elementary matrices. Cambridge: Cambridge University Press 1938
- Jeffreys, H.: The Earth, 6th edn. New York: Cambridge University Press 1976
- Lambeck, K., Nakiboglu, S.M.: Seamount loading and stress

- in the ocean lithosphere 2. Viscoelastic and elastic-viscoelastic models. *J. Geophys. Res.* **86**, 6961–6984, 1981
- 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 response to surface loading with application to Lake Bonneville. *Geophys. J. R. Astron. Soc.* **70**, 577–620, 1982
- 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
- Sneddon, I.A.: *Fourier transforms*. New York: McGraw-Hill 1951
- Walcott, R.I.: Flexural rigidity, thickness, and viscosity of the lithosphere. *J. Geophys. Res.* **75**, 3941–3954, 1970
- 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.: An improved estimate of lithospheric thickness based on a re-interpretation of tilt data from Pleistocene Lake Algonquin. *Can. J. Earth Sci.* 1985c (in press)
- Wu, P., Peltier, W.R.: Viscous gravitational relaxation. *Geophys. J. R. Astron. Soc.* **70**, 435–485, 1982

Received June 5, 1984; Revised version November 5, 1984
Accepted November 6, 1984

Magnetic expression of the continent-ocean boundary between the western margin of Australia and the eastern Indian Ocean

J.J. Veevers, J.W. Tayton, B.D. Johnson, and L. Hansen

School of Earth Sciences, Macquarie University, North Ryde, N.S.W. 2113, Australia

Abstract. A comprehensive review of the Early Cretaceous seafloor-spreading magnetic anomalies (M0 to M10) in the eastern Indian Ocean leads to the isolation of a distinctive magnetic anomaly at the continent-ocean boundary (COB). This anomaly is traceable 2000 km southward from the rifted margin of the magnetically smooth central Exmouth Plateau, through the transform-faulted and rifted margins of the Cuvier Abyssal Plain and Carnarvon Terrace and the set of narrow spreading segments south of the Zenith-Wallaby Fracture Zone to the area west of Perth. The anomaly corresponds to the COB as indicated by: 1. the lower part of the continental slope in a mean water depth of 3.75 km for the rifted margin and 4.5 km for the transform-faulted margin and 2. a change in seismic-reflection character from the faulted breakup unconformity on the continent to the smooth but hyperbolic oceanic layer 2. The COB anomaly at the rifted margin is modelled by modifying the magnetization of the oldest oceanic block of the seafloor-spreading sequence adjacent to the continental crust; in places, the COB anomaly is flanked by smaller anomalies modelled as rift-related dykes in the adjacent continental crust. The amplitude of the COB anomaly, commonly twice or more that of the adjacent oceanic magnetic anomalies, is due either to a thicker or a more intensely magnetized source. The COB anomaly at the transform-faulted margin is modelled by a thick vertical body that extends 10 km seaward of the COB.

The Wallaby Plateau is probably underlain by oceanic crust, as shown by the continuity of the abandoned spreading ridge of the Sonne Ridge southwestward from the Cuvier Abyssal Plain; the shape of the Wallaby Plateau and the volcanic composition of dredge-hauls indicate that it is probably a volcanic upgrowth of the oceanic crust, as exemplified by Iceland today. Furthermore, like Iceland, the Wallaby Plateau is crossed by magnetic anomalies that are possibly degraded seafloor-spreading anomalies. The Zenith Plateau lacks magnetic lineations and its crustal structure, like that of the Naturaliste Plateau to the south, remains unknown.

Key words: Seafloor-spreading magnetic anomalies – Continent-ocean boundary magnetic anomaly – Spreading pattern of the eastern Indian Ocean – Volcanic upgrowths on oceanic crust

Introduction

The 2000-km-long western margin of Australia, from the northwestern Exmouth Plateau to the southern Naturaliste Plateau (Figs. 1 and 2), is marked by a series of marginal plateaus indented by abyssal plains. The abyssal plains are crossed by Early Cretaceous seafloor-spreading magnetic anomalies, as described by Powell (1978) and Powell and Luyendyk (1982) for the Gascoyne Abyssal Plain, by Larson (1977), Larson et al. (1979) and Johnson et al. (1980) for the Cuvier Abyssal Plain and by Markl (1974a, b; 1978a, b) and Larson et al. (1978) for the Perth Abyssal Plain. As revised here, the anomalies range from M10 (122[125] Ma) through M0 (108 [110.7] Ma) into the Cretaceous interval of normal polarity and, calibrated by DSDP drilling, define the pattern of seafloor spreading. To provide direct comparison with previous work, ages are expressed here in terms of the old radiometric decay constants and,

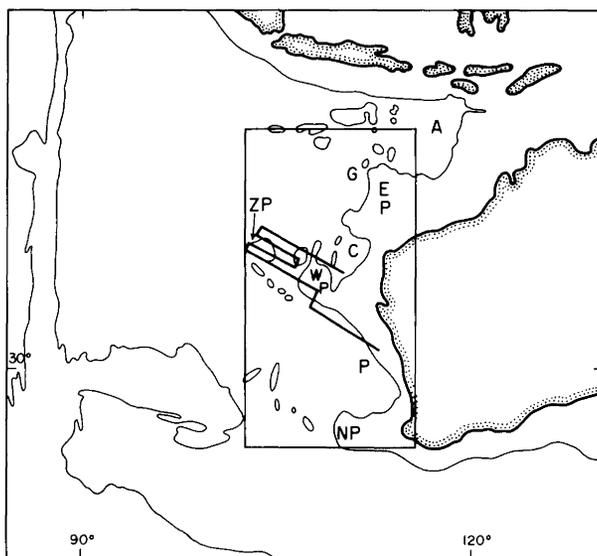


Fig. 1. Western margin of Australia, located within box that shows outline of Fig. 2. Also shown are 4-km isobaths and the track of the 1983 *Cook* cruise off the western margin. From north to south, A = Argo Abyssal Plain, G = Gascoyne Abyssal Plain, EP = Exmouth Plateau, ZP = Zenith Plateau, C = Cuvier Abyssal Plain, WP = Wallaby Plateau, P = Perth Abyssal Plain, NP = Naturaliste Plateau

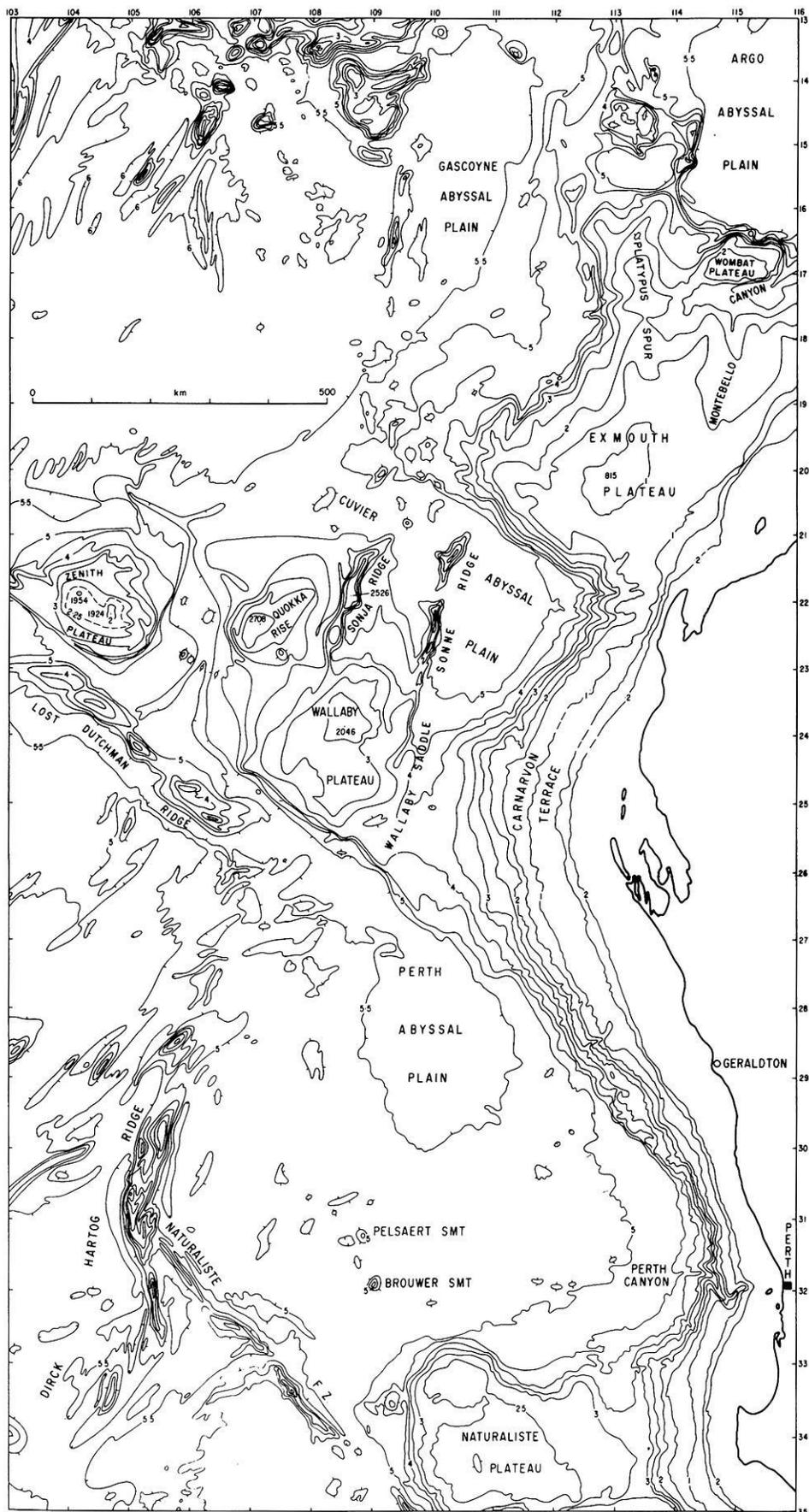


Fig. 2. Bathymetry of the western margin (0.5-km contour interval and, additionally, the 0.2-km isobath) from Fisher (1982), amended by new data: for the Zenith Plateau and the area southwest of the Wallaby Plateau from the 1983 cruise of *Cook* plotted on GEBCO sounding sheets; for the Exmouth and Wallaby Plateaus from von Stackelberg et al. (1980); for the Joey Rise from Heirtzler et al. (1978); and for the 2-km isobath of the Naturaliste Plateau from Coleman et al. (1982). *Bar* refers to scale at 19° S. Nomenclature: Heezen and Tharp (1966) named the complex of plateaus northwest of the Carnarvon Terrace the Wallaby Plateaus; Tomoda et al. (1968) called the western one the Zenith Seamount; various authors then confined the word Wallaby Plateau to the eastern one; Fisher (1982) called the western one the Wallaby Plateau and the eastern one the Cuvier Plateau. We resolve the confusion by calling the western one the Zenith Plateau (it is too big for a seamount), the eastern one the Wallaby Plateau and recognise a third as the Quokka Rise, named after a small wallaby common on islands offshore from Perth

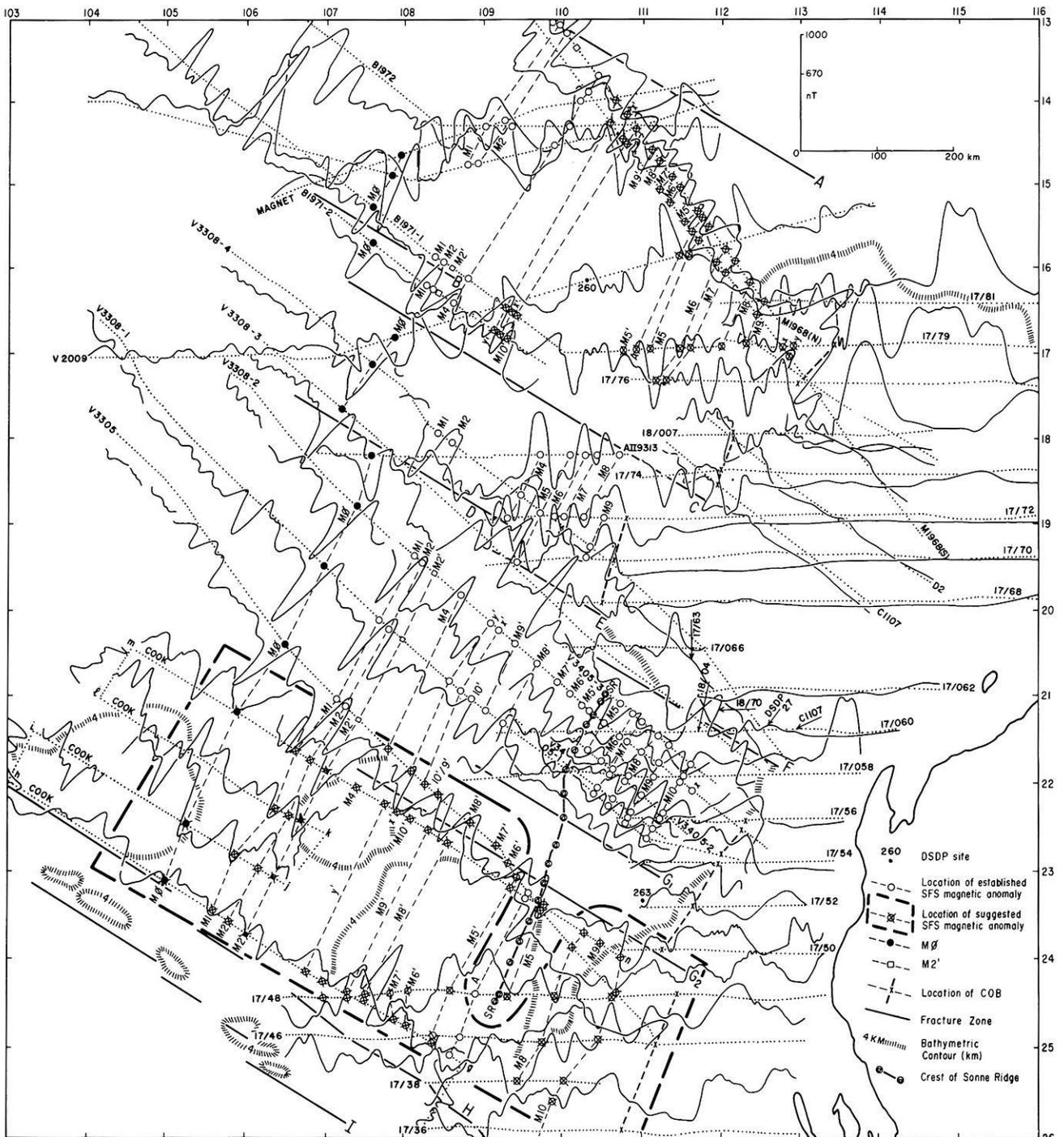


Fig. 3. Selected ships' tracks and magnetic profiles for the northern area. For the sake of clarity, ships' tracks across the Cape Range Fracture Zone (F) are indicated by *arrows*. New profiles for the Zenith Plateau area are from *Cook*, reduced by the Australian Geomagnetic Reference Field (AGRF) of Petkovic and Whitworth (1975); rest of the profiles were compiled from Powell (1978), Larson (1977), Larson et al. (1979), Falvey (1972a, b) and BMR Margins Survey, published by the Australian Government Publishing Service, Canberra. Magnetic anomaly determinations in the Cuvier Abyssal Plain from Larson (1977), Larson et al. (1979) and Johnson et al. (1980), were tentatively extended southwestward into the Wallaby-Quokka area (tentatively interpreted anomalies enclosed by *heavy broken line*) by us; while those in the Gascoyne Abyssal Plain from Powell (1978), cited in Larson et al. (1978) and also in Powell and Luyendyk (1982), were amended by us to include a reflected set M9'-A-M9 and the COB. Fracture zones determined by bathymetry are F (Cape Range FZ), H (Zenith-Wallaby FZ) and I (fitted to Lost Dutchmann Ridge); the rest are postulated from offsets in the magnetic anomaly pattern

in square brackets, converted to the new constants (Dalrymple, 1979).

M10 lies within the Jurassic-Cretaceous M-series of mixed polarity so that, unlike the central North Atlantic

which is bounded by the Jurassic quiet zone, a continuous set of anomalies extends up to the continental margin. The central and southern parts of the Exmouth Plateau are underlain by continental crust, as shown most clearly by the

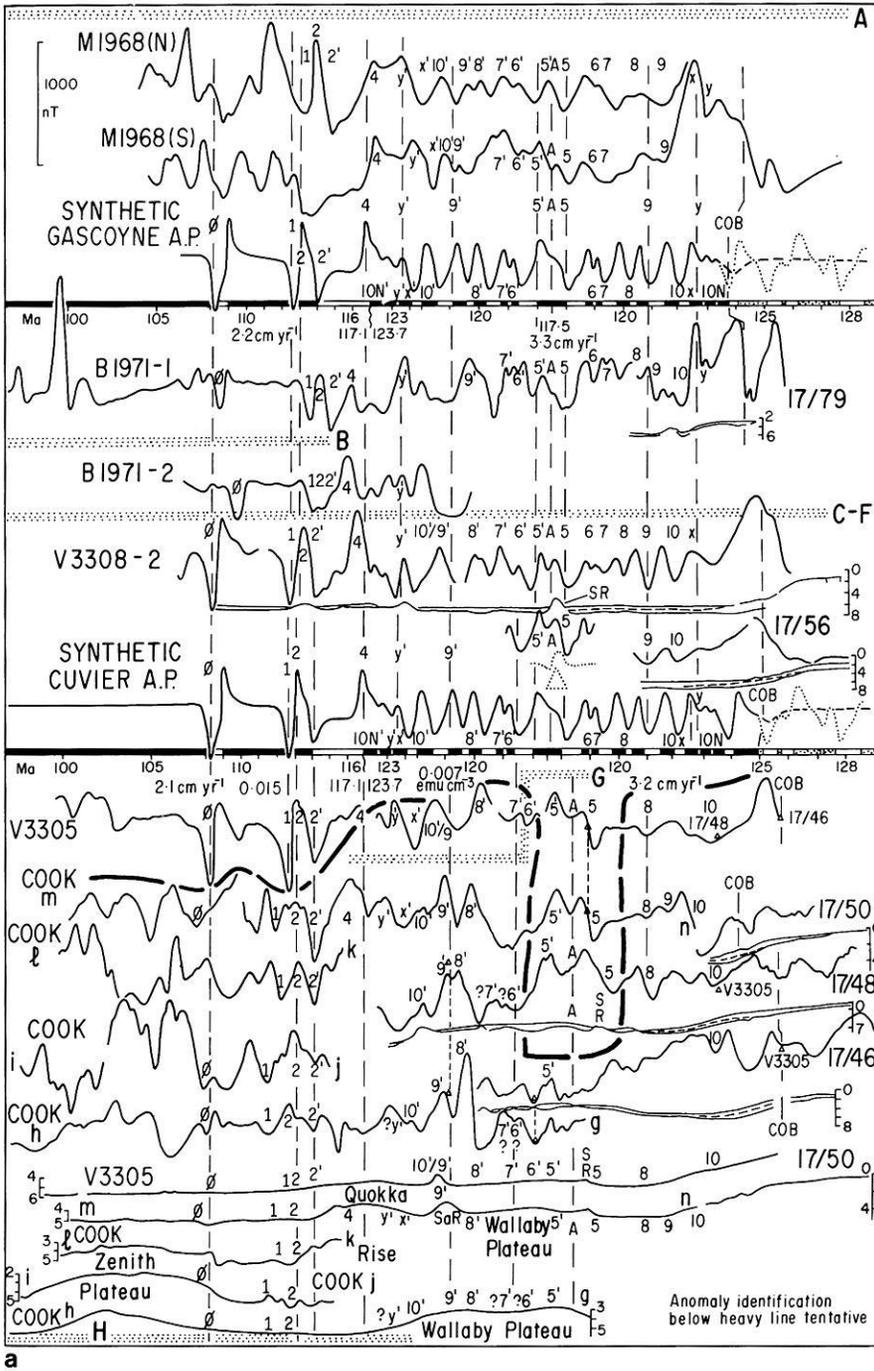
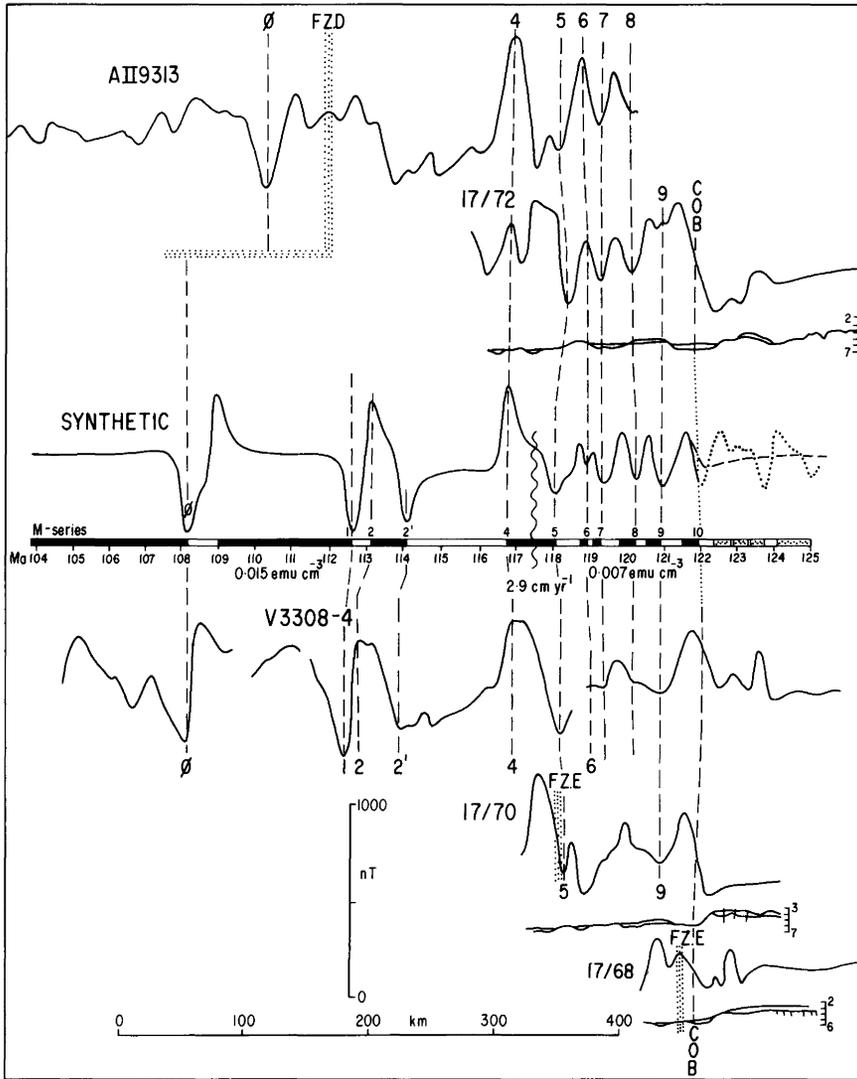


Fig. 4. a Selected magnetic, seismic and bathymetric profiles, located in Fig. 3, projected on an azimuth of 302°. Fracture zones A–H shown by *stippled lines*. Block model of M-series (Larson and Hilde, 1975) with normal polarity in *black*, reversed polarity *clear*—note that ages are expressed in the old K/Ar decay constants for direct comparison with previous work — and synthetic magnetic anomalies, involving reflection of blocks about axis (A), after Larson (1977), Larson et al. (1979) and Johnson et al. (1980) for the Cuvier Abyssal Plain, extended by us to the Gascoyne Abyssal Plain and tentatively to the Wallaby/Quokka region, distinguished from area of established anomalies by a *heavy broken line*. Model parameters not shown on the figure are: model source is 5.5–6.0 km below sea level and trend is 030°; present-day magnetic field parameters are inclination 55° up and declination 2° west. Remanent-magnetization parameters are inclination 55° up and declination 40° west. Remanent-magnetization intensity is expressed in old units of emu/cm^3 ($=10^{-3}$ A/m) to provide direct comparison with previous work. The positive anomaly at 122.4 Ma is labelled *x*, and the negative anomaly at 122.7 Ma, *y*. Lines through modelled anomalies afford direct comparison with observed anomalies; note offset of M7', A and M8 on V3305 and southward. SR = Sonne Ridge, whose magnetic anomaly is modelled by adding the magnetic response of a ridge with intensity $0.00075 \text{ emu}/\text{cm}^3$ to the block model (*dotted line*). Sa R – Sonja Ridge. The COB is located at 123.7 Ma in the Gascoyne Abyssal Plain and at 124.9 Ma in the Cuvier Abyssal Plain. The block model is extended (*dotted blocks*) past 125 Ma, and its magnetic profile shown by a *dotted line*. Track crossings are shown by *triangles*. Depths of seismic profiles in seconds of reflection time. Bathymetric profiles (km) of tracks across Zenith Plateau, Quokka Rise and Wallaby Plateau are concentrated at the bottom. Scales: 150 km are equivalent to 1000 nT. **b** Block model and synthetic magnetic anomalies in segment CE, extended from determinations by Larson (1977). Depth of seismic profiles in seconds of reflection time. Model parameters not indicated in the figure are as in Fig. 4a



b
Fig. 4b

occurrence of Triassic (200–230 Ma) sediment (Barber, 1982) in drill-holes. These parts of the Exmouth Plateau are magnetically featureless or smooth so that the boundary between the central rifted and southern sheared Exmouth Plateau and the adjacent ocean floor is marked by a continent-ocean boundary (COB) magnetic anomaly that can be modelled as the edge of an oceanic magnetic body, associated in places with rift-related dykes in the adjacent continental crust. South of the Exmouth Plateau, at the edge of the Cuvier and Perth Abyssal Plains, a similar anomaly is found but it is less distinctive because the continental crust is not magnetically smooth. Hence, all the seafloor-spreading magnetic anomalies must be identified so that the COB anomaly can be isolated.

In contrast, Veevers et al. (1985) show that in the region to the northeast, along the COB between the northwestern margin of Australia and the Jurassic Argo Abyssal Plain (Heirtzler et al., 1978), a prominent magnetic anomaly is due to a 40–80-km-wide body, most of which lies beneath the continental margin.

During the first cruise of HMAS *Cook* round Australia, measurements of the total magnetic field and bathymetry, supplemented in places by seismic reflection profiles, were

made along the western margin and across the poorly known Zenith and Wallaby Plateaus and provide critical data for a review of the magnetic pattern.

Bathymetry (Fig. 2)

The main change to the recently published GEBCO sheet (Fisher, 1982) and earlier charts (Falvey and Veevers, 1974; Veevers and Cotterill, 1978) is in and about the Zenith Plateau, which is now seen to be some 35 km wider, as defined by the 4.5-km isobath. We found a deeper (5.5 km) trough between the Zenith Plateau and the Quokka Rise, which is seen to be more elongate along a north-east trend. Sonja Ridge, detached from the Quokka Rise by a 4.5-km-deep trough, and its companion, Sonne Ridge, continue southward into the Wallaby Plateau.

The bathymetry of the region has two trends: 1. a trend of 300°, expressed by the southwest flank of Zenith and Wallaby Plateaus, Lost Dutchman Ridge and the southwest flank of Exmouth Plateau, which are recognised as fracture zones; other fracture zones are inferred from the offsetting of magnetic anomalies; the Naturaliste Fracture Zone

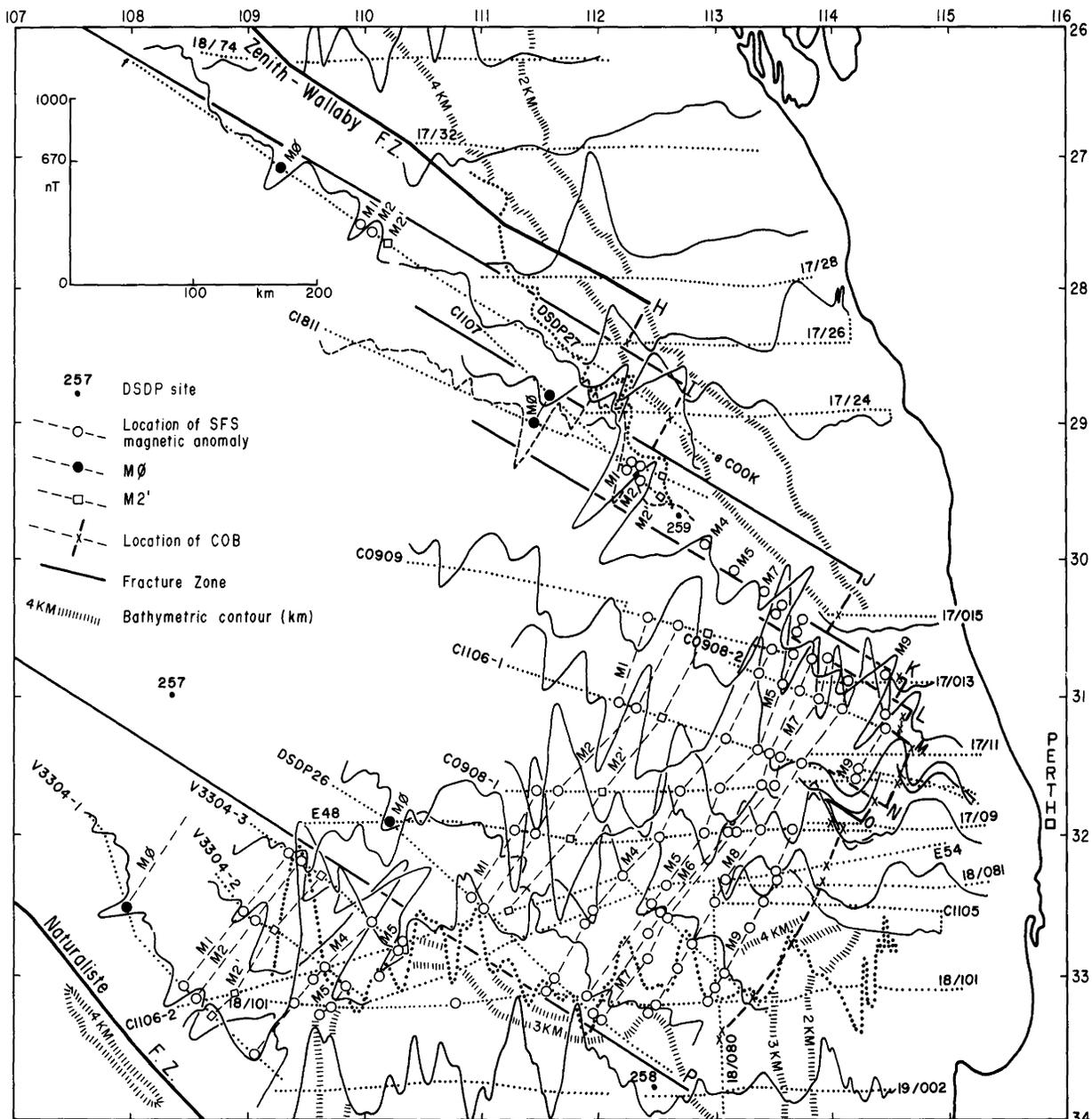


Fig. 5. Selected ships' tracks and magnetic profiles for the southern area. New profile for segment IJ from Cook; rest of profiles compiled from Markl (1974a, b; 1978a, b), Davies et al. (1973), Veevers et al. (1974), Hawkins et al. (1965) and BMR Margins Survey. Magnetic-anomaly determinations south of latitude 31°S slightly amended from Markl; north of 31°S, in segments J to M, amended after Larson et al. (1978). Determinations in IJ and of COB anomaly by us

(Markl, 1978b) intersects this trend at an angle of 20°; and 2. at right-angles, a trend of 030°, expressed by the continental slope at the southeast margin of the Cuvier Abyssal Plain, the northwest Exmouth Plateau, the northwest flank of the Wallaby Plateau and Sonja Ridge and the northern part of Sonne Ridge; athwart this are the trends of the southern Sonne Ridge (010°) and Dirck Hartog Ridge (020°).

As shown by the seafloor-spreading magnetic anomalies, the 030°-trend is the strike of the spreading and the 300°-trend the orthogonal strike of the fracture zones. This can be modelled by a pole of rotation at 56.5° S, 2.7° W of India from Australia, relative to Australia, for the interval 125 [128]–92.7 [95.0] Ma ago (Johnson and Veevers, 1984). The linear features athwart these trends – the Natur-

aliste Fracture Zone, Dirck Hartog Ridge and the southern part of the Sonne Ridge – are anomalous.

Seafloor-spreading magnetic anomalies (Figs. 3–7)

West and south of the Exmouth Plateau

Following Powell (1978) and Larson (1977), we identify the anomalies in the area west and south of the Exmouth Plateau as shown in Figs. 3 and 4. The best-mapped set of anomalies in the region is that in the Cuvier Abyssal Plain (Larson, 1977; Larson et al. 1979; Johnson et al. 1980) which, as shown by V3308-2 (Fig. 4a), comprises M10 through M5 reflected about the axis(A) of an abandoned spreading ridge marked by Sonne Ridge, succeeded

by M4 through M0. The magnetic effect of Sonne Ridge, which rises 2 km above the level of layer 2 along V3308-2 (dotted lines in Fig. 4a), is small compared with the sea-floor-spreading magnetic anomalies and suggests that ridges and depressions in the oceanic basement contribute only a small part of the total magnetic anomaly in this region.

The same model is compared with two pairs of profiles in the northernmost spreading segment, between fracture zones A and C. All profiles have a broad and deep magnetic depression northwest of M4: M1968(N) and B1971-1 contain M1, M2 and the adjacent trough, designated M2'; all contain y' (a characteristic anomaly between M10' and M10N'); and M1968(N) contains an almost complete set of M5 through M9 reflected about an axis (A). Correlation of the other complete profile [M1968(S)] with the model is weak, except for M4, y' , y and M5'-A-M5. In isolation, few of these determinations are acceptable and the pattern of the poorly correlated profiles of the Gascoyne Abyssal Plain only palely reflects the good correlation of observed and modelled anomalies in the Cuvier Abyssal Plain, but it satisfies the other constraint on the spreading pattern: the comparable width of separation of M0 and the COB (600 km), which is not explicable without duplication.

In the intervening segment, between fracture zones C and D (Fig. 4b), in which the distance from M0 to the COB is only 400 km, the magnetic-anomaly sequence is modelled from V3308-4 (after Larson, 1977) as a clear M0 through M4, and on AII9313 and 17/72 as tentatively M5 through M9 so that the difference of 200 km in the width from M0 to the COB is accommodated by an unreflected set. Larson (1977) remarked that the positive anomaly we identify on V3308-4 as M4 "is probably not M-4, but an anomaly associated with the edge of the Exmouth Plateau". We disagree because the anomaly matches the model and the edge of the Exmouth Plateau, as indicated by the lower continental slope seen in 17/70, is about 250 km away. We differ with Powell and Luyendyk (1982) in identifying their M4 as y' in B1971-1 (Fig. 4a), and M6 as M9'.

In the Cuvier area (Fig. 4a), the COB anomaly is isolated as the large anomaly shown in V3308-2 and 17/56 that corresponds to the position of the COB as independently determined by Roots et al. (1979) from the contrast in seismic character and by modelling the free-air gravity anomaly. The COB-anomaly straddles an extrapolated spreading age of 125 [128] Ma ago and would be modelled by an expanded (higher intensity or thicker) source block of normal polarity between 124.07 and 124.9 Ma ago. The spreading block model and its anomalies (dotted lines) are extended to the east in Fig. 4a to show the non-correlation with the observed anomaly beyond 125 Ma ago. The oldest modelled age in the reflected set to the west of Sonne Ridge is 123.7 Ma, so that we must postulate removal westward of M0 of the reflected block between 123.7 and 125 Ma ago, there being no indication of a symmetrically arranged set of this age to the east of Sonne Ridge. In the Gascoyne Abyssal Plain (Fig. 4a), as shown by M1968(N), the COB anomaly is isolated as a large anomaly with a modelled age of 123.7 Ma, corresponding to an expanded (higher intensity or thicker) block between y (122.7 Ma) and M10N (123.7 Ma), which coincides with the age of the reflected set on the west. In both the Cuvier and Gascoyne Abyssal Plains the age of A is 117.5 Ma and the oldest isochron in the resumed spreading after the ridge jumped to the west

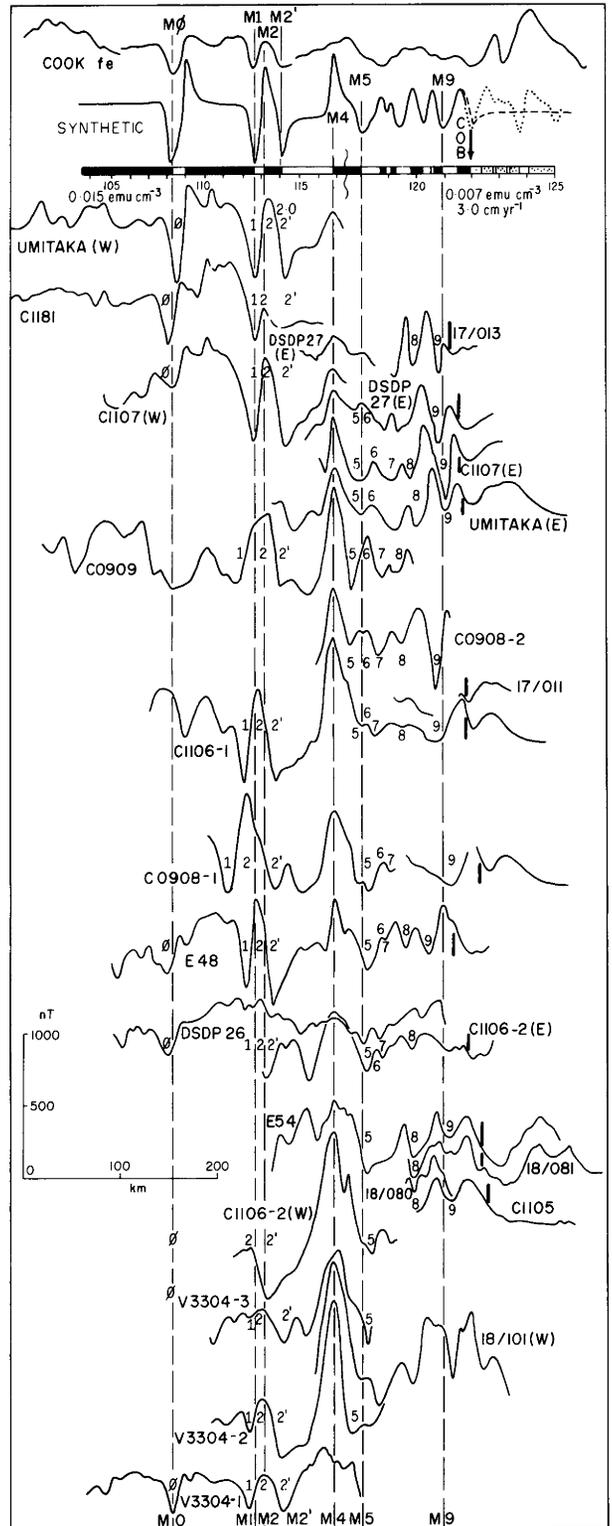


Fig. 6. Selected magnetic profiles, located on Fig. 5, projected on an azimuth of 302° and aligned on M4, compared with a synthetic model extended past the COB (dotted blocks and dotted line). Lines through modelled anomalies afford direct comparison with observed anomalies. COB anomaly indicated by heavy broken line on right. Model parameters not shown in the figure are as for Fig. 4

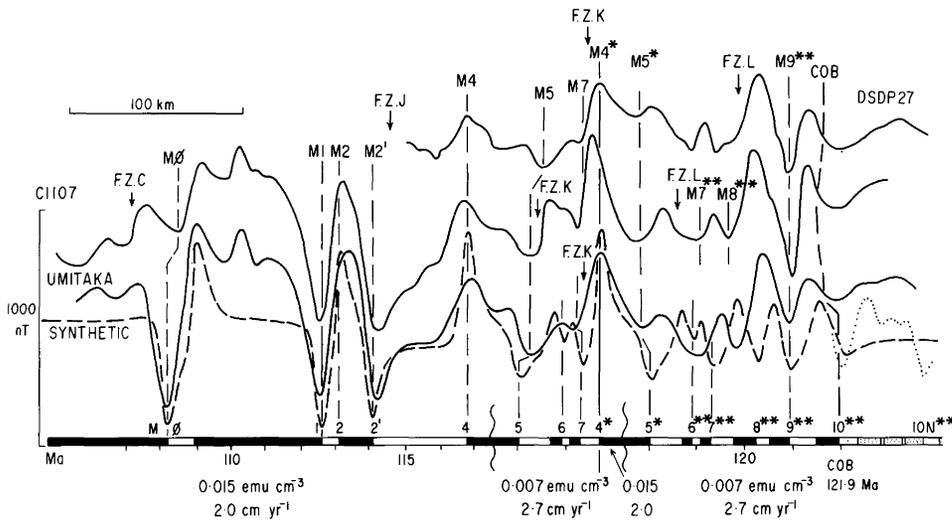


Fig. 7. Magnetic profiles of Leg 27 and C1107, located on Fig. 5, and *Umitaka* aligned on M9 across the segments between J and M, projected on 302°, and synthetic model (broken line) incorporating a repetition of M4 and M5 (asterisked) in segment KL, and M6 through M10 (double asterisked) in LM. COB at 121.9 Ma. Block model extended beyond 121.9 Ma by dots. Model parameters not shown in the figure are as in Fig. 4

is 117.1 Ma, immediately before M4. In the segment between fracture zones C and D (Fig. 4b), the COB anomaly is isolated as the western flank of a negative anomaly (M10) at a modelled age of 121.9 Ma, best seen on 17/70, with the smooth zone of the Exmouth Plateau beyond. On the other profiles, some smaller anomalies at the edge of the Exmouth Plateau are later modelled (Fig. 10) as due to dykes.

Southwest of the Cuvier Abyssal Plain: Wallaby Plateau, Quokka Rise, Zenith Plateau

The magnetic lineation in the Cuvier Abyssal Plain (Fig. 3) cannot be traced south of V3305. Fracture zone G2 marks a discontinuity in the regularity of magnetic anomalies, from the set of good anomalies exemplified by V3308-2 to the irregular anomalies of *Cook* mn, 17/48, 17/46, and *Cook* hg. The only link between the oceanic structure of the Cuvier Abyssal Plain and the Wallaby-Quokka region is the continuity of the Sonne Ridge as seen in seismic profiles (Figs. 3 and 4a), with its characteristic magnetic signature of M5'-A-M5 in V3305, *Cook* mn and 17/48. A possible link in the western part of the region is a positive magnetic anomaly on *Cook* mn, lk, ij and hg, tentatively identified as M2, flanked by M1 and M2'. The irregular magnetic anomalies of this region are interpreted in Figs. 3 and 4a as a degraded extension of the set of seafloor-spreading magnetic anomalies of the Cuvier Abyssal Plain, but a detailed magnetic survey would be required to establish the validity or not of this interpretation. Accordingly, the interpreted anomalies are distinguished in Figs. 3 and 4a from the established ones to the north. Later, in the discussion, we suggest a mode of spreading for the Wallaby-Quokka region like that in the area between the Reykjanes Ridge and the southern shelf of Iceland. The only possible COB anomaly is that on V3305 and 17/46, which cross at this point; a model age of 125 [128] Ma, as in the Cuvier Abyssal Plain, is indicated.

West of Perth

In the region northwest of Perth (Fig. 5), a broad segment between fracture zones L and P is defined by a set of sub-parallel anomalies that range from M0 through M9 (Fig. 6). Between latitude 31° and 28° S, parts of this set are offset progressively to the northwest to define three narrow

spreading segments and a fourth, between I and H (the Zenith-Wallaby Fracture Zone), is postulated to account for the offset in the COB. South of P, which marks the northeast flank of the Naturaliste Plateau, a set of anomalies (M0 through M4) discovered by Markl (1978a), and possibly also M5, lie northwest of the Naturaliste Plateau; as noted by Markl (1978a), M4 in V3304-1 and -3 is perturbed by the topography of the plateau margin.

All these anomalies are aligned along M4 in Fig. 6 and are modelled with the same parameters as apply in the north. Except northwest of Naturaliste Plateau, where M5 is the oldest recognised anomaly, the anomaly set starts with M9 and the COB anomaly is isolated as a modified M10 at 121.5 [124.7] Ma in segment JK, 121.3 [124.3] Ma in KL and 122.5 [125.5] Ma in LP, the same age as in CD. By extrapolation, the COB anomaly in IJ is 125 [128] Ma. The strong parts of the interpretation are: 1. the near-parallelism of the anomalies along an azimuth of 032°, 2. the well-defined M0 through M4, and 3. a well-defined M9.

In segment IJ, M0 through M2', hinted at in the profile collected on *Umitaka Maru* by Tomoda et al. (1968), are identified along the track of *Cook*. In JK, KL and LM, the tracks of C1107, DSDP Leg 27 and *Umitaka*, whose track (not shown) is close to that of the others, are interpreted as crossing the three segments (Fig. 7) as first mapped for JK by Larson et al. (1978). Within this segment, DSDP site 259 lies midway between M2' and M4 with a predicted magnetic age of 115.5 [118.3] Ma. The oldest fossils recovered from this site were dated as late Neocomian by Morgan (1980, p. 63) and more precisely by Dr. R. Helby (1983, personal communication) as middle-late Valanginian to early-middle Hauterivian, which has a midpoint of 115 [117.8] Ma, thus marking a precise tie of the magnetic and biostratigraphic time-scales.

The tracks of the three profiles shown in Fig. 7 lie close to each other in compartment JK, and that of C1107 diverges slightly from the others in KL. The anomaly sequence in segment JK is interpreted as extending to M7 on Leg 27 and *Umitaka*, and to M5 on C1107, and in segment KL (anomalies asterisked in Fig. 7) from M4* to M5* on Leg 27 and C1107. M7** and M8** on C1107 and M9** on Leg 27 and C1107 are in segment LM. The short offset along L is expressed in these profiles by the much broader anomaly between M8** and M9**.

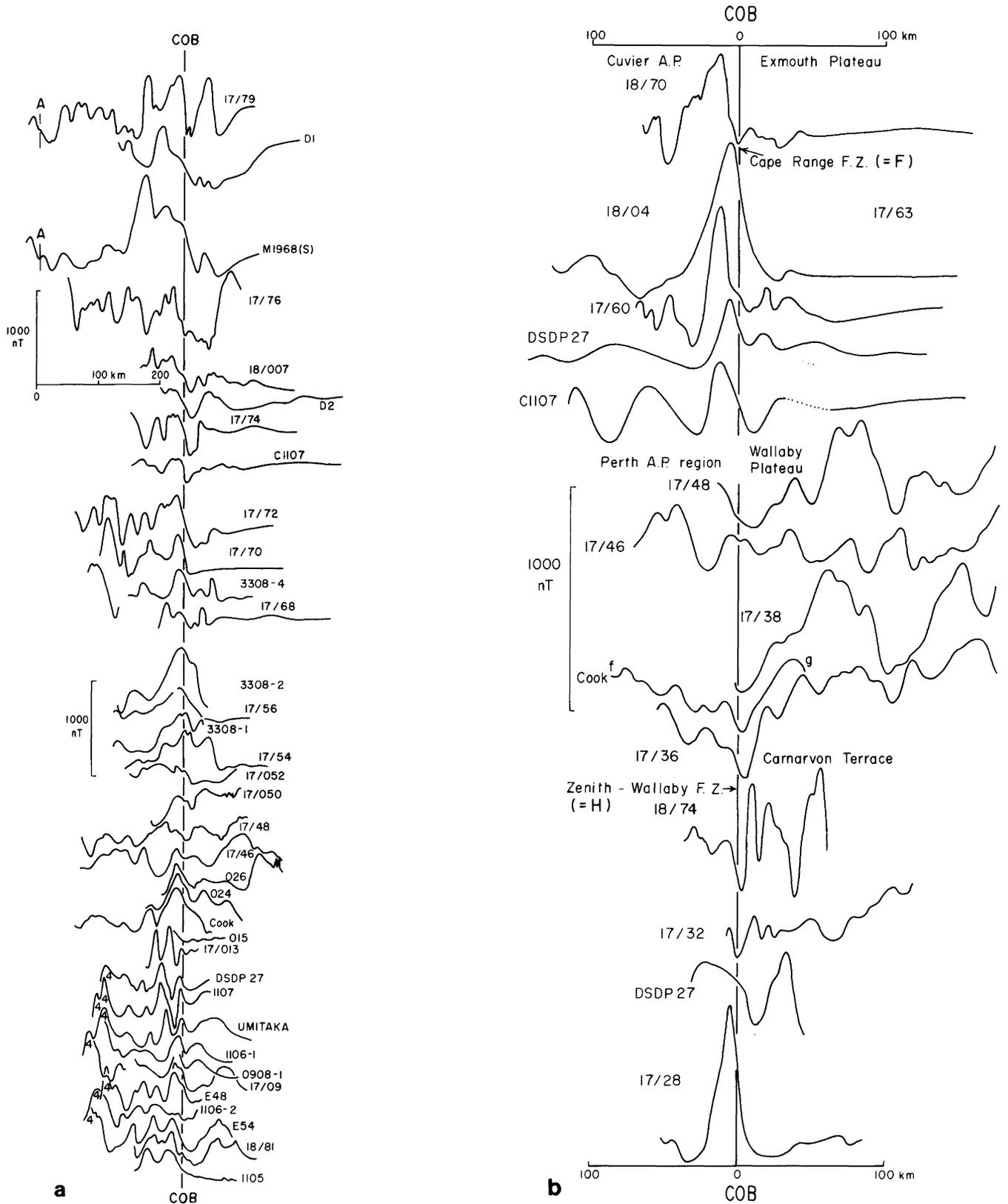
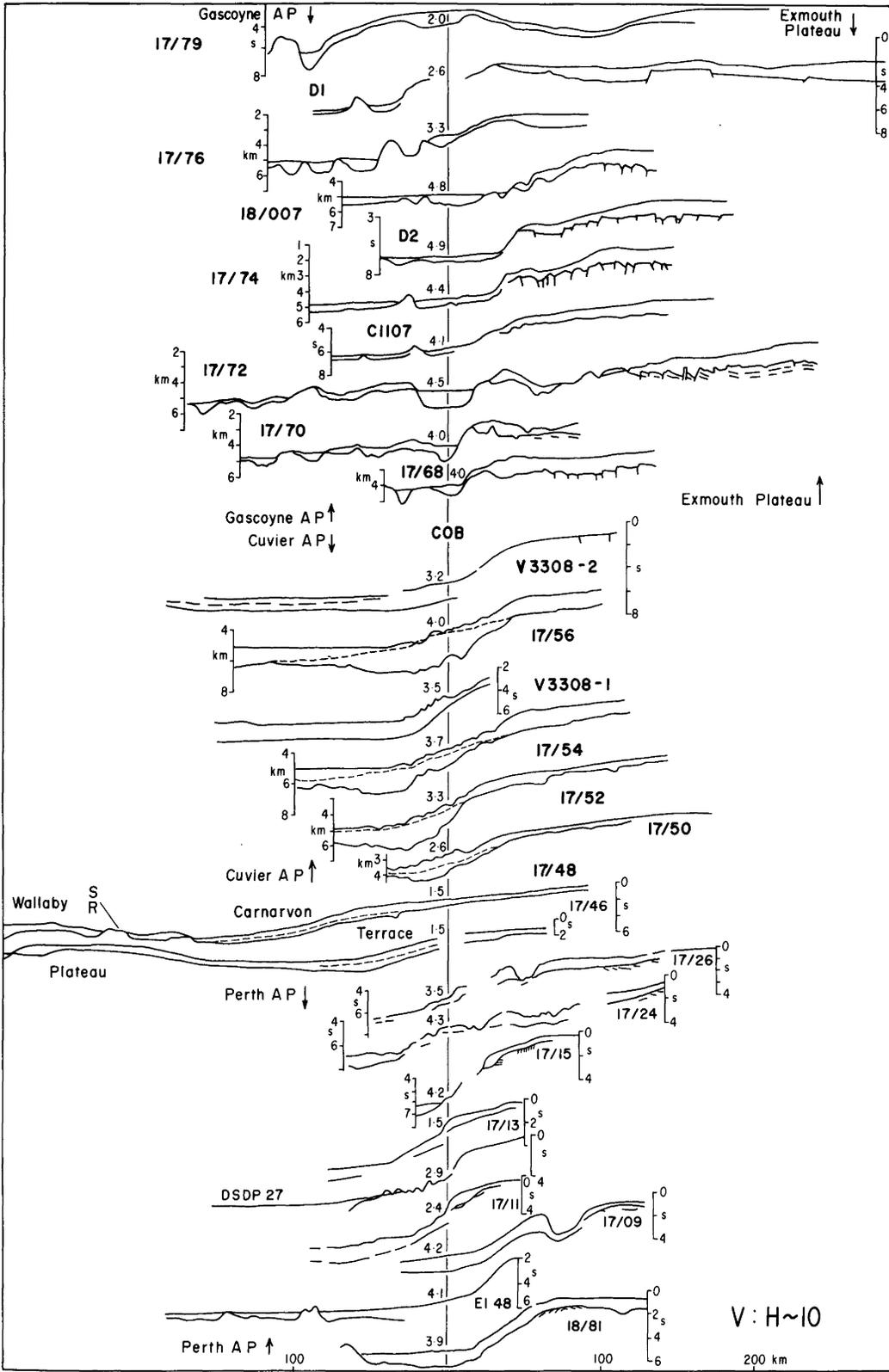


Fig. 8. a Magnetic profiles projected on 302° across the rifted margin from north to south and aligned along the COB. **b** Magnetic profiles projected on 032° across the transform-faulted margin from north (Cape Range Fracture Zone) to south (Zenith-Wallaby FZ) and aligned along the COB, except the northern part of the Zenith-Wallaby FZ (17/48 to 18/74) which we believe to be a boundary between normal and epilithic oceanic crust



a

Fig. 9. a Seismic profiles (reflection time in seconds), some converted to depth (km) sections, projected on 302° across the rifted margin from north to south and aligned along the COB, all corresponding to the magnetic profiles of Fig. 8a. Water depth at COB in km. Lower line is layer 2 on the oceanic side and the breakup unconformity on the continental side; broken line is boundary between clay below and carbonate above. **b** Seismic profiles (reflection time in seconds), projected on 032° across the transform-faulted margin from north (Cape Range FZ) to south (Zenith-Wallaby FZ). We believe the boundary between the Perth Abyssal Plain and the Wallaby Plateau to be that between normal and epilithic oceanic crust. Water depth at COB in km

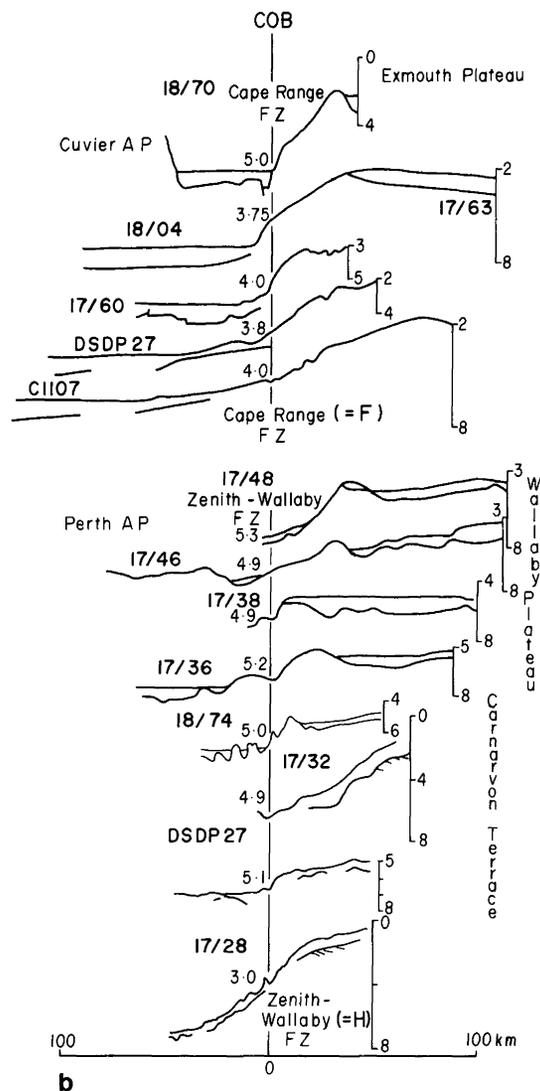


Fig. 9b

Magnetic anomaly at the COB

The COB anomalies, isolated from the seafloor-spreading anomalies, are aligned in Fig. 8a for profiles projected along an azimuth of 302° , i.e. normal to the isochrons and the trend of the rifted margin; and in Fig. 8b for profiles projected along 032° , i.e. parallel to the isochrons, across the transform-faulted margins of the Cape Range Fracture Zone (F) and part of the Zenith-Wallaby Fracture Zone (H). Corresponding seismic profiles (Fig. 9a and b) show that, with the few exceptions discussed later, the COB anomaly lies at or within 40 km of the foot of the lower continental slope at water depths of 2.4–4.9 km (mean 3.75 km) for the rifted margin and of 3.0–5.3 km (mean 4.5 km) for the transform-faulted margin, corresponding to a change in seismic-reflection character from the faulted breakup unconformity on the continent to the smooth but hyperbolic oceanic layer 2 (Veevers and Cotterill, 1978).

From the few available profiles, Falvey (1972b) identified the magnetic anomaly at the lower slope as the COB anomaly and modelled it by assuming magnetic induction and remanence in both oceanic and continental "high susceptibility 'basement' at an intermediate crustal depth".

Here, we follow the assumption of Talwani and Eldholm (1973) that the COB anomaly is due to the contrast between the remanently magnetized oceanic crust and the non-magnetic continental crust. The most distinctive magnetic anomaly across the COB of the rifted margin is in 17/70, a simple downslope that meets the smooth field of the Exmouth Plateau at a shallow magnetic trough. This anomaly is modelled (Fig. 10) as a normal remanently magnetized oceanic block (at M10) that terminates at the COB against the non-magnetic continent. The amplitude of 430 nT requires the oceanic block to have an intensity or thickness double that of the other blocks. The neighbouring profiles in segment CD (17/72, V3308-4, 17/68) have anomalies of similar shape except for two small peaks that rise above the smooth Exmouth Plateau flank. These peaks, in 17/68, are modelled (Fig. 10) by the same contrast between the oceanic block and the adjacent continental crust with, in addition, three dykes of reversed polarity and each 0.2 km wide in the basement of the outer Exmouth Plateau, following a general suggestion by Exon and Willcox (1980, p. 40). North and south of segment CD, the COB anomaly maintains the same shape on the southeast flank of a downward-sloping anomaly. As shown by the 35 complete COB profiles in Fig. 8a, the amplitude of the COB anomaly has a mean and mode of 300 nT. The model of the COB anomaly in 17/70, with an amplitude of 430 nT, represents those profiles with an amplitude greater than the mean and that in 17/68 represents those less than the mean.

Across the COB at the transform fault of the Cape Range Fracture Zone (F in Fig. 3), between the Cuvier Abyssal Plain and the southwest Exmouth Plateau, the magnetic anomaly (upper part of Fig. 8b) is a nearly symmetrical positive feature of variable width and amplitude, with the COB on the northeast-sloping flank. Across the Zenith-Wallaby Fracture Zone (lower part of Figs. 8b and 9b), the COB anomaly again corresponds to a down-sloping flank but with an irregular shape except in 17/28, which resembles the anomaly across the Cape Range Fracture Zone.

The COB anomaly across the Cape Range Fracture Zone is modelled along 18/04–17/63 (Fig. 10) by a two-dimensional magnetic body 10 km wide, extending from the surface to the base of the oceanic crust. The landward edge of the body lies along the transform fault, which marks the COB. This body also satisfies the FAA (Willcox, 1977). The COB magnetic anomaly across the Cape Range Fracture Zone is everywhere positive and the body must have been intruded after the Cuvier Abyssal Plain was generated and during an interval of normal polarity. If the body were simply a thick marginal phase of the normal and reverse polarized ocean floor, then it would have changed polarity along the Cape Range Fracture Zone to produce an alternation of positive and negative anomalies.

Talwani and Eldholm (1973) and Rabinowitz (1976) modelled the COB positive anomaly across the Agulhas Fracture Zone off South Africa as due to the unadorned edge of oceanic layer 2, but did so only by assigning it a high intensity (a susceptibility of 0.04 cgs units for magnetic induction; and substantially the same for remanent magnetization, according to Rabinowitz [1976]). This kind of interpretation, involving a high-intensity source, is not applicable to the Cuvier Abyssal Plain, whose floor has a modelled remanent intensity of 0.007 emu/cm^3 (Larson, 1977).

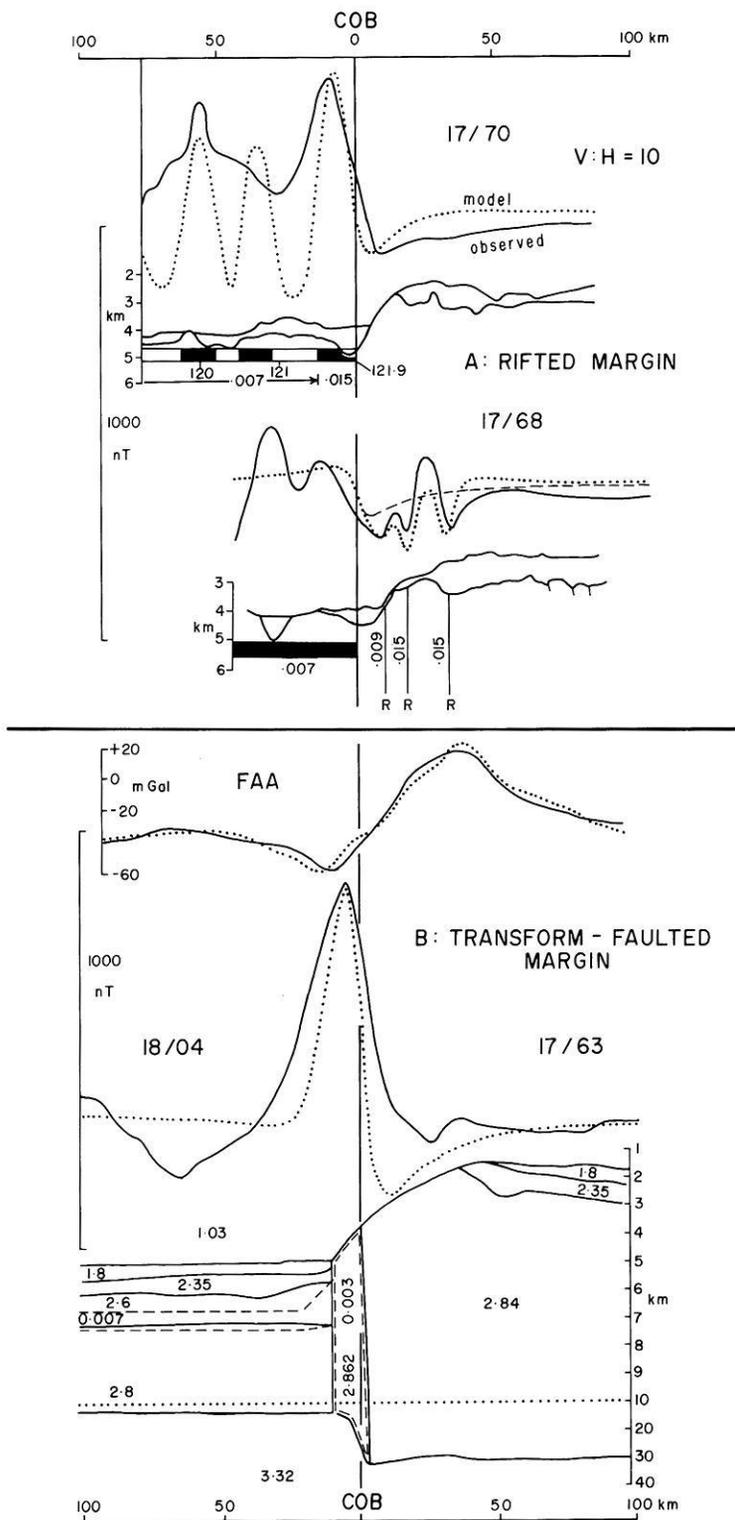


Fig. 10. A, models of the COB anomaly along the rifted margin on the western side of the Exmouth Plateau (17/70 and 17/68) projected on an azimuth of 302° and B, across the transform-faulted southern side (18/04-17/63) projected on an azimuth of 032°. The simple COB anomaly in 17/70 at the rifted margin is modelled (with the field parameters as given in Fig. 4) by modifying the magnetization of the oldest oceanic block of the seafloor-spreading sequence so that it has double the remanent-magnetization intensity of the others (as shown) or double the thickness. In 17/68, also at the rifted margin, a simple oceanic magnetic block of uniform intensity against non-magnetic continental crust produces the anomaly shown by the *dotted line* on the oceanic side and the *broken line* on the continental side; the addition of magnetic dykes of reversed polarity (R) near the edge of the continent simulates the complete set of anomalies. In 18/04-17/63, at the transform-faulted margin, a dense (2.862 t/m³) vertical body at the COB is required to account for the free-air gravity anomaly (FAA) (Willcox, 1977; Roots et al. 1979; Exon and Willcox, 1980), implying a vertical body emplaced along the Cape Range FZ, interpreted from seismic diffractions at the foot of the slope (Exon and Willcox, 1980). Vertical exaggeration is x 10 above 10 km depth, x 1 below. Other densities from refraction velocities found by Larson et al. (1979). The magnetic anomaly is modelled by two bodies, outlined by *broken lines*: (1) part of oceanic layer 2 (density = 2.6 t/m³, remanent-magnetization intensity = 0.007 emu/cm³) and (2) a nearly vertical body (density = 2.862 t/m³, intensity = 0.003 emu/cm³). Other parameters as in Fig. 4. The width of the model anomaly is less than that observed, but is close to that of the other crossings of the Cape Range Fracture Zone

With the exception of the symmetrical positive anomaly in 17/28 (Fig. 8b), the magnetic anomaly across the Zenith-Wallaby Fracture Zone cannot be modelled as for the Cape Range Fracture Zone. Four crossings of the Zenith-Wallaby Fracture Zone – across the COB along the Carnarvon Terrace (18/74, 17/32) and across the boundary between the normal oceanic and thick (presumed epilithic) crust along the Wallaby Plateau (Cook fig. 17/36) – are marked by a negative magnetic anomaly, which implies an oceanic

block of reversed polarity at the boundary. This is the only hint of a reversely polarized block along the COB of the western margin.

The dominantly, if not exclusively, normal polarity of the oldest oceanic block at the COB of the western margin is unexpected because the magnetic reversal sequence in the range of interest, from 120 to 130 Ma, contains an equal number of short (< 1 Ma) normal and reverse polarity intervals.

– by von Stackelberg et al. (1980) by sampling of seismic sequences.

On the eastern and southern Wallaby Plateau, and on the Sonne Ridge which extends northward into the Cuvier Abyssal Plain, the layered sequence [seen in processed seismic profiles to be at least 3000 m thick] beneath the main Neocomian unconformity consists of interbedded weathered tholeiitic and differentiated alkali basalts, tuffs, basalt breccias and thick volcanoclastic sandstones and conglomerates. A minimum mid-Cretaceous age (K/Ar age: 89 m.y.) was determined for a somewhat altered basalt from the southern Wallaby Plateau. This suggests that intense volcanism and associated deposition of volcanoclastic debris flows formed the plateau, during or after the Neocomian breakup of the region.

Von Rad and Exon (1983) note further that the composition of the basalt supports an oceanic origin. This work countered the interpretation of Symonds and Cameron (1977) that the Wallaby Plateau has a basement of Palaeozoic rocks, implying a continental crustal structure, but did not rule out the possibility that the volcanic pile was a thick superstructure on continental crust. This possibility is now questioned by: (a) the continuity in seismic and magnetic profiles of the Sonne Ridge southward from the Cuvier Abyssal Plain along the eastern flank of the Wallaby Plateau; (b) the possible magnetic lineation of M1–M2–M2' traced southward from the Cuvier Abyssal Plain into the trough between Quokka Rise and Zenith Plateau; and (c) the further interpretation from (a) and (b) of a reflected set of magnetic anomalies, as in the Cuvier Abyssal Plain, in the Wallaby-Quokka region which satisfies the geometrical constraints of spreading. The degraded shape of the anomalies compared with the regularly shaped anomalies in the Cuvier Abyssal Plain requires explanation.

A similar transition from clear to degraded seafloor-spreading magnetic anomalies is found from the normal oceanic crust of the Reykjanes Ridge and Kolbeinsey Ridge segments of the mid-Atlantic Ridge to the thick (Icelandic-type) oceanic crust of Iceland (Nunns et al. 1983). The aeromagnetic survey of the region by Serson et al. (1968) (who only showed the vertical component) – see Rutten (1975) for the total-field anomalies of this survey – has parameters similar to those of the marine survey of the Wallaby Plateau region (a line spacing of about 40 km and a separation of 3–5 km between the source and receiver) and affords a direct comparison with the Wallaby region. The present latitude of Iceland (65°) compares with the Early Cretaceous palaeolatitude of the Wallaby Plateau of 40°–50° (Embleton, 1984). Talwani et al. (1971) and Nunns et al. (1983) trace the prominent axial Reykjanes Ridge anomaly northeastward into a degraded anomaly along the axial rift zone of the Reykjanes Peninsula. Von Stackelberg et al. (1980) envisage the Early Cretaceous environment of the Wallaby Plateau as similar to that of Iceland, including shallow-water to subaerial extrusion of amygdaloidal basalt. Volcanism ceased with the jump of the spreading ridge westward and the plateau submerged below sea level. The volcanoclastic sediments were probably deposited either as sediment wedges that built out the volcanic island shelf or as deep marine deposits on the proto-abyssal plain. In this environmental setting, the transition from the northern part of Sonne Ridge progressively offset eastward into the Wallaby Plateau corresponds to that of the Reykjanes Ridge into Iceland. In the magnetic setting, the central

anomaly of the northern Sonne Ridge (M5'-A-M5) is traceable southward into the Wallaby Plateau, as the central anomaly (A1) of the Reykjanes Ridge is traceable to the shelf of Iceland, with the central anomaly continuing northeastward into the neovolcanic zone of Iceland. According to Nunns et al. (1983), "In view of the many unusual features of the Icelandic spreading environment (such as multiple and shifting axes, laterally extensive lava flows, prevalent central volcanoes and modifications due to erosion and alteration) the absence [outside the neovolcanic zone] of persistent, identified linear anomalies is not surprising", and this is the probable explanation of the blurred anomalies of the Wallaby Plateau. The only other parts along the margin with inferred epilithic structure are the adjacent outer Carnarvon Terrace and the outer Platypus Spur. It will require further work to ascertain if these areas are examples of the thick oceanward-dipping layers described by Hinz (1981) and of the subaerial seafloor spreading of Mutter et al. (1982).

Thick oceanic layer at the COB

According to Roots et al. (1979), the oceanic crust in the Cuvier Abyssal Plain (and elsewhere) thickens from 6 or 7 km to 16 km at the COB. Some of the profiles in the Cuvier Abyssal Plain and elsewhere on the western margin contain a COB anomaly with an amplitude as much as double that of the adjacent seafloor-spreading magnetic anomalies, suggesting double the intensity or thickness of the magnetic source layer. If, on the one hand, the magnetic source layer includes the entire oceanic crust then such a thick crust at the COB would produce the large observed anomaly. If, on the other hand, only the uppermost 500 m or so of the oceanic crust are significantly magnetic, then the bigger anomaly would be attributable to a surface layer with an increased intensity. COB anomalies of smaller amplitude, common in the region northwest of Perth, suggest thin or weakly magnetized crust.

Crustal structure of Zenith and Naturaliste Plateaus

Magnetic lineations were not found over either of these plateaus, so their crustal structure, whether oceanic or continental, remains unknown. From a study of dredged material, Coleman et al. (1982) view the Naturaliste Plateau as probably oceanic, but conclusive evidence is lacking, as it is from the Zenith Plateau.

Acknowledgements. We thank: Dr. I.S.F. Jones, R.A.N. Research Laboratory, for the opportunity of taking part in the R.A.N. Research Laboratory cruise on HMAS *Cook*; Cmdr P. Cooke-Russell, captain, and the crew of *Cook* for unstinting cooperation; our colleague Mr K.S. Gibbons for help in collecting the geophysical data; the R.A.N. Hydrographic Office for sounding sheets; the N.S.W. Geological Survey for the use of their magnetometer; Mr R. Whitworth and Dr D.A. Falvey of the Bureau of Mineral Resources for a copy of the AGRF routine; Dr D.A. Falvey, Mr T.S. Powell and Dr R.G. Markl for a copy of their theses; and Dr H.A. Roeser, Dr D.K. Smythe and Mr P.A. Symonds for critical reviews of the manuscript. This work was supported by a grant from the Funding Advisory Panel of the Australian Marine Sciences and Technologies Advisory Committee and Marine Sciences and Technologies Grants Scheme of the Department of Science and Technology.

References

- Barber, P.M.: Palaeotectonic evolution and hydrocarbon genesis of the central Exmouth Plateau. *Aust. Pet. Expl. Assoc. J.* **22**, 131–144, 1982
- Coleman, P.J., Michael, P.J., Mutter, J.C.: The origin of the Naturaliste Plateau, SE Indian Ocean: implications from dredged basalts. *Geol. Soc. Aust. J.* **29**, 457–468, 1982
- Dalrymple, G.B.: Critical tables for the conversion of K-Ar ages to new constants. *Geology* **7**, 558–560, 1979
- Davies, T.A., Luyendyk, B.P., et al.: Initial reports of the Deep Sea Drilling Project, **26**. Washington: U.S. Govt. Printing Office 1973
- Embleton, B.J.J.: Continental palaeomagnetism. In: Phanerozoic earth history of Australia, J.J. Veivers (ed): pp. 11–16. Oxford: Oxford University Press 1984
- Exon, N.F., Willcox, J.B.: The Exmouth Plateau: stratigraphy, structure, and petroleum potential. *Aust., Bur. Miner. Resour., Geol. Geophys., Bull.* **199**, 1980
- Exon, N.F., von Rad, U., von Stackelberg, U.: The geological development of the passive margins of the Exmouth Plateau off northwest Australia. *Marine Geol.* **47**, 131–152, 1982
- Falvey, D.A.: Sea-floor spreading in the Wharton Basin (northeast Indian Ocean) and the breakup of eastern Gondwanaland. *Aust. Pet. Expl. Assoc. J.* **12**, 86–88, 1972a
- Falvey, D.A.: The nature and origin of marginal plateaux and adjacent ocean basins off northern Australia. Ph.D. thesis, University of N.S.W., 1972b
- Falvey, D.A., Veivers, J.J.: Physiography of the Exmouth and Scott Plateaus, Western Australia, and adjacent Wharton Basin. *Marine Geol.* **17**, 21–59, 1974
- Fisher, R.L.: General bathymetric chart of the oceans (GEBCO), Sheet 5.09, 5th Edn. *Canad. Hydro. Service. Ottawa*, 1982
- Hawkins, L.V., Hennion, J.F., Nafe, J.E., Thyer, R.F.: Geophysical investigations in the area of the Perth Basin, Western Australia. *Geophysics* **30**, 1026–1052, 1965
- Heezen, B.C., Tharp, M.: Physiography of the Indian Ocean. *Philos. Trans. R. Soc. London, Ser. A*: **259**, 137–149, 1966
- Heirtzler, J.R., Cameron, P., Cook, P.J., Powell, T., Roeser, H.A., Sukardi, S., Veivers, J.J.: The Argo Abyssal Plain. *Earth Planet. Sci. Lett.* **41**, 21–31, 1978
- Hinz, K.: A hypothesis on terrestrial catastrophes: wedges of very thick oceanward dipping layers beneath passive continental margins. *Geol. Jahrb., Ser. E*, **22**, 3–28, 1981
- Johnson, B.D., Veivers, J.J.: Oceanic palaeomagnetism. In: Phanerozoic earth history of Australia, J.J. Veivers (ed): pp. 17–38. Oxford: Oxford University Press 1984
- Johnson, B.D., Powell, C.McA., Veivers, J.J.: Early spreading history of the Indian Ocean between India and Australia. *Earth Planet. Sci. Lett.* **47**, 131–143, 1980
- Larson, R.L.: Early Cretaceous breakup of Gondwanaland off Western Australia. *Geology* **5**, 57–60, 1977
- Larson, R.L., Hilde, T.W.C.: A revised time scale of magnetic anomalies for the Early Cretaceous and Late Jurassic. *J. Geophys. Res.* **80**, 2586–2594, 1975
- Larson, R.L., Carpenter, G.B., Diebold, J.B.: A geophysical study of the Wharton Basin near the Investigator Fracture Zone. *J. Geophys. Res.* **83**, 773–782, 1978
- Larson, R.L., Mutter, J.C., Diebold, J.B., Carpenter, G.B., Symonds, P.: Cuvier Basin: a product of ocean crust formation by Early Cretaceous rifting off Western Australia. *Earth Planet. Sci. Lett.* **45**, 105–114, 1979
- Markl, R.G.: Bathymetry, sediment distribution, and sea-floor spreading history of the southern Wharton Basin, eastern Indian Ocean. Ph.D. thesis, Univ. Connecticut, 1974a
- Markl, R.G.: Evidence for the breakup of eastern Gondwanaland by the Early Cretaceous. *Nature* **251**, 196–199, 1974b
- Markl, R.G.: Further evidence for the Early Cretaceous breakup of Gondwanaland off southwestern Australia. *Marine Geol.* **26**, 41–48, 1978a
- Markl, R.G.: Basement morphology and rift geometry near the former junction of India, Australia and Antarctica. *Earth Planet. Sci. Lett.* **39**, 211–225, 1978b
- Morgan, R.: Eustasy in the Australian Early and Middle Cretaceous. *N.S.W., Geol. Surv., Bull.* **27**, 1980
- Mutter, J.C., Talwani, M., Stoffa, P.L.: Origin of seaward-dipping reflectors in oceanic crust off the Norwegian margin by “sub-aerial sea-floor spreading”. *Geology* **10**, 353–357, 1982
- Nunns, A.G., Talwani, M., Lorentzen, G.R., Vogt, P.R., Sigurgeirsson, T., Kristjansson, L., Larsen, H.C., Voppel, D.: Magnetic anomalies over Iceland and surrounding seas. In: Structure and development of the Greenland-Scotland Ridge, M.H.P. Bott, S. Saxov, M. Talwani, J. Thiede (eds): pp. 661–678, New York: Plenum Press 1983
- Petkovic, J.J., Whitworth, R.: Problems in secular variation in the Australian region. *EOS* **56**, 547–548, 1975
- Powell, T.S.: The sea-floor spreading history of the eastern Indian Ocean. M.A. thesis, Univ. California, Santa Barbara, 1978
- Powell, T.S., Luyendyk, B.P.: The sea-floor spreading history of the eastern Indian Ocean. *Marine Geophys. Res.* **5**, 225–247, 1982
- Rabinowitz, P.D.: Geophysical study of the continental margin of southern Africa. *Bull. Geol. Soc. Amer.* **87**, 1643–1653, 1976
- Roots, W.D., Veivers, J.J., Clowes, D.F.: Lithospheric model with thick oceanic crust at the continental boundary: a mechanism for shallow spreading ridges in young oceans. *Earth Planet. Sci. Lett.* **43**, 417–433, 1979
- Rutten, K.: Two-dimensionality of magnetic anomalies over Iceland and Reykjanes Ridge. *Marine Geophys. Res.* **2**, 243–263, 1975
- Serson, P.H., Hannaford, W., Haines, G.V.: Magnetic anomalies over Iceland. *Science* **162**, 355–357, 1968
- Symonds, P.A., Cameron, P.J.: The structure and stratigraphy of the Carnarvon Terrace and Wallaby Plateau. *Aust. Pet. Expl. Assoc. J.* **17**, 30–41, 1977
- Talwani, M., Windisch, C.C., Langseth, M.G.: Reykjanes Ridge crest: a detailed geophysical study. *J. Geophys. Res.* **76**, 473–517, 1971
- Talwani, M., Eldholm, O.: Boundary between continental and oceanic crust at the margin of rifted continents. *Nature* **241**, 325–330, 1973
- Tomoda, Y., Ozawa, K., Segawa, J.: Measurements of gravity and magnetic field on board a cruising vessel. *Bull. Ocean Res. Inst., Univ., Tokyo*, **3**, 1–170, 1968
- Veivers, J.J., Cotterill, D.: Western margin of Australia: evolution of a rifted arch system. *Bull. Geol. Soc. Amer.* **89**, 337–355, 1978
- Veivers, J.J., Heirtzler, J.R. et al.: Initial reports of the Deep Sea Drilling Project, **27**. Washington, U.S. Govt Printing Office, 1974
- Veivers, J.J., Tayton, J.W., Johnson, B.D.: Prominent magnetic anomaly along the continent-ocean boundary between the northwestern margin of Australia (Exmouth and Scott Plateaus) and the Argo Abyssal Plain. *Earth Planet. Sci. Lett.* 1985 (in press)
- von Rad, U., Exon, N.F.: Mesozoic-Cenozoic sedimentary and volcanic evolution of the starved passive margin off northwest Australia. *Amer. Assoc. Pet. Geol., Mem.* **34**, 253–281, 1983
- von Stackelberg, U., Exon, N.F., von Rad, U., Quilty, P., Shafik, S., Beiersdorf, H., Seibert, E., Veivers, J.J.: Geology of the Exmouth and Wallaby Plateaus off northwest Australia: sampling of seismic sequences. *BMR J. Aust. Geol. Geophys.* **5**, 113–140, 1980
- Willcox, J.B.: Some gravity models of the continental margin in the Australian region. *Bull. Aust. Soc. Expl. Geophys.* **8**, 118–124, 1977

Domain state of Ti-rich titanomagnetites deduced from domain structure observations and susceptibility measurements

E. Appel and H.C. Soffel

Institut für Allgemeine und Angewandte Geophysik, Theresienstraße 41, D-8000 München 2, Federal Republic of Germany

Abstract. Domain structure observations and measurements of the temperature dependence of susceptibility on synthetic and natural titanomagnetites are reported. At room temperature Ti-rich titanomagnetite particles ($x \approx 0.6$) of MD grain size normally develop a very complicated domain structure. The experimental results of our investigations, in addition to simple theoretical calculations, indicate that internal stress is the dominant cause of the observed anomalous domain patterns. Part of the results have already been published by Appel and Soffel (1984). The paper presented here, however, is a far more extended summary of the actual state of our research.

Key words: Rock magnetism – Titanomagnetites – Domain structure

Introduction

The acquisition and high stability of thermoremanent magnetization (TRM) in single domain (SD) particles is theoretically well explained by the theory of Néel (1949). In basaltic rocks, however, the volume fraction of SD particles is, in general, negligible (e.g. Winhard, 1983). Therefore, the micro-magnetic structure of the larger multidomain (MD) particles is of fundamental interest for the understanding of magnetization processes in rocks. There is already a broad spectrum of explanations of the pseudo-single domain (PSD) behaviour, especially the high stability of TRM: SD moments caused by dislocations (Verhoogen, 1959), subregions with submicroscopic grain boundaries (Ozima and Ozima, 1965), Barkhausen discreteness (e.g. Stacey, 1963), surface domains and subdomains (e.g. Banerjee, 1977), domain wall moments (e.g. Dunlop, 1977), absence of domain structure at all (e.g. Radhakrishnamurty et al., 1982) and domain wall nucleation (e.g. Halgedahl and Fuller, 1980). However, none of the theories is really satisfying and the problem of stable PSD remanence is still an open question in rock magnetism.

In ocean floor basalts the ferrimagnetic mineral phase normally consists of Ti-rich titanomagnetites $\text{Fe}_{3-x}\text{Ti}_x\text{O}_4$ with x around 0.6 (TM60) and some amounts of other cations, mainly Al and Mg. The typical composition is given by



(Appel and Moll [1980], calculated from microprobe analysis of Petersen et al. [1979], other cations neglected). For this reason, our studies were concentrated on Ti-rich compositions (partly substituted with Al, Mg). Domain structure observation provides the most direct analysis of the domain state. Earlier reported experiments (e.g. Soffel et al., 1982) proved the existence of domain structures on titanomagnetites up to TM75, but all studies have been carried out only with particles showing a nearly classical domain structure (constant specific wall energy, dominance of magnetocrystalline anisotropy as the source of anisotropy). However, such particles have rarely been observed. Non-classical, very complicated domain configurations – by far the most frequently occurring on surfaces of Ti-rich titanomagnetites – have generally been disregarded because of the assumption of incomplete removal of the irregular strained surface layer (Beilby layer) due to mechanical polishing. The latest investigations, however, provide good arguments for the existence and dominance of internal stress, suggesting that these anomalous domain structures are in fact the typical ones in Ti-rich titanomagnetites. This new view of stress-controlled domain state is based on the following experiments and calculations:

- qualitative domain structure observations,
- domain wall behaviour in external magnetic fields,
- domain wall behaviour under external uniaxial pressure,
- balancing of magnetocrystalline and magnetostrictive energies,
- measurements of the temperature dependence of susceptibility.

The possible origin and amount of internal stress in titanomagnetite particles are discussed in a later section.

Sample description

Most of the experiments were carried out with synthetic titanomagnetites (sintered at 1,300°C in an equilibrium oxygen atmosphere near the spinel–spinel + wüstite phase boundary). The polycrystalline specimens were tested with optical methods, X-ray analysis, Curie temperatures, dependence of saturation magnetization J_s on temperature and hysteresis measurements. According to the results of the tests, all synthetic samples are homogeneous and stoichiometric. The Curie temperature T_C , cell edge a_0 and coerciti-

Table 1. Curie temperature T_C , cell edge a_0 and coercivity H_C of the analysed synthetic and natural titanomagnetite samples

Composition	T_C , °C	a_0 , Å	H_C , Oe
Synthetic titanomagnetites			
Fe _{2.90} Ti _{0.10} O ₄ (TM10)	499	8.404	6
Fe _{2.76} Ti _{0.24} O ₄ (TM24)	424	8.421	4
Fe _{2.74} Ti _{0.26} O ₄ (TM26)	411	8.424	10
Fe _{2.60} Ti _{0.40} O ₄ (TM40)	314	8.442	4
Fe _{2.48} Ti _{0.52} O ₄ (TM52)	234	8.461	6
Fe _{2.40} Ti _{0.60} O ₄ (TM60)	154	8.476	5
Fe _{2.28} Ti _{0.72} O ₄ (TM72)	58	8.491	2
Fe _{2.25} Ti _{0.75} O ₄ (TM75)	40	8.492	10
Fe _{2.18} Ti _{0.62} Al _{0.20} O ₄ (ATM20/62)	56	8.450	4
Fe _{2.18} Ti _{0.62} Mg _{0.20} O ₄ (MTM20/62)	127	8.472	10
Fe _{2.315} Ti _{0.62} Al _{0.04} Mg _{0.025} O ₄ (AMTM4/2.5/62)	120	8.475	4
Fe _{2.075} Ti _{0.60} Al _{0.20} Mg _{0.125} O ₄ (AMTM20/12.5/60)	46	8.451	10
Basalt			
Fe _{2.26} Ti _{0.52} Al _{0.06} Mg _{0.13} Mn _{0.03} O ₄ (Basalt of Triebendorf TR)	220		65

vity H_C data are listed in Table 1. For experiments with small isolated particles, the polycrystalline material was crushed and embedded in a nonmagnetic matrix.

Further observations were carried out with natural Ti-rich titanomagnetites in a basaltic rock specimen. $J_s(T)$ measurements and optical analysis indicate that the titanomagnetite grains are unexsolved and only moderately oxidized. The composition was determined by microprobe analysis. Values of T_C and H_C are given in Table 1.

Domain structure studies

Observation technique

Domain structure observations were carried out with the Bitter pattern technique. The domain walls near the observed surface are represented by dark lines (accumulation of colloid particles). An ester-based magnetite colloid ("Ferrofluid") was used, enabling us to observe the dynamics of magnetization processes (domain wall displacements) directly. Strain-free surfaces were achieved by ionic polishing (Soffel and Petersen, 1971).

Qualitative domain structure observations

As already pointed out, recent domain observations were concentrated on the complicated non-classical domain structures which are seen, as a rule, on the major parts of the surfaces of Ti-rich titanomagnetite grains.

Following Syono (1965), the magnetocrystalline anisotropy constant K_1 at 290 K is negative for Ti-poor titanomagnetites and changes sign for $x > 0.65$ (linear interpolation of Syono's data). $K_1 < 0$ means that the spontaneous magnetization is orientated in the [111] direction and, there-

fore, a subdivision into domains with 180°, 109° and 71° domain walls is expected if a classical domain structure (Fig. 1) exists.

On synthetic Ti-poor titanomagnetites (up to TM26) classical domain structures comparable to Fig. 1 partially occur if the observed surface is nearly a (110) plane (Fig. 2). Typical closure domains (as already known from nickel) are developed if the observed surface is slightly inclined (perhaps about 10°) to a (110) plane (Fig. 2). If the angle between the directions of the spontaneous magnetization and the observed surface increases, more complicated surface domain patterns occur. In most cases the Bitter lines on Ti-poor titanomagnetites are nearly straight, not interrupted in the crystal's interior and fully developed up to the particle edges. However, some differences from the classical concept were observed: more strongly bent domain walls and weak and varying distinctness of Bitter lines. Furthermore, domain walls seem to dissolve inside the particles.

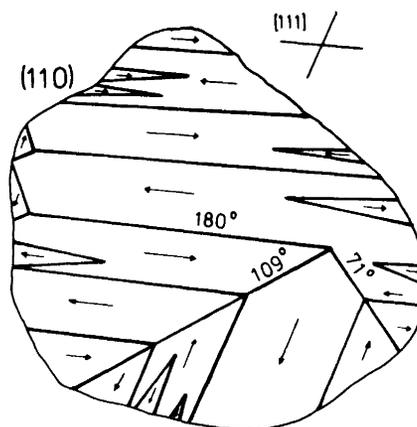


Fig. 1. Classical domain structure on a (110) plane for $K_1 < 0$ (71°, 109° and 180° domain walls, closure domains and spike domains)

Synthetic Ti-rich titanomagnetites around TM60 (compositions of greatest importance for rock- and palaeomagnetism) exhibit quite different domain structure configurations. The above-mentioned non-classical features partially occurring on Ti-poor titanomagnetites are typical for the Ti-rich compositions (examples are shown in Fig. 3):

- 180° domain walls with large domain spacing,
- narrowly spaced, more complicated domains,
- faint fine structures,
- zones without visible domain structures,
- bent domain walls,
- varying intensity of Bitter lines, dissolution of Bitter lines near the particle edges and even inside the crystal.

Naturally occurring Ti-rich titanomagnetites show similar domain structure characteristics. The examples of Fig. 4, however, represent domain configurations still of simple form. Normally, extremely complicated domain configurations (Fig. 5) were observed.

All observed features can be explained by an inhomogeneous anisotropy field caused by internal stress. Apart from the width of the domains, similar domain structures have been reported from metallic glasses (e.g. Kronmüller et al., 1979; Salzmann and Hubert, 1981). Metallic glasses are characterized by the absence of magnetocrystalline anisotropy. Stress is believed to be responsible for the local anisotropy.

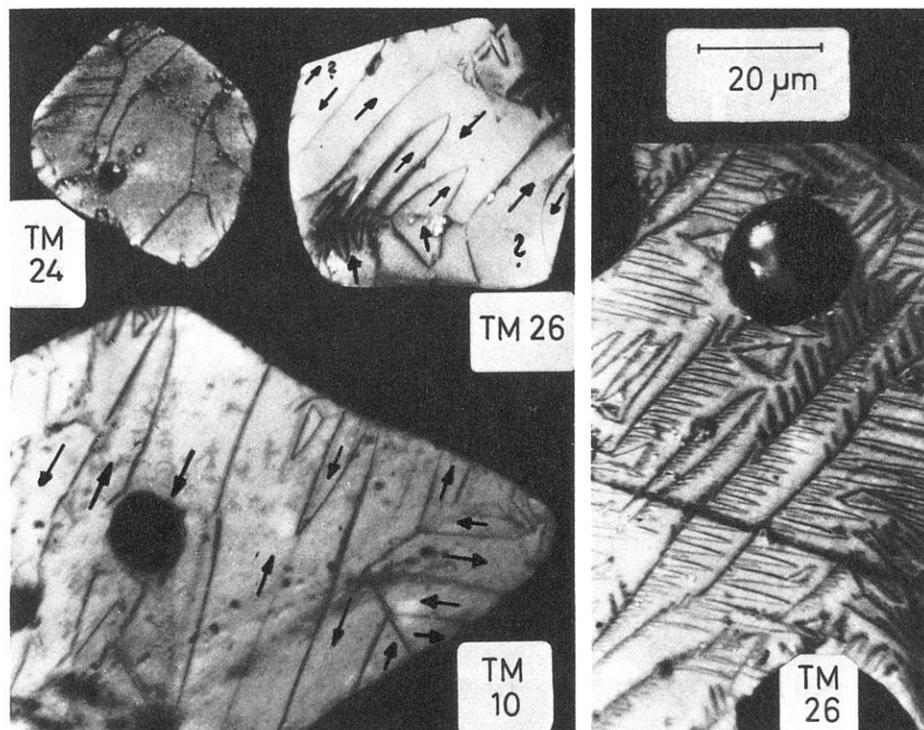


Fig. 2. Domain structures on synthetic polycrystalline Ti-poor titanomagnetites (on surfaces near to (110) plane)

Domain wall behaviour in external magnetic fields

A coil system was attached to the microscope allowing external fields H up to 500 Oe (40,000 A/m). The domain wall displacements were observed directly on small isolated synthetic MD particles showing simple domain configurations (Fig. 6). There are good reasons to assume that these particles are in a nearly strain-free state (see discussion section).

The example of Fig. 7 documents two characteristics:

- With increasing external field, the Bitter line (on the right side of the particle) is getting faint and finally disappears without changing its position. This may be explained by spin rotations inside the domains associated with reduced stray field variations. If the stray field variations are too low, the colloid particles cannot be attracted and the Bitter pattern technique becomes inadequate for the analysis of the real domain structure.

- Small external fields of only 10 Oe (800 A/m) or even less are able to change the domain pattern of the isothermal remanence (IRM) state drastically.

Coercivities of isolated Ti-rich particles were determined from balancing the domain areas while undergoing hysteresis (magnetic field H parallel or antiparallel to the assumed direction of spontaneous magnetization). The following basic suppositions were made for the method:

- Only particles with nearly straight and parallel domain walls (e.g. Fig. 6) were selected for the experiments.
- The observed Bitter lines represent the volume domains.
- The spontaneous magnetization of the domains is homogeneous and antiparallel in neighbouring domains.
- Domain wall displacements are the only magnetization process; spin rotations do not occur (the external magnetic field is parallel to the domain walls).

Assuming that $\Delta A/A = [A - A']/[A + A']$ (where A, A' are the total areas of the domains with antiparallel orienta-

tion of the spontaneous magnetization) is identified with the normalized magnetization of the particle J/J_s (J_s spontaneous magnetization), we get hysteresis curves $\Delta A/A(H)$. An example is given in Fig. 8. Values of coercivity H_c can be determined from these curves and allow at least an estimate of H_c . Nine particles of different grain sizes (8.5–35 μm) of the Ti-rich compositions TM72 and AMTM4/2.5/62 were tested. All values of H_c are below 10 Oe (Fig. 9). They are about 10 times smaller than H_c of TM40 of the same grain size (taken from macroscopic hysteresis measurements of Day [1977]). The difference is probably even more evident in a comparison of specimens of the same composition because the critical diameter of the SD-MD transition is smaller for TM40 than for TM72 (e.g. Butler and Banerjee, 1975).

As the minimum statement we can conclude that Ti-rich titanomagnetite particles with simple domain structure (lamellae-shaped domains) cannot provide high coercivity by domain wall blocking.

Domain structure behaviour under external uniaxial pressure

Earlier domain structure observations on titanomagnetites with applied pressure are known only for pure magnetite (Bogdanov and Vlasov, 1966; Soffel, 1966; Kean et al., 1976). These experiments proved that preferentially orientated domains grow with increasing uniaxial pressure or, if not previously existing, preferentially orientated domains nucleate. No domain rotations were observed up to a pressure of about 850 bar. The amount of internal stress of the Beilby layer was determined to be $\sigma_{\text{pol}} \approx 2.5$ kbar by Soffel (1966).

Pressure experiments of the present work were carried out with synthetic polycrystalline and natural Ti-rich titanomagnetites. A pressure apparatus attached to the microscope was used. The value of σ was measured with a ma-

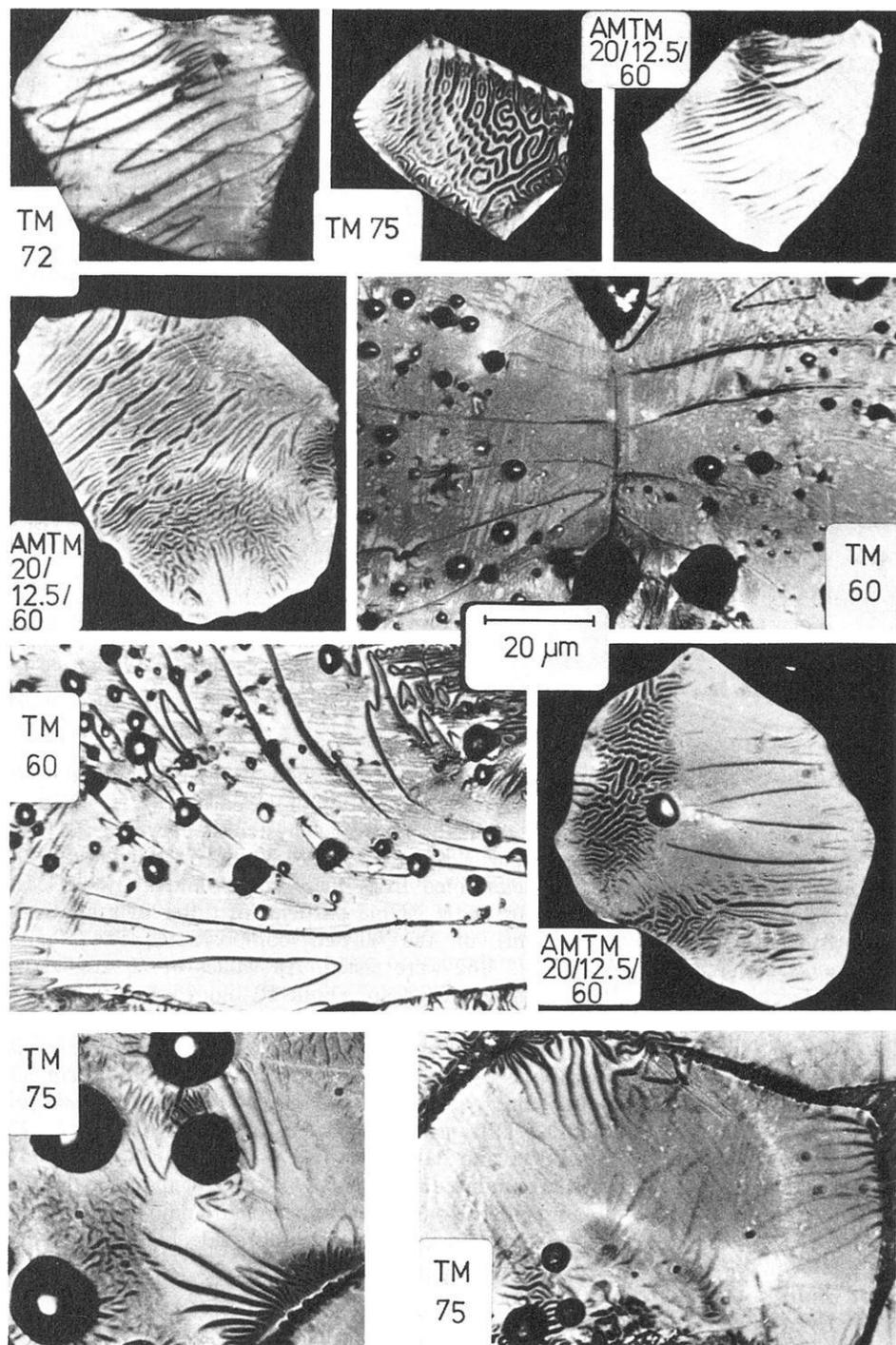


Fig. 3. Domain structures on Ti-rich synthetic polycrystalline titanomagnetites (partially + Al, Mg)

nometer. Problems arise from the porosity of the synthetic polycrystalline samples because the direction of the external pressure σ_a might be changed by the pores.

What kind of transitions of the domain configuration appear with increasing σ_a ? First of all, only domain structures with wide domain spacing (volume parts assumed to be free of high internal stress or high inhomogeneity of stress) will be regarded. Three principal processes are possible (Fig. 10):

— dissolution of the primary domain walls and re-forma-

tion of preferentially orientated domains (spin rotation or reduction of anisotropy?, Fig. 10a),

— domain rotation as a whole (Fig. 10b),

— nucleation of preferentially orientated domains growing with increasing σ_a (Fig. 10c).

For Ti-rich titanomagnetites (and probably for Al, Mg substituted Ti-rich samples too) the magnetostriction constants λ_{111} , λ_{100} , λ_s are positive (Syono, 1965). This means that the preferred orientation of the spontaneous magnetization is perpendicular to the direction of pressure. Nuclea-

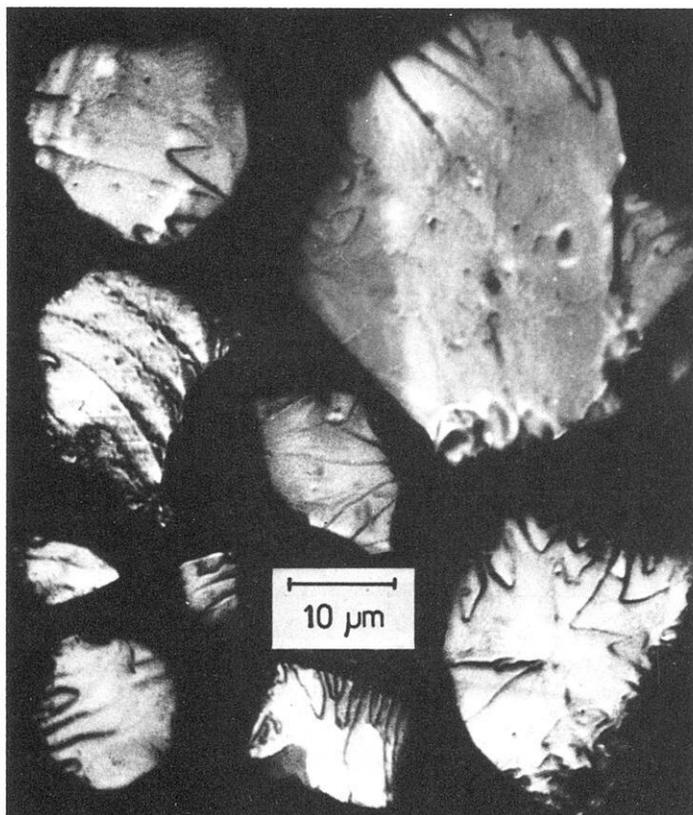


Fig. 4. Domain structures on natural Ti-rich titanomagnetite particles in the Basalt TR of composition $\text{Fe}_{2.26}\text{Ti}_{0.52}\text{Al}_{0.06}\text{Mg}_{0.13}\text{Mn}_{0.03}\text{O}_4$

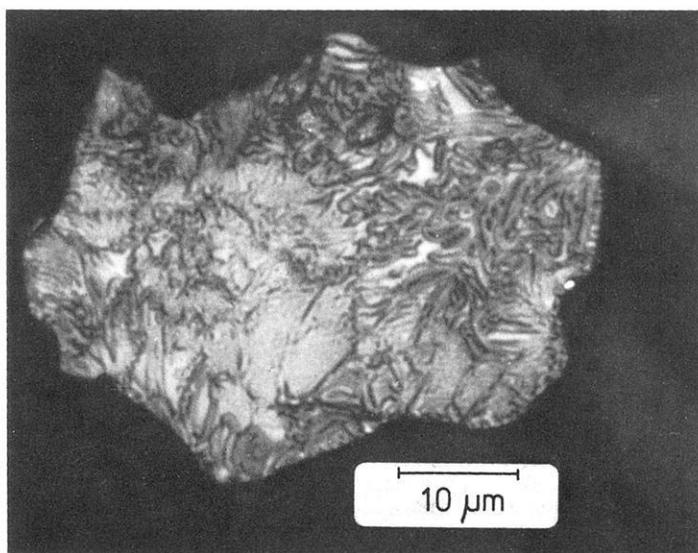


Fig. 5. Typical domain structure on natural Ti-rich titanomagnetite (Basalt TR)

tion was never observed on Ti-rich titanomagnetites, but previously existing preferentially orientated domains grow with increasing σ_a . On the contrary, both dissolution and re-formation of preferentially orientated domains (Fig. 11, left side), as well as domain rotation (Fig. 11, right side), were observed frequently.

Further attention was focussed on the transition of extremely complicated domain patterns induced by an inhomogeneous anisotropy field (presumably caused by internal stress σ_i). Applied uniaxial pressure σ_a superposes the original anisotropy and reduces the inhomogeneity of the result-

ing total intrinsic anisotropy. This will be associated with the transition of the complicated domain structures into simpler ones (if σ_a is high enough). For E_{σ_a} lower than E_i (E_{σ_a} stress anisotropy energy caused by external pressure, E_i intrinsic anisotropy energy), no drastic changes of the domain pattern are expected. For E_{σ_a} exceeding E_i , an almost complete reorganization of the domain structure should occur. An example is shown in Fig. 12. Appel and Soffel (1984) demonstrated the variation of the amount of the intrinsic anisotropy. In most cases, low σ_a ($\sigma_a \approx 0.1$ – 0.2 kbar) changes the domain configuration drastically and

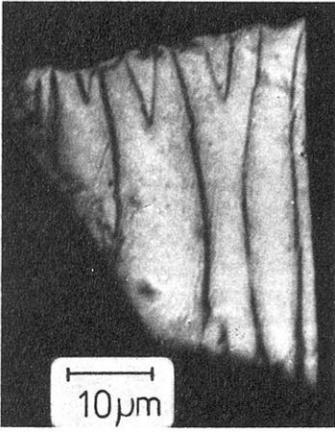


Fig. 6. Domain structure on an isolated synthetic particle of AMTM4/2.5/62

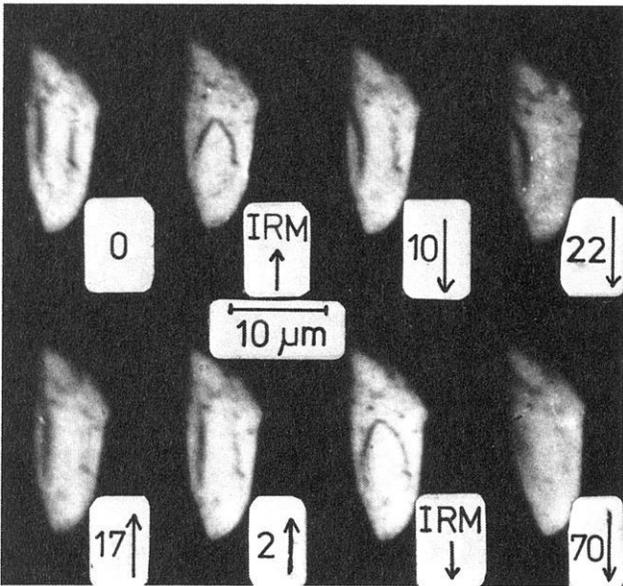


Fig. 7. Domain structures on an isolated synthetic TM72 particle under external magnetic fields H parallel to the observed surface. Values of H are given in Oe

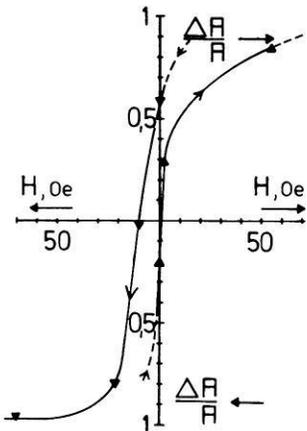


Fig. 8. Hysteresis curve from planimetry of the total domain areas with antiparallel magnetization. \blacktriangle Magnetic field from $H_{\max} \leftarrow$ to $H_{\max} \rightarrow$; \blacktriangledown Magnetic field from $H_{\max} \rightarrow$ to $H_{\max} \leftarrow$

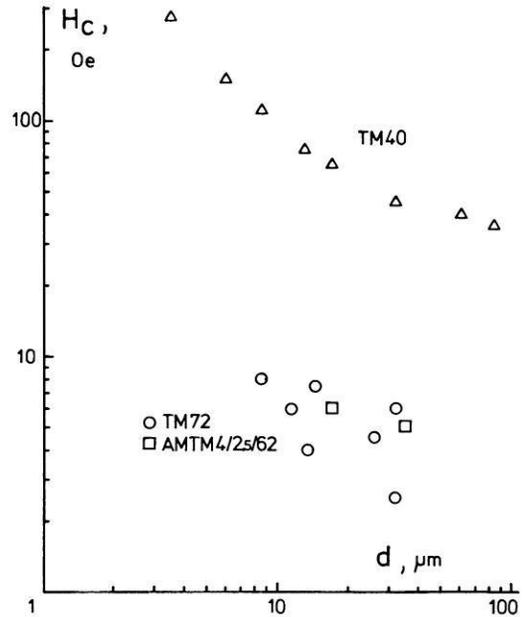


Fig. 9. Coercivity H_c as a function of grain diameter d . Values of TM72 and AMTM4/2.5/62 are from domain area hysteresis, whereas H_c of TM40 are from macroscopic measurements of Day (1977)

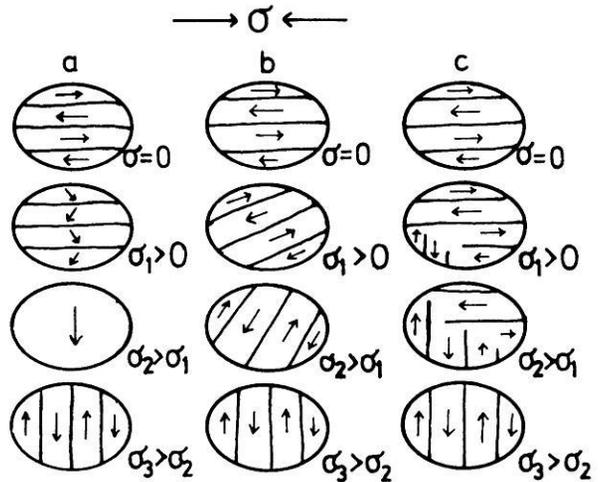


Fig. 10. Possible transitions of domain structures with increasing external uniaxial pressure σ ($\lambda > 0$). The spontaneous magnetization is marked by arrows

only on a minor part of the grain surface does the domain pattern remain complicated up to almost 1 kbar. The essential result is the high stress sensitivity of the material.

On natural Ti-rich titanomagnetites the observed processes are principally the same, but usually higher σ_a values are required to obtain a simple domain configuration. Frequently, $\sigma_a = 1$ kbar does not simplify the complicated domain pattern substantially. The particle shown in Fig. 13 represents an example for presumably low intrinsic anisotropy (lamellae-shaped domain pattern for $\sigma_a > 0.5$ kbar).

Temperature dependence of susceptibility ($\chi - T$)—inference of domain state

$\chi(T)$ curves of synthetic MD titanomagnetites (powders with particle sizes of about 10–100 μm) were measured between liquid nitrogen and Curie temperature using an AC

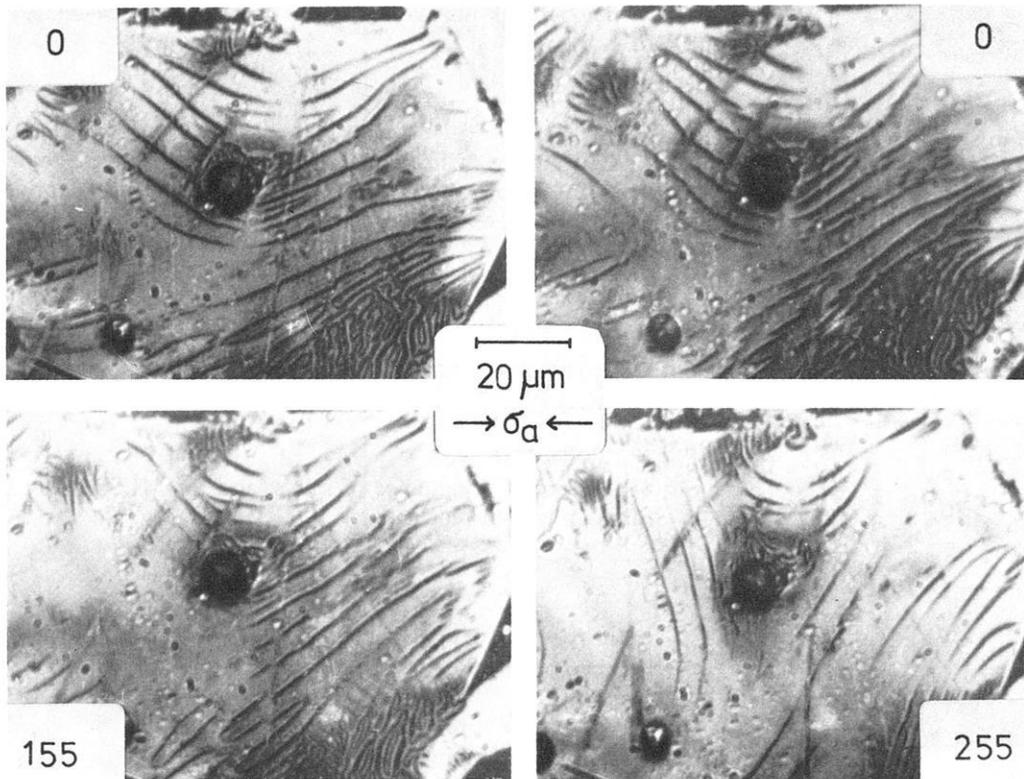


Fig. 11. Domain structure transition under external uniaxial pressure σ_a on synthetic polycrystalline ATM20/62. Values of σ_a are given in bar

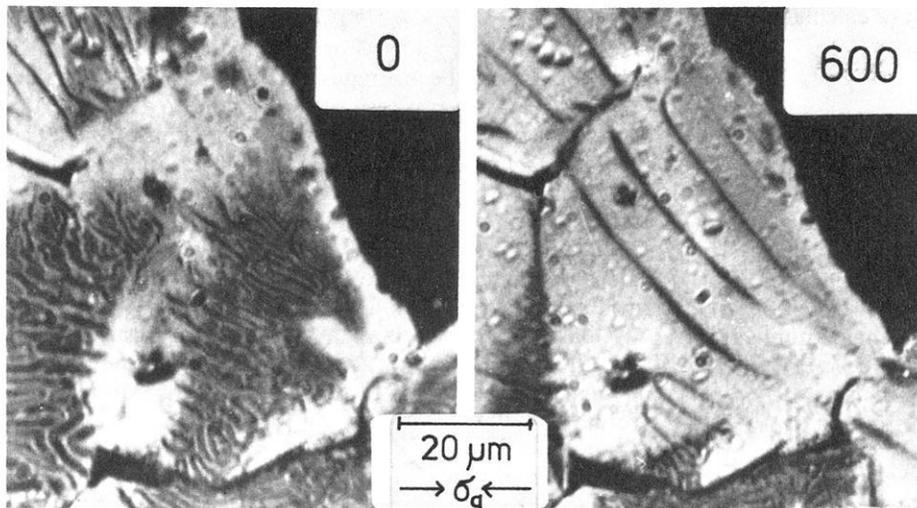


Fig. 12. Domain structure transition of complicated domain configurations into simpler ones by external uniaxial pressure σ_a on synthetic polycrystalline MTM20/62. Values of σ_a are given in bar

bridge with a frequency of 1 kHz and a magnetic field amplitude of 1.5 Oe (120 A/m).

The measured $\chi_0(T)$ curves (Fig. 14) show some systematic variations with the composition of the samples. The first one is the increasing enhancement of the Hopkinson peak with increasing Ti content. However, more important for the inference of the domain state is the peak (or change in the steepness of the curve) at low temperature, which is quite distinct for TM10 and TM24, less pronounced for TM40, only very weak for TM52 and no longer recognizable for the Ti-rich TM60 and TM75 samples. The peaks of TM10, TM40 and TM52 are in rather good agreement

with $T(K_1=0)$ (interpolation of the data of Syono [1965]) and only in the case of TM24 is there a discrepancy (but this is probably a consequence of the insufficiency of the interpolation of Syono's data for TM18 and TM31).

Interpolating Syono's data shows that K_1 of TM60 should be zero at about -30°C . The disappearance of the magnetocrystalline anisotropy, however, is not expressed in the measured $\chi_0(T)$ curve of TM60. For this reason, we suppose that there is no essential influence of the magnetocrystalline anisotropy in Ti-rich (around TM60) MD titanomagnetite particles.

For comparisons of theoretical and experimental $\chi_i(T)$

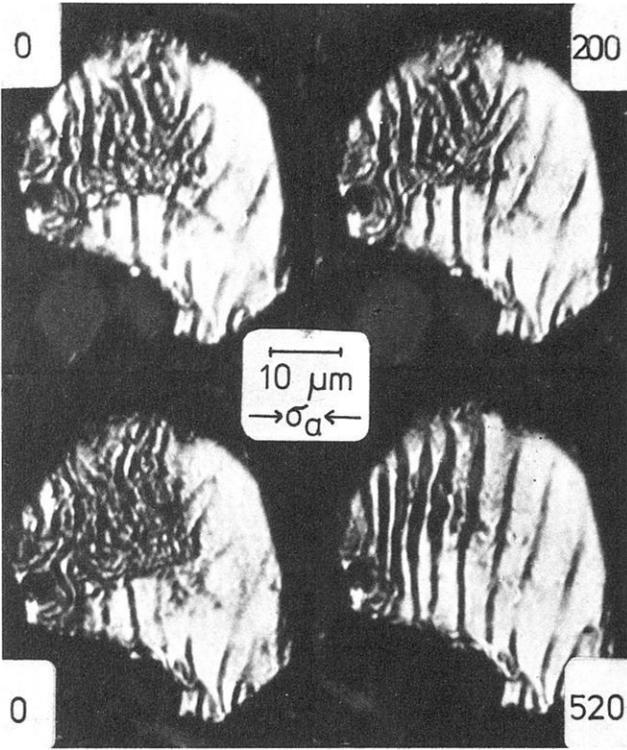


Fig. 13. Domain structure transition under external uniaxial pressure σ_a on natural titanomagnetite $\text{Fe}_{2.26}\text{T}_{0.52}\text{Al}_{0.06}\text{Mg}_{0.13}\text{Mn}_{0.03}\text{O}_4$. Values of σ_a are given in bar

curves, the intrinsic susceptibility χ_i must be calculated from χ_0 with the equation

$$\chi_i = \frac{\chi_0}{1 - N\chi_0}, \quad (1)$$

where N is the demagnetizing factor. The shape of $\chi_i(\chi_0, N) - T$ is strongly dependent on the value of N . Unfortunately, the demagnetizing factor N is unknown. N of MD particles is not only a function of the particle size but strongly dependent on the domain configuration. For MD particles of cubic size, values of N between about 1.6 and $4\pi/3$ are possible (Dunlop, 1983). The maximum value N_{\max} can be derived from the maximum value $\chi_{0, \max}$ of the measured susceptibility with the equation $N_{\max} = 1/\chi_{0, \max}$ ($\chi_i \rightarrow \infty$).

Theoretical normalized $\chi_i(T)$ curves were calculated from the relation

$$\chi_i \sim \frac{J_s^2}{f_1(K_1, \lambda_s) + f_2(N, J_s)}. \quad (2)$$

$f_1(K_1, \lambda_s)$ arises from the magnetocrystalline and magnetostriction anisotropy, $f_2(N, J_s)$ from the shape anisotropy. Neglecting f_2 , we get (Träuble, 1966) for SD processes (spin rotations):

$$\chi_i \sim \frac{J_s^2}{K_1} \quad (3)$$

in the case of dominance of magnetocrystalline anisotropy

$$\chi_i \sim \frac{J_s^2}{\lambda_s} \quad (4)$$

in the case of dominance of stress, and for MD processes (domain wall displacements):

$$\chi_i \sim \frac{J_s^2}{K_1^{7/4}} \quad (5)$$

in the case of dominance of magnetocrystalline anisotropy

$$\chi_i \sim \frac{J_s^2}{\lambda_s K_1^{1/4}} \quad (6)$$

in the case of dominance of stress.

Figures 15 and 16 show the theoretical $\chi_i(T)$ curves for TM52 and TM60. The expected peaks at low temperatures (caused by $K_1 \rightarrow 0$) are partly suppressed by stress and shape anisotropy (see also Clark and Schmidt [1982]). The peak of $\chi_i(T)$ from relation (6) is only weakly pronounced (weak dependence with $\chi_i \sim K^{-1/4}$) and may not be detected in measurements of $\chi_0(T)$. For this reason, SD and MD behaviour cannot be distinguished in the case of dominance of stress. But, in any case, it is possible to diagnose the dominance of stress or magnetocrystalline anisotropy. The theoretical $\chi_i(T)$ curves calculated from relation (4), of both TM52 and TM60, are in good agreement with the $\chi_i(T)$ curves determined from $\chi_0(T)$ with $N=2.0$ (TM60) and $N=2.8$ (TM52), respectively (Figs. 15 and 16). The small values of N are confirmed by the values of $N_{\max} (\approx 3)$ derived from the $\chi_{0, \max}$ values.

Consequently we infer, as the essential result of this section, the dominance of stress for magnetization processes in small magnetic fields (at least in the temperature range around $K_1 \rightarrow 0$).

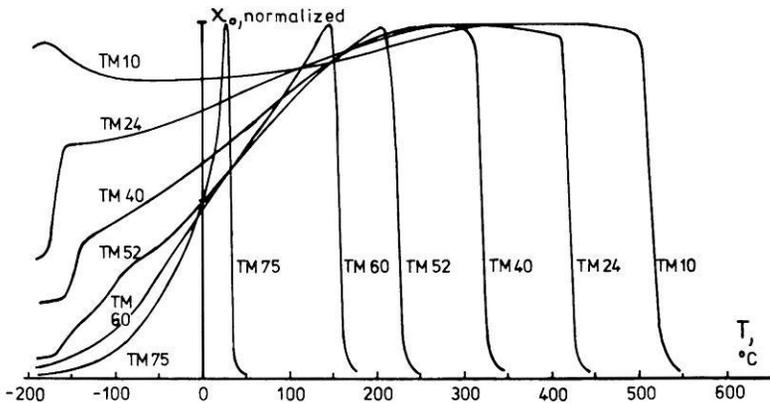


Fig. 14. Temperature dependence of susceptibility χ_0 of synthetic MD titanomagnetites TM10, TM24, TM40, TM52, TM60 and TM75 (grain size 10–100 μm)

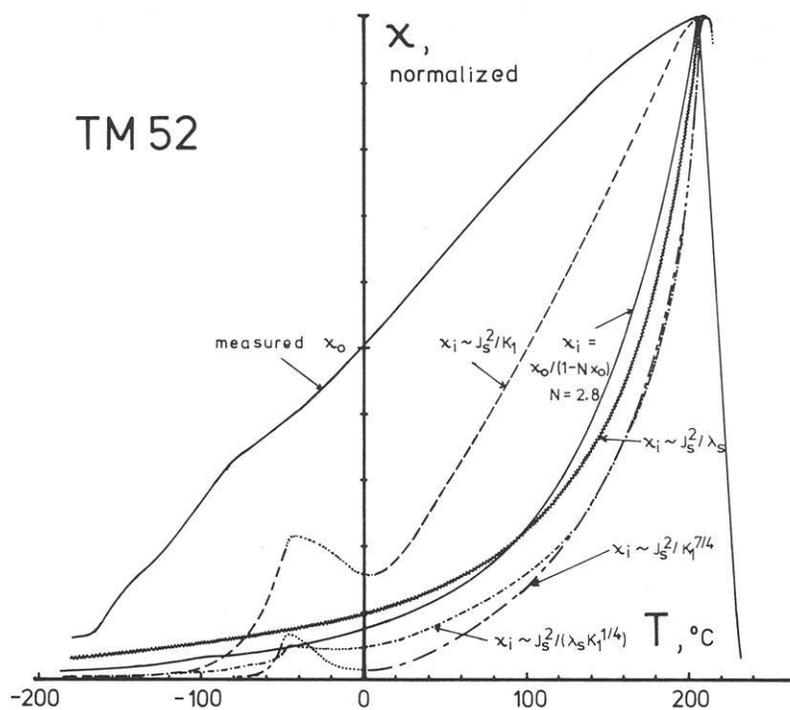


Fig. 15. Theoretical and experimental $\chi_i(T)$ curves of TM52. For details, see text

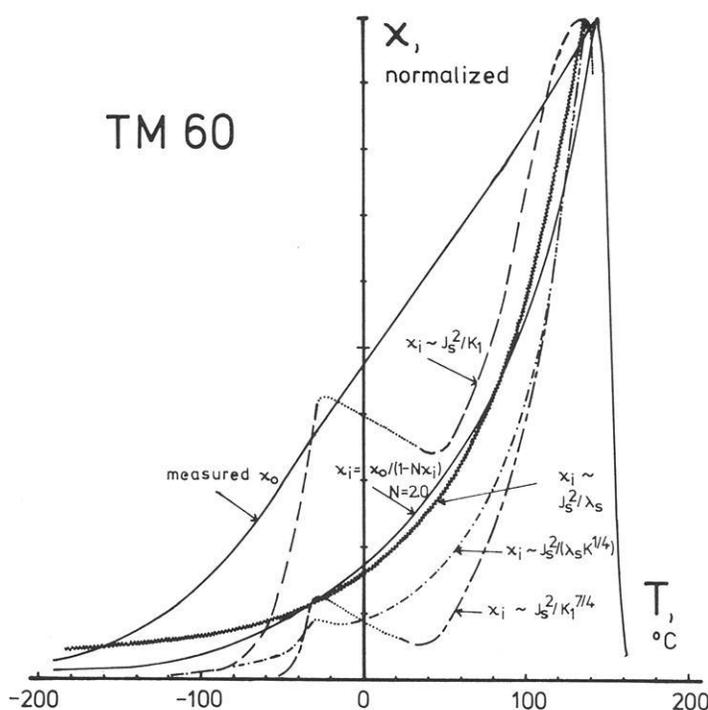


Fig. 16. Theoretical and experimental $\chi_i(T)$ curves of TM60. For details, see text

Comparisons of theoretical and experimental $\chi_i(T)$ curves are not possible for compositions TM10, TM24 and TM40 because $(1 - N\chi_0) \rightarrow 0$ for temperatures near T_c (plateau of $\chi_0 - T$ because of high χ_i -value). For TM75, only a extremely rough extrapolation of λ_s would be possible from Syono's data and no comparison was made for this reason.

Discussion and conclusions

Domain structure observations showed that at room temperature Ti-rich titanomagnetites normally develop a non-

classical domain structure characterized by a very complicated domain pattern. Possible explanations are:

- mineralogical inhomogeneity of the specimens,
- inhomogeneous anisotropy field, causing a locally varying preferential direction of the spontaneous magnetization.

Possible sources of anisotropy are stress, magnetocrystalline and shape anisotropy.

Mineralogical inhomogeneity was not detected by the various analysis methods.

The general occurrence of bent domain walls, the varying intensity of the Bitter lines and the local sensitivity on external stress indicate the existence of an inhomogeneous

anisotropy field, in amount and in direction. Consequently, the preferential direction of the spontaneous magnetization and the specific domain wall energy should fluctuate and prevent the development of a classical domain configuration. Complicated domain patterns will arise from the locally varying anisotropy conditions. Mainly, the spontaneous magnetization should be orientated in the direction defined by the local anisotropy (apart from the domain walls). Depending upon certain anisotropy conditions, the development of true domain structures may even be impossible in part of the crystals.

The most probable cause for the formation of complicated domain patterns is the existence of an inhomogeneous stress field. Of course, the magnetocrystalline and shape anisotropy will not vanish, but superposes the stress anisotropy. However, from the pressure experiments we can derive the dominance of stress because of

- the high difference of sensitivity on external stress between natural and synthetic titanomagnetites (of about the same composition) and
- the usually low sensitivity on external stress in natural Ti-rich titanomagnetites.

Magnetocrystalline anisotropy and anisotropy of magnetostatic origin are limited by material constants. The maximum values of the magnetostatic anisotropy energy E_M and magnetocrystalline anisotropy energy E_K can be calculated from the equations

$$E_K = |K_1|/3 \quad (\text{for cubic anisotropy with } K_1 < 0) \quad (7)$$

$$E_M = 0.5 (N_1 - N_2) J_s^2. \quad (8)$$

Taking values for TM60 (magnetocrystalline constant K_1 from Syono [1965], saturation magnetization J_s from Akimoto et al. [1957]) and $(N_1 - N_2) = 2\pi$, we obtain $E_K \approx 0.1 \times 10^5$ erg/cm³ and $E_M \approx 0.6 \times 10^5$ erg/cm³. Both are too low to explain the frequently observed stability of the complicated domain structures in natural specimens up to 1 kbar because, with the equation

$$E_\sigma = 1.5 |\lambda_s \sigma|, \quad (9)$$

the stress anisotropy energy E_σ becomes several times higher ($\approx 2 \times 10^5$ erg/cm³, λ_s from Syono [1965], $\sigma = 1$ kbar). It seems that anisotropy of magnetostatic origin cannot explain the amount of observed intrinsic anisotropies, although it might contribute considerably to the total anisotropy. Magnetocrystalline anisotropy is even lower.

Considering only E_K and E_M we cannot find a satisfactory explanation for the difference of sensitivity to external stress in natural and synthetic specimens. Strong specific differences of mineralogical or crystalline and geometrical structure would be required. Stress-controlled anisotropy, however, agrees with our observations because E_σ depends not only on material constants but high and low state of intrinsic anisotropy can be easily interpreted by varying σ_i .

Calculations balancing the maximum energies E_K and E_σ (Eqs. 7 and 9) confirm the importance of stress. E_σ exceeds E_K for ≈ 100 bar (290 K) in the case of TM60 (Fig. 17). Stress becomes dominant above this value. Only in the case of very low stress (or more precisely, stress inhomogeneity) a simple domain structure (like Figs. 6 and 7) is possible. In Fig. 18 ($E_K = E_\sigma$) is plotted against temperature for TM56. Above $\approx -50^\circ\text{C}$, stress below 200 bar is sufficient for $E_\sigma > E_K$. With decreasing temperature E_K becomes more and more dominant for TM56.

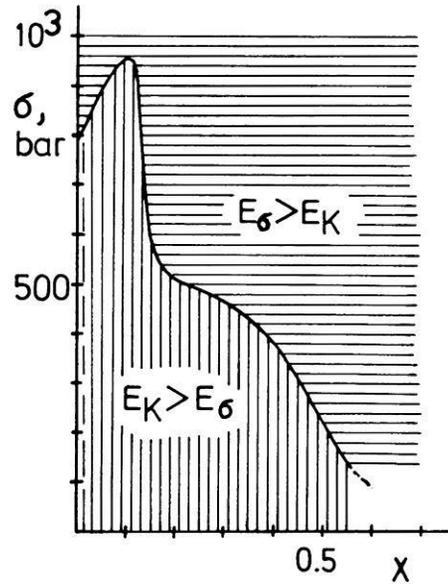


Fig. 17. Required stress σ for $E_K = E_\sigma$ of titanomagnetites (composition x) at 290 K

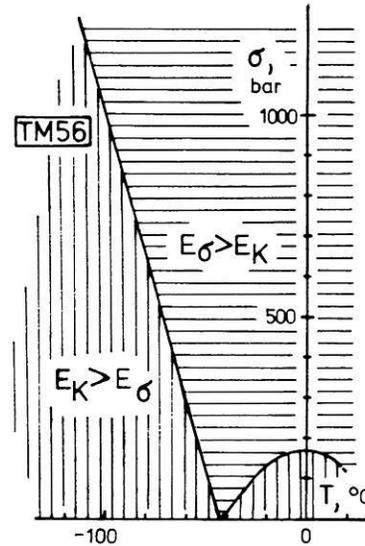


Fig. 18. Required stress σ for $E_K = E_\sigma$ of TM56 at various temperatures

The stress control of magnetization processes in Ti-rich titanomagnetites near room temperature is indicated by the hysteresis experiments—if we identify almost classical domain structures with low internal stress—and by the measurements of temperature dependence of susceptibility. The dominance of stress is also inferred from Hodych (1982a, 1982b) by measuring coercivity. The author finds that H_C is well correlated with the magnetostriction constant λ_s in the case of large (≈ 200 μm) natural MD titanomagnetites ($x \approx 0.6$, 295 K to 190 K) but also for natural MD magnetite (2.7–65 μm , 300 K to 130 K).

The existence of internal stress σ_i up to several hundred bars and the inhomogeneity of the amount and direction of internal stress is plausible. Internal stress might arise from:

- Crystallization pressure: there is only poor knowledge about the amount of crystallization pressure.

– Magnetoelastic energy: when basaltic rock is cooling below the Curie temperature of the titanomagnetites the minerals are already crystallized. Therefore, no free deformation (due to magnetostriction) is possible and internal stress will be produced. The amount of σ_i can be approximately calculated from

$$1.5|\lambda_s\sigma_i| = E'_K(c_{ij}, \lambda) \quad (10)$$

(E'_K magnetoelastic energy, c_{ij} elastic constants). For TM60 a maximum value $\sigma_i \approx 140$ bar is calculated (c_{ij} from interpolation of values for magnetite and ulvospinel, after Hearmon [1956] and Ishikawa and Syono [1971]).

– Vacancies in the spinel lattice caused by oxidation: the cell parameter a_0 decreases with oxidation. Based on a model of an inner sphere with degree of oxidation z and a surrounding shell with $z + \Delta z$, σ_i (radial stress) can be calculated from

$$\sigma_i = 0.03 k \Delta z \quad (11)$$

(k bulk modulus) with $\Delta a_0/a_0 \approx -0.01 \Delta z$ (Readman and O'Reilly, 1972; Moll, 1980). For $\Delta z = 0.01$ (very small value) σ_i becomes 310 bar ($k \approx 10^{12}$ dyn/cm² is calculated from c_{ij}). High values of σ_i from oxidation are confirmed by the occurrence of shrinkage cracks.

– Dislocations: the maximum mean value of σ_i around dislocation can be estimated from

$$\sigma_i = (2Gb)/(\pi d) \quad (12)$$

(G shear modulus, b Burger's vector, d critical diameter for the superparamagnetic–single domain transition). $d = 0.1 \mu\text{m}$ (Butler and Banerjee [1975], for TM60), leads to $\sigma_i \approx 1.1$ kbar ($G \approx 0.6 \times 10^{12}$ dyn/cm² from c_{ij} , $b \approx 3 \times 10^{-8}$ cm). However, the range width of σ_i caused by dislocations is probably too small for creating PSD behaviour.

Crystallization pressure and magnetoelastic stress cause internal stress with inhomogeneous direction because of the irregular shape of the particles, whereas the degree of oxidation itself is locally varying (and therefore causes internal stress with locally varying direction). The consequence is an inhomogeneous direction of anisotropy.

A preliminary simple model for the domain state of Ti-rich titanomagnetites has been published by Appel and Soffel (1984). Following the idea of the model, particles of MD grain size should consist of multidomain, single domain and spin cluster regions. The consequence is the decrease of the "effective grain size" (division into subvolumes) and an enhanced importance of spin rotation processes. The micromagnetic state will be comparable with the conditions in magnetic sponges (cold-pressed and partially sintered magnetic powders: e.g. Weil [1953]).

Acknowledgements. We wish to thank Dipl.-Min. W. Bloch for the analysis of the natural titanomagnetite sample and Prof. A. Schult for his help. The numerous constructive remarks of the reviewers are also gratefully acknowledged. Furthermore, we thank the Deutsche Forschungsgemeinschaft for the financial support of this research.

References

- Akimoto, S.T., Katsura, T., Yoshida, M.: Magnetic properties of Ti–Fe₂O₄–Fe₃O₄ system and their change with oxidation. *J. Geomagn. Geoelectr.* **9**, 165–178, 1957
- Appel, E., Moll, A.: Synthese von reinen und Aluminium- sowie Magnesiumdotierten Titanomagnetiten und Messungen charakteristischer Eigenschaften an diesem Material. Dipl.-Arbeit Teil I, Inst. Allg. Angew. Geophys. Univ. München, 1980
- Appel, E., Soffel, H.C.: Model for the domain state of Ti-rich titanomagnetites. *Geophys. Res. Lett.* **3**, 189–192, 1984
- Banerjee, S.K.: On the origin of stable remanence in pseudo-single domain grains. *J. Geomagn. Geoelectr.* **29**, 319–329, 1977
- Bogdanov, A.A., Vlasov, A.Y.: On the effect of elastic stresses on the domain structure of magnetite. *Izv. Earth Phys.* **1**, 42–46, 1966
- Butler, R.F., Banerjee, S.K.: Theoretical single-domain grain-size range in magnetite and titanomagnetite. *J. Geophys. Res.* **80**, 4049–4058, 1975
- Clark, D.A., Schmidt, P.W.: Theoretical analysis of thermomagnetic properties, low-temperature hysteresis and domain structure of titanomagnetites. *Phys. Earth Planet. Inter.* **30**, 300–316, 1982
- Day, R.: TRM and its variation with grains size. *J. Geomagn. Geoelectr.* **29**, 233–265, 1977
- Dunlop, D.J.: The hunting of the 'Psark'. *J. Geomagn. Geoelectr.* **29**, 293–318, 1977
- Dunlop, D.J.: On the demagnetizing energy and demagnetizing factor of a multidomain ferromagnetic cube. *Geophys. Res. Lett.* **10**, 79–82, 1983
- Halgedahl, S., Fuller, M.: Magnetic domain observation of nucleation processes in fine particles of intermediate titanomagnetite. *Nature* **288**, 70–72, 1980
- Hearmon, R.F.S.: The elastic constants of anisotropic materials. *Adv. Phys.* **5**, 323, 1956
- Hodych, J.P.: Magnetic hysteresis as a function of low temperature for deep-sea basalts containing large titanomagnetite grains – inference of domain state and controls of coercivity. *Can. J. Earth Sci.* **19**, 144–152, 1982a
- Hodych, J.P.: Magnetostrictive control of coercive force in multidomain magnetite. *Nature* **298**, 542–544, 1982b
- Ishikawa, Y., Syono, Y.: Giant magnetostriction due to Jahn-Teller distortion in Fe₂TiO₄. *Phys. Rev. Lett.* **26**, 1335–1338, 1971
- Kean, W.F., Day, R., Fuller, M., Schmidt, V.A.: The effect of uniaxial compression on the initial susceptibility of rocks as a function of grain size and composition of their constituent titanomagnetites. *J. Geophys. Res.* **81**, 861–872, 1976
- Kronmüller, H., Fähnle, M., Domann, M., Grimm, H., Grimm, R., Gröger, B.: Magnetic properties of amorphous ferromagnetic alloys. *J. Magn. Magn. Mat.* **13**, 53–70, 1979
- Moll, A.: Untersuchung der Tieftemperatur-Oxidation von synthetischen – z.T. mit Aluminium und Magnesium dotierten – Titanomagnetiten und Messung charakteristischer Eigenschaften der Oxidationsprodukte. Dipl.-Arbeit Teil II, Inst. Allg. Angew. Geophys. Univ. München, 1980
- Néel, L.: Théorie du trainage magnétique des ferromagnétiques en grains fins avec applications aux terres cuites. *Ann. Géophys.* **5**, 99–136, 1949
- Ozima, M., Ozima, M.: Origin of thermoremanent magnetization. *J. Geophys. Res.* **70**, 1363–1369, 1965
- Petersen, N., Eisenach, P., Bleil, U.: Low temperature alteration of magnetic minerals in ocean floor basalts. Maurice Ewing Series 2, M. Talwani, C.G. Harrison, D.E. Hayes, eds.: pp 169–209. Am. Geophys. Union, Washington D.C., 1979
- Radhakrishnamurty, C., Likhite, S.D., Deutsch, E.R., Murthy, G.S.: On the complex magnetic behaviour of titanomagnetites. *Phys. Earth Planet. Inter.* **30**, 281–290, 1982
- Readman, P.W., O'Reilly, W.O.: Magnetic properties of oxidized (cation-deficient) titanomagnetites (Fe,Ti,□)₃O₄. *J. Geomagn. Geoelectr.* **24**, 69–80, 1972
- Salzmann, P., Hubert, A.: Local measurement of magnetic anisotropy in metallic glasses. *J. Magn. Magn. Min.* **24**, 168–174, 1981
- Soffel, H.C.: Stress dependence of the domain structure of natural magnetite. *Z. Geophys.* **32**, 63–77, 1966
- Soffel, H.C., Petersen, N.: Ionic etching of titanomagnetite grains in basalts. *Earth Planet. Sci. Lett.* **11**, 312–316, 1971

- Soffel, H.C., Deutsch, E.R., Appel, E., Eisenach, P., Petersen, N.: The domain structure of synthetic stoichiometric TM10-TM75 and Al-, Mg-, Mn-, V-doped TM62 titanomagnetites. *Phys. Earth Planet. Inter.* **30**, 336–346, 1982
- Stacey, F.D.: The physical theory of rock magnetism. *Adv. Phys.* **12**, 45–133, 1963
- Syono, Y.: Magnetocrystalline anisotropy and magnetostriction of Fe_3O_4 – Fe_2TiO_4 series with special application to rock magnetism. *Jap. J. Geophys.* **4**, 71–143, 1965
- Träuble, H.: *Moderne Probleme der Metallphysik*. Bd. II, A. Seeger, ed.: pp 420–459. Heidelberg: Springer 1966
- Verhoogen, J.: The origin of thermoremanent magnetization. *J. Geophys. Res.* **64**, 2441–2449, 1959
- Weil, L.: The texture of fine ferromagnetic powders. *Rev. Mod. Phys.* **25**, 324–326, 1953
- Winhard, H.: *Bestimmung der Art der Platznahme von Basalten mit Hilfe der Anisotropie der magnetischen Suszeptibilität*. Thesis, Inst. Allg. Angew. Geophys., Univ. Munich, 1983

Received May 2, 1984; Revised August 17, 1984

Accepted October 31, 1984

Large-scale studies of Pi-2's associated with auroral breakups

J.C. Samson

Institute of Earth and Planetary Physics, Department of Physics, University of Alberta, Edmonton, Alberta, Canada T6G 2J1

Abstract. In the interval 22–27 October 1979, an observational campaign was conducted to study the correlation of substorm-triggered geomagnetic pulsations (Pi 2's) and auroral breakups. This campaign took advantage of the North American magnetometer networks which were operated during the International Magnetospheric Study. Data from 28 magnetometer stations ranging from 73.9°–38.9° N (centred dipole) were used in this study, 20 of which yielded high time resolution data which permitted proper analysis of Pi 2's.

An analysis of two substorm onsets indicates that the Pi 2's occurred in conjunction with the brightening of a quiet arc, near the equatorward border of visible auroral activity. One breakup occurred within the region of the ambient, eastward ionospheric electrojet, suggesting that the onset of the Pi 2's and the substorm current systems can occur equatorward of the Harang discontinuity. Plots of the intensities and polarizations of the Pi 2's indicate that a substantial part of the Pi 2's magnetic field comes from field-aligned and ionospheric currents associated with the substorm onset.

A comparison of the polarizations of the Pi 2's with those computed from a simple model of ionospheric and field-aligned currents indicates that the Pi 2's in this study might be caused by the motion of a transient current system with a westward velocity of 20–50 km/s.

Key words: Aurora – Substorm – Pi 2 pulsation – Polarization – Electrojets

Introduction

Pi 2 geomagnetic pulsations are directly connected with auroral brightening and the onset of polar magnetic substorms (Troitskaya and Gul'elmi, 1967; Afanasyeva et al., 1970; Pytte and Trefall, 1972; Pytte et al., 1976). At the onset of the substorm, expansive phase enhanced field-aligned currents (FAC) flow into and out of the auroral ionosphere. Although changes in this FAC propagate, in part, as shear Alfvén waves (e.g. Mallinckrodt and Carlson, 1978), much of the upward current over an auroral arc is carried by electrons with energies of several keV (Vondrak, 1975) and, consequently, some of the magnetic transients associated

with Pi 2's might be produced by precipitating high-energy electrons.

In this study, I would like to show that the morphology of Pi 2 magnetic fields is dominated by FACs and ionospheric currents near brightening arcs. In an attempt to explain the data I shall also propose a simple model for the Pi 2 currents near the brightening arc. This heuristic model gives a reasonable prediction of the Pi 2 polarizations and fields, suggesting that the Pi 2's might be associated, in part, with the currents (and possibly high-energy electrons) which cause the arc brightening.

There is already considerable evidence that even mid- and low-latitude Pi 2's are a direct result of high-latitude FACs, with two dominant FAC regions, associated in part with the substorm current wedge (Björnsson et al., 1971; Samson and Harrold, 1983; Lester et al., 1983). The above studies are either statistical in nature or limited to one geomagnetic region (e.g. midlatitudes). The study presented here gives relatively complete coverage of high- and mid-latitude Pi 2's for individual substorm expansive phase onsets and arc brightenings.

The experiment

The auroral campaign was conducted in the interval 22–27 October 1979. The campaign was designed to take advantage of the data from the IMS magnetometer array and from the University of Alberta magnetometer array before it ceased operation in late autumn, 1979. These data were complemented by magnetometer data from the Air Force Geophysics Laboratory (AFGL) stations in the United States. Because of data transmission and noise problems, only nine IMS stations yielded data which were suitable for the analysis of substorm electrojets and Pi 2 polarizations. Data from seven standard Canadian observatories were also used to determine the configuration of the substorm electrojets, but not the polarizations of the Pi 2's.

The centred dipole coordinates (Hakura, 1965) of all the stations and observatories used in this study are given in Table 1. Figure 1 presents a map of the centred dipole coordinates of the 20 stations used to determine the spectral characteristics and polarizations of the Pi 2's.

Table 1. Coordinates of stations and observatories

Code	Station	Centred dipole coordinates	
		North	East
University of Alberta stations			
PROV	Fort Providence	67.5	292.2
SMIT	Fort Smith	67.3	300.1
LEDU	Leduc	60.6	303.0
FTCH	Fort Chipewyan	66.1	301.0
Standard Canadian Observatories			
	Victoria	54.3	294.2
	Meanook	61.9	301.1
	Yellowknife	69.1	294.6
	Cambridge Bay	76.8	296.8
	Great Whale River	66.6	348.8
	Ottawa	56.7	352.8
	Baker Lake	59.9	325.3
	Fort Churchill	68.7	322.8
AFGL stations			
NEW	Newport	55.2	299.3
RPC	Rapid City	53.3	319.0
CDS	Camp Douglas	54.5	334.5
MCL	Mt. Clemens	53.6	343.9
SUB	Sudbury	53.6	358.4
LOC	Lompoc	41.4	302.1
TPA	Tampa	38.9	345.8
IMS stations			
CPY	Cape Perry	73.9	274.1
AVI	Arctic Village	68.1	284.7
COL	College	64.8	282.9
TLK	Talkeetna	63.0	283.1
GIM	Gilliam	66.3	325.3
NOW	Norman Wells	69.2	279.6
FSP	Fort Simpson	67.2	287.7
TUC	Tucson	40.5	313.3
BOU	Boulder	48.7	317.9

Throughout the discussion, coordinates are centred dipole coordinates, and time is universal time (UT), unless otherwise indicated.

A manually operated, wide-angle camera was located at Fort Smith for the auroral photography. This camera had a field of view of 105° and exposure time of 5–20 s. In addition, an observer's log was kept of the observations at this site.

Reduction of data

All the magnetometer data from the Alberta, IMS and AFGL array were recorded in digital form and then sampled at a 10 s interval (each channel) for this study, even though the recorded data often had a smaller sample interval. This sample interval (10 s) gives a Nyquist frequency of 50 mHz which is well above Pi 2 frequencies (5–15 mHz). All the data presented in this study have been rotated to correspond to centred dipole coordinates, with H magnetic north (centred dipole), D magnetic east, and Z downward.

Five representations of the magnetic field data were used in this study. These representations include:

- standard magnetograms
- high-pass-filtered-(5 mHz) magnetograms
- latitude profiles
- difference-equivalent currents and
- spectral representations with estimates of power spectra and polarization parameters.

The standard magnetograms show the recorded magnetic fields in H, D, Z (centred dipole) coordinates, with *no absolute baseline*. The high-pass-filtered data are useful for locating Pi 2's in the data (low-frequency substorm fields are removed), and these data are also used as input to the program which determines the spectral representation.

Both the latitude profiles and the equivalent currents used difference values of the magnetic field in order to estimate the configurations of the substorm currents. The latitude profiles were computed only near the Alberta array, and used the stations NEW, LEDU, FTCH, SMIT as well as the standard observatories at Meanook, Yellowknife and Cambridge Bay. For these profiles, a quiet-time baseline (H, D and Z) was subtracted from the fields recorded during a substorm or substorms. The difference values were then plotted as a function of latitude. These profiles are often easy to interpret as the ionospheric currents typically flow east-west. However, near the front of the westward-travelling surge (WTS), the profiles can be difficult to interpret. Note also that since these profiles used a quiet-time baseline, the inferred currents contain other existing current systems in addition to the desired substorm currents.

Conversely, the equivalent current representations used in this study attempted to isolate the currents for single substorms. Difference values of the magnetic fields in H and D were computed by subtracting base values from 2 min before the substorm onset (as determined by the Pi 2 onset) from field values during the substorm. The vector perturbations were then rotated 90° clockwise (viewed downward) in order to estimate equivalent "sheet" currents flowing in the ionosphere.

The polarization characteristics of the Pi 2's were computed in the spectral (Fourier transform) domain because of the band limited nature of these pulsations (see the examples to be given below). Also, since the magnetic data are typically noisy, the polarization parameters were calculated using statistical methods to determine estimators. Complete details are given in Samson and Harrold (1983) or Samson (1983a). The parameters presented in this study are:

- The power spectra $\mathbf{z}^\dagger(j)\mathbf{z}(j)$, where

$$\mathbf{z}(j) = \sum_{t=0}^{N-1} \mathbf{x}(t) \exp(-2\pi i j t N^{-1})$$

and

$$\mathbf{x}^\top(t) = [H(t), D(t), Z(t)].$$

The power spectra are the sums of the frequency dependent power on all three components of the magnetometer.

- The pure state power spectra, $\hat{a}^2(k)\hat{\mathbf{u}}^\dagger(k)\hat{\mathbf{u}}(k)$ (Samson and Harrold, 1983, Eqs. A7 and A8). The

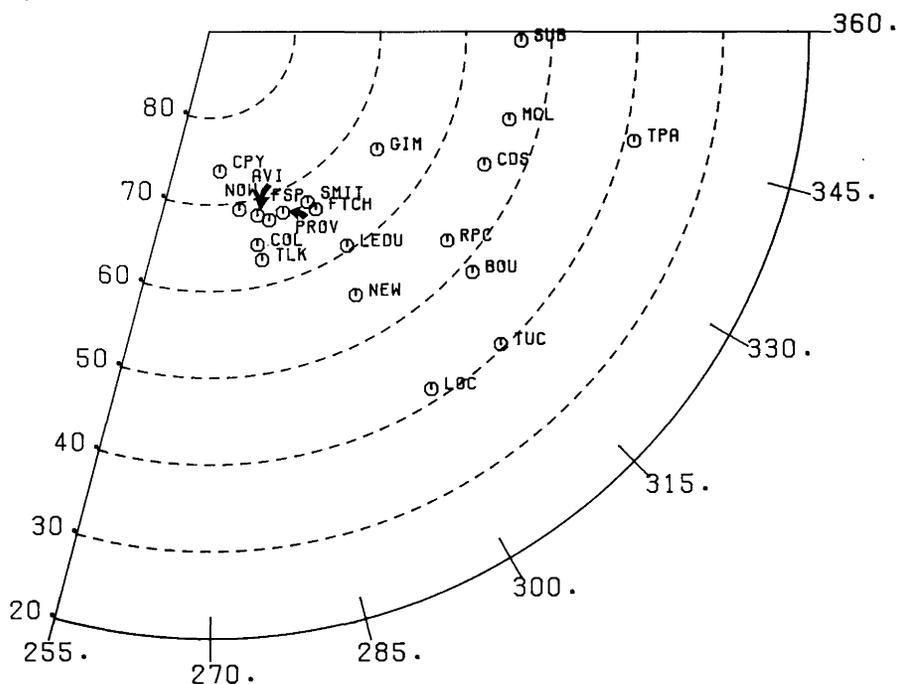


Fig. 1. A map of the positions of the magnetic stations used in this study. Standard observatories are not included. Centred dipole coordinates are used

pure state power spectra estimate the power in the “coherent” waves like Pi 2’s, with “noise” power removed.

c) The degree of polarization estimator \hat{B}_1^2 (Samson and Harrold, 1983, Eq. A6).

d) The parameters of the polarization ellipse in the horizontal ($H-D$) plane, including the ellipticity (ratio minor axis to major axis) and orientation of the polarization ellipse (direction of major axis) [Samson and Harrold, Eqs. (1), (2) and (3)]. The polarization ellipses are estimated from the spectra and cross spectra (of the pure states).

Note that the power spectra a) are unsmoothed. The estimators of the pure states were calculated using a spectral window with seven degrees of freedom (complex Wishart sense). If the signal is a pure state or totally polarized, then $\hat{B}_1^2=1$. Conversely, if data comprise isotropic noise, then $\xi\{\hat{B}_1^2\}\approx 0.06$ (Samson, 1983b), where ξ denotes expectation.

In processing the auroral photographs, both slides and enlarged prints were made. The borders of the aurorae were digitized by tracing over the photographs on a flat bed digitizer. These data were then mapped to centred dipole coordinates by assuming the height of the auroral luminosity to be 100 km.

Observations

Auroral activity

The nights of October 22, 23 and October 23, 24 had considerable cloud cover at Fort Smith (SMIT) and consequently no magnetic data from these nights were analysed. The night of October 24, 25 did, however, yield an excellent series of photographs of auroral breakups. The descriptions which follow are based on the observer’s log and the photographs from SMIT.

The first visible aurora occurred at $\sim 2:30$ UT October 25 (20:24 local time, October 24), with a long arc

north of zenith (at SMIT), running east-west. The growth of this arc was followed by patchy brightening to the east. At 2:38 a faint arc formed near the southern horizon. By 3:12 the northern arc had moved overhead. Between 3:12 and $\sim 6:32$ the arc remained approximately overhead with variable intensity and some brightening and structure.

At $\sim 6:32$ a diffuse arc began forming near the southern horizon. This arc brightened considerably at 6:33 and then expanded northward, indicating the beginning of an auroral breakup. Unfortunately, this event was out of the field of view of the camera. However, the magnetic data showed a clear substorm onset and consequently Pi 2 data for this event were analysed.

By 6:56 the auroral activity had died down and only a quiet arc could be seen near the southern horizon. At $\sim 7:04$ the southern arc began to brighten again, and a second breakup followed. A mapping of the brightened arc is given in Fig. 2. As shown later, this breakup event started near 60° – 61° N and by 7:15 the auroral activity had expanded to $\sim 65^\circ$ N. A note of caution should be included here, however, as there was also considerable auroral activity far to the west after the beginning of the breakup. This breakup seems to be only part of a much larger increase in auroral and magnetic activity.

The rest of the night had variable auroral activity, most consisting of patchy and structured arcs overhead. No distinct breakup events were seen.

Magnetic activity

Selected magnetograms from Newport (NEW) and the Alberta array for the intervals near the two breakups discussed above are given in Fig. 3. Inspection of the trace at LEDU shows clear negative H excursions at both 6:33 and 7:04 (arrows) in conjunction with the auroral breakups. This indicates that the Alberta array

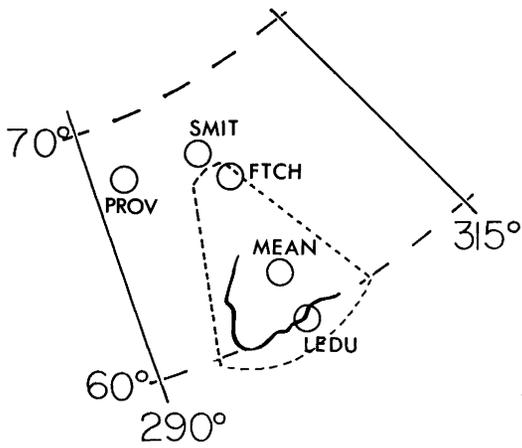


Fig. 2. A map in centred dipole coordinates of the position of the border of the aurora at 7:15 UT. The dotted line indicates the field of view of the camera

was situated over the substorm-enhanced westward electrojet, to the east of any WTS. The negative Z component at LEDU and the positive Z component at FTCH indicate that both breakups occurred to the north of LEDU but south of FTCH (see the latitude profiles to be presented later).

Before continuing in a discussion of the two events, some comment should be made about the ionospheric electrojets before the substorm onsets. A latitude profile from the Alberta array and nearby observatories for 6:00 is given in Fig. 4. To the south of approximately 64° N, all the H components are positive, indicating the presence of an *eastward* electrojet flowing south of a large westward electrojet (negative H). Near the boundary between these electrojets there is a level shift of about 90 nT in the D component. This level shift might

be caused by upward FAC. This upward FAC and the transition from eastward to westward electrojets are all signatures of the Harang discontinuity (Rostoker et al., 1975). Consequently, the southern arcs mentioned in the auroral observations were south of the Harang discontinuity.

The eastward electrojet remained stable until the breakup event at 6:33 (note the stable H component at LEDU in Fig. 3). Consequently, this auroral breakup appears to have begun just south of the Harang discontinuity, driving a wedge of westward electrojet into the region south of the discontinuity and then moving the discontinuity westward. These observations are consistent with those of Baumjohann et al. (1981).

Figure 5 (top) shows the equivalent currents for the first breakup event. A strong electrojet extended from $\sim 345^\circ$ to 300° E at $\sim 61^\circ$ – 62° N. These equivalent currents indicate that the event was fairly well localized near the Alberta array.

The equivalent currents for the second event (Fig. 5, bottom) show that the electrojets in this event were far more widespread, with the largest electrojet strengths at the westernmost stations near $\sim 68^\circ$ N (particularly AVI). The westward electrojet was centred near 61° – 62° N at 300° E, shifting northward to $\sim 68^\circ$ N at $\sim 275^\circ$ E. This event also had considerably more structure in the directions of the equivalent currents, with some suggestion of a developing eastward electrojet to the south of the strong westward electrojet.

Latitude profiles at 7:03 and 7:14 (Fig. 6) show that a well-defined westward electrojet was associated with the second breakup near the Alberta array. A slight D level shift near the northern border of the westward electrojet at 7:03 indicates that these stations were near a region of *downward* FAC, in contrast to the upward FAC at 6:00 (Fig. 4). These features, and auroral obser-

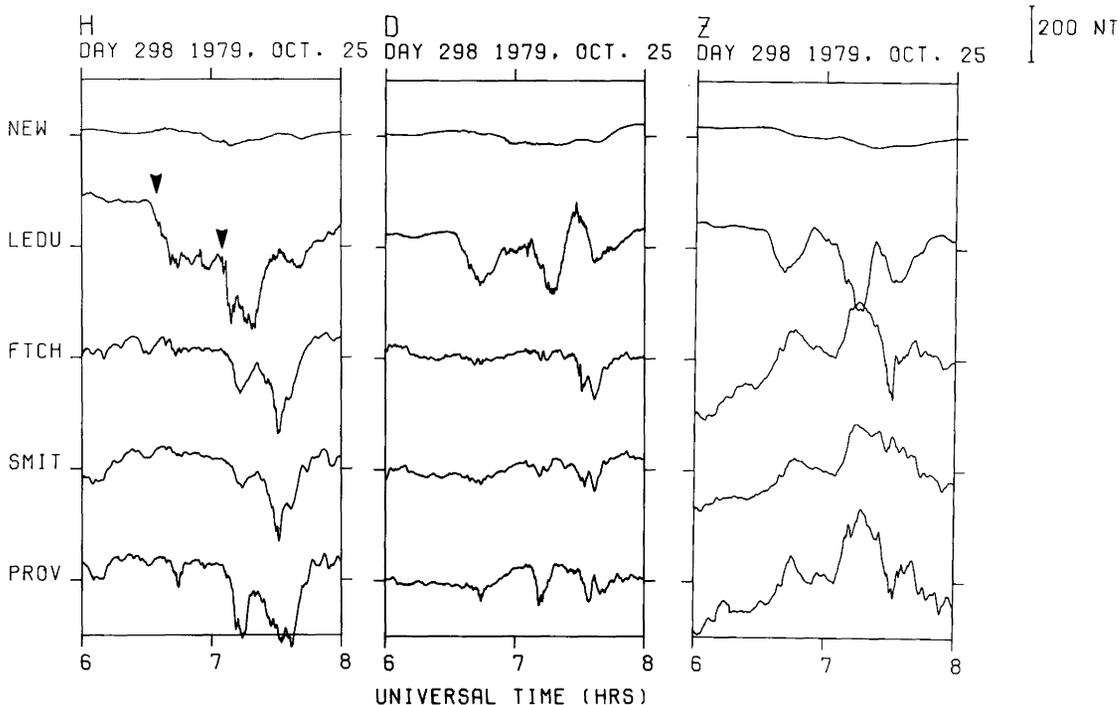


Fig. 3. Magnetograms from the Alberta array and NEW. Centred dipole coordinates are H – magnetic north, D – magnetic east and Z – downward

DAY298/79 600: 0 UT

BASE TIME 2000

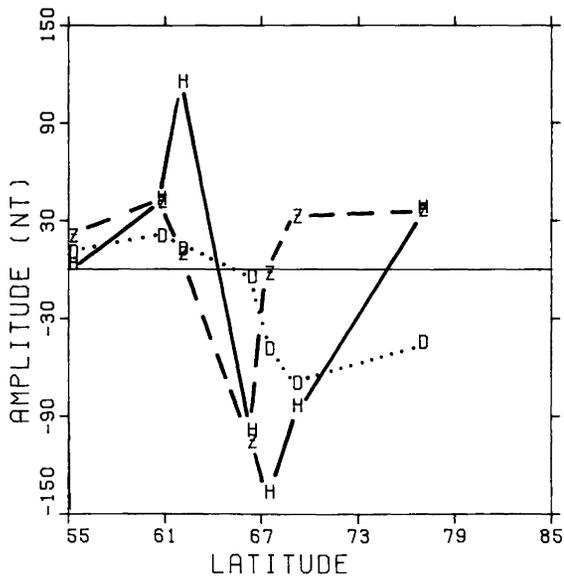
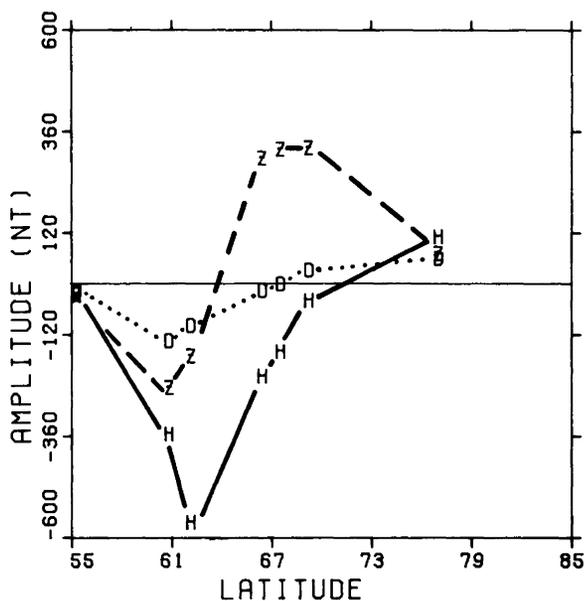


Fig. 4. Latitude profiles at 06:00 October 25. Base time was 20:00 October 24

DAY298/79 714: 0 UT

BASE TIME 2000



DAY298/79 703: 0 UT

BASE TIME 2000

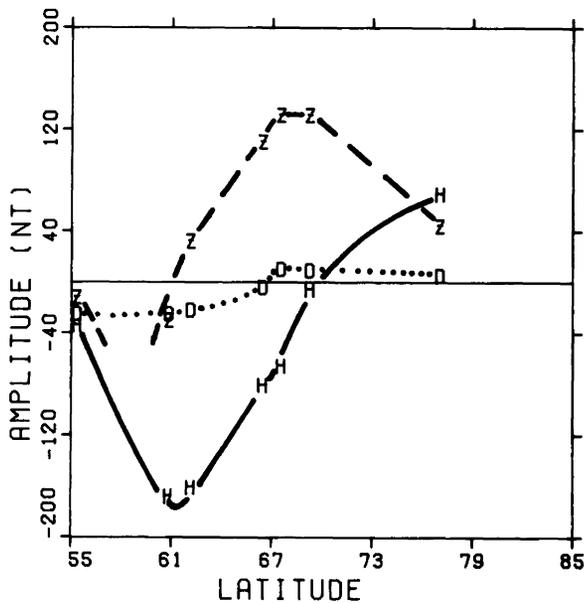


Fig. 6. Latitude profiles at 7:03 and 7:14

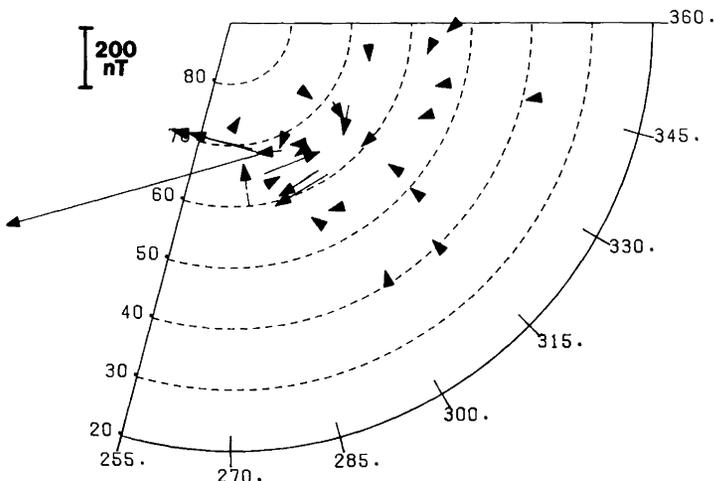
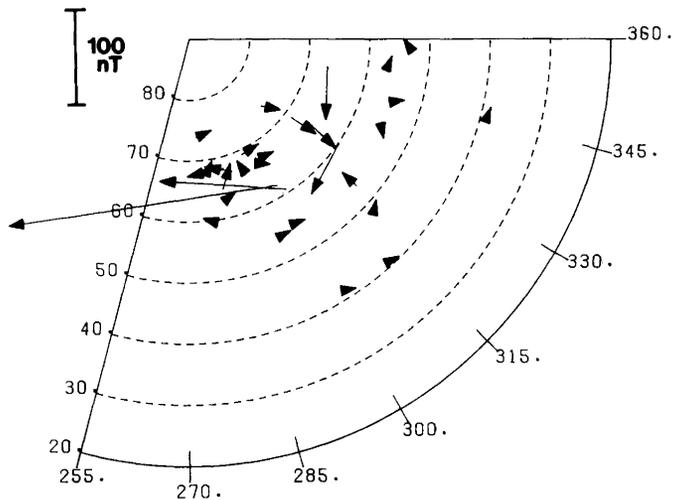


Fig. 5. Difference equivalent currents at 6:36 (top) and 7:08 (bottom)

ations, suggest that the centre of the breakup activity may have been considerably to the west (see Fig. 2 in Samson and Rostoker, 1983).

Even though the auroral map (Fig. 2) shows a large north-south feature to the west of the Alberta array, these latitude profiles are compatible with an electrojet flowing almost directly westward and centred near 60°-63°. Thus the magnetic fields at the Alberta array appear to be entirely due to currents flowing near the

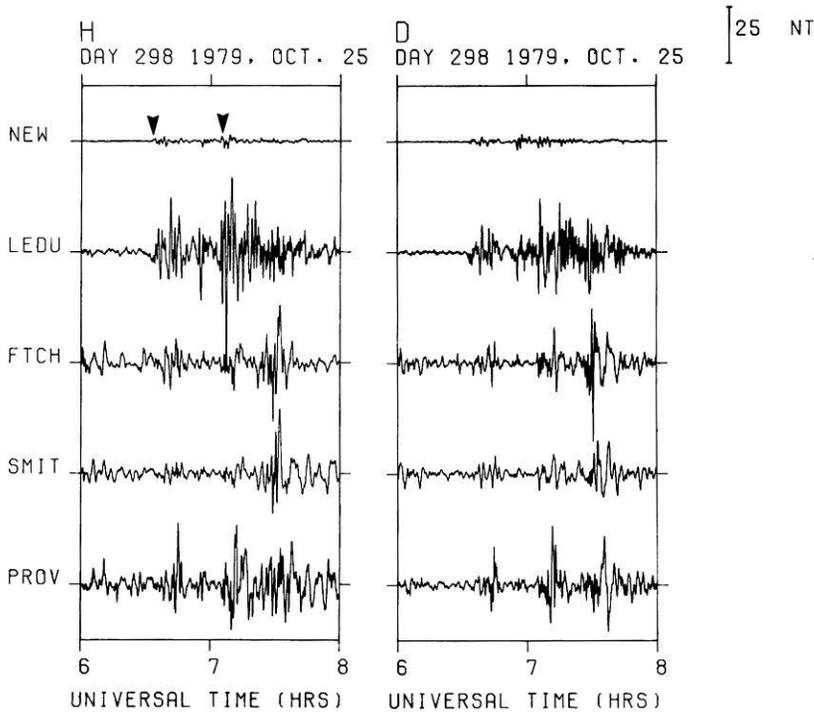


Fig. 7. High pass (5 mHz) filtered data for the *H* and *D* components of the Alberta array and NEW. The arrows indicate the Pi 2 onsets at NEW

low-latitude (60° – 61°) part of the arc, which extended in an east-west direction.

Pi 2 pulsations

High-pass-filtered data for the interval 06:00–08:00 are given in Fig. 7. Clear Pi 2's were associated with both breakups and are easiest to see at NEW (see arrows on the diagram). The extreme complexity of the wave forms at FTCH, SMIT and PROV emphasizes the difficulties in estimating the parameters of Pi 2's recorded by stations near the breakup.

Figure 8 summarizes the characteristics of the Pi 2 associated with the 6:33 breakup (no auroral photos are available). Power spectra over the interval 6:30–6:45 indicate that this Pi 2 pulsation train had a mean frequency of about 10 mHz. A comparison of the contours of the pure state power in Fig. 8 (top) with the equivalent currents in Fig. 5 (top) shows that the region of the substorm westward electrojet, and the peak in the Pi 2 power, overlapped. In this study we do not have sufficient latitudinal resolution to determine whether the Pi 2 peak was actually south of the centre of the electrojet as found by Rostoker and Samson (1981). The data in Fig. 8 also indicate that the centre of the Pi 2 activity did not extend beyond the westernmost stations.

Values of the degree of polarization, \hat{B}_1^2 (10 mHz), are contoured in Fig. 8 (middle). Near the western edge of the enhanced westward electrojet, \hat{B}_1^2 drops to very low values ($\hat{B}_1^2 \approx 0.2$) indicating that the Pi 2 here was obscured by noise. This is perhaps not surprising since this might be near the region of the formation of the WTS. The FAC and electrojets associated with the WTS contribute large amounts of noise to the ULF magnetic spectrum (Samson and Rostoker, 1983; Samson and Harrold, 1983).

The highest value of \hat{B}_1^2 (~ 0.8) occurred just equatorward of the centre of activity of the Pi 2. This characteristic most likely arises from the fact that stations at these positions see only the integrated effects of FAC associated with the Pi 2, and not the more detailed spatial and temporal structures in the electrojets and high-latitude FAC. Also, the high-latitude regions see more dynamic motions of the Pi 2 sources (Samson and Rostoker, 1983) and this can make the high-latitude Pi 2's look unpolarized.

The polarization ellipses in the *H*–*D* plane for the 6:33 event are plotted in Fig. 8 (bottom). There are only two regions with CW (clockwise viewed down) polarization; one centred near the maximum in the Pi 2 activity and the other at the two easternmost stations. Except possibly for the data at the two easternmost stations, these data are consistent with the statistical picture given by Samson and Harrold (1983). The two easternmost stations are much further from the substorm onset than any of the data used by Samson and Harrold. A more detailed evaluation of these polarization data will be given later in the discussion section of the manuscript.

The Pi 2 associated with the 7:04 breakup has been analysed in somewhat more detail because of the availability of the auroral photographs for correlative purposes. Some representative power spectra ($\mathbf{z}^{\dagger}\mathbf{z}$) for this event are given in Fig. 9 (top). The spectral peak of the Pi 2 (~ 9 mHz, marked by arrow) is quite prominent at the mid-latitude station NEW. At LEDU, the peak is less prominent, but the peak power increased. The station LEDU is near the latitude of the onset. At the highest latitude station PROV, the Pi 2 peak cannot be clearly identified except by correlating it with the peaks at the other two stations.

The pure state powers are shown in Fig. 9 (middle). Now the Pi 2 peaks at NEW and LEDU are even more

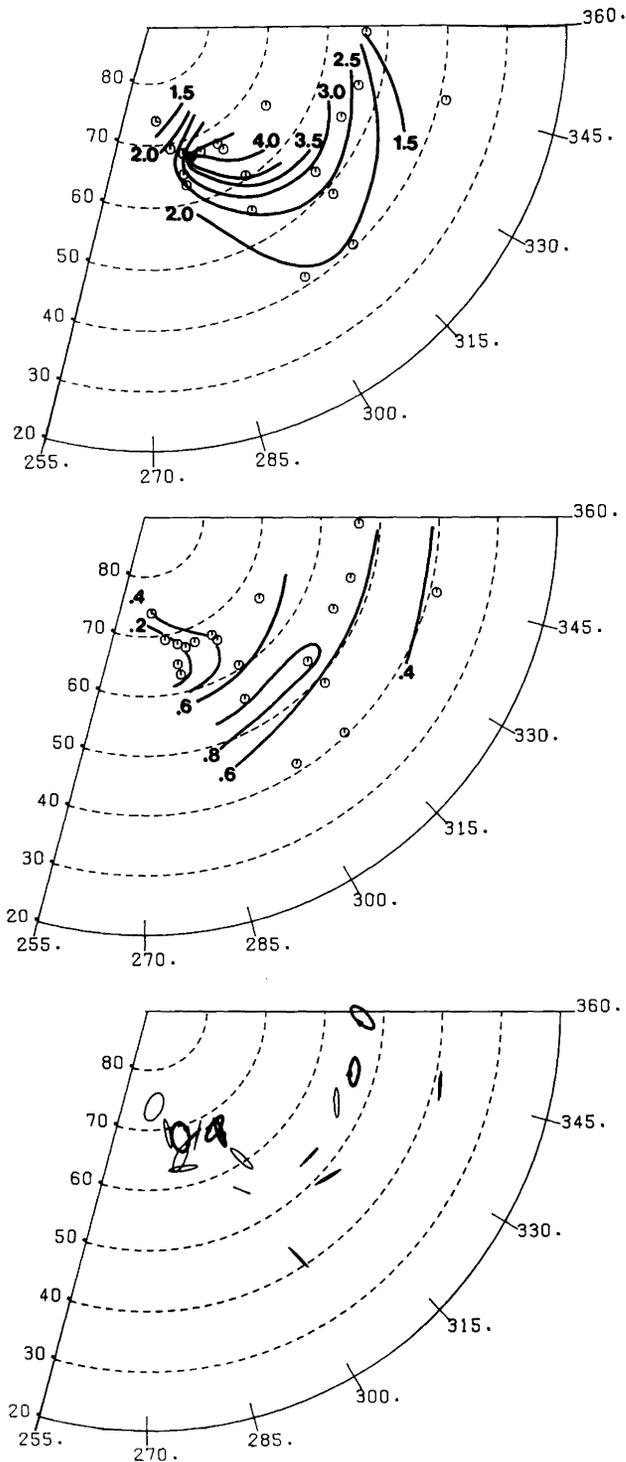


Fig. 8. Power and polarization parameters for the Pi2 in the interval 6:30–6:45, at 10 mHz. *Top:* log pure state power ($\log_{10} \hat{u}^{\dagger} \hat{u}$). *Middle:* degree of polarization estimates, \hat{B}_1^2 . *Bottom:* the polarization ellipses in the horizontal plane. A dark ellipse with an arrow indicates CW polarization, otherwise the polarization is CC

prominent, since the Pi 2's appear highly polarized. The peak at PROV is more clearly resolved, though it is still difficult to identify. The Pi2 is somewhat easier to find by inspecting plots of \hat{B}_1^2 against frequency. These plots for the three stations are given in Fig. 9 (bottom). At NEW the Pi2 had a \hat{B}_1^2 in excess of 0.8 and the

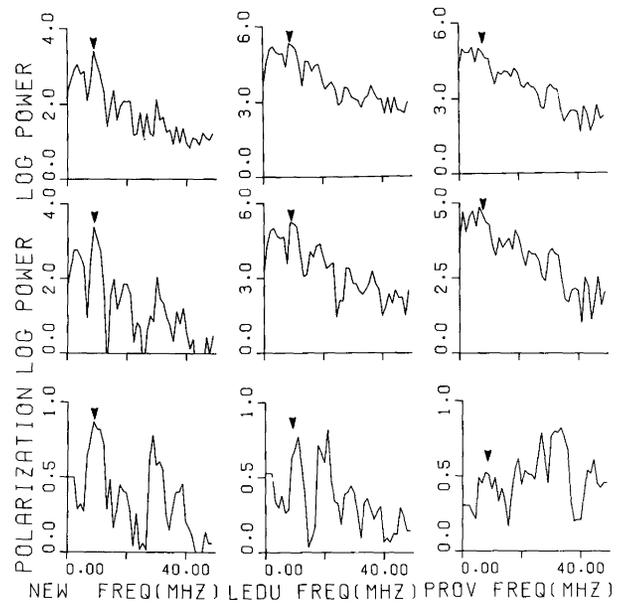


Fig. 9. *Bottom:* polarization spectra \hat{B}_1^2 for the second Pi2 (7:02–7:17). *Middle:* pure state power spectra. *Top:* power spectra ($\mathbf{z}^{\dagger} \mathbf{z}$)

peak is very prominent. At LEDU, the Pi2 was also highly polarized ($\hat{B}_1^2 \approx 0.7$) and is clearly visible on the plot. A pulsation with a frequency near 20 mHz shows even higher degrees of polarization. At PROV, the Pi2 had $\hat{B}_1^2 \approx 0.5$, but the Pi2 peak is quite visible. Note once again the very highly polarized waves occupying a band from ~ 20 –30 mHz. These high-frequency pulsations probably are not directly connected to the Pi2 source mechanism and consequently I shall not consider them here.

The contours of the powers of the Pi2 associated with the second breakup, given in Fig. 10 (top), show that this Pi2 had a very large longitudinal extent. Also, the maximum did not occur over the area of magnetometer coverage, but possibly further to the west. The maximum was well localized latitudinally and seems to coincide with the substorm-enhanced electrojet (Fig. 5, bottom) and with the position of the brightened arc in the vicinity of the Alberta array (Fig. 2). These data show that the centre of activity was west of the stations (west of 270° E). This agrees with the interpretation of the latitude profile in Fig. 6 (bottom), where the D level shift indicated downward FAC. Thus the auroral features in Fig. 2 are only part of a complex and extended breakup.

The contours of \hat{B}_1^2 (Fig. 10, middle) show that this event has many similarities to the previous event. The highest polarizations are found to the south of the intensity maximum. In this case, however, the lowest values are found not near the western edge of the array of stations, but to the southwest of the intensity maximum. This region of low polarity coincides with the northward and eastward equivalent currents shown in Fig. 5 (bottom) and is due west of the extended northward arc in Fig. 2. Consequently, this region of low polarization might have been caused by a localized WTS.

The polarization ellipses for the Pi2 associated with

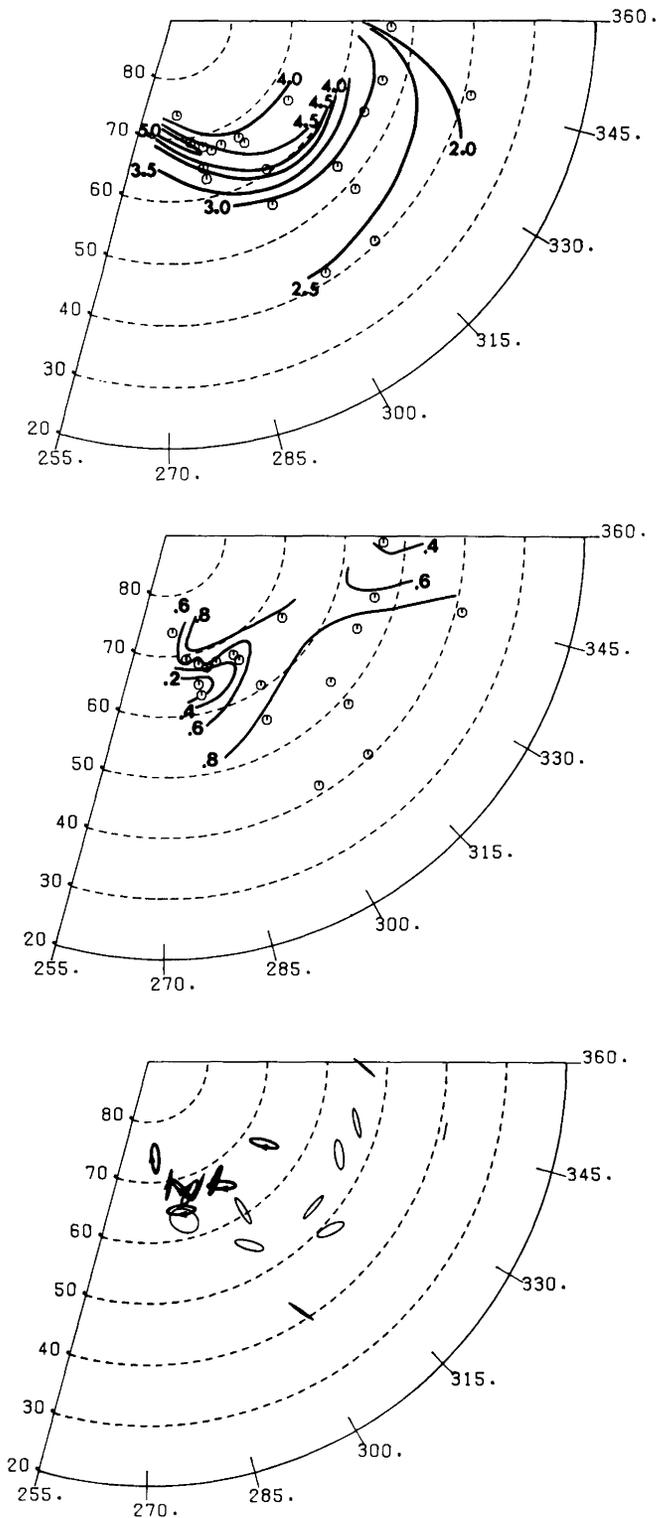


Fig. 10. Power and polarization parameters for the Pi 2 in the interval 7:02–7:17, at 9 mHz. *Top*: log pure state power. *Middle*: \bar{B}_1^2 . *Bottom*: the polarization ellipses

the second event are plotted in Fig. 10 (bottom). All stations south of $\sim 64^\circ\text{N}$ show CC (counterclockwise) polarization, indicating, as shown later, westward propagation of a region of FAC at high latitude. A more detailed evaluation of the polarization data is given in the discussion to follow.

Discussion

Before continuing the appraisal of the data in this study, it is essential that we consider existing, theoretical constructs for the generation of Pi 2's. A substantial number of theories have been developed, but the complexity of the FAC and ionospheric currents associated with the substorm expansive phase has proven to be a formidable obstacle to the development of complete and self-consistent theories.

It is now generally accepted that the enhanced FAC associated with the expansive phase must propagate, at least initially, as shear Alfvén waves (Maltsev et al., 1974; Mallinckrodt and Carlson, 1978; Southwood and Stuart, 1979). However, nearer the ionosphere, this enhanced FAC can lead to the formation of electrostatic shocks, double layers, anomalous resistivity, of kinetic Alfvén waves (see e.g. Kindel and Kennel, 1971; Kan, 1975; Shawhan et al., 1978; Mozer et al., 1980; Goertz, 1981). Any of these mechanisms can lead to the formation of the beams of electrons (in keV range) which cause the auroral arc brightening associated with the substorm expansive phase. These beams of energetic electrons often carry a large part of the upward FAC associated with an auroral arc (Vondrak, 1975). Consequently these high-energy electrons might contribute to a large part of the Pi 2's magnetic field.

The precipitating electrons in the arc cause increases in the ionospheric conductivities. These increases in the conductivity lead, in turn, to rapid changes in the configuration of ionospheric currents associated with the auroral arc. In addition, changes in the horizontal gradients of the conductivities lead to changes in the FAC associated with the arc. Transient changes in the ionospheric currents and FAC contribute to part of the Pi 2's magnetic field near the brightening arc (Samson and Rostoker, 1983).

It appears unlikely that the mechanism for the formation of Pi 2's can be separated from the mechanism leading to FAC and arc brightening during the substorm expansive phase. Consequently, we are led to consider a complex sequence of interactions, some using kinetic theory, others using magnetohydrodynamic (MHD) theory. It is doubtful that MHD theory alone can give an answer to all the observed features of Pi 2's, at least at high latitudes. Boström (1975) has reviewed the complexity of the interactions in magnetospheric-ionospheric coupling. All of the complexities inherent in ionospheric current and FAC systems associated with the substorm expansive phase probably apply to Pi 2's as well.

To discuss the transient FAC and ionospheric currents associated with Pi 2's it is perhaps most informative to begin with the cold plasma equations in the magnetosphere. Then the changing electric field $\mathbf{E}e^{i\omega t}$ is governed by the vector-equation (see Stix, 1962)

$$\nabla \times \nabla \times \mathbf{E} - \frac{\omega^2}{c^2} \mathbf{K} \mathbf{E} = \mathbf{0}, \quad (1)$$

where, in the low frequency limit ($\omega \ll$ ion cyclotron frequency), the equivalent dielectric tensor \mathbf{K} is given by

$$\mathbf{K} \approx \text{diag}[c^2/v_a^2, c^2/v_a^2, -\omega_{pe}^2/\omega^2]. \quad (2)$$

In Eq. (2), c is the speed of light, v_a is the Alfvén speed, ω_{pe} is the electron plasma frequency and the geomagnetic field is $\mathbf{B}_0 = [0, 0, B_3]$.

The problem now is to choose suitable boundary conditions and geometries for the sources of the changing electric fields and FAC associated with the substorm expansive phase and Pi 2's. Some evidence now suggests that the source of FAC near the Harang discontinuity, and the region 1 currents of Iijima and Potemra (1976), is a velocity shear zone associated with the low-latitude boundary layer (LLBL) and the outer region of the central plasma sheet (CPS) (Sonnerup, 1980; Rostoker, 1983, 1984; Rostoker and Samson, 1984). The anti-sunward convection in the LLBL leads to electric fields which map to equatorward, ionospheric electric fields, poleward of the Harang discontinuity in the evening sector. The sunward convection in the CPS leads to poleward, ionospheric electric fields in regions equatorward of the Harang discontinuity.

The data from this campaign and the observations of Baumjohann et al. (1981) indicate that at least some of the onsets of substorm expansive phases can occur in association with quiet arcs which are equatorward of the Harang discontinuity. Consequently, the Pi 2's and substorm expansive phases might be connected with changes in the convective velocity of plasma in the CPS, and possibly near the LLBL. Rostoker et al. (1984) have suggested that some of the substorm energy might be derived from a slowing of the sunward convective velocity in the CPS.

A slowing of the sunward convective velocity will be accompanied by a transient electric field, which maps to an *equatorward transient* electric field, superposed on the poleward convection field in regions of the ionosphere, equatorward of the Harang discontinuity. For convenience and simplicity here, I shall initially adopt a Cartesian geometry and return to a dipolar geometry for numerical modelling. In the magnetotail, near the LLBL, vectors in the x_1 direction are perpendicular to \mathbf{B}_0 , pointing in the direction (dawn to dusk) of the cross-tail electric field. In the ionosphere, I shall assume that x_1 is magnetic north, x_2 is magnetic east, and x_3 is in the direction of \mathbf{B}_0 . I shall also assume that the arc is aligned east-west and consequently gradients $\frac{\partial}{\partial x_1} \gg \frac{\partial}{\partial x_2}$. This inequality of the gradients is also true for regions near the LLBL, where the convective velocities show considerable shear in the x_1 direction (Eastman et al., 1976). Also, in association with the convective flow $E_1 \gg E_2$, ($\mathbf{E} \approx -\mathbf{v} \times \mathbf{B}_0$).

Now, neglecting terms that are second order in E_2 and $\partial/\partial x_2$, the third component of Eq. (1) is

$$\frac{\partial^2 E_3}{\partial x_1^2} - \frac{\partial^2 E_1}{\partial x_1 \partial x_3} - \frac{\omega_{pe}^2}{c^2} E_3 = 0. \quad (3)$$

Equation (3) indicates that the parallel electric field E_3 is coupled to E_1 through electron plasma oscillations and the FAC is carried by electrons. Parallel electric fields associated with localized Alfvén waves were noted by Fejer and Lee (1967). Normally $E_3 \ll E_1$. However, if $\partial/\partial x_1$ is large, then the parallel electric fields at the leading edge of a wavefront can accelerate electrons to hundreds of eV (Goertz and Boswell, 1979).

Before the wavefront reaches the ionosphere, any FAC carried by electrons is fed by polarization currents (perpendicular to \mathbf{B}_0)

$$j_p = (\mu_0 v_a^2)^{-1} \frac{\partial \mathbf{E}_\perp}{\partial t}. \quad (4)$$

When the wave reaches the ionosphere, the polarization currents are replaced by Hall and Pedersen currents (see e.g. Nishida, 1978, Chap. III.5) and the impedance mismatch leads to reflected waves propagating up the field lines. In the model I am considering here, the gradients in the Hall and Pedersen conductivities are perpendicular to the quiet arc and consequently $\nabla \sigma_{H,P} \times \mathbf{E}_\perp = \mathbf{0}$, where σ_H and σ_P are the Hall and Pedersen conductivities, respectively. Then

$$E_{1R}(x_1) = \left[\frac{(\mu_0 v_a)^{-1} - \Sigma_P(x_1)}{(\mu_0 v_a)^{-1} + \Sigma_P(x_1)} \right] E_{1I}(x_1), \quad (5)$$

where R and I indicate reflected and incident, respectively, and Σ_P is the height-integrated Pedersen conductivity (Glassmeier, 1984). Since Σ_P is normally much greater than $(\mu_0 v_a)^{-1}$, the reflected wave has E_{1R} positive or poleward. (Note that the incident wave is a transient with equatorward \mathbf{E}_{1I} .)

The FAC associated with the waves is given by the relation (see e.g. Kan et al., 1982)

$$\begin{aligned} j_3(x_1) &= (\mu_0 B_0 v_a)^{-1} \left[\hat{\mathbf{n}}_R \cdot \mathbf{B}_0 \frac{\partial E_{1R}(x_1)}{\partial x_1} \right. \\ &\quad \left. + \hat{\mathbf{n}}_I \cdot \mathbf{B}_0 \frac{\partial E_{1I}(x_1)}{\partial x_1} \right] \\ &\approx (\mu_0 B_0 v_a)^{-1} \left[\hat{\mathbf{n}}_R \cdot \mathbf{B}_0 \frac{\partial E_{1R}(x_1)}{\partial x_1} \right], \end{aligned} \quad (6)$$

where $\hat{\mathbf{n}}$ is a unit vector in the direction of propagation. In making the approximation in Eq. (6), I have assumed $\frac{\partial E_{1R}}{\partial x_1} \gg \frac{\partial E_{1I}}{\partial x_1}$, due to the large gradients in conductivities near the arc. Since $E_{1R}(x_1)$ is positive everywhere (i.e. points poleward near the ionosphere), $\frac{\partial E_{1R}(x_1)}{\partial x_1}$ is negative

on the field lines threading the poleward border of the arc and positive on the equatorward border. Consequently FAC flows downward on field lines threading the poleward side of the arc and the region of enhanced conductivity and upward on equatorward field lines.

To model the Pi 2 magnetic fields, the finite scale sizes in the east-west (x_2) directions must, in reality, be considered. These finite east-west dimensions lead to east-west offsets in the FAC, due to the Hall conductivity in the arc (normally $\Sigma_H > \Sigma_P$). Consequently the downward FAC will tend to be somewhat to the east of the upward FAC (see e.g. Samson, 1982, Fig. 8). Note also that the configuration of FAC that is discussed here is compatible with the features of the substorm current "wedge" (McPherron et al., 1973) and the formation of a transient westward electrojet (see Fig. 11).

In order to compute the magnetic fields and polarizations associated with the transient FAC, the FAC

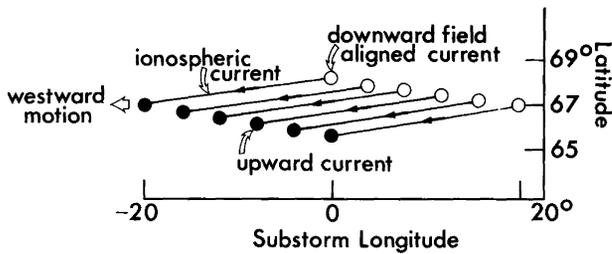


Fig. 11. A model current system for the numerical calculation of Pi 2 magnetic fields. The ionospheric currents flow at a height of 110 km

and ionospheric current model in Fig. 11 have been used. This model is meant to retain the essential features of the reflected electric fields and FAC. The Alfvén velocity near the ionosphere is very high ($\sim 1,000\text{--}3,000$ km/s), and since Pi 2's have time scales on the order of tens of seconds or greater, the high latitude Pi 2 magnetic field can be derived from a model with no phase changes along a current loop.

Ionospheric currents were deduced by assuming a conductivity ratio $\Sigma_H/\Sigma_P = 5$ (Brekke et al., 1974), $\nabla \times \mathbf{E} \approx 0$, with \mathbf{E} (sum of incident and reflected) in the ionosphere pointing equatorward. The FAC sheets at the edge of the region of enhanced conductivity were approximated by the set of six discrete FAC systems (for numerical computations). A dipolar geometry was used for the FAC. The magnetic fields associated with the currents were calculated using programs similar to those discussed by Kisabeth (1979). The FAC were integrated to the magnetic equator (dipole field). Although the model is somewhat nonphysical in the geometry of the ionospheric currents, this feature will not compromise the interpretation as I tend to use this only as an heuristic model to approximate the Pi 2 oscillations.

Many observations indicate that Pi 2 fields show westward apparent phase velocities (Lester et al., 1983 and references therein). Also, since the Pi 2 currents at a given location are transient in nature, the current system in Fig. 11 moves westward to produce the oscillating Pi 2 fields seen on the ground. This westward motion correspond to a *sunward* propagation of the slowing of convection in the CPS. The entire current system moves westward at a constant velocity but with diminishing strength. Note that the current system does not oscillate in the moving frame. Pashin et al. (1982) used a similar principle in modelling Pi 2 magnetic fields. For the simulation, the velocity was chosen to be 4° longitude/time step, and the longitudinal attenuation was $\exp(-\text{longitude}/7)^2$. The scale size (-20° to 20° longitude) was based on the observations of Samson and Harrold (1983).

To visualize how this model can produce oscillating fields, it is best to consider a point to the southwest of the start of the current system. An observer at that station will see an "apparent" substorm current wedge moving westward, with oscillations in the H component; negative, positive, negative in that sequence. The D component will be 90° out of phase (D is zero when H is maximum) leading to the elliptical polarizations seen at mid latitudes.

To maintain compatibility with the experimental results for the Pi 2 polarizations in this study and that of Samson and Harrold (1983), I have calculated the polarizations of the model Pi 2 fields by using the same procedure as that used for the observations (see Samson and Harrold, 1983, Appendix). Figure 12 gives the estimated polarizations for the model. Near and to the west of the initial onset the polarization pattern is fairly complex because of the mixture of magnetic fields from FAC and ionospheric currents (see the description in Samson and Rostoker [1983]). Clockwise polarization occurs directly equatorward of the initial upward FAC, and CC directly poleward. This reversal is caused by a latitudinal change in the direction of the magnetic field of the FAC, whereas the fields from the electrojet show little latitudinal phase shift. In most other regions (away from onset), the magnetic fields are dominated by those from the FAC and show CC polarization equatorward of 64° and CW poleward of 68° latitude.

The orientations of the major axis clearly show the substorm "wedge" effect, with the axis oriented to be compatible with two regions of FAC near the onset. This pattern results from the attenuation of the Pi 2 currents with distance westward. Also note that in Fig. 11, the sheets of FAC are slightly tilted from east-west orientation. This tilt was added to give better agreement with the results of Samson and Harrold (1983).

A detailed comparison of the polarization data in Figs. 8 and 10 with the model parameters in Fig. 11 indicates that the model is compatible with many of the polarization features of the Pi 2's in this study. The maximum powers of the Pi 2's occur in and near the brightening arc and the enhanced substorm expansive phase electrojets (at least to the resolution of the station spacing of this study). I could find no secondary maxima in the total pure state power ($H+D+Z$) in regions equatorward of the brightening arc and the primary maximum of the Pi 2 power. In the model, the northern border of the region of CW polarization, which is centred near the breakup, is near the northern border of the enhanced westward electrojet associated with the Pi 2. Accordingly, the two northernmost stations (NOW, CPY) in Fig. 8, which are far north of the substorm-enhanced electrojet (see Fig. 5), had CC polarization while those closer to the electrojet (AVI, FSP, PROV, SMIT, FTCH) had CW polarization. Also, all stations which were well south of the enhanced electrojet have CC polarization as the model predicts (except MCL and SUB). Unfortunately, no data could be obtained for stations in the region $60^\circ\text{--}75^\circ\text{N}$, $315^\circ\text{--}345^\circ\text{E}$, in order to determine whether the region of CW polarization predicted for high latitudes does in fact occur in this event.

The polarizations of the second Pi 2 (Fig. 10) also appear to match those in the model. All stations south of $\sim 64^\circ\text{N}$ show CC polarization, indicating westward propagation of the high-latitude FAC associated with the Pi 2. Near the region of the intensity maximum of the Pi 2's and the region of the substorm westward electrojet, the polarizations are predominately CW. In this case, the breakup electrojet and Pi 2 maximum were too far north for the high-latitude region of CC polarization to be seen.

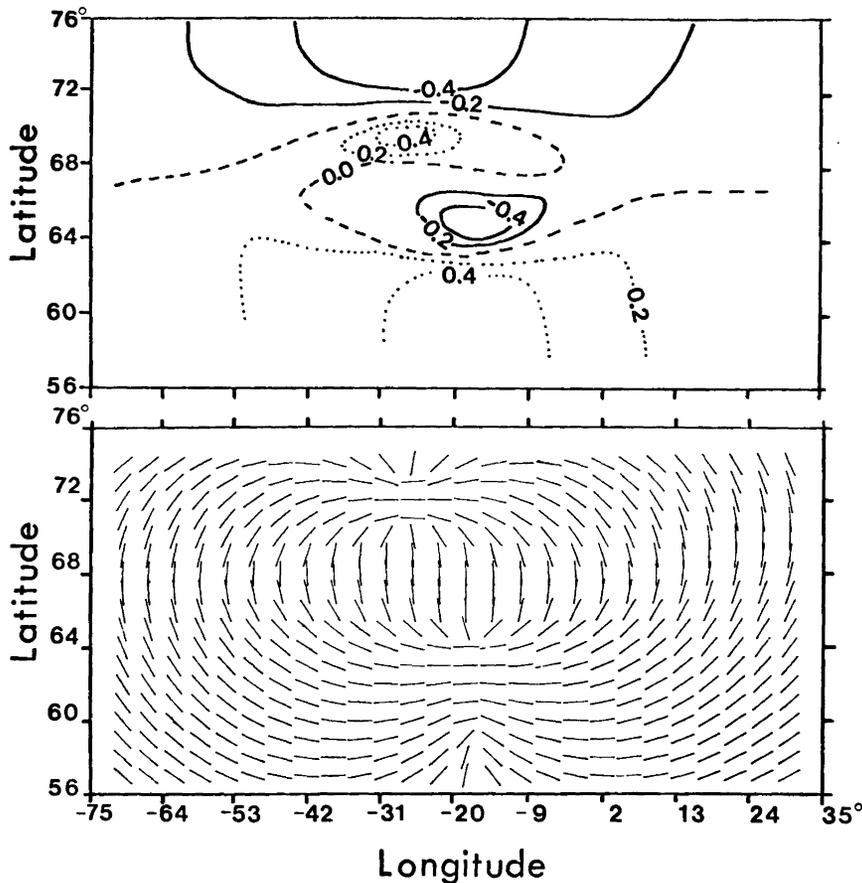


Fig. 12. Polarizations for the Pi 2 model. *Top:* Ellipticity (ratio minor to major axis). Positive and negative ellipticities denote CC and CW polarization, respectively. *Bottom:* Orientations of the major axes of the polarization-ellipses

Turning to the orientations of the polarization ellipses, the major axis in both events appear to be aligned in a manner that suggests a region of FAC somewhere near $\sim 65^\circ$ N. For the first Pi 2 (Fig. 8) the FAC occurs over the longitudinal range $\sim 300^\circ$ to 330° , whereas for the second Pi 2 (Fig. 10) the longitudinal range is larger, $\sim 270^\circ$ to 340° . Inspection of the model polarizations (Fig. 12, bottom) indicates that this must be the region of net downward FAC.

The region of CW polarization near the location of the onset of the substorm expansive phase shows that both the ionospheric electrojets and FAC contribute to the polarizations in this region. South of the intensity-maximum, all polarizations are CC and this is consistent with a westward moving FAC system at high latitudes. Unfortunately, this campaign was limited to events which showed only the net downward FAC, because the cameras were placed near the western part of the array.

To produce the observed mid-latitude frequencies (9–10 mHz) of the Pi 2's, the westward velocity of the FAC associated with the Pi 2 must be near 20–50 km/s. I have computed this velocity by using an FAC system in which the net FAC cells (upward to the west, downward to the east) are 20° – 50° apart in longitude (this study; Samson and Harrold, 1983; Lester et al., 1983). The westward velocity of the Pi 2's FAC is much larger than typical westward velocities of the WTS (typical maxima are ~ 2 – 3 km/s, see e.g. Tighe and Rostoker, 1981). Consequently the westward motion of the Pi 2 currents is most likely caused by a mechanism sub-

stantially different from that causing the formation of the WTS.

Summary and conclusions

The data from this campaign indicate that Pi 2's are associated with the brightening of an auroral arc at the onset of the substorm expansive phase. In at least some cases, which this study included, the brightening and onset of the expansive phase occur in the region of the eastward electrojet, equatorward of the existing Harang discontinuity.

A comparison of the polarization data from the Pi 2's in this study suggests that the Pi 2's might be associated with a westward-moving, transient reflection of an incident pulse of equatorward electric fields (at ionospheric levels). Although only one pulse was used to model the Pi 2's in this study, subsequent reflections from the magnetospheric source regions (Nishida, 1979; Kan et al., 1982), or the southern polar ionosphere, could lead to a series of westward-moving pulses, giving the longer Pi 2 trains often seen on the ground. However, the westward motion must be an integral part of the model as westward "apparent" phase velocities are seen at mid-latitude stations. In any case, the single transient model presented here should give the simplest possible Pi 2 pulses, with only ~ 1 cycle observed at mid latitudes in the H component as the FAC "wedge" moves westward.

If my conjecture that the source of the electric field transient is in the CPS near the LLBL (see the Discussion), then the Pi 2's might indicate that the distur-

bance leading to the substorm expansive phase is propagating sunward from the outer magnetotail.

If the Pi 2 is typically associated with the brightening of an arc, the westward motion of the Pi 2 might also be accompanied by a rapid westward expansion of the brightening. Because of the high velocities of the expansion (20–50 km/s), typical all-sky camera exposure times (5–20 s) are much too long and it may be necessary to use a television system with an image intensifier in order to observe the dynamic features of this brightening. Westward-propagating brightening with velocities in excess of 10 km/s have already been reported (Opgenoorth et al., 1983) giving some support to the predictions of the Pi 2 model which I have proposed here. A careful correlation of high time resolution (5 s or better) auroral photographs with Pi 2 magnetic signature would help considerably in an evaluation of the transient model.

Acknowledgements. I wish to thank Prof. G. Rostoker for many interesting discussions and talks on the topic of Pi 2's, and Mr. M.G. Connors for his collaboration in the computation of the numerical results for the Pi 2 model. The original programs for computation of the magnetic fields were written by Dr. J.L. Kisabeth. I am grateful to Transport Canada (Telecommunications Branch), Dept. of Environment and Fisheries (Atmospheric Environmental Service – Western Regional Headquarters) and to Mr. and Mrs. Ted Malewski of Fort Providence, N.W.T., for their help in operating the magnetometer array.

The research described in this report was supported in part by the Boreal Institute for Northern Studies of the University of Alberta, the Natural Sciences and Engineering Research Council of Canada and by the U.S. Air Force Geophysics Laboratory (Contract No. F19628-82-k-0032). The author is an NSERC University Research Fellow.

References

- Afanasyeva, L.T., Raspopov, O.M., Schepetrov, R.V., Koshel'evskiy, V.I., Hazarov, M.D.: Relationship between geomagnetic pulsations of the Pi 2 type and the parameters of the auroral zone. *Geomagn. Aeron.* **10**, 600–601, 1970
- Baumjohann, W., Pellinen, R.J., Opgenoorth, H.J., Nielsen, E.: Joint two-dimensional observations of ground magnetic and ionospheric electric fields associated with auroral zone currents: current systems associated with local auroral breakups. *Planet. Space Sci.* **29**, 431–447, 1981
- Björnsson, A., Hillebrand, O., Voelker, H.: First observational results of geomagnetic Pi 2 and Pc 5 pulsations on a north-south profile through Europe. *Z. Geophys.* **37**, 1031–1042, 1971
- Boström, R.: Mechanisms for driving Birkeland currents. In: *Physics of the hot plasma in the magnetosphere*, B. Hultqvist and L. Stenflo, eds., 341–362. New York Plenum Press 1975
- Brekke, A., Doupnik, J.R., Banks, P.M.: Incoherent scatter measurements of the E region conductivities and currents in the auroral zone. *J. Geophys. Res.* **79**, 3773–3790, 1974
- Eastman, T.E., Hones, E.W. Jr., Bame, S.J., Asbridge, J.R.: The magnetospheric boundary layer: site of plasma, momentum and energy transfer from the magnetosheath into the magnetosphere. *Geophys. Res. Lett.* **3**, 685–688, 1976
- Fejer, J.A., Lee, L.F.: Guided propagation of Alfvén waves in the magnetosphere. *J. Plasma Physics* **1**, 387–406, 1967
- Glaßmeier, K.-H.: On the influence of ionospheres with non-uniform conductivity distribution on hydromagnetic waves. *J. Geophys.* **54**, 125–137, 1984
- Goertz, C.K.: Discrete breakup arcs and kinetic Alfvén waves. In: *Physics of auroral arc formation*, S.I. Akasofu and J.R. Kan, eds.: pp. 451–455. American Geophysical Union 1981
- Goertz, C.K., Boswell, R.W.: Magnetosphere-ionosphere coupling. *J. Geophys. Res.* **84**, 7239–7246, 1979
- Hakura, Y.: Tables and maps of geomagnetic coordinates corrected by the higher order spherical harmonic terms. *Report of Ionospheric and Space Res. in Japan* **19**, 121–157, 1965
- Iijima, T., Petemra, T.A.: The amplitude distribution of field aligned currents at northern high latitudes observed by Triad. *J. Geophys. Res.* **81**, 2165–2174, 1976
- Kan, J.R.: Energization of auroral electrons by electrostatic shock waves. *J. Geophys. Res.* **80**, 2089–2095, 1975
- Kan, J.R., Longenecker, D.U., Olson, J.V.: A transient response model of Pi 2 pulsation. *J. Geophys. Res.* **87**, 7483–7488, 1982
- Kindel, J.M., Kennel, C.F.: Topside current instabilities. *J. Geophys. Res.* **76**, 3055–3078, 1971
- Kisabeth, J.L.: On calculating magnetic and vector potential fields due to large scale magnetospheric current systems and induced currents in an infinitely conducting earth. In: *Quantitative modeling of magnetospheric processes*, W.P. Olson, ed.: pp. 473–498, *Geophys. Monogr. Ser.*, **21**. Washington, D.C., 1979
- Lester, M., Hughes, W.J., Singer, H.J.: Polarization patterns of Pi 2 magnetic pulsations and the substorm current wedge. *J. Geophys. Res.* **88**, 7958–7966, 1983
- Mallinckrodt, A.J., Carlson, C.W.: Relations between transverse electric fields and field-aligned currents. *J. Geophys. Res.* **83**, 1426–1432, 1978
- Maltsev, Yu.P., Leontyev, S.V., Lyatskiy, V.B.: Generation and natural frequencies of Pi 2 pulsations. *Geomagn. Aeron.* **14**, 101–107, 1974
- McPherron, R.L., Russell, C.T., Aubry, M.P.: Satellite studies of magnetospheric substorms on August 15, 1968. 9, Phenomenological model for substorms. *J. Geophys. Res.* **78**, 3131–3149, 1973
- 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
- Nishida, A.: Possible origin of transient dusk-to-dawn electric field in the nightside magnetosphere. *J. Geophys. Res.* **84**, 3409–3412, 1979
- Nishida, A.: *Geomagnetic diagnosis of the magnetosphere*. New York: Springer-Verlag, 1978
- Opgenoorth, H.J., Pellinen, R.J., Baumjohann, W., Nielsen, E., Marklund, G., Eliasson, L.: Three dimensional current flow and particle precipitation in a westward travelling surge (observed during the barium-GEOS rocket experiment). *J. Geophys. Res.* **88**, 3138–3152, 1983
- Pashin, A.B., Glaßmeier, K.H., Baumjohann, W., Raspopov, O.M., Yahnin, A.G., Opgenoorth, H.J., Pellinen, R.J.: Pi 2 magnetic pulsations, auroral break-ups, and the substorm current wedge: a case study. *J. Geophys.* **51**, 223–233, 1982
- Pytte, T., Trefall, H.: Auroral-zone electron precipitation event observed before and at the onset of negative magnetic bays. *J. Atmos. Terr. Phys.* **34**, 315–337, 1972
- Pytte, T., McPherron, R.L., Kokubun, S.: The ground signatures of the expansion phase during multiple onset substorms. *Planet. Space Sci.* **24**, 1115–1132, 1976
- Rostoker, G.: Dependence of the high-latitude ionospheric fields and plasma characteristics on the properties of the interplanetary medium. In: *High-latitude space plasma physics*, B. Hultqvist and T. Hagfors, eds.: p. 189, Plenum Publ. Co., New York, 1983

- Rostoker, G.: Implications of the hydrodynamic analogue for the solar terrestrial interaction and the mapping of high latitude correction pattern into the magnetotail. *Geophys. Res. Lett.* **11**, 251–254, 1984
- Rostoker, G., Samson, J.C.: Polarization characteristics of Pi 2 pulsations and implications for their source mechanisms; location of source regions with respect to the auroral electrojets. *Planet. Space Sci.* **29**, 225–247, 1981
- Rostoker, G., Samson, J.C.: Can substorm expansive phase effects and low frequency Pc magnetic pulsations be attributed to the same source mechanism? *Geophys. Res. Lett.* **11**, 271–274, 1984
- Rostoker, G., Armstrong, J.C., Zmuda, A.J.: Field-aligned current flow associated with intrusion of the substorm-intensified westward electrojet into the evening sector. *J. Geophys. Res.* **25**, 3571–3579, 1975
- Rostoker, G., Spadinger, I., Samson, J.C.: Local time variation in the response of Pc 5 pulsations in the morning sector to substorm expansive phase onsets near midnight. *J. Geophys. Res.* 1984 (in press)
- Samson, J.C.: Pi 2 pulsations: high latitude results. *Planet. Space Sci.* **30**, 1239–1247, 1982
- Samson, J.C.: Pure states, polarized waves, and principal components in the spectra of multiple, geophysical time-series. *Geophys. J.R. Astron. Soc.* **72**, 647–664, 1983a
- Samson, J.C.: The reduction of sample bias in polarization estimators for multichannel geophysical data with anisotropic noise. *Geophys. J.R. Astron. Soc.* **75**, 289–308, 1983b
- Samson, J.C., Harrold, B.G.: Maps of the polarizations of high latitude Pi 2's. *J. Geophys. Res.* **88**, 5736–5744, 1983
- Samson, J.C., Rostoker, G.: Polarization characteristics of Pi 2 pulsations and implications for their source mechanism; influence of the westward travelling surge. *Planet. Space Sci.* **31**, 435–458, 1983
- Shawhan, S.D., Fälthammer, C.G., Block, L.P.: On the nature of large auroral zone electric fields at $1-R_E$ altitude. *J. Geophys. Res.* **83**, 1049–1054, 1978
- Sonnerup, B.U.Ö.: Theory of the low latitude boundary layer. *J. Geophys. Res.* **85**, 2017–2026, 1980
- Southwood, D.J., Stuart, W.F.: Pulsations at the substorm onset. In: *Dynamics of the magnetosphere*, S.-I. Akasofu, ed.: pp. 341–355, D. Reidel, Hingham, MA, 1979
- Stix, T.H.: *The theory of plasma waves*. New York, McGraw-Hill, 1962
- Tighe, W.G., Rostoker, G.: Characteristics of westward travelling surges during magnetospheric substorms. *J. Geophys. Res.* **50**, 51–67, 1981
- Troitskaya, J.A., Gul'elmi, A.V.: Geomagnetic micropulsations and diagnostics in the magnetosphere. *Space Sci. Rev.* **7**, 689–768, 1967
- Vondrak, R.R.: Model of Birkeland currents associated with an auroral arc. *J. Geophys. Res.* **80**, 4011–4014, 1975

Received December 20, 1983; Revised October 12, 1984

Accepted October 25, 1984

Book reviews

Problems and Prospects in Long and Medium Range Weather Forecasting. Edited by D.M. Burridge and E. Källén. Springer-Verlag, Berlin, Heidelberg, New York, Tokyo, 74 figs, 290 pp, 1983

Das Problem der Wettervorhersage steht seit Jahrzehnten im Mittelpunkt der meteorologischen Forschung. Dabei ist zu sehen, daß Ertel bereits 1941 den Nachweis brachte, daß eine Prognose exakt nur über ein Gebiet in der Größe der gesamten Erdoberfläche möglich ist. Bei der Auswahl der jeweiligen Randwerte ist das Ertel'sche Resultat zu beachten, daß „die sich aus der prinzipiellen Unbestimmtheit einer Prognose für ein Teilgebiet der Atmosphäre ergebenden Mängel unter das für die Forderungen des praktischen Lebens zulässige Maß herabzudrücken (sind)“ (Ertel, Meteorol. Z. 61 (1944), 181. Charney, Fjortoft und Neumann mußten bei ihren numerischen Wettervorhersagen stets (im Vergleich zur gesamten Erdoberfläche) kleinere Teilgebiete einbeziehen. In diese Gesamtproblematik der Theoretischen Meteorologie und den damit zusammenhängenden Fragen der Praxis führt das vorliegende Buch mit ausgewählten Vorträgen einer Tagung des European Centre for Medium Range Weather Forecasts ein. Dabei werden von den Autoren verschiedene Teilaspekte aufbereitet, die dem Meteorologen, der an diesen Fragen interessiert ist, sowohl einen Einblick in die Gesamtproblematik als auch in offene Fragen vermitteln. Der an diesen Fragen der Wettervorhersage interessierte Leser wird in den Fachaufsätzen insoweit weitergehende Informationen finden. Ausstattung und Druck des Buches sind gut. Das Studium des Bandes erfordert hinreichende Kenntnisse der theoretischen Meteorologie.

W. Schröder

Max Steenbeck, Hans-Jürgen Treder: Möglichkeiten der experimentellen Schwerkrafftorschung. Akademie-Verlag, Berlin, 1984

In diesem für den an fundamentalen Fragen interessierten Physiker gut lesbaren Bändchen geht es nicht bzw. nicht unmittelbar um die Frage der experimentellen Überprüfung von Ergebnissen der

Einsteinschen Gravitationstheorie oder alternativer Schwerkrafttheorien. Vielmehr schürfen die Autoren tiefer und stellen die Frage der experimentellen Prüfung der Bausteine von Schwerkrafttheorien in den Mittelpunkt ihrer Betrachtungen. Die faszinierende Diskussion jener grundsätzlichen Hypothesen führt den Leser auf die oft problematischen Konsequenzen, die sich aus ihrer Annahme ergeben.

Ausgehend von drei verschiedenen Massebegriffen, nämlich der „trägen Masse“ im Sinne einer Menge der Materie, einer „passiven schweren Masse“, unter der die Fähigkeit, auf Schwerkraft zu reagieren, verstanden wird und der „aktiven schweren Masse“, die die Fähigkeit, Schwerkraft zu erzeugen, beschreibt, stehen folgende zwei Fragen zur Diskussion: 1. Ist die träge Masse der passiven schweren Masse gleich? 2. Sind passive und aktive schwere Masse zueinander proportional (Proportionalitätskonstante ist die universelle Gravitationskonstante)?

Daraus ergeben sich eine Reihe von bedeutsamen Gesichtspunkten, insbesondere in kosmologischen Dimensionen. Zum Beispiel, ob die Schwerkraft bei zunehmender Materieansammlung unbegrenzt wachsen kann oder ob sie einer Sättigung zustrebt. Im engen Zusammenhang damit wird u.a. die mögliche Absorption der Schwerkraft durch Materie und die Suppression der Schwerkraft durch ein bereits vorhandenes Schwerepotential besprochen. Dabei werden immer die Möglichkeiten und Grenzen des experimentellen Zugangs zu diesen Fragen aufgezeigt.

Die Anschaulichkeit der Darstellung ist den Autoren gut gelungen; allerdings hätte die Gliederung des Aufsatzes manchmal klarer sein können, um zu vermeiden, daß gelegentliche gedankliche Überschneidungen zwischen einzelnen Kapiteln den Leser gerade wegen der Fülle faszinierender Gedanken verwirren. Empfehlenswert ist dies Heft auch für den Geophysiker, und sei es nur zur ideellen Bereicherung, gehört doch der Schluß von der passiven schweren Masse (Gravimeter) auf die Verteilung der aktiven schweren Massen im Erdkörper in die Kompetenz seiner Fachkenntnisse.

L. Engelhard

Source orientation from grid test and synthetic seismograms and an application to the Ibbenbüren earthquake of July 1981

K.-G. Hinzen* and H. Krummel

Institut für Geophysik, Ruhr-Universität Bochum, Universitätsstraße 150, D-4630 Bochum, Federal Republic of Germany

Abstract. A fault-plane solution technique for regional seismic events is developed. A flexible grid test procedure is used to make combined use of *P*-wave first-motion readings and angles of polarisation of *S* waves. All solutions, which are compatible with the measurements on the basis of a double-couple source are found in this way. Synthetic seismograms calculated with the reflectivity method are used to obtain the most reliable among these solutions.

The procedure is tested with data from the 13 July 1981 Ibbenbüren event ($M_L=4.1$). Seventeen *P*-wave first-motion readings and one angle of polarisation leave 102 out of 93,312 solutions which are compatible with the double-couple model when a 5° increment of strike-, dip- and slip-angle is used. The synthetic seismograms show that a mechanism of strike-slip type is a reasonable model. This result is in contrast to a normal-fault mechanism from a former graphical fault-plane solution.

Key words: Fault-plane solution – Angle of polarisation – Grid test – Synthetic seismograms

Introduction

Fault-plane solutions are a common and basic tool in the determination of seismic source orientations. For local and regional events the focal sphere is often poorly covered with stations reporting clear *P*-wave first-motion readings (Engdahl and Kisslinger, 1977). In order to compensate for this lack of information parameters other than *P*-wave first motions have been introduced to fault-plane solution programs. Shapira and Båth (1978) use the amplitudes of *P* waves; Kisslinger (1980) and Kisslinger et al. (1981) introduced the amplitude ratio of *SV* and *P* waves recorded on the vertical component of near stations. Most of these methods use only one parameter derived from the seismogram. In many cases the quality of data is heterogeneous. Determination of *P*-wave first motions may be possible from analog seismograms; while amplitudes, amplitude ratios or angles of polarisation of *S* waves are more easily measured with three-component digital recording stations. Therefore, it is desirable to use a combination of parameters from seismograms of different quality in one calculation.

* Present address: Bundesanstalt für Geowissenschaften und Rohstoffe, Stilleweg 2, D-3000 Hannover

Offprint requests to: K.-G. Hinzen

This desire supposes a flexible algorithm for the numerical application.

In least-squares algorithms, which are often used in fault-plane solution techniques (Kisslinger et al., 1981; Stevens, 1964; Udias, 1964), it is difficult to combine different parameters, especially when *P*-wave first motions are concerned (Dillinger et al., 1972). Due to the nonlinearity of the problem, a bad choice of starting solution for the least-squares fit may lead to an incorrect orientation.

The implementation of a systematic trial-and-error method, or so-called 'grid method', can avoid these difficulties. Langston (1982) uses a grid method to model teleseismic records of *P* and *SH* waveforms to get the orientation of the source. Pearce (1977; 1979a, b; 1981) searches, in a systematic way, for orientations which are compatible with the amplitude ratios of *P*, *pP* and *sP* waves from teleseismic events. In the following, a flexible procedure is described which uses different parameters in the search for possible source orientations. Special emphasis is given to the angle of polarisation of *S* waves. The procedure is applied to a regional earthquake of 07.13.1981 at Ibbenbüren, which has already been examined by a common graphic fault-plane solution technique (Harjes et al., 1983).

The results of the grid test are used for the calculation of synthetic seismograms with the reflectivity method (Fuchs, 1968; Fuchs and Müller, 1971; Kind, 1978; 1979) in order to show which of the possible orientations gives the best fit to the data.

Description of the grid method

The orientation of a pure double-couple source is completely described by three angles:

- ϕ – strike of fault plane
- δ – dip of fault plane
- λ – slip angle.

These angles define the space of orientation in which the possible solutions have to be searched for. They are measured in the same sense as described by Aki and Richards (1980, Fig. 4.13). The angular boundaries for the search are influenced by symmetry properties of *P*- and *S*-wave radiation patterns. The radiation patterns for a double-couple source in terms of ϕ , δ , λ , as well as azimuth θ and take-off angle i of the ray are given, for example, by Aki and Richards (1980, p. 115).

If the symmetry properties caused by the interchange of fault and auxiliary plane, are neglected, the angular

boundaries necessary to include all orientations are given by Pearce (1979a):

$$\begin{aligned} d &\leq \phi \leq 2\pi \\ d &\leq \delta \leq \pi \\ d &\leq \lambda \leq \pi, \end{aligned} \quad (1)$$

where d is the increment of the angles of orientation which controls the total number of possible orientations.

Figure 1 shows a schematic flow chart of the implemented procedure. The input parameters are the number of stations, azimuth and take-off angles for observations consisting of P -wave first motions or angles or polarisation of S waves. For every angle of polarisation a confidence interval is needed. This is calculated from the rectilinearity of the hodogram of the S phase (Hinzen, 1984). The greater the deviation from linear polarisation, the smaller is the number of rejected solutions. The program starts with the initialisation of a field containing all possible solutions. That means all orientations are possible at the start of the grid test. The number of possible solutions grows rapidly with decreasing increments d . Therefore, 15 solutions were encoded in one 16-bit integer word.

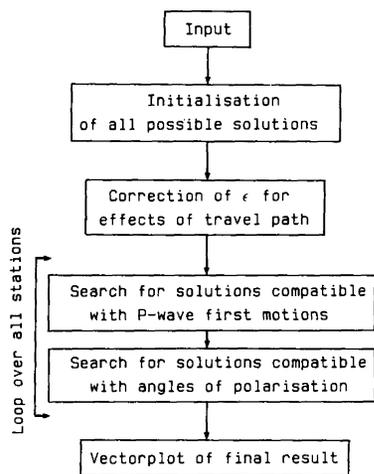


Fig. 1. Schematic flow chart of the grid test procedure for fault-plane solutions

The concept of focal sphere, introduced by Koning (1942) and Ritsema (1958), requires all data to be reduced to the values at the source. The data must be corrected for the influence of the travel path. The polarity of direct or refracted P waves does not change for regional or local distances in this reduction.

The influence of the free surface and reflection and transmission at boundaries on the angle of polarisation must be corrected. Båth (1961) solved the equations of motion for a plane wave incident on a plane boundary for several boundary conditions. In this way he could calculate the change of the angle or polarisation. The angle of polarisation, which is defined as

$$\varepsilon = \arctan \frac{A_{SH}}{A_{SV}} \quad (2)$$

is measured clockwise from the vertical in the interval

$$0 \leq \varepsilon < \pi. \quad (3)$$

A_{SH} and A_{SV} are the amplitudes of SH and SV components, respectively.

The ratio of the angle of polarisation of the reflected wave and the incident wave is:

$$\frac{\tan \varepsilon_R}{\tan \varepsilon} = \frac{A_{SH}^R/A_{SH}}{A_{SV}^R/A_{SV}} \quad (4)$$

A_{SH}^R/A_{SH} is the reflection coefficient for the SH component; A_{SV}^R/A_{SV} the reflection coefficient for the SV component. In the case of refracted waves, the ratio of the angles of polarisation is:

$$\frac{\tan \varepsilon_T}{\tan \varepsilon} = \frac{A_{SH}^T/A_{SH} \beta_2}{A_{SV}^T/A_{SV} \beta_1}. \quad (5)$$

A_{SH}^T/A_{SH} and A_{SV}^T/A_{SV} are the transmission coefficients and β_1 and β_2 denote the S -wave velocities in media 1 and 2, respectively. If the elastic parameters of the media are known, the changes of the angle of polarisation can be calculated.

After correction of the input data, the grid test procedure is made for every first-motion direction and angle of polarisation. The theoretical radiation value is compared with the measured value. If the first motion reading does not agree with the theoretical value or if the calculated angle of polarisation is not within the limits for the measured value, the orientation is rejected from the set of possible solutions. To avoid a strong influence of P readings close to nodal planes, which may be uncertain, solutions are kept as possible orientations if the theoretical radiation value for the P amplitude is smaller than a chosen fraction of the maximum amplitude. Normally a fraction of 5% is used. At the end of the test only those solutions, which are compatible with all measurements, remain.

The solutions are printed and displayed in a vectorplot (Pearce, 1977). In this plot, every combination of dip and slip angle is represented by a point in a grid. A possible solution is then indicated by a vector originating there and pointing in the direction of the fault-plane strike.

Application of the grid method

The grid method is applied to an earthquake ($M_L=4.1$) which occurred on the 13 July 1981 near Ibbenbüren in the north-west part of the Teutoburger Wald ($7^{\circ}42.5'$ E, $52^{\circ}15.7'$ N). The source orientation of this event has already been examined by the common graphical method in an earlier paper (Harjes et al., 1983), hereafter called Paper 1. In this solution, 19 P -wave first motions of European stations were used. The mechanism was interpreted as a normal fault on a $N207^{\circ}$ E-striking and either 60° NW- or 30° SE-dipping plane. Due to the lack of near stations, the position of the SE-dipping plane remained uncertain. The first-motion readings of two French stations disagreed with this solution.

In order to fit synthetic seismograms (see next section) to the recording of BUG, the crustal model of Paper 1 was modified. The wave velocities of the new model are given in Fig. 2. Take-off angles and corrections for ε were calculated on the basis of this model.

Input data for the grid test were the same first-motion readings as in Paper 1. If all 19 readings are used, none of the 93,312 solutions ($d=5^{\circ}$) remain possible, indicating that a pure double couple can not explain the observations. The two compressional readings at SSF and MZF have been omitted in a second test, as it was in the interpretation of Paper 1. In this case, 2152 solutions are compatible with

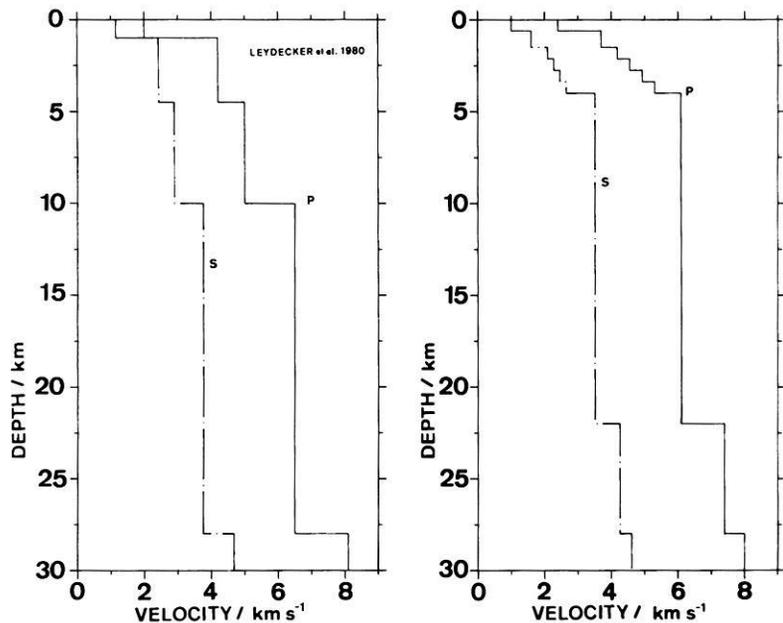


Fig. 2. Distribution of velocities of seismic waves. The *left model* (Leydecker et al., 1980) was used in Paper 1 (Harjes et al., 1983) and for the calculation of the synthetic seismogram in Fig. 7. The *right part* shows the modified model. This was used for the grid tests and synthetic seismograms of Fig. 8

the remaining observations. The solutions are displayed in the vectorplot of Fig. 3. The solution of Paper 1 is indicated in this figure by a small arrow and a capital 'A'. From this plot it becomes obvious that the solution of Paper 1 was subjectively selected by the interpreter from a great variety of possible solutions. In the vectorplot, most orientations with $\delta < 80^\circ$ in the range of thrust mechanisms and normal-fault mechanisms with $125^\circ \leq \delta \leq 150^\circ$ are rejected.

In order to use more information from the digital three-component registration at BUG, the angle or polarisation of the *S* wave was determined. Figure 4 shows the hodogram of the ground movement with a duration of 1.6 s. By calculating the eigenvalues of the covariance matrix (Kanasewich, 1975) the angle of polarisation is found to be $\varepsilon_{\text{BUG}} = 136^\circ$ with a rectilinearity of $R = 0.97$. ε_{BUG} was interpreted to belong to the S_g phase with a take-off angle $i = 36.5^\circ$. The angle $\varepsilon_{\text{BUG}} = 136^\circ \pm 15^\circ$ leaves 14,928 compatible solutions, which are shown in the vectorplot of Fig. 5.

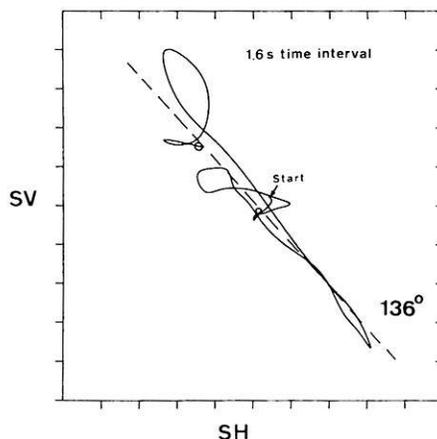


Fig. 4. Hodogram of the *S* wave of the 13 July, 1981, Ibbenbüren event derived from the seismogram of station BUG. The time interval is 1.6 s. The angle of polarisation is $\varepsilon = 136^\circ$ with a rectilinearity of $R = 0.97$

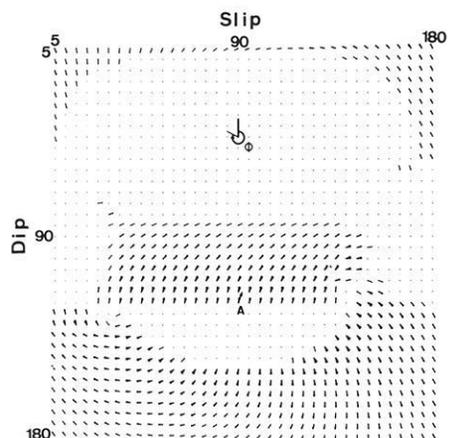


Fig. 3. Vectorplot of 2152 solutions which are compatible with 17 *P*-wave first-motion readings. The *small arrow* indicates the graphical solution of Paper 1 (see text for details). The *insert* shows the sense in which the strike of the fault planes, ϕ , is measured

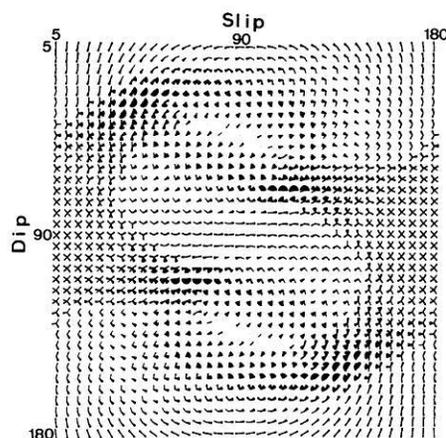


Fig. 5. Vectorplot of 14,928 solutions which are compatible with the angle of polarisation at BUG ($136^\circ \pm 15^\circ$)

A great variability of orientations becomes obvious. Fig. 6 shows the 102 solutions which agree with the first-motion readings and the angle of polarisation at BUG. The orientations form three ensembles in the vectorplot. The orientation from Paper 1 is indicated by a small arrow and a capital 'A'. Synthetic seismograms are calculated for the orientations B, C and D in the next section. A small deviation of take-off angle i does not change the results significantly. Even if the S phase is interpreted as S_n ($i=27^\circ$), the solutions B, C and D remain possible. The theoretical values of ϵ are still within the interval of $\Delta\epsilon = \pm 15^\circ$ centred at the measured value. The inserts in Fig. 6 give the projections of the lower focal hemisphere of the labelled solutions. Areas of compressional P -wave first motions are black, as usual.

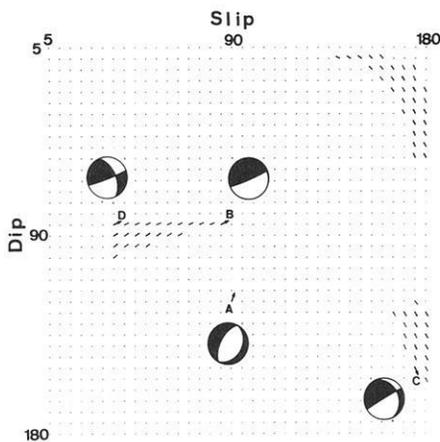


Fig. 6. Vectorplot of 102 solutions which are compatible with 17 P -wave first-motion readings and the angle of polarisation of the S wave recorded at BUG. Synthetic seismograms of those source orientations indicated by *small arrows* and *capital letters* are shown in Fig. 8. The *inserts* give the projections of the lower focal hemispheres for the labelled solutions. The *black quadrants* have compressional P -wave first motions

Synthetic seismograms

Synthetic seismograms of the 13 July, 1981, Ibbenbüren event have been calculated using the reflectivity method (Kind, 1978; 1979). The initial velocity model (Fig. 2) was that of Paper 1 after Leydecker et al. (1980). Additional information was provided by logs of the nearby Münsterland I well (Andres and Lichtenberg, 1963) and results of seismic refraction experiments (Mooney and Prodehl, 1978). The seismograms calculated with the velocity struc-

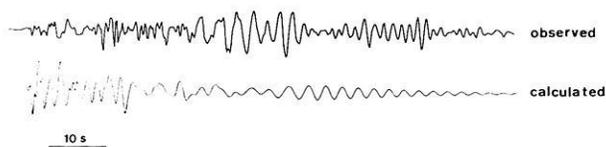


Fig. 7. The *upper trace* is the vertical component of the bandpass-filtered (0.05–2.0 Hz) displacement seismogram observed at BUG. The *lower trace* is a synthetic seismogram calculated with the source orientation from Paper 1 and the velocity structure after Leydecker et al. (1980). Misfits of amplitude ratios and within the surface waves are obvious

ture from Paper 1 showed reasonable agreement of theoretical and measured travel times of P and S waves at the station BUG. However, the surface wave trains observed at BUG could not be explained by this model. In Fig. 7, the bandpass-filtered (0.05–2.0 Hz) vertical component of the displacement seismogram from station BUG is compared with the synthetic seismogram calculated for the initial model. The filter was applied in order to get the same frequency content in measured and calculated seismograms. The reading of the polarity of P and the measurement of the angle or polarisation were made on the original seismograms with an upper frequency limit of 25 Hz (-3dB). The synthetic seismogram in Fig. 7 is dominated by body waves. The two surface wave groups of the observation are not visible.

From information provided by the log, the P -wave velocities were reduced to a value of 2.4 km/s in the upper 600 m and five layers were introduced in the depth range from 0.6 km to 4.0 km to form a velocity gradient. The depth of the layer with P -wave velocity greater than 6 km/s was lowered to 22 km. By making these changes, the observed surface waves could be modelled.

In order to fit the time interval between the P phase and the surface wave arrivals, the velocity ratio α/β was increased from $\sqrt{3}$ to values between 2.4 and 2.0 in the first 4 km and was left at a value of $\sqrt{3}$ below this depth. The shape and duration of surface waves turned out to be very sensitive to these changes. Finally, by making the described changes to the velocity model (Fig. 2) the observed arrival times of P and S and the surface waves could be explained quite reasonably.

Source depth and source orientation were varied in order to fit amplitude ratios within the seismogram. The amplitudes of the surface waves are strongly influenced by the depth of the source. A set of seismograms was calculated with source depths in the range from 0.1 km to 10 km. Hypocentres shallower than 0.5 km and deeper than 5.0 km could be excluded by these calculations. The best fits were achieved with a source in the second layer between 0.6 km and 1.5 km. This agrees with the results of the macroseismic survey described in Paper 1, where a source depth of 2.0 ± 0.5 km was found.

The amplitude ratio of P and S waves is a good indicator for a correct orientation of the double-couple source. In the grid test procedure, 102 orientations agreed with the measured P -wave first motions of 17 stations and the angle of polarisation of S at BUG. About 20 orientations were chosen from the three ensembles in the vectorplot of Fig. 6 to calculate synthetic seismograms. Three examples are shown in Fig. 8. The upper and lower traces in this figure show the registration of station BUG (same as in Fig. 7) for comparison. P and S arrival times are given by arrows. The horizontal bar in the upper trace indicates the time interval of the hodogram in Fig. 4. The traces denoted B, C and D are calculated with the modified model and the orientations of the double couple given for each seismogram within the figure. These orientations are indicated in the vectorplot (Fig. 6) by small arrows and the corresponding focal spheres. The seismogram A is calculated with the model from this paper and the source orientation from Paper 1. The misfit of the amplitude ratio of P and S is obvious in cases A, B and C. The orientations A, B and C show amplitude ratios of P to S which are greater than or equal to one. The measured ratio from the

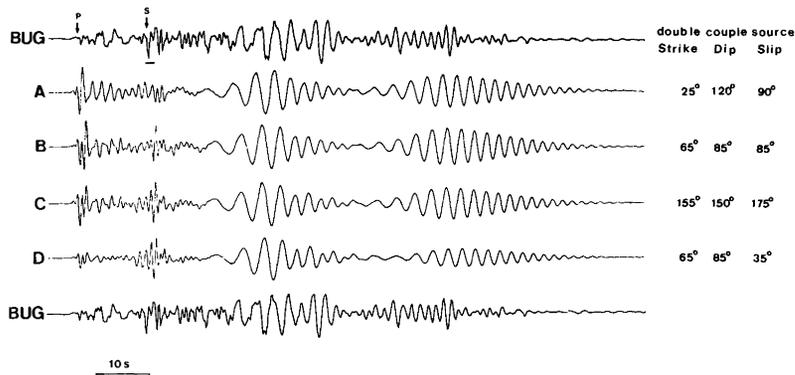


Fig. 8. Comparison of the seismogram recorded at BUG and synthetic seismograms for the 13 July, 1981, Ibbenbüren event. The orientation of the double-couple source for seismograms A, B, C and D is given next to each trace. A is the orientation derived in Paper 1 (see text for detail). The best fit is achieved with orientation D. The bar in the first trace shows the time interval of the hodogram in Fig. 4. The polarity of the traces A, B, C and D is reversed with respect to the measured seismogram

BUG seismogram is smaller than 0.5. Orientation D demonstrates the best fit within the possible solutions from the grid test. The projection of the nodal planes for solution D is shown in Fig. 9 (continuous lines). The dashed lines in the same figure give the position of nodal planes from the solution of Paper 1.

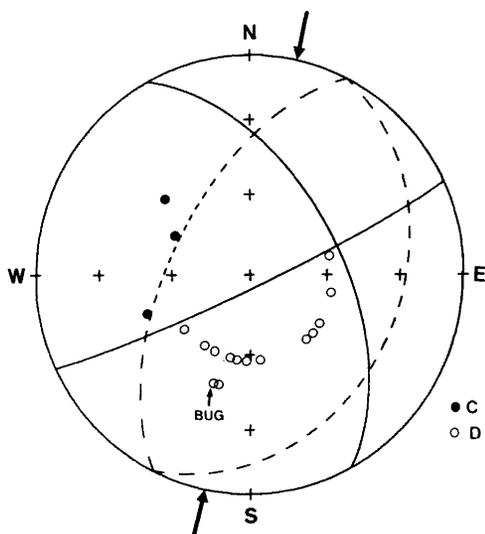


Fig. 9. Equal area projection of the lower focal hemisphere for the 13 July, 1981, Ibbenbüren event. Compressions and dilatations are shown by closed and open circles, respectively. The dashed lines are the nodal planes of the solution from Paper 1. The continuous lines give the orientation of the nodal planes of solution D from this paper. The arrows show the direction of the horizontal component of maximum compressive stress from this solution

Discussion and conclusions

The determination of fault-plane solutions for regional seismic events is problematic if only a few reliable *P*-wave first-motion readings are available. In such cases the solutions are often affected by the subjective influence of the interpreter.

A systematic search for solutions in the space of source orientations gives all orientations which agree with the data. The parameters derived from the seismograms may be polarities as well as amplitude ratios or angles of polarisation. In this paper, for example, polarities of *P*-wave first motions and angles of polarisation of *S* waves are used. The selected source orientations can be used for calculation of

synthetic seismograms. In this way, the amount of calculation time is normally smaller than if the source orientation is searched for only by forward modelling with synthetic seismogram techniques in a trial-and-error sense.

An $M_L=4.1$ event (13 July, 1981) at Ibbenbüren (FRG) is used to test the applicability of the procedure. Seventeen *P*-wave first motions and one angle of polarisation leave 102 solutions out of a possible 93,312 source orientations, if the increment of the angles of orientation is 5°. An earlier interpretation of the double-couple orientation on the basis of the classical graphic fault-plane solution technique resulted in a normal fault mechanism on a N207°E-striking and 60°NW- or 30°SE-dipping plane. This orientation is one of 2152 solutions which agree with the 17 *P*-wave first motions, but it disagrees with the angle of polarisation of *S* at BUG. A disagreement is also obvious between the measured seismogram and the calculated seismogram for the orientation from Paper 1. While the amplitude ratio of *P* to *S* is greater than one in the synthetic seismogram, it is smaller than 0.5 in the measured one. The amplitude ratio of *P* to *S* is strongly influenced by the orientation of the double-couple source. The best fit is achieved with solution D, which is a sinistral strike-slip motion on the ENE-WSW-striking plane or a dextral motion on the NNW-SSE-striking plane.

In Paper 1, an activation of the graben structure within the 'Ibbenbürener Bergplatte' was assumed to be the tectonic explanation for the earthquake. The main argument for this interpretation was the parallel trend of the assumed fault plane of solution A and the strike of the tectonic elements in the region. Solution D from this paper seems more trustworthy to the authors, due to the grid test and synthetic seismograms. This makes the former interpretation questionable. The direction of the horizontal component of maximum compressive stress is N13°E in solution D, indicated by two arrows in Fig. 9. This direction deviates significantly from the value of the mean direction for central Europe (N142°E ± 20°), which was derived by Ahorner (1975) from earthquake mechanisms and in situ stress measurements. In situ stress measurements during the last decade show directions approximately EW for boreholes within the Lower Saxonian Block (Rummel et al., 1983). Among these, a hydraulic fracturing experiment close to Bielefeld (Rummel and Baumgärtner, 1982) was the nearest to the 'Ibbenbürener Bergplatte'. The direction of the maximum horizontal stress was found to be N108°E. Therefore one would expect, from the fault-plane solution, a direction closer to the regional stress field direction for central Europe. Since it seems unlikely that the present local stress

field is still dominated by the tectonic evolution of the 'Ibbenbürener Bergplatte' and considering the shallow focal depth (1.5–2.0 km), a possible connection between mining activity in the area and earthquake occurrence can not be fully ignored. Mineworks within the 'Ibbenbürener Bergplatte' range down to a depth of more than 1 km. The influence of mineworks on the local stress field components could explain the deviating direction of maximum compressive stress.

Acknowledgements. The authors would like to thank H.-P. Harjes, F. Rummel and an anonymous reviewer for helpful comments. We are grateful to R. Kind for providing the reflectivity program. The grid tests were calculated on a HP 1000 computer of the Institute of Geophysics, while the synthetic seismograms were calculated with a CDC 175 of the 'Rechenzentrum der Ruhr-Universität'.

References

- Ahorner, L.: Present day stress field and seismotectonic block movements along major fault zones in Central Europe. *Tectonophysics*, **29**, 233–249, 1975
- Aki, K., Richards, P.G.: *Quantitative seismology*. Vol. 1. San Francisco: W.H. Freeman, 1980
- Andres, J., Lichtenberg, K.: Die Auswertung der seismischen Voruntersuchungen auf Grund der durch die Bohrung Münsterland I gewonnenen Daten. *Fortschr. Geol. Rheinl. und Westfalen*, **11**, Krefeld, 1963
- Båth, M.: Polarisation of transverse seismic waves. *Geophys. J. Roy. Astron. Soc.* **4**, 106–123, 1961
- Dillinger, W.H., Harding, S.T., Hope, A.J.: Determining maximum likelihood body wave focal plane solutions. *Geophys. J. Roy. Astron. Soc.* **30**, 315–329, 1972
- Engdahl, E.R., Kisslinger, C.: Seismological precursors to a magnitude 5 earthquake in the central Aleutian islands (supplement). *J. Phys. Earth* **25**, 243–250, 1977
- Fuchs, K.: Das Reflektions- und Transmissionsvermögen eines geschichteten Mediums mit beliebiger Tiefenverteilung der elastischen Parameter und der Dichte für schrägen Einfall. *Z. Geophys.* **34**, 389–413, 1968
- Fuchs, K., Müller, G.: Computation of synthetic seismograms with the reflectivity method and comparison with observations. *Geophys. J. Roy. Astron. Soc.* **23**, 417–433, 1971
- Harjes, H.-P., Hinzen, K.-G., Cete, A.: Das Erdbeben bei Ibbenbüren am 13. Juli 1981. *Geol. Jb.* **E26**, 65–76, 1983
- Hinzen, K.-G.: Vergleich von Herdflächenlösungen und Momententensoren. *Ber. d. Inst. f. Geophysik d. Ruhr-Universität Bochum*, Reihe A15, 1984
- Kanasewich, E.R.: *Time sequence analysis in geophysics*. Edmonton: Univ. of Alberta Press 1975
- Kisslinger, C.: Evaluation of S to P amplitude ratios for determining focal mechanisms from regional network observations. *Bull. Seismol. Soc. Amer.* **70**, 999–1014, 1980
- Kisslinger, C., Bowman, J.R., Koch, K.: Procedures for computing focal mechanisms from local (*SV/P*)_z data. *Bull. Seismol. Soc. Amer.* **71**, 1719–1729, 1981
- Kind, R.: The reflectivity method for a buried source. *J. Geophys.* **44**, 373–383, 1978
- Kind, R.: Extensions of the reflectivity method. *J. Geophys.* **45**, 373–383, 1979
- Koning, L.P.G.: On the mechanism of deep focus earthquakes. *Gerl. Beitr. Geophys.* **58**, 159–197, 1942
- Langston, C.A.: Single station fault plane solutions. *Bull. Seismol. Soc. Amer.* **72**, 729–744, 1982
- Leydecker, G., Steinwachs, M., Seidl, D., Kind, R., Klassmann, J., Zerna, W.: Das Erdbeben vom 2. Juni 1977 in der norddeutschen Tiefebene bei Soltau. *Geol. Jahrb. Reihe E* **18**, 3–18, 1980
- Mooney, W.D., Prodehl, C.: Crustal structure of the Rheinisch Massif and adjacent areas, a reinterpretation of existing seismic refraction data. *J. Geophys.* **44**, 573–601, 1978
- Pearce, R.G.: Fault plane solutions using relative amplitudes of *P* and *pP*. *Geophys. J. Roy. Astron. Soc.* **50**, 381–394, 1977
- Pearce, R.G.: Earthquake focal mechanisms from relative amplitudes of *P*, *pP* and *sP*: method and computer program. *AWRE Report No 041/79*, 1979a
- Pearce, R.G.: Fault plane solutions using relative amplitudes of *P* and surface reflections: further studies. *Geophys. J. Roy. Astron. Soc.* **60**, 459–487, 1979b
- Pearce, R.G.: Complex *P*-waveforms from a Gulf of Aden earthquake. *Geophys. J. Roy. Astron. Soc.* **64**, 187–200, 1981
- Ritsema, A.R.: (I, Delta)- curves for bodily seismic waves of any focal depth. *Lemb. Meteor. dan Geof. Djakarta*, **54**, 1–25, 1958
- Rummel, F., Baumgärtner, J.: Spannungsmessungen im östlichen Bereich der südwestdeutschen Scholle. *Rep. Contr. No. RUB 7084408-82-3*, Bochum, 1982
- Rummel, F., Baumgärtner, J., Alheid, H.J.: Hydraulic fracturing stress measurements along the eastern boundary of the SW-German Block. In: *Hydraulic fracturing stress measurements, proceedings of a workshop, Dec. 2–5, 1981*, National Academy Press, Washington, 1983
- Shapira, A., Båth, M.: Some mechanism determinations of short distance microearthquakes. *Seis. Inst. Uppsala Rep.*, 1/78, 1–25, 1978
- Stevens, A.E.: Earthquake mechanism determined by *S*-wave data. *Bull. Seismol. Soc. Amer.* **54**, 2034–2047, 1964
- Udias, A.: A least squares method for earthquake mechanism determination using *S*-wave data. *Bull. Seismol. Soc. Amer.* **54**, 2037–2047, 1964

Received August 2, 1984; revised version December 6, 1984

Accepted December 7, 1984

Active seismic layers and crustal structure in some Italian regions

R. Cassinis, P. Mazzoni and A. Ranzoni

University of Milano, Dept. of Earth Science, Geophysical division, 7, Via L. Cicognara, 20129 Milano, Italy

Abstract. Following a general comparison between the hypocentres and crustal structure derived mainly from deep seismic soundings, a more detailed analysis is attempted in some areas of the Italian region. This study is undertaken by defining a region as homogeneous when the same type of velocity-depth functions has been obtained. All available earthquake hypocentres within the area are used for the correlation with the V_p distribution in order to obtain a sufficiently large number of data. Where available, low magnitude events and aftershock sequences have also been used. The conclusions already derived (Cassinis et al., 1984) are more clearly confirmed. In the south-eastern Alps (Friuli), the main earthquake sources are in the upper crust, between 5 and 15 km, at the top of the first velocity inversion zone. In Tuscany (thin continental, transitional crust), the seismicity is weak and very shallow. In the Calabrian arc, the main seismicity seems to originate in the lower crust or at the top of the lithospheric “lid” which appears to be decoupled from the overlying, highly mobilized formations of the crust. Tentative explanations of these different regimes are given with regard to the influence of factors that can determine the transition from brittle to quasi-plastic behaviour.

Key words: Seismogenesis – Brittle-plastic transition – Velocity functions – Inversion zones

Introduction

In a previous paper (Cassinis et al., 1984) the available hypocentres of earthquakes with $M > 3.5$ were correlated with individual functions of V_p obtained using wide angle reflection seismic profiles (deep seismic soundings = DSS).

The correlation is made with respect to hypocentres contained in a vertical cylinder with a radius of 40 km around the average position of the calculated velocity function. The seismic activity is described by histograms of the distribution with depth of the number of foci, the maximum magnitude and the released energy during the historical period (1000–1981 A.D.).

There are strong limitations and constraints in this general analysis due to the inequality of the data. The number of available hypocentres is often too small to allow reliable statistics and the accuracy of the focal depth is very poor for the old data (before 1970). Also, the interpretation of the velocity functions is sometimes debatable. However, the results led to some preliminary considerations, namely (Fig. 1):

- In north and central Italy the most active seismic layer seems to be located in the upper crust, near the top of the first zone of velocity inversion. The seismic activity in the middle-lower crust is very low everywhere or, at least, high magnitude earthquakes do not originate there.

- In the south, especially along the Calabrian arc, the main activity seems to originate at the “M” discontinuity, while the low seismicity in the upper crust could correspond to a low strength material and, possibly, a decoupling could exist between the crust and upper mantle.

- In other words, in north and central Italy the upper crust seems more brittle than in the southern section. Here, in turn, the upper mantle appears to be stronger or, at least, it seems that stresses are not transmitted to the upper crust.

- Another point is that, considering the interpretation of the crustal structure according to the DSS results (Fig. 2), the main seismicity in northern and central Italy takes place on the lifted edge of the foreland crust (Adriatic), while in south Italy and in Sicily it occurs along the lifted edge of the hinterland crust (South Tyrrhenian) (see also Cassinis, 1983; Cassinis et al., 1979; Giese and Morelli, 1975).

In view of these preliminary results (Cassinis et al., 1984), in this study we have considered the need for more detailed surveys in particular areas where adequate data are available. The models of the crust, designed according to the behaviour of the seismic velocity V_p , have been improved by the introduction of hypotheses on other parameters (temperature and type of material). On the other hand, aftershock series and microseismicity have also been used, when available, for the correlation with the velocity functions in order to increase the number of foci. The behaviour of the main shocks has been compared to that of the microseisms and aftershocks.

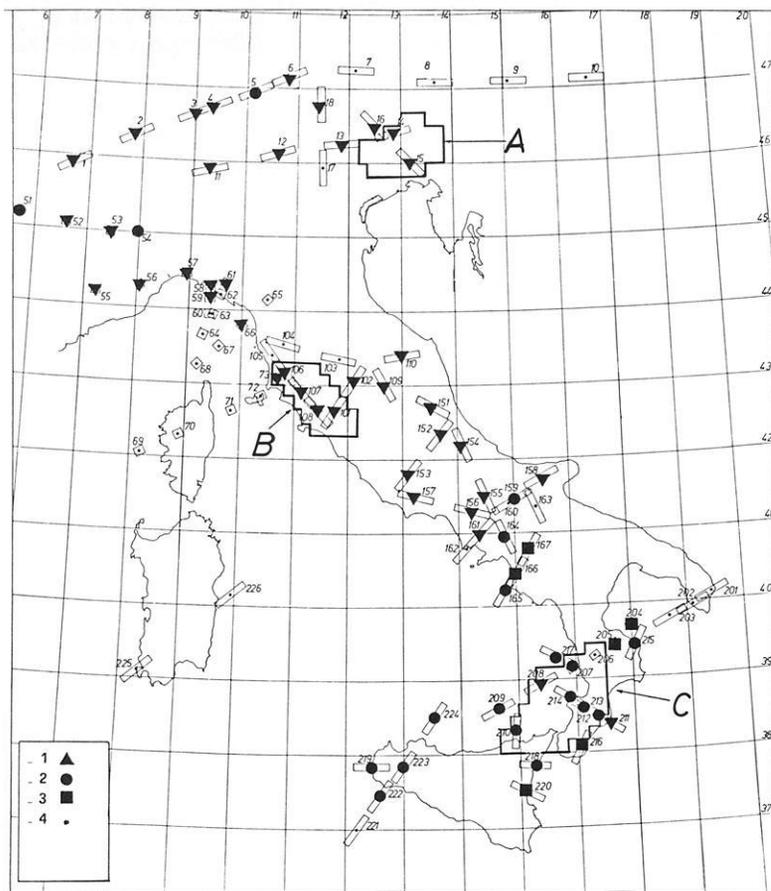


Fig. 1. Position map of velocity functions derived from seismic profiles. The small *dots* represent the position of the half-critical distance for the "M" surface; the *rectangles* (oriented according to the strike of the seismic profile) indicate the linear range of velocity distribution; the *squares* correspond to a vertical distribution. Comparison between seismic activity and crustal structures. The most active seismic layers are located: 1) at the top of the first zone of velocity inversion (upper crust); 2) near the "M" discontinuity; 3) near the upper transition to the "M" (in the case of "double" transition); 4) lack of information on seismicity (Cassinis et al., 1984). The three areas examined in detail in the present study are indicated: A) Friuli; B) south Tuscany; C) Calabrian arc

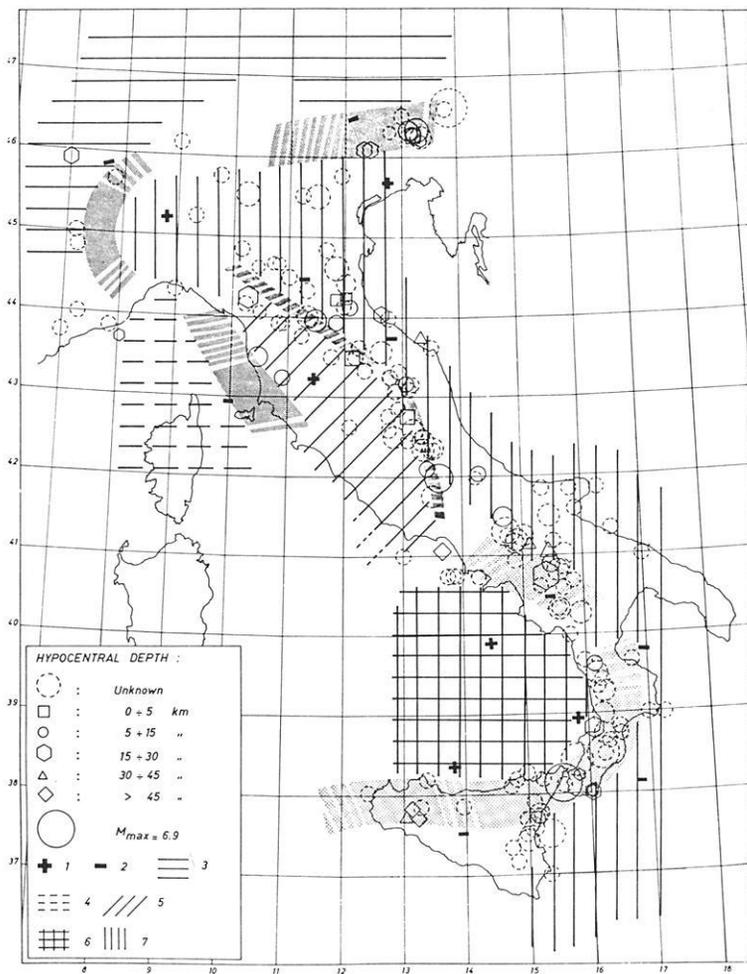


Fig. 2. The major lateral discontinuities and crustal domains compared with the map of strong earthquakes from 1000 to 1975 A.D. ($M > 5.5$). 1) and 2) Lifted or depressed "M". Domains: 3) European-Alpine; 4) Corsican-Ligurian; 5) "Tuscan" intermediate crust; 6) South Tyrrhenian; 7) Adriatic-African. The *broad bands* indicate the zones of "double" "M" transition; the *narrow strip* corresponds to a "flexure" of the "M" (Cassinis et al., 1984)

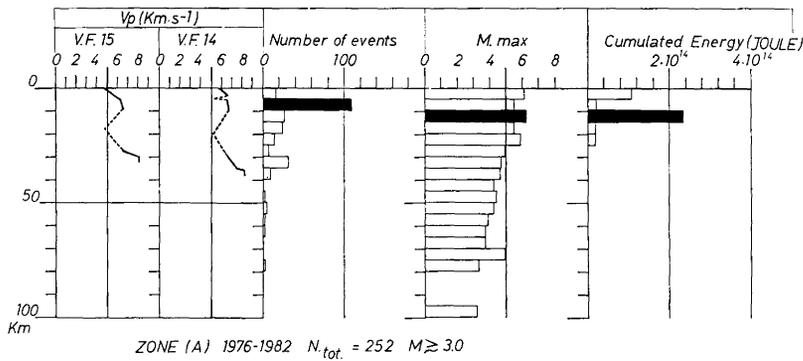


Fig. 3. Velocity functions of zone A (Friuli) (Italian Group for explosion seismology) compared with the seismic activity during the period 1976-1982 (hypocentres with $M > 3.0$ recorded by the Istituto Nazionale di Geofisica), 252 events. From left to right: a) velocity functions; b) number of events in steps of 5 km; c) highest magnitude recorded in each step; d) cumulated energy in each step (Joule), linear scale. *Black bars*: depth interval with maximum values. Energy values calculated after the Bath formula

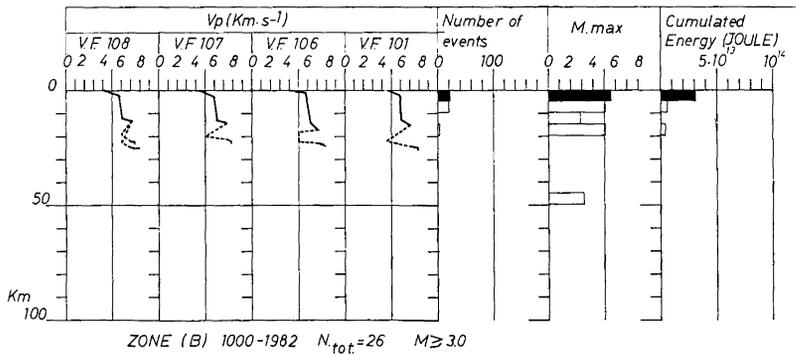


Fig. 4. Same as in Fig. 2 for zone B (south Tuscany). Period of observation: 1000-1982 A.D. Total number of events: 26

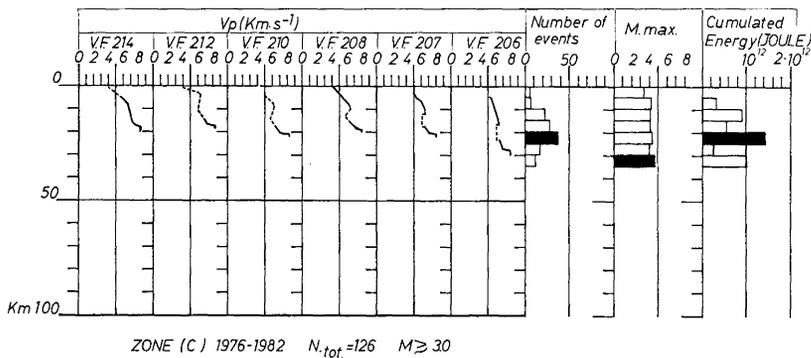


Fig. 5. Same as in Fig. 2 for zone C (Calabrian arc). Period of observation: 1976-1982. Total number of events: 126

Investigation in particular areas

Three areas have been selected for more detailed analysis (Fig. 1): Friuli (A) on the margin of the southern Alps; central and southern Tuscany (B); the Calabrian arc (C). The velocity functions as well as the foci falling within the area are used. For the analysis. Each area is defined as a homogeneous structural unit, when the velocity functions within its boundaries show a similar behaviour.

In Fig. 3, two velocity functions and histograms of the observed seismic activity during the period 1976-1982 (252 events) are plotted for zone A (Friuli). The depth interval is 5 km. The hypocentres considered here ($M \geq 3$) do not include aftershock sequences or small magnitude events. The correlation seems clear; the number of events shows a maximum between 5 and 10 km while the M_{\max} and the cumulated energy reach the highest value one step below, at the top of the velocity inversion. The seismic activity is very weak in the middle and lower crust. The shift of 5 km between

the first and the other two maxima can be explained by the high number of small energy events recorded in the period (Siro and Slejko, 1982).

In Fig. 4, four velocity functions in central-southern Tuscany (zone B, a geothermal area) are plotted. All data indicate a thinned continental crust (total thickness about 22 km) with a low velocity gradient in the upper crust and a velocity inversion in the lower crust. Unfortunately, local seismographic networks do not operate there, so only the catalogue and bulletins of the low density national observatories are available. The number of plotted events is very small. However, the focal depths of the observed events agree with the macroseismic data of the area, indicating very shallow earthquakes (ENEL catalogue, unpublished). The most active layer seems to be located within the uppermost 5 km.

Finally, Fig. 5 illustrates six velocity functions of zone C (Calabrian arc); all functions show a large thickness of low-velocity material overlying a zone of strong velocity gradient. The velocity inversion seems

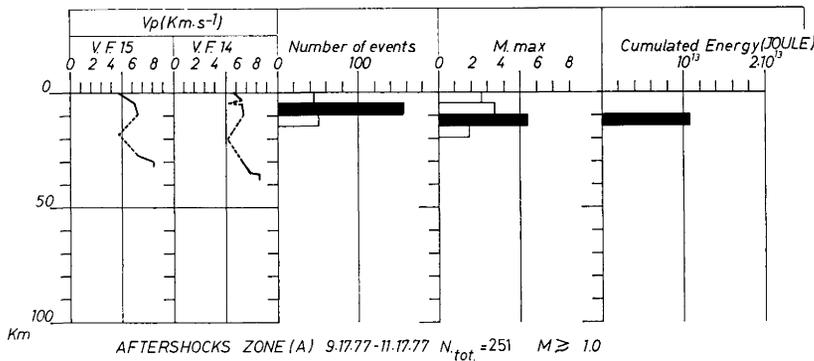


Fig. 6. Histograms of an aftershock sequence in zone A, following a main shock of $M = 5.4$ (251 events). Same symbols as in Fig. 3 and the following figures. Compare with Fig. 3

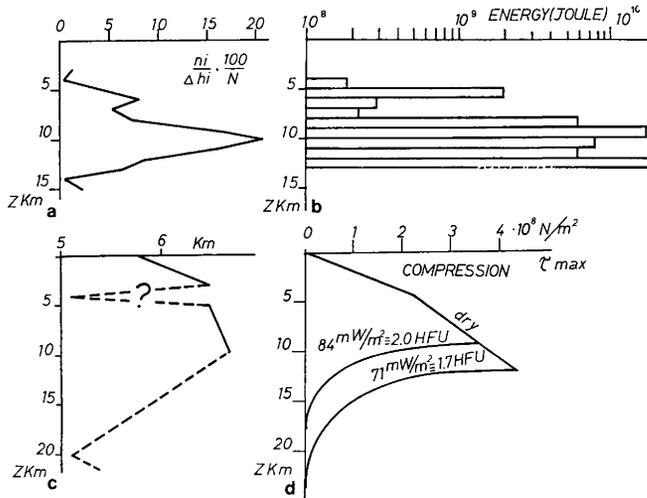


Fig. 7a-d. Aftershock sequence of 16 September-17 November, 1977 (same as Fig. 6). **a** normalized frequency-depth curve (depth interval: 1 km); **b** cumulated energy in each interval (logarithmic scale); **c** velocity function 14 (upper part - Fig. 6); **d** maximum stress curves (after Meissner and Strehlau, 1982). Dry quartz rheology, compression mechanism

to be located in the lower crust and is generally weak. The transition to the Moho is fast and is found at depths ranging between 18 and 25 km. The main seismicity seems to be concentrated between 20 and 25 km, i.e. at the “ M ” transition zone or in the uppermost mantle. The foci with depths > 30 km (subduction slab) were not considered.

Study of Aftershocks in zone A (Friuli)

An aftershock sequence containing 251 events, recorded by the local seismographic network (O.G.S., Trieste) (Suhadolc, 1982) during the 60 days following a moderate event ($M = 5.4$) that occurred on 16 September 1977 at a depth of 14 km (± 2 km), was analysed. The sequence took place in the same area of the destructive shock of 16 May 1976 with $M_s = 6.4-6.5$ at an estimated depth of about 10 km. The minimum recorded magnitude is about 1.0. Only those events with an estimated accuracy of local depth better than ± 3 km were used for the analysis.

In Fig. 6 the correlation between the velocity functions and the seismic activity is shown again, using the aftershock sequence. The peaks of the histograms correspond exactly to those of Fig. 3 (1976-1982).

The number of hypocentres has also been represented (Fig. 7) with a normalized frequency-depth curve (“defreq” diagram, Meissner and Strehlau, 1982) using depth intervals of 1 km. The peak of the curve is found at a depth of 10 km, corresponding to the top of the inversion zone.

The depth range of the aftershocks is between about 14 (main shock) and 4 km. According to Shebalin (1976) and other authors, the thickness of the active layer increases very slightly with magnitude. A magnitude increase from 6.0 to 6.5 corresponds to an increase in thickness of about 2 km (from 10 to 12 km).

Models of the continental crust and the seismic regime

The tentative models of the crustal structure in Italy are based mainly on the velocity distribution of P waves as interpreted from DSS profiles. Other complementary data, such as gravity or magnetic anomalies, are considered when significant. Unfortunately, one key factor, the heat flow value, is generally not available, especially inland, except in central and southern Tuscany.

For the following considerations, it is assumed, in a general sense that, in the continental crust, the upper part of the velocity functions (gradient decreasing with depth) corresponds to the upper crust (the top of the “granitic” layer corresponding to $V_p = 5.8-6.0$ km/s). In general, the first velocity inversion is considered as marking the transition from the upper to a “middle” crust. The last part of the curve, where an increasing velocity gradient is generally observed, corresponds to the lower crust and the transition to the upper mantle, where V_p reaches about 8 km/s.

The Italian region is a transitional area where the structure and, consequently, the velocity distribution with depth are subjected to strong lateral variations. Therefore, the crustal type is often very different from the model proposed for the “continental” crust (see, for instance, Mueller, 1977).

In all the proposed models, the upper and middle crust is considered as “granitic” or having a behaviour similar to quartzitic rocks. On the basis of this assumption, and taking experimental and theoretical results into account, several authors (Byerlee, 1968; Brace and Kohlstedt, 1980; Sibson, 1982; Meissner and Strehlau, 1982) have calculated the limits of stresses in the con-

tinental crust. The variables are temperature and pressure; their increase determines a transition from a frictional regime to a quasi-plastic flow that induces stresses in the rigid formation above.

Crustal strength-depth curves are calculated for different quartz rheologies and considering both dry and wet conditions (no pore pressure or hydrostatic pore pressure). Also, the different types of faulting (compression, extension or strike slip) are taken into consideration. According to these model parameters, the peak of the maximum stress marks the transition between the brittle and ductile behaviour. The peak range is quite wide: in the case of a wet crust, heat flow = 2 HFU and extension faulting, the peak is at 5 km and $\tau_{\max} = 10^7 \text{ N/m}^2$. For a dry crust, heat flow = 1.2 HFU and compression, the peak is at 32 km and $\tau_{\max} = 9 \times 10^8 \text{ N/m}^2$ (see for example Meissner and Strehlau, 1982).

Let us apply these "stress max" curves to the model of the upper-middle crust in zone A (Fig. 7).

Considering the independent information on the focal mechanism of the earthquakes in this area, we can assume a compressive stress. It is more difficult to choose the other parameters involved (wet or dry rheology, creep rate and heat flow).

Only on the basis of the depth of the peak of the depth-frequency diagram (Fig. 7a) and of its value (consistent with the largest known earthquakes in the area) do we assume a dry quartz rheology and a creep rate $\dot{\epsilon} = 10^{-17} \text{ s}^{-1}$.

As far as the heat flow is concerned, and also considering the following section, we estimate a range between 1.7 and 2.0 HFU (corresponding to 71 and 84 mW/m^2 , respectively) as reasonable. With these assumptions, the peak of τ_{\max} ranges between about $3.8 \times 10^8 \text{ N/m}^2$ at $Z = 10 \text{ km}$ and $4.5 \times 10^8 \text{ N/m}^2$ at $Z = 14 \text{ km}$ (Fig. 7d).

To make only a rough estimate, we can transform these two values of the peak of τ_{\max} into the maximum distortional strain energy per unit volume, E_d (Sibson, 1982):

$$E_d = \frac{[(\sigma_1 - \sigma_3)^2 \cdot (1 - K + K^2)]}{6\mu}, \quad (1)$$

where

$$0 < K = \frac{(\sigma_2 - \sigma_3)}{(\sigma_1 - \sigma_3)} < 1.$$

If we take

$$\sigma_2 = \frac{\sigma_1 + \sigma_3}{2},$$

then $E_d = \frac{\Delta\sigma^2}{8\mu}$, where μ is the rigidity.

According to the Mohr formula (see Terzaghi and Peck, 1948)

$$\tau_{\max} = \frac{\sigma_1 - \sigma_3}{2} = \frac{\Delta\sigma}{2},$$

we obtain

$$E_d = \frac{\tau_{\max}^2}{2\mu}. \quad (2)$$

Assuming an average density for granitic rocks of $\rho = 2.7 \text{ gr/cm}^3$ and $V_p = 6.2 \text{ km/s}$, we have the rigidity $\mu = 3.45 \times 10^{11} \text{ dyne/cm}^2$. Substituting these values in Eq. (2) we find, for the two peaks:

$$E_{d1} = 2.93 \times 10^6 \text{ Joule/m}^3 \quad \text{for } \tau_{\max} = 4.5 \times 10^8 \text{ N/m}^2$$

and

$$E_{d2} = 2.09 \times 10^6 \text{ Joule/m}^3 \quad \text{for } \tau_{\max} = 3.8 \times 10^8 \text{ N/m}^2.$$

In order to estimate the radiated energy, we have to account for the seismic efficiency η . Assuming a coefficient η of 0.1 (Lomnitz, 1974) and taking the average value between E_{d1} and E_{d2} we arrive at $E_{\text{d rad}} = 2.51 \times 10^5 \text{ Joule/m}^3$. This value seems to be compatible with the size of the largest earthquake recorded in historical times (16 May 1976). Taking a focal volume for this earthquake:

$$V_f = A \cdot \bar{u} = 400 \text{ km}^2 \times 33 \text{ cm} = 1.32 \times 10^{14} \text{ cm}^3$$

(Cipar, 1980) we calculate a total energy of:

$$E_{\text{tot}} = 3.31 \times 10^{13} \text{ Joule} = 3.31 \times 10^{20} \text{ erg},$$

a value which corresponds fairly well to the published values (Cipar, 1980).

In conclusion, the choice of the "stress max" curve seems well supported by three independent sets of data, namely: the depth of the bottom of the active layer, the energy produced by the largest known earthquakes and the depth of the velocity inversion. Of course, the above considerations do not allow us to forecast the M_{\max} to be expected in the area; however, they contribute to the definition of the active layer and of the most probable depth of the nucleation of the strongest earthquakes. According to both the V_p functions and to the hypocentres, this depth should be between about 10 and 14 km.

Heat flow and velocity inversion in zone A

The selection of a heat flow ranging between 1.7 and 2.0 HFU (71 and 84 mW/m^2) is supported by the velocity-depth functions. The interpretation of velocity inversions based on the DSS data contains many ambiguities. However, the top of the first inversion is generally fairly well determined by the termination of the P_g travel-time curve, while the thickness and velocity distribution within the inversion zone are both unknown. Christensen (1979) has calculated the critical thermal gradient (i.e. the thermal gradient that balances the increase of velocity due to the increase of pressure) for various types of rocks.

If we consider a homogeneous granitic upper crust, we see from Christensen's values that at 10 km depth a very high heat flow is required to produce a velocity inversion, unless other factors such as pore pressure or

change of material properties are involved. Thus, it seems reasonable to consider a heat flow higher than the average in order to explain, at least partially, the low-velocity zone. Also, geological considerations support these high values, the area being located on the margin of the Alpine orogenic belt (Mongelli, 1983).

Zone B (South Tuscany)

The velocity functions of Fig. 4 do not show inversions in the upper crust. The velocity of 6.0 km/s is reached at a relatively small depth (between 1 and 3 km). Instead, a strong inversion is seen in the lower crust, probably at the "M" transition. This has been interpreted as evidence of high temperatures in the lower crust (Giese et al., 1980).

As already mentioned, the number of foci is too small to attempt an interpretation as for zone A. One possibility seems to be that the high temperature at the "M" transition could lower the strength of the material of the upper mantle thus producing a ductile behaviour there also. According to Calcagnile and Panza (1981), the lithospheric "lid" seems to be very thin and "weak" here.

Zone C (Calabrian Arc)

The results, already illustrated in Fig. 5, seem to support the idea of a decoupling of the upper crust which is formed by "crystalline" nappes overlying the sedimentary layers, both being highly mobilized, allochthonous formations.

The strong seismicity concentrated in the lower crust or at the top of the mantle lid, suggests a strong and relatively cold lower lithosphere; this could also be related to the subduction environment.

Discussion and conclusions

The increase of available information, both in hypocentres and crustal structure, stimulates new and more precise studies on the seismogenesis of shallow earthquakes in the continental crust. The crustal models that are proposed nowadays, in some way repeat the model of the dynamic relationship between the lithosphere and the asthenosphere, the former being considered as a "sandwich" formed by several "slices" alternatively brittle and ductile (Chen and Molnar, 1983; Sibson, 1982).

The earthquakes should originate at the transitions between one slice and another, the primary source being located in the deepest levels, probably at the lithosphere-asthenosphere boundary.

Several factors are responsible for this behaviour. For a fairly uniform type of material (as e.g. supposing the upper crust to be formed of quartzitic rocks), the main variables seem to be temperature and pressure as well as the ratio between the pore pressure and external pressure. When the type of material changes, as in olivine rocks, the models should be modified accordingly. So far, suitable models have been proposed only for the "granitic" rocks of the upper crust: shear resistance versus depth profiles have been calculated, also considering the type of rupture (compression, extension, strike-slip mechanism).

In the Italian region this type of investigation faces many difficulties. First of all, the lithospheric structure is subjected to very strong lateral discontinuity (see Fig. 2); second, the direction of stress is often variable from the upper to the lower crust; third, accurate hypocentral locations by dense seismographic networks are available only from the past two or three years and only in some areas; in most areas the accuracy is not sufficient to observe reliable "fluctuations" of the thickness of the active layer. The use of aftershock sequences would help considerably and the results obtained in zone A seem encouraging. However, further experiments are needed.

Nevertheless, we believe that the experimental findings that have been shown are a good starting point. We consider the comparison with the distribution of *P*-wave velocities with depth to be especially useful. In spite of many constraints, the crustal structure derived from the *P* velocity functions, seems to agree with the behaviour of the seismic activity, provided that the latter is accurately determined. On the other hand, these results support the interpretation of DSS by adding proof of the existence of velocity inversions, which have in the past, been the subject of very long controversies.

It is also apparent that the negative velocity gradient within these low-velocity zones is generally larger than that predicted by models (Christensen, 1979) which consider only the increase of temperature and pressure. This is probably due, in part, to the ambiguity involved in the interpretation, but could also be additional evidence for the need to take other factors, such as pore pressure and change of material properties, into consideration.

It is clear that other types of data would be useful, as, for instance, *S*-wave velocities obtained both from the inversion of passive seismic data and from seismic active profiles.

In conclusion, the above results have confirmed that the shallow seismicity in the Italian area is strongly dependent on crustal and lithospheric structure. In zone A, the seismic active layer is clearly directly correlated to the boundary between a brittle upper crust and a weaker (or more plastic) middle-lower crust. In zone B the regime (very shallow foci) seems dependent on the high temperature of the lower crust or of the mantle itself. Finally, in zone C, the seismicity seems to originate in the lower crust or in the lithospheric lid and, therefore, it appears as more directly related to the upper mantle structure.

The study will be extended to other areas, as soon as suitable data become available, in order to verify the general trend described in Fig. 1.

Acknowledgements. Thanks are due to Istituto Nazionale di Geofisica - ING - (E. Boschi and C. Gasparini) for releasing the data on the recent seismic activity in Italy and to Osservatorio Geofisico sperimentale of Trieste (D. Slejko) for the aftershock sequence of Friuli.

References

- Brace, W.F., Kohlstedt, D.I.: Limits of lithospheric stress imposed by laboratory experiments. *J. Geophys. Res.* **85**, 6248-6252, 1980

- Byerlee, J.D.: Brittle-ductile transition in rocks. *J. Geophys. Res.* **73**, 3629–3648, 1968
- Calcagnile, G., Panza, G.F.: The main characteristics of the lithosphere-asthenosphere system in Italy and surrounding regions. *Pure Appl. Geophys.* **119**, 866–879, 1980–81
- Cassinis, R.: Seismicity and crustal structure in the Italian region: a preliminary zoning. *Boll. Geofisica Teor. Appl.* **25**, 3–26, 1983
- Cassinis, R., Franciosi, R., Scarascia, S.: The structure of the Earth's crust in Italy: a preliminary typology based on seismic data. *Boll. Geofisica Teor. Appl.* **21**, 105–126, 1979
- Cassinis, R., Scarascia, S., Zini, E.: Shallow seismicity and seismic velocity distribution as determined by DSS in the Italian region. *Boll. Geofisica Teor. Appl.* **26**, 49–59, 1984
- Chen, W.P., Molnar, P.: Focal depth of intracontinental and intraplate earthquakes and their implications for the thermal and mechanical properties of the lithosphere. *J. Geophys. Res.* **88**, 4183–4214, 1983
- Christensen, N.I.: Compressional wave velocities in rocks at high temperatures and pressures, critical thermal gradients and crustal low-velocity zones. *J. Geophys. Res.* **84**, 6849–6857, 1979
- Cipar, J.: Teleseismic observations of the 1976 Friuli, Italy, earthquake sequence. *Bull. Seismol. Soc. Amer.* **70**, 963–983, 1980
- Giese, P., Morelli, C.: Crustal structure in Italy. *Quaderni della ricerca scientifica, CNR*, 90, 1975
- Giese, P., Morelli, C., Nicolich, R., Wigger, P.: Seismic studies for the determination of crustal structure in the area of the geothermal anomaly in Tuscany. 2nd Int. seminar of the results of E.C. geothermal energy research, 4–6 March, 1980
- Lomnitz, C.: *Global tectonics and earthquake risk*. Elsevier, 1974
- Meissner, R., Strehlau, J.: Limits of stresses in continental crusts and their relation to depth – frequency distribution of shallow earthquakes. *Tectonics* **1**, 73–89, 1982
- Mongelli, F.: *Calore terrestre*. In: *Enciclopedia delle Scienze*, pp. 206–217, Istituto Geografico De Agostini, Novara, 1983
- Mueller, S.: *A new model of the continental Crust*. Geophysical monograph 20, American Geophysical Union, Washington D.C., 1977
- Shebalin, N.V.: Earthquake origin zones and distribution of maximum expected seismic intensity for the Balkan region. *Proc. of the seminar on seismic zoning map, UNDP/UNESCO survey of the seismicity of the Balkan region*, vol. II, Skopje, 1976
- Sibson, R.H.: Fault zone models, heat flow, and the depth distribution of earthquakes in the continental crust of the United States. *Bull. Seismol. Soc. Amer.* **72**, 151–163, 1982
- Siro, L., Sleiko, D.: Space-time evolution of the 1977–1980 seismicity in the Friuli area and its seismotectonic implications. *Boll. Geofisica Teor. Appl.* **24**, 67–77, 1982
- Suhadolc, P.: Time evolution of the aftershock sequence relative to the September 16, 1977 Friuli earthquake. *Boll. Geofisica Teor. Appl.* **24**, 57–65, 1982
- Terzaghi, K., Peck, R.B.: *Soil mechanics in engineering practice*. J. Wiley and Sons Inc.: New York: 1948

Received October 15, 1984; Revised version January 15, 1985
Accepted January 25, 1985

Nonlinear inversion of local seismic travel times for the simultaneous determination of the 3D-velocity structure and hypocentres – application to the seismic zone Vrancea

M. Koch

Geophysikalisches Institut, Universität Karlsruhe, Hertzstraße 16, D-7500 Karlsruhe, Federal Republic of Germany

Abstract. A new method for solving the full nonlinear problem of simultaneous inversion for 3D structure and hypocentres (SSH method) has been developed. In comparison with the simplified linear inversion procedure of Aki and Lee (1976), the method proposed here exhibits several extensions and improvements:

a) The seismic forward problem is solved exactly by a newly developed 3D ray tracing method. This procedure has been conceived as a shooting method. The 3D heterogeneous medium is parametrized by rectangular blocks.

b) The full nonlinear inverse problem is solved. The method works iteratively and was conceived in terms of the Levenberg – Marquardt (LM) algorithm. In each iteration step the seismic forward problem is solved with the ray tracing procedure and a linear inversion is performed.

c) In order to improve the stability and uniqueness of the inverse solution, a priori information on the model space may be used in the inversion process. For example, a known crustal structure may be exploited for the retrieval of the deeper lithosphere.

The SSH method has been tested on various theoretical 1D and 3D models and its capabilities and limitations are discussed. The results demonstrate the existence of a trade-off between hypocentral depths, origin times and seismic velocity, particularly when the horizontal dimensions of the recording network are small relative to the depths of the earthquakes.

The SSH method has been applied to real earthquake data of the Vrancea region, Romania. About 50 crustal and intermediate depth events were used in the inversion process. First, the crustal events were inverted to infer the crustal structure of the region. This crustal information was then used as an a priori constraint in the inversion of the intermediate depth events. The 3D-velocity model found for the Vrancea region reduces the RMS value of the travel-time residuals from 0.5 to 0.3 s. The most essential feature of the model is a high-velocity anomaly of about 4%–6% beneath the eastern Carpathian foredeeps, extending from about 80 to 160 km depth. This appears to be in agreement with results of teleseismic studies of other authors and may be understood in terms of the plate tectonic concept proposed for the Vrancea region. Due to lack of spatial resolution and high standard errors, the results obtained above should, nevertheless, be taken with some care and have to be substantiated by further investigations with high-quality data.

Key words: 3D-inversion – Lateral heterogeneities – Vrancea region

Introduction

The determination of lateral velocity heterogeneities of the lithosphere by inversion of seismic travel-time data has become of major importance in seismology in recent years. There are essentially two different methods of inversion used to date for the retrieval of the 3D structure of the lithosphere. One method of inversion (ACH method), which is known as the method of Aki et al. (1977), uses teleseismic arrival-time data to infer the structure under a seismic array (see e.g. Koch [1983a; b] for a numerical study and improvements of this method). The other method of inversion, which is the subject of the present paper, has been developed for the inversion of arrival-time data from local earthquakes. This procedure is, consequently, restricted to the retrieval of the 3D structure of the lithosphere in seismically active regions. In contrast to the ACH method, the hypocentres themselves are now unknown and a function of the unknown velocity structure. Thus, one is faced with the problem of simultaneous structure and hypocentre determination (SSH method).

The theoretical bases of the SSH method have been laid for a vertical inhomogeneous medium (1D medium) by Crosson (1976a, b). For a 3D seismic medium a SSH method has been formulated by Aki and Lee (1976) and has been widely applied to different seismically active regions of the world for the determination of lateral velocity heterogeneities in the upper crust (Aki and Lee, 1976; Hawley et al. 1981) or in the lithosphere down to 250-km depth (Horie and Aki, 1982; Roecker, 1982; Takanami, 1982).

Despite its numerous applications, the SSH method of Aki and Lee (1976) has some severe limitations inherent in its construction which may considerably bias the solution obtained. First of all, it is a linear inverse method and, thus, only valid in the presence of small lateral velocity perturbations. On the other hand, lateral velocity inhomogeneities up to 10% contrast have been found, particularly for seismically active subduction zones (Spencer and Gubbins, 1980; Huppert and Froehlich, 1981; Roecker, 1982). In such a case the linear approach may lead to a systematic bias in the solution. In fact, for teleseismic arrival-time data, numerical simulations by Koch (1983a, b) show systematic

errors greater than 20% in the reconstruction of lateral velocity perturbations with 5%–6% contrast. These errors could be eliminated with a nonlinear iterative inversion procedure using exact 3D ray tracing. The strong nonlinearity of the SSH problem has also been pointed out by Pavlis and Booker (1983).

Another important approximation in the SSH method of Aki and Lee (1976) is the oversimplified procedure for calculating the ray paths in the different blocks of the model. In fact, Aki and Lee (1976) use a homogeneous initial model, i.e. the ray path segments for the different blocks are summed up along a straight line between source and receiver. As Aki and Lee (1976) admit, this primitive block-sampling may bias the precision of the calculated hypocentres (see also Engdahl and Lee, 1976). On the other hand, as a consequence of the strong nonlinear coupling between hypocentres and velocity structure, the latter may also be strongly biased.

In the present paper a new SSH method will be proposed. Its essential characteristics are:

a) It uses an exact 3D ray tracing program, developed especially for this purpose, in order to overcome the unreliable block-sampling and allowing for better hypocentre and velocity determination.

b) It is conceived as a fully nonlinear inversion procedure and thus tries to eliminate the systematic bias in the linear inverse solution of Aki and Lee (1976). In this sense, the method proposed here appears to be similar to the one proposed by Hawley et al. (1981) for the retrieval of crustal structure, but uses a different ray tracing which avoids a priori smoothing of the lateral velocity structure and so giving a higher spatial resolution of the model. An iterative improvement of the linear inverse solution has also been obtained by Horie and Aki (1982) and Takanami (1982) for the structure under the Japanese arc using a simplified ray tracing technique of Thurber and Ellsworth (1980).

c) It can be applied in such a way that a priori information on the velocity structure of the lithosphere can be handled conveniently to reduce the instabilities and the non-uniqueness in the solution. In fact, the use of explicit a priori information in inversion theory appears to be the most efficient method for a unique retrieval of the seismic structure of the earth (Jackson, 1979; Koch, 1983a).

The nonlinear SSH method developed here will be applied to earthquake data of the Vrancea region, Romania, to retrieve the 1D and 3D structure of this tectonically complicated zone. In fact, the Vrancea region has been put into the concept of plate tectonics (McKenzie, 1972; Fuchs et al., 1979). That is, the occurrence of seismic activity, which is confined to a rather small volume at intermediate depths of about 100–160 km, appears to be a consequence of the subduction of the continuation of the Black Sea plate under the Eurasian plate. Thus, one is led to the assumption that there must be strong lateral velocity inhomogeneities in this region. Some seismic structure investigations, using teleseismic (Vinnik and Lenartovich, 1975; Hovland and Husebye, 1982) and local (Koch, 1982) arrival-time data, essentially show positive lateral velocity anomalies in the Eastern Carpathian foredeeps. This seems to substantiate the plate tectonic hypothesis mentioned above, but further quantitative studies in this field must be continued to elucidate the, as yet, unclear picture of this region.

The present paper is organized as follows. In the next section the essential mathematical foundations of the non-

linear SSH method are formulated. An exact 3D ray tracing program will be presented for solving the seismic forward problem. In the following section the results of some numerical computations are given, which demonstrate the potentials and limitations of the nonlinear SSH method proposed here. The stability and uniqueness of the solution will be discussed. Finally, the SSH method will be applied to retrieve the 1D and 3D seismic structure of the Vrancea region, Romania.

Theory of the SSH method

Formulation of the nonlinear inverse problem

In the following, a short outline of the mathematical foundations of the nonlinear SSH method is given. A derivation of the SSH method for a vertically inhomogeneous medium (1D case) can be found in Crosson (1976a, b). The formulation for a 3D medium in terms of optimization theory is almost identical to the 1D case. However, it will be shown that the formulation is different in the manner in which the Fréchet derivatives of the travel-time function are calculated with respect to the model parameters.

The principle of the 3D SSH method can be summarized as follows. The earth volume to be modelled is divided in layers. Each layer is divided in blocks which need not have the same size in the different layers as it is the case with the original SSH method of Aki and Lee (1976). This modification proved to be computationally very advantageous, as will be shown later.

The theoretical arrival-time vector is a nonlinear function of the hypocentral coordinates and the velocity parameters

$$\mathbf{t}_{th} = \mathbf{F}(\mathbf{H}, \mathbf{v}) \quad (1)$$

where $\mathbf{t}_{th} = (t_1, t_2, \dots, t_n)$ and $n = p \cdot q$ is the maximum possible number of arrival-time data. q is the number of events and p is the number of stations. $\mathbf{H}(x_{01}, y_{01}, z_{01}, t_{01}, \dots, x_{0q}, y_{0q}, z_{0q}, t_{0q})$ is the vector of the hypocentral parameters for the q events, i.e. its dimension is $4 \cdot q$. $\mathbf{v} = (v_1, v_2, \dots, v_l)$ is the vector of the l parametrized velocities, l is the number of blocks in the model. Thus, one has $m = 4q + l$ unknown model parameters to be determined by the nonlinear inversion of Eq. (1). For the observed arrival times a similar relationship to Eq. (1) also applies, but now with an error vector \mathbf{e} including measurement errors and errors resulting from an insufficient parametrization of the model

$$\mathbf{t}_{ob} = \mathbf{F}(\mathbf{x}) + \mathbf{e} \quad (2)$$

with $\mathbf{x} = (\mathbf{H}, \mathbf{v})$.

The purpose now is to minimize the error \mathbf{e} in Eq. (2) in the least squares sense:

$$\|\mathbf{t}_{ob} - \mathbf{F}(\mathbf{x})\|^2 = \|\mathbf{e}\|^2 \rightarrow \text{minimum} \quad (3)$$

The solution of Eq. (3) by methods of optimization theory (e.g. Gill et al., 1981) can be obtained through linearization of the nonlinear function $\mathbf{F}(\mathbf{x})$ around a starting value $\mathbf{x}^{(0)}$, differentiating Eq. (3) with respect to \mathbf{x} and equating the resulting expression to zero. In this way one gets the classical normal equations for the correction vector $\Delta \mathbf{x}$

$$(\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \cdot \Delta \mathbf{x} = \mathbf{r}. \quad (4)$$

\mathbf{A} is the Jacobian matrix of the Fréchet derivatives of the travel-time function F with respect to the model vector \mathbf{x} ,

i.e. $\mathbf{A}=\mathbf{F}$ and $\mathbf{r}=\mathbf{t}_{ob}-\mathbf{F}(\mathbf{x}^{(0)})$, i.e. the travel-time residual between the observed data \mathbf{t}_{ob} and the theoretical data computed with the initial model vector $\mathbf{x}^{(0)}$.

Because of the linearization of Eq. (3), the new vector $\mathbf{x}^{(1)}=\mathbf{x}^{(0)}+\Delta\mathbf{x}$ will not locate the minimum of Eq. (3), but will only define a new starting vector for the next iteration. In each iteration the normal equations (4), with the new Jacobian matrix calculated for $\mathbf{x}^{(i+1)}$, have to be solved. The iterations are stopped when either $\Delta\mathbf{x}$ or the residual sum of squares $s(\mathbf{x})=\mathbf{r}^T\mathbf{r}$ is smaller than a prescribed bound.

The solution of the linear system of equations (4) is by no means trivial and comprises all the pitfalls of linear inverse theory. In fact, because of errors in the data and insufficient model specifications (bad distribution of events and stations), the inverse problem is highly 'illposed'. This means that the solution of Eq. (4) is unstable and nonunique and possesses high covariances. To remove the instabilities and to reduce the covariances, one has to regularize the inverse problem (Tykhonov and Arsenine, 1976), which, on the other hand, introduces a degradation of the spatial resolution. There is a trade-off between these parameters (e.g. Backus and Gilbert, 1968; 1970; Aki and Richards, 1980), or a statistical bias of the inverse estimator (Hoerl and Kennard, 1970) and finally a degradation of the fit of the model to the observed data. For an extensive review and application of regularization theory for the recovery of the 'ill-posedness' of the geophysical inverse problem, see Koch (1983 a).

In terms of nonlinear inversion theory, instabilities in $\Delta\mathbf{x}$ in Eq. (4) will cause overshooting of the new starting vector over the minimum searched and thus prevent the iteration converging to the minimum of Eq. (3). For improvement, Levenberg (1944) and Marquardt (1963) proposed damping the correction vector \mathbf{x} in Eq. (4) by adding a constant k on the diagonal of the normal matrix $\mathbf{A}^T\mathbf{A}$

$$(\mathbf{A}^T\mathbf{A}+k\mathbf{I})^{-1}\mathbf{A}^T\Delta\mathbf{x}=\mathbf{r}. \quad (5)$$

It can be shown (e.g. Marquardt, 1963; Gill et al., 1981) that k will not only shorten the step vector $\Delta\mathbf{x}$, but also turn it towards another search direction which will often lead to a better convergence to the minimum of Eq. (3).

Among the different algorithms proposed so far for the optimal choice of k in Eq. (5) (see Beck and Arnold, 1977; Gill et al., 1981), the original version of Marquardt (1963) appears to be the most convenient to use in practice. Here, k is only subjected to the condition for the residual sum

$$s(\mathbf{x}^{(i+1)}) < s(\mathbf{x}^{(i)}), \quad (6)$$

i.e. k must be chosen in such a way that the step vector from one iteration to the following ensures convergence. This is accomplished by increasing k within an iterative loop until Eq. (6) is satisfied. This algorithm (see Koch [1983a] for details) is used in the present paper and proves to give generally good convergence after 2–4 iterations.

The solution of the seismic forward problem (3D ray tracing)

According to Eq. (1), the nonlinear inversion procedure requires the knowledge of the nonlinear travel-time function $\mathbf{F}(\mathbf{x})$. Thus, the solution of the seismic forward problem must be found by 3D ray tracing.

The 3D ray tracing procedure used in the present paper is exact in the sense that it eliminates the approximations in the SSH method of Aki and Lee (1976). The ray tracing

method has been adapted to block parametrization, i.e. in each block the velocity is assumed to be constant. From each seismic source, rays with variable azimuths and take-off angles are shot towards the respective station. The initial ray direction is modified in an iterative process, until the ray hits the station within a prescribed vicinity (<100 m was found to be sufficient for obtaining travel-time errors <0.05 s). The calculation of the ray paths and travel times in the different blocks is performed by applying Snell's law at the block boundaries (see Koch [1983 a, b] for further details).

The ray tracing procedure proposed here is sort of an extension of the ray tracing method of Thurber and Ellsworth (1980), in the sense that the travel times along the true ray paths are calculated. In the method of Thurber and Ellsworth, the computed ray paths represent only average paths through the lateral heterogeneous structure and are assumed to be straight lines in one layer. Thus, lateral refraction at the vertical block boundaries is neglected. Numerical simulations showed that, for lateral velocity perturbations of 5%, travel-time differences up to 0.1 s exist between the travel times calculated along the true ray paths and the ones computed along the laterally averaged paths of Thurber and Ellsworth.

The present ray tracing program is computationally very efficient. Numerical problems may arise when a ray is critically refracted or totally reflected at the vertical block boundaries. As a result, ray geometrical shadow zones are created at the earth's surface. Due to wave seismic effects such shadow zones will barely be seen in the real earth, so that seismic stations situated in a numerical shadow zone will also record a seismic onset. Numerical simulations showed that due to these effects about 10% of the rays are lost for the inversion process. To retrieve these rays the travel time is calculated approximately along a laterally unrefracted ray path. In such a case the ray tracing procedure corresponds to the method of Thurber and Ellsworth (1980).

Another way to overcome the ray geometrical effects, due to vertical block discontinuities, consists of a horizontal smoothing of the velocity field, e.g. by bicubic spline functions (Červený et al., 1977; Hawley et al., 1981). This introduces a degradation of the lateral structural resolution and, moreover, makes the ray tracing procedure more time consuming as the differential equations for the propagation of seismic rays in a 3D medium must be solved in this case (Červený et al., 1977).

Computation of the Fréchet derivatives

The nonlinear inversion procedure requires that the Jacobian matrix $\mathbf{A}=\nabla\mathbf{F}$ of the Fréchet derivatives in Eq. (4) is computed in each iterative step. For the SSH problem, \mathbf{A} has the following form (Crosson, 1976 a)

$$\mathbf{A}=\begin{pmatrix} \mathbf{C}_1 & 0 & \cdots & 0 & \vdots \\ 0 & \mathbf{C}_2 & & & \vdots \\ \vdots & & \ddots & & \vdots \\ 0 & & & \mathbf{C}_q & \vdots \end{pmatrix} \mathbf{B} \quad (7)$$

where \mathbf{C}_i are the submatrices of the partial derivatives of the arrival times with respect to the four hypocentral parameters x, y, z, t of the seismic event i . \mathbf{C}_i is the matrix generally used for the classical problem of hypocentre loca-

tion in a laterally homogeneous earth (Lee and Lahr, 1972; Buland, 1976).

The matrix \mathbf{B} in Eq. (7) represents the partial derivatives of the arrival time t_{ij} with respect to the block velocities v_s of the model. For an element b of \mathbf{B} in column s and row r [$r=(i-1)*p+j$] one has

$$b_{rs} = \partial t_{ij} / \partial v_s = -L/v_s^2 \quad (8)$$

where L is the length of the ray path in block r calculated by ray tracing. Thus, the Fréchet derivatives with respect to the velocity parameters can be computed analytically.

The same is no longer true for the partial derivatives with respect to the hypocentral parameters, x , y , z . Only in the 1D, vertically inhomogeneous medium can the C_i be computed analytically (Crosson, 1976a). In 3D media one has to approximate the partial derivatives in Eq. (8) by a finite difference quotient (DQ). Here the one-sided DQ has been used, which requires only one additional computation of the travel-time function for the hypocentre shifted by an increment h . Thus one gets, for example for x ,

$$\partial t_{ij} / \partial x_{oi} \approx \Delta t_{ij} / \Delta x_{oi} = [t_{ij}(x_{oi} + h) - t_{ij}(x_{oi})] / h \quad (9)$$

and similar equivalent expressions for the DQ with respect to y and z .

The approximation of the partial derivatives by a finite DQ, Eq. (9), depends strongly on the choice of the increment h (Gill et al., 1981). If one chooses h too large, considerable errors of linearization form in Eq. (9), whereas with a very small h the time increment Δt_{ij} for the shifted hypocentre will be very small and numerical rounding errors arise. A comparison of the finite DQ with the analytically computed, exact partial derivative in a 1D medium showed (Koch, 1983a) that, for $20 \text{ m} < h < 100 \text{ m}$, these approximation errors can be neglected.

The matrix \mathbf{B} in Eq. (7) introduces a coupling of the different submatrices C_i of the hypocentre determination problem, which, in the case of $\mathbf{B} = 0$ (no structure determination), could be solved independently of each other. The coupling of the C_i by \mathbf{B} makes the matrix \mathbf{A} in Eq. (7) very sizeable. Despite its high sparsity, the effective number of unknown parameters is exclusively limited by the amount of computer storage available. A different approach to break off the coupling between hypocentres and seismic velocities, and thus to reduce the size of the matrix \mathbf{A} in Eq. (7), has been proposed by Pavlis and Booker (1980) for a 1D medium by spectral decomposition of the data set.

Numerical method of inversion

The numerical solution of Eq. (5) can be performed by construction of the generalized Inverse (g-Inverse or Lanczos-Inverse) of the matrix \mathbf{A} (see e.g. Lanczos, 1961; Aki and Richards, 1980). The singular value decomposition of \mathbf{A} may be written as

$$\mathbf{A} = \mathbf{U} \cdot \mathbf{S} \cdot \mathbf{V}^T \quad (10)$$

where \mathbf{S} is the matrix of the eigenvalues (singular values) of \mathbf{A} and ordered in a decreasing series, and \mathbf{U} , \mathbf{V} are the corresponding eigenvectors. Putting Eq. (10), into Eq. (5) and solving for $\Delta \mathbf{x}$ one gets, with the definition of the g-Inverse $\mathbf{A}^+ = \mathbf{V} \mathbf{S}^{-1} \mathbf{U}^T$,

$$\Delta \mathbf{x} = \mathbf{V} \cdot \mathbf{S}(\mathbf{S}^2 + k\mathbf{I})^{-1} \mathbf{U}^T \mathbf{r}. \quad (11)$$

Thus the parameter k has the effect of filtering out the very small eigenvalues of \mathbf{A} (which in fact are the numerical manifestation of the ill-posedness of the inverse problem) and so preventing the solution $\Delta \mathbf{x}$ from blowing up and overshooting. This filtering or tapering process may be compared with the sharp 'cut-off' strategy proposed by some authors (Wiggins, 1972; Jackson, 1972) where all eigenvalues smaller than a prescribed bound are eliminated completely from the g-Inverse. Several investigations showed (e.g. Marquardt, 1970) that the tapering of the spectrum by the parameter k is more flexible to use than the 'cut-off' strategy and thus mostly preferred in nonlinear optimization theory.

The numerical computations showed, however, that the spectral decomposition of \mathbf{A} is very time consuming. The direct solution of the system, Eq. (5), proved to be numerically much more efficient. The normal matrix $(\mathbf{A}^T \mathbf{A} + k\mathbf{I})$ is triangularized by applying Householder transformations (QR-decomposition) (e.g. Lawson and Hanson, 1974) and subsequent back-substitution. Moreover, this direct approach has the advantage that, for the classical least squares problem ($n > m$), only the $m \times m$ matrix $\mathbf{A}^T \mathbf{A}$ and not the $n \times m$ matrix \mathbf{A} has to be stored in the central core. Thus, problems with a bigger number of unknowns can be solved.

Numerical computations

Method

Here, the results of the reconstruction of some numerical test models by the nonlinear SSH method will be presented. Having the application of the SSH method to real earthquake - data from the Vrancea region, Romania in mind, the test models were chosen in such a way as to simulate the conditions of this region fairly well. These are: the rather confined seismic activity at 100-160 km depth, the possible presence of a subduction plate (McKenzie, 1972; Fuchs et al., 1979) and the seismic station array used for the recording of the earthquake data (Fig. 1).

Most of these events were recorded at a maximum of 10-15 stations. This is about the number of recordings to be used in the numerical simulations. Technical drop outs of seismic stations that exist in reality were simulated by a random generator.

The numerical model experiments were performed in the following way. An earth volume with a specified seismic structure (layers, blocks and seismic velocities) was chosen initially. Then, randomly distributed hypocentres were generated within this earth volume and theoretical arrival times were computed. To simulate noise in the real data, normally distributed travel-time errors with a standard deviation $\sigma = 0.1-0.2 \text{ s}$ were added.

For the application of the SSH method, a starting solution for the hypocentral and velocity parameters is required. The starting hypocentres were obtained by applying the standard hypocentre localization program, HYPO, of Lee and Lahr (1972) for a laterally homogeneous earth model.

The reconstruction of 1D models

In Fig. 2 the results of the reconstruction of three vertically inhomogeneous models are shown. The number of iterations is $n=4$. Numerous numerical computations showed that fairly good agreement between the reconstructed model

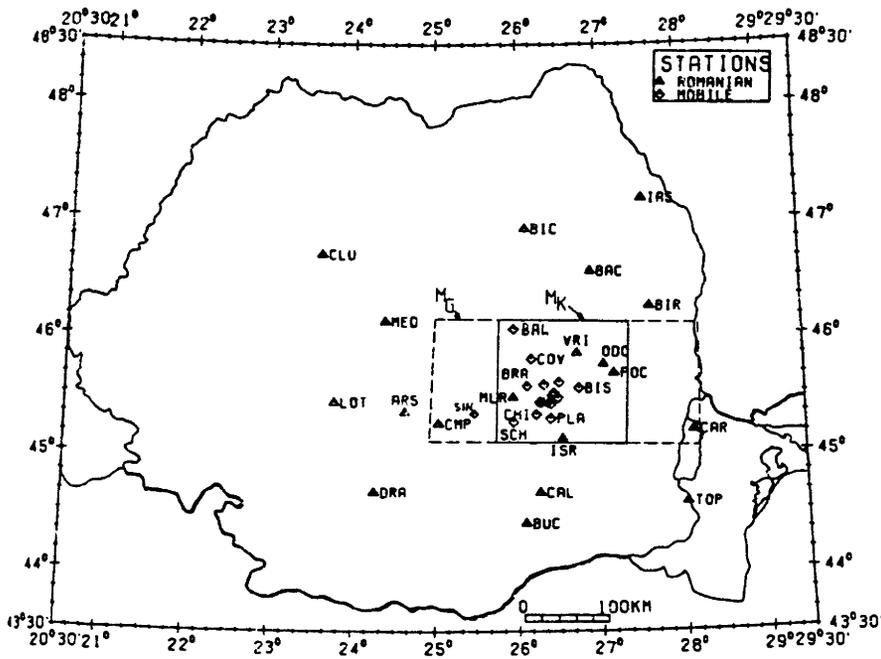


Fig. 1. Map of Romania, showing the seismic station array used for the location of the Vrancea earthquakes. The two rectangles delineate the horizontal extensions of two earth models, which will be discussed later

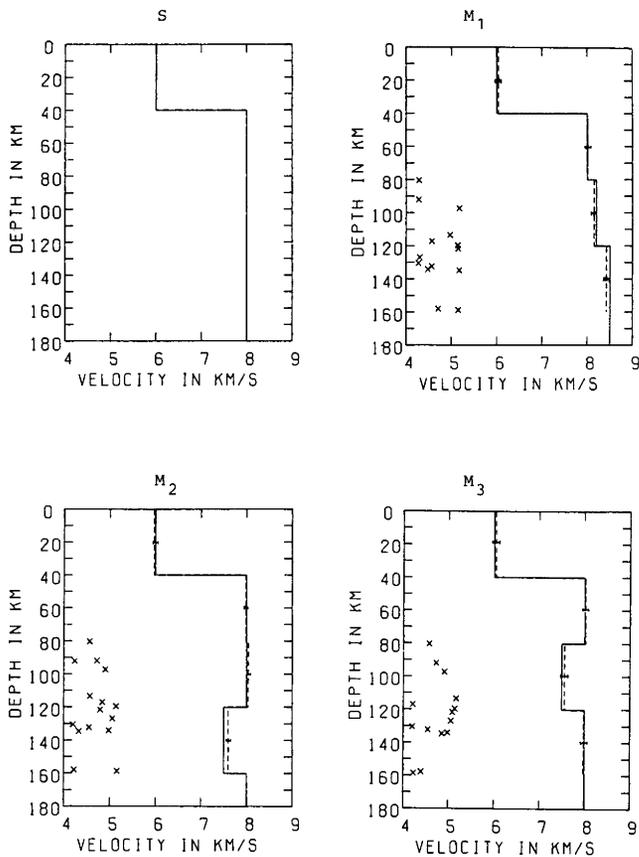


Fig. 2. Results of the nonlinear inversion for three 1D-velocity models M_1 , M_2 , M_3 . S is the starting velocity model, also used for the initial hypocentre location with the HYPO program (Lee and Lahr, 1972). The original model is represented by the full line and the reconstructed one by the dashed line. The error bars indicate the standard deviations of the inverted layer velocity. The crosses represent the depths of the theoretical hypocentres used in the simulation procedure

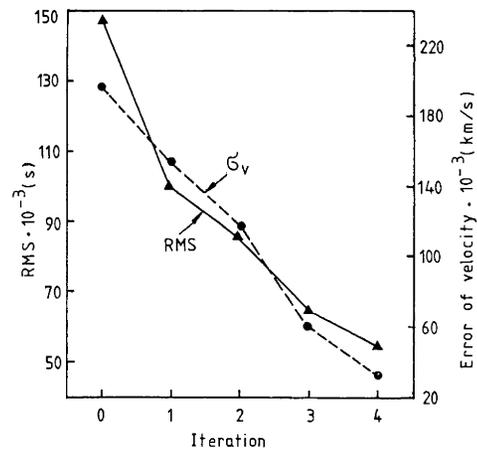


Fig. 3. The RMS of the residuals (left ordinate) and the reconstruction of the velocity (right ordinate) as a function of the number of iteration steps for model M_1 in Fig. 2

and the original one (errors $< 1\%$) could generally be found after 3–4 iterations. Moreover, the computations showed that, by careful choice of the LM-parameter k in Eq. (5), the final solution is almost independent of the starting model, thus demonstrating the excellent convergence properties of the LM-algorithm.

This can be seen in Fig. 3 for model M_1 of Fig. 2, where the convergence behaviour for the SSH inversion is shown as a function of the number of iterations. Figure 3 clearly shows the high nonlinearity of the inverse problem, since a significant improvement in the reconstruction can be achieved by further iteration of the inverse solution, found by one step linear inverse theory ($n=0$).

Because of the coupling of the hypocentres with the velocity structure, one can also assume a shift of the SSH hypocentres, relative to the starting ones. For the model

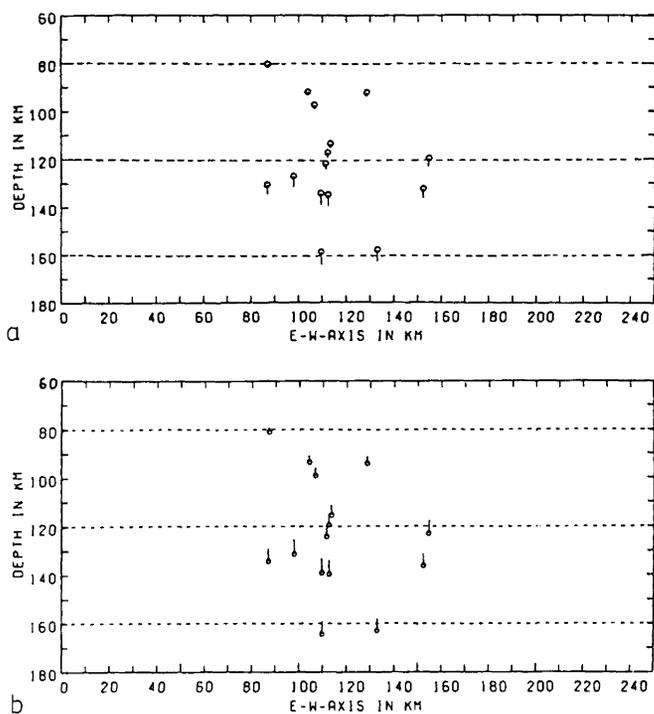


Fig. 4. Cross-sections in the EW-Z-plane showing the hypocentres used in the inversion of model M_1 . *Top:* The circles mark the true hypocentres and the lines the shifts of the hypocentres computed with the HYPO program using the wrong starting model S (Fig. 2). *Bottom:* Shift of the hypocentres in the SSH inversion with respect to the starting hypocentres. The circles now denote the starting hypocentres, i.e. correspond to the endpoints of the lines in Fig. 4 top. The endpoints of the lines mark the hypocentres obtained, simultaneously with model M_1 in Fig. 2, by the SSH procedure

M_1 , these shifts are demonstrated in Fig. 4. The two diagrams are cross-sections in the EW-Z-plane and show the distribution of the hypocentres. In the top diagram the circles represent the true hypocentres, whereas the endpoints of the lines mark the hypocentres located with HYPO using the starting model S in Fig. 2. It can be seen that, due to the velocity of S being too low relative to the true velocity of M_1 , the HYPO hypocentres are shifted systematically about 3–5 km deeper. Along with this, there is an advance of the origin time of up to 0.5 s. On the other hand, due to the lateral homogeneity of the model, the epicentres do not show significant deviations from the true ones.

Figure 4b now shows the shifts of the SSH hypocentres (endpoints of the lines) with respect to the starting hypocentres (circles). That is, the circles in Fig. 4b are situated at the endpoints of the lines of Fig. 4a. The SSH method corrects most of the hypocentres located at too great a depth in Fig. 4a to a shallower depth, thus nearer to the true ones, where the true errors in depth are between 1–2 km. Since this is also approximately the statistical error in depth, the SSH method gives a significant improvement in hypocentre location.

The results for the models M_2 and M_3 are quite similar to those of M_1 . As a consequence of the higher average velocity of the starting model S with respect to the velocities of M_2 and M_3 (see Fig. 2), the hypocentres located with HYPO are situated at shallower depths than the true ones.

In this case the SSH method was able to correct the erroneous starting hypocentres in the direction of the deeper, true ones.

A detailed investigation of the statistical properties of the solution (i.e. correlation, covariance and resolution matrices) showed, however, that the deeper hypocentres in particular are not very well resolved. This means that the computed depth of the hypocentre is highly correlated both with the origin time and with the seismic layer velocity of the model. There is a trade-off for these three parameters, i.e. they can be varied one against the other without a significant change in the fit of the model to the observed data. This shows the strong ill-posedness of the SSH problem, which can be remedied by regularization and use of a priori information on the model space (Jackson, 1979). A more detailed discussion of this problem can be found in Koch (1983).

A significant break in the coupling between these three parameters can be obtained by additional use of S phases, which better constrain the hypocentral depths (Lee and Lahr, 1972; Buland, 1976), and by use of arrival-time data from a seismic array with a larger aperture. In the latter case the seismic rays propagate less steeply through the model, thus giving more independent information on the hypocentres and on the velocity structure. The foregoing simulations were performed with a relatively small station array, similar to the one used for the recording of the real Vrancea earthquakes (Fig. 1). Therefore, the practical application of the SSH method to these events will not be free of problems.

The reconstruction of 3D models

In this section the results of the reconstruction of a seismic model will be presented, which approximately simulates the hypothesized plate-tectonic structure of the Vrancea region.

For the simulation of a high-velocity subducting plate, the test model of Fig. 5a was used. For simpler modelling, the plate was assumed to be sinking vertically, as proposed by Fuchs et al. (1979) for the Vrancea region. The velocities are $v = 8.5$ km/s within the plate and $v = 8.0$ km/s outside.

For the starting model in the inversion, the lateral model S in Fig. 2 was also used in the reconstruction. Similarly, the hypocentres computed with HYPO using this starting model are taken as the starting hypocentres in the SSH method.

As in the case of Fig. 4, the systematical mislocations, obtained with HYPO and the lateral homogeneous velocity model, are discussed first (Fig. 6a). It can be observed that the hypocentres located with HYPO are systematically shifted to greater depths and in the direction of the subducting plate. This is an important result because it shows that if one tries to retrieve the geometry of a subduction zone by looking solely at the distribution of the located seismic events (which in fact was the principal way in which Benioff zones were inferred), one will get a subduction which appears to be steeper than it is in reality (see also Engdahl, 1973; Huppert and Froehlich, 1981). Thus, in the presence of lateral heterogeneities, a systematic mislocation occurs when the earthquakes are located using a laterally homogeneous model. This mislocation can only be remedied by using 3D ray tracing (Engdahl and Lee, 1976).

The final hypocentres obtained with the SSH method exhibit a somewhat smaller mislocation (Fig. 6b). The cor-

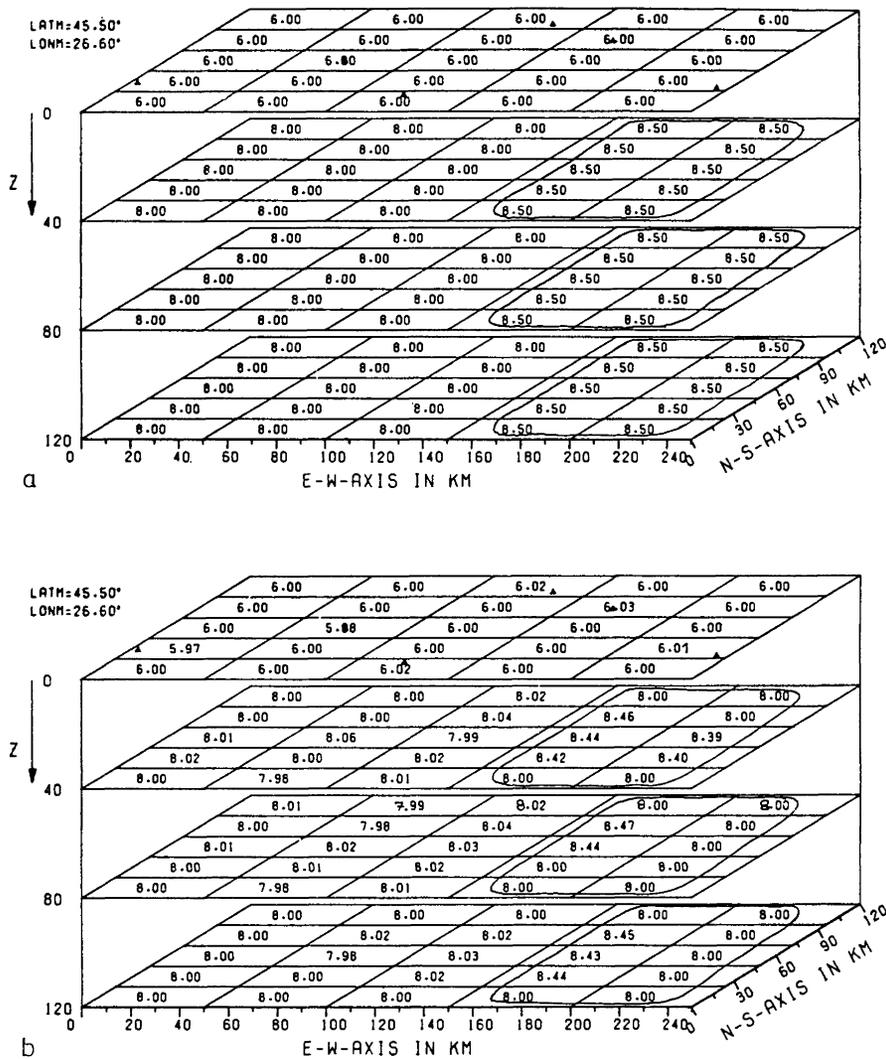


Fig. 5a and b. a 3D-velocity model to approximately simulate the subduction zone hypothesized for the Vrancea region. The location of the earth model corresponds to the dashed rectangle in Fig. 1. The velocities chosen for the model are given within the blocks. The high-velocity zone associated with the subducting plate is marked with a *line*. **b** The reconstruction of the model of Fig. 5a

rections of the epicentres are significant, but the errors in the depths are still quite large. This is due to the strong trade-off between hypocentral depth and origin time. The arrival-time errors due to the over-estimated depths of the computed hypocentres of about 5 km are compensated by an advance of the origin time of about 0.5 s. The results showed that the 'trade-off' could be somewhat decreased by further iterations, but these improvements proved to be statistically insignificant. Therefore, with the present data it is not possible to remove the inherent 'trade-off' between hypocentral depth and origin time.

Better results could also be achieved by using *S* phases in the arrival-time data. On the other hand, this requires a precise knowledge of the v_p/v_s ratio of the earth volume under study. The v_p/v_s ratio is known to vary significantly in lateral direction in the Vrancea region, which would cause additional intricacies (Koch, 1982).

In Fig. 5b the final results for the reconstruction of the velocities, in comparison with the original ones of Fig. 5a, are shown. As a consequence of the sparse station coverage of the model, not all blocks are resolved by seismic rays. The initial velocities (i.e. 6.00 and 8.00 km/s) were assigned to these blocks. Figure 5b shows that the original model and particularly the anomalies in the subducting plate have

been reconstructed fairly well. The standard deviations of the velocity anomalies vary between 0.01 and 0.04 km/s, depending on the location of the corresponding blocks. On the other hand, the diagonal elements of the resolution matrix (Koch, 1983a), in general, have values between 0.7 and 0.9 for blocks within the centre of the model, but values of only 0.6 for the two blocks located in the most eastern part of the second layer of the model.

For a more detailed computation of the geometrical resolution of the structure obtained, a 3D extension of the method of Backus and Gilbert (Backus and Gilbert, 1967; 1968; 1970) has been applied to the present model. The results showed (Koch, 1983a) that the spread of the resolved velocity in both the lateral and vertical direction is about 1.5–2.0 times larger than the block size for the eastern part of the model, resulting in a substantial leaking of the resolved anomalies from one block into adjacent ones. The reason for the strong smoothing of the velocity structure is the poor station coverage in the Eastern Carpathian foredeeps, where in fact there is only one station, CAR (see Fig. 1). These results thus show that a unique estimation of the geometry of the hypothesized subduction plate will barely be possible with the presently existing station network.

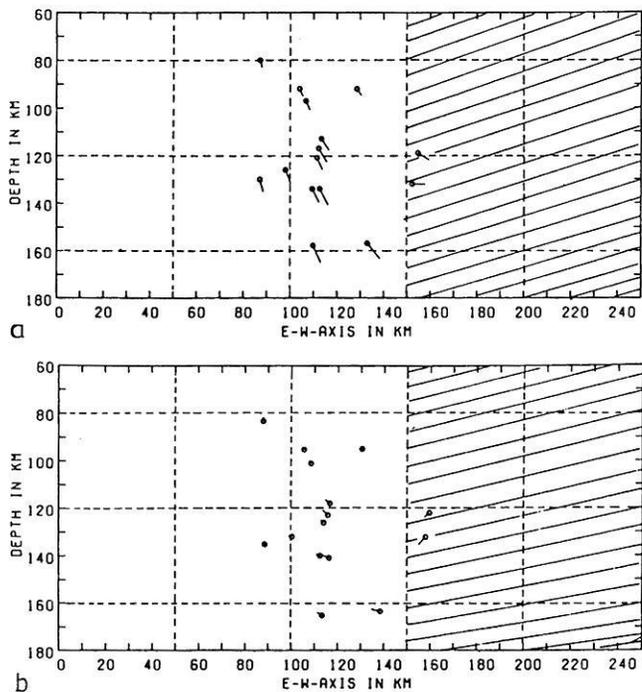


Fig. 6a and b. As Fig. 4, but now for the 3D-velocity model of Fig. 5a. **a** The shifts of the hypocenters computed with HYPO relative to the true ones (circles). **b** The shifts of the SSH hypocenters relative to the starting hypocenters (circles). The shaded region delineates the high-velocity subduction zone

Application of the SSH method to earthquake data of the seismic zone Vrancea, Romania

Evaluation of the data

In the following sections the SSH method will be applied to real earthquake data to infer the 1D- and 3D-seismic

structure of the Vrancea region, Romania. The data used in the inversion may be divided into three different groups:

1) Aftershocks of the big seismic event of 4 March 1977, recorded at a seismic network installed temporarily for the two months following the main event (see Fuchs et al. [1979] for details). About 15 crustal and 15 intermediate depth events, for which average residuals (RMS values) of 0.5–0.8 s were obtained with the localization program HYPO (Lee and Lahr, 1972), are used in the SSH inversion. The events were generally recorded at about 8–12 stations.

2) Microevents (Surduc-events) which were recorded at a temporarily installed seismic network for three months in 1979 (Jung, 1983). Approximately 20 crustal- and a similar number of intermediate-depth events were detected. As a consequence of denser station coverage (with 10–15 arrival times per event) and better technical recording, this data set has higher reliability than the set in 1). For these events the RMS values are 0.2–0.3 s.

3) Seismic events of intermediate depths as recorded by the Romanian seismic station network in the years 1964–1981. Although this data set comprises about 50 events, its reliability must be taken with reservation as the accuracy of the arrival times is only moderate. The RMS values for these events range between 0.5 and 1.0 s.

In Fig. 7 the epicentral distribution of the seismic events is shown (see Koch [1982] for a discussion of the seismicity and the results of preliminary travel-time investigations). These epicentres have been relocated with the HYPO program, using the vertically inhomogeneous velocity model of Fuchs et al. (1979).

Inversion of the crustal events

1D-inversion. First, the results of the inversion of the crustal events with the SSH method, in order to retrieve the 1D- and 3D-crustal structure of the Vrancea region, will be presented.

Figure 8 (left) shows the refined vertically inhomoge-

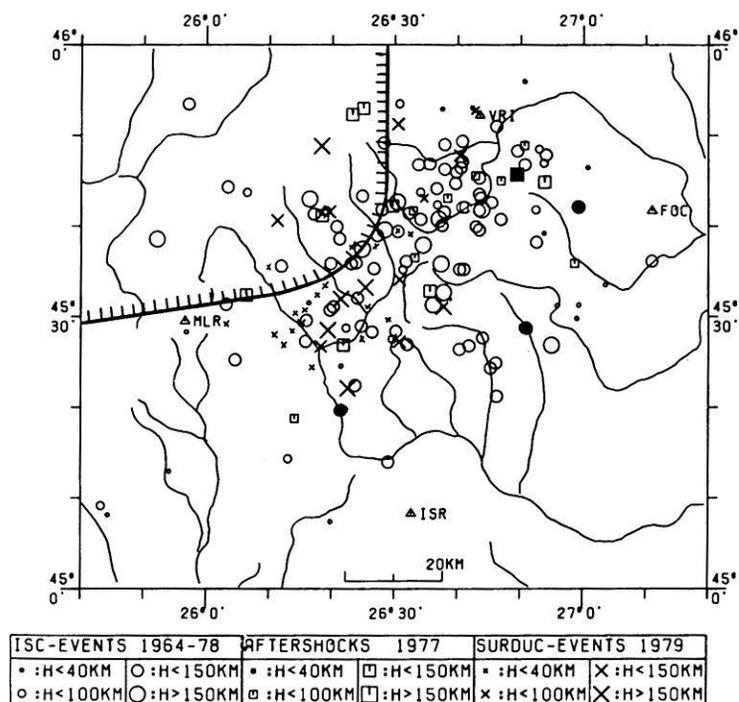


Fig. 7. The epicentral distribution of the seismic events of the three data sets used in the SSH inversion to retrieve the structure of the Vrancea region

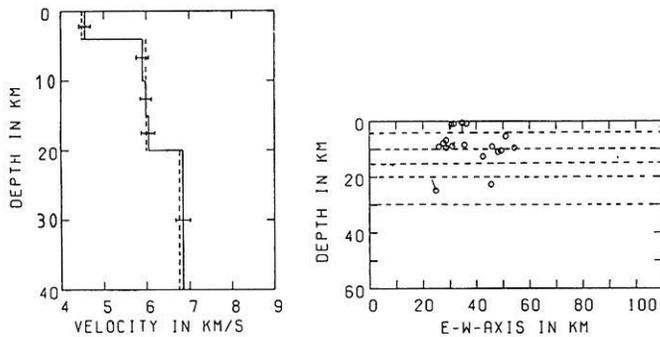


Fig. 8. Results of the SSH inversion of the crustal events for the refinement of the 1D-velocity structure of the crust of the Vrancea region. *Left:* The final velocity model obtained (solid line) with standard errors. The hatched line denotes the initial velocity model. *Right:* EW-Z-cross-section showing the shifts of the final SSH hypocentres relative to the initial ones (circles)

neous (1D) velocity model for the crust of the Vrancea region. There appears to be no significant refinement of the standard model of Fuchs et al. (1979) possible with the present data. The RMS value of the fit of the model to the data could be improved by about 0.1 s with the final velocity model and the final hypocentres obtained (Fig. 8 right). Nevertheless, there remain about 0.4 s unexplained by the data. The reasons for the high RMS values are mainly systematic errors in assigning the correct seismic phases, but may also be due to the impossibility of a 1D-velocity model to account for travel-time residuals due to the lateral velocity heterogeneities present in this region.

3D-inversion. In this section we will present a 3D-velocity model for the crustal structure of the Vrancea region, as it resulted from various inversion runs, with a varying number of layers and blocks. As a consequence of the decreasing coverage of the model by seismic rays with increasing depth, the horizontal block dimensions were chosen to be larger in the deeper layers than in the shallower ones (Fig. 9). In comparison with the original SSH method of Aki and Lee (1976), which works with a fixed number of blocks in the different layers, this modification proved to be very advantageous. In fact, it allows the number of blocks to be chosen according to the resolution by seismic rays and to reduce it, for example in those parts of the model which are barely hit by rays anyway.

The 3D-velocity model finally obtained (Fig. 9) reduces the RMS value for the travel-time residual from 0.5 s to about 0.3 s. A statistical *F*-test (Draper and Smith, 1966) showed this improvement to be significant with more than 95% certainty over the initial vertically inhomogeneous velocity model.

In the following, the different blocks will be subscripted by (*i, j*) going from the SW- corner of the model in an EW and NS direction, respectively.

The most striking feature of the 3D model is the decrease of the seismic velocity in the second layer with respect to the starting model ($v = 6.0$ km/s). This is particularly true for the two resolved blocks (4,3) and (4,4) in the eastern part of the model. To some extent these anomalies project into the third layer. These negative velocity anomalies, with standard errors of 0.03–0.05 km/s and diagonal elements of the resolution matrix between 0.6 and 0.7, may be interpreted as the thick sedimentary layers in the eastern Car-

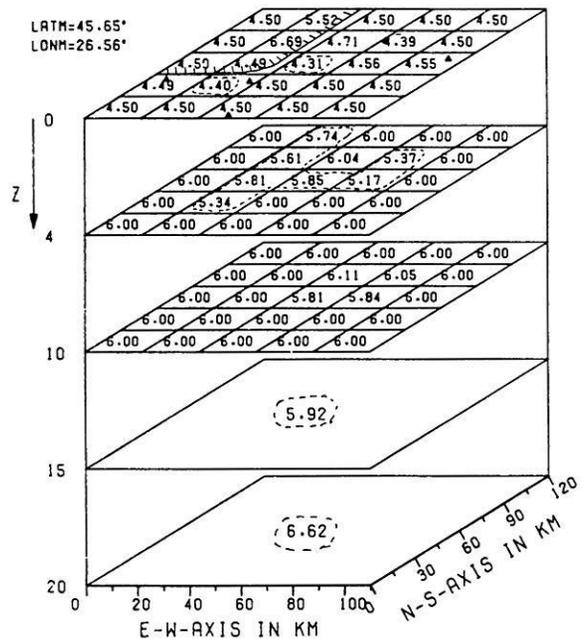


Fig. 9. 3D model of the crust of the Vrancea region. The horizontal extensions of the model correspond to the rectangle M_K in Fig. 1. Zones of negative anomalies are delineated with a hatched line. The thick hatched line approximately marks the crest of the Carpathian chain

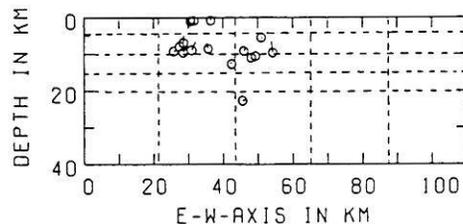


Fig. 10. Cross-section in the EW-Z-plane, showing the shifts of the hypocentres, obtained simultaneously with the crustal structure of Fig. 9, relative to the initial ones (circles)

pathian foredeeps which, according to refraction seismology, increase up to a depth of 15 km under the station FOC (Sollugub, 1969).

The shift of the SSH hypocentres with respect to the original hypocentres computed with HYPO is shown in Fig. 10 in the EW-Z-cross-section. Whereas the epicentral shift appears to be negligible, the depths of the hypocentres are generally shallower than the initial ones (circles). This shows the strong nonlinear coupling of the velocity structure and the hypocentre relocation.

Inversion of the intermediate-depth events

The 3D model for the crustal structure of the Vrancea region will be partly used to retrieve a model for the deeper lithosphere by inverting the arrival-time data of the intermediate-depth events. This means that the crustal structure will be used as an *a priori* constraint in the inversion for the structure of the lithosphere, which reduces the degree of nonuniqueness in the inverted model (Jackson, 1979).

In the following sections, two 3D models will be proposed which are different in their horizontal dimensions and are denoted by M_G and M_K , (see Fig. 1). This proves

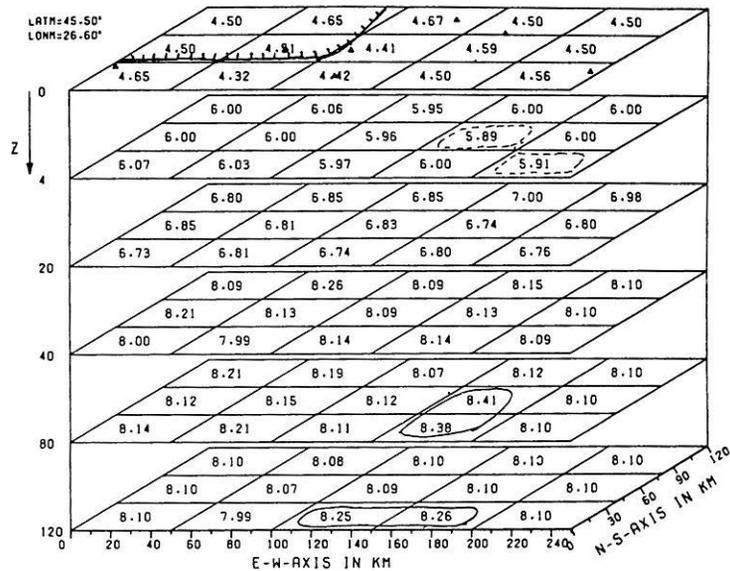


Fig. 11. 3D structure of the lithosphere of the Vrancea region obtained by inversion of the intermediate-depth events. The horizontal extensions of the model correspond to the rectangle M_G in Fig. 1. Zones of positive velocity anomalies and of negative ones are delineated with a solid or a dashed line, respectively. The triangles mark the locations of the Romanian seismic stations (see Fig. 1). The hatched line delineates the crest of the Carpathian chain

to be advantageous for optimally exploiting the different data sets used in the inversion process and taking into account the limited computer storage available. Thus, the larger model M_G was obtained by using mainly the data sets 1) and 3) which were recorded at stations covering a larger area also outside the Vrancea region. Model M_G thus mainly reveals the coarse structure of this region. The smaller model M_K is based on data set 2), which was recorded at a dense station array within the Vrancea region itself (Jung, 1983), and parts of data set 2). Model M_K therefore gives a finer resolution of this part of the lithosphere than model M_G .

Before performing the 3D-inversion, several numerical experiments were executed, to get a refined 1D-velocity model, starting from the vertically inhomogeneous velocity model of Fuchs et al. (1979). The results were negative. Because of the rather small aperture of the array, relative to the depths of the events in the Vrancea region which go down to 160 km depth, there appears to be a strong trade-off between hypocentral depths, origin times and layer velocities. Thus, a better 1D-velocity model than that proposed by Fuchs et al. (1979) could not be retrieved uniquely with the present data.

We will first discuss the larger model M_G (see Fig. 1). Figure 11 shows the 3D model obtained after several numerical test inversions with varying layers and blocks. The 3D model in Fig. 11 reduces the original RMS value significantly from 0.5 s in the laterally homogeneous starting model of Fuchs et al. (1979) to 0.3 s. The essential features of the 3D-model in Fig. 11 can be summarized as follows:

The blocks (4,2) and (5,1) in the eastern part of layer 2 show a slightly reduced velocity compared with the starting model. Although these blocks have not been constrained a priori, due to lack of crustal information from Fig. 9, they show the same tendency as in Fig. 9. Thus, once more, the effect of the sedimentary layers can be seen in the solution.

In layer 5 (80–120 km), a positive velocity anomaly of about 4%–5% can be observed in blocks (4,1) and (4,2). The standard errors of the velocities are between 0.2 and 0.3 km/s, whereas the diagonal elements of the resolution

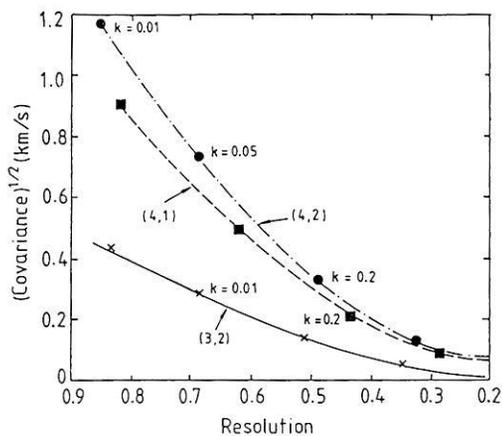


Fig. 12. Trade-off curves between the diagonal elements of the resolution matrix and the standard deviations for three blocks in layer 5 of Fig. 11 as a function of the regularization parameter k

matrix range between 0.4 and 0.6, depending on the choice of the damping parameter k in Eq. (5). In fact, from a more detailed investigation of the results, there appears to be a strong trade-off between the covariance and the resolution of the solution, so that an optimal point on the trade-off curve (Aki and Richards, 1980) is difficult to find. This can be seen from Fig. 12 where the trade-off curves between diagonal elements of the resolution matrix and the standard deviations are shown for three blocks in layer 5. Since no distinguished knee in the trade-off can be observed, the choice of an optimal k to reduce the standard deviations without decreasing the resolution too much is somewhat arbitrary. Nevertheless, for $k=0.2$ a good compromise appears to be obtained. This value of k proved to be appropriate also in ensuring convergence in the LM-procedure according to Eq. (6).

Despite the moderate quality of the inverse solution, these positive velocity anomalies can be identified as the zone of the hypothesized subducting plate (McKenzie, 1972; Fuchs et al., 1979). This appears to be coincident

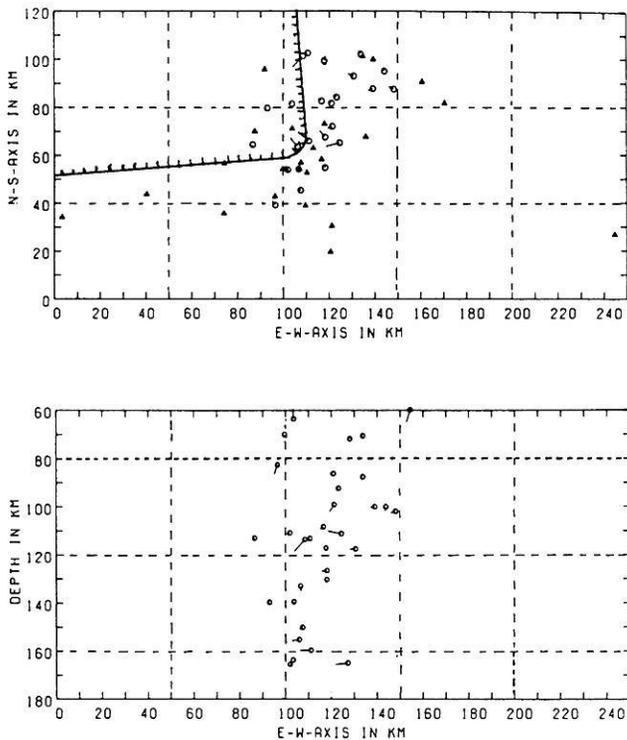


Fig. 13. Shifts of the intermediate-depth hypocentres, inverted simultaneously with the 3D structure of Fig. 11, relative to the initial ones (circles)

with the results of teleseismic investigations of Vinnik and Lenartovich (1975), Hovland and Husebye (1982), and Oncescu et al. (1983) who inferred positive velocity anomalies under the eastern Carpathian region.

Figure 13 shows the shift of the intermediate-depth hypocentres inverted simultaneously with the 3D structure of Fig. 11 relative to the original ones. One can observe a systematic westward drift of the final SSH hypocentres, of about 5–10 km, away from the zone of high velocity. A comparison of this drift to the drift obtained for the theoretical plate model in Fig. 6b shows the same trend. Thus, this appears to corroborate the high-velocity zone found in Fig. 11. Furthermore, it shows the systematic errors were introduced in the relocation of the hypocentres of the Vrancea region when using a laterally homogeneous velocity model, instead of using ray tracing in the true laterally heterogeneous structure (e.g. Engdahl, 1973).

The situation is somewhat similar for the second model M_K (see Fig. 1) used for the retrieval of the 3D structure of the Vrancea region (see Fig. 14). In the inversion for this model, the upper crust has been constrained a priori with the crustal model of Fig. 9, thus reducing the degree of nonuniqueness of the model. Nevertheless, as a consequence of a complete lack of events between 40 and 80 km depth, there is still considerable ambiguity in the final model which requires further investigations. The essential features of the 3D model in Fig. 14, which significantly reduces the original RMS value from 0.5 s to 0.3 s, can be summarized as follows:

In layer 5 (40–80 km depth), the blocks (3,1) and (3,2) are characterized by negative velocity anomalies of 0.2–0.3 km/s. Since the standard deviations are larger than

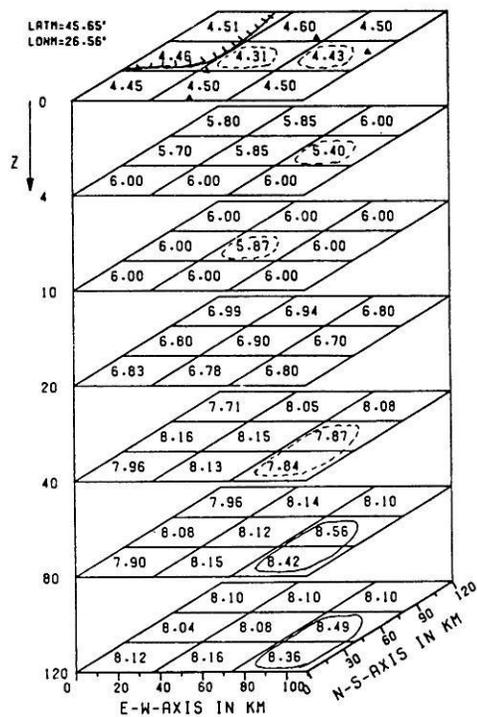


Fig. 14. 3D-structure for the model M_K (see Fig. 1) using the crustal structure in Fig. 9 as an a priori constraint in the inversion process

0.1 km/s, with diagonal elements of the resolution matrix of about 0.5 (which may be a result of the lack of coverage of this depth range with seismic events), no statistical inference can be made. Nevertheless, as these negative anomalies have been obtained in different inversion experiments, one physical explanation would be to assume this depth range to be the zone of low viscosity postulated by Fuchs et al. (1979) to explain its aseismicity.

The layers 6 (80–120 km) and 7 (120–160 km) show high-velocity anomalies of 8.4–8.6 km/s in the blocks (3,1) and (3,2), which correspond to a velocity contrast of 4%–6%. In favour of an improved spatial resolution with diagonal elements of about 0.6, the regularization parameter in Eq. (5) might have been chosen too low resulting, according to Fig. 12, in slightly exaggerated velocity perturbations and covariances or standard errors, the latter ranging between 0.2 and 0.3 km/s. Nevertheless, with the reservation of these high errors, the general trend is the same as in the model M_G in Fig. 11. Consequently, this high-velocity region appears to outline the hypothesized subducting plate in the Vrancea region (McKenzie, 1972; Fuchs et al., 1979). However, with the present data it is not possible to delineate exactly the geometry of the subducting plate and, in particular, to answer the question of whether the plate is steeply inclined, as in other Benioff zones (Engdahl, 1973; Huppert and Froehlich, 1981), or even sinking vertically as a detached slab, as proposed by Fuchs et al. (1979). In fact, calculation of the 3D-spread functions with a 3D extension of the Backus and Gilbert method (Backus and Gilbert, 1967; 1968; 1970) showed that, for the denoted blocks (3,1) and (3,2), the block volume effectively resolved is about 1.5–2 times larger than the initially chosen block

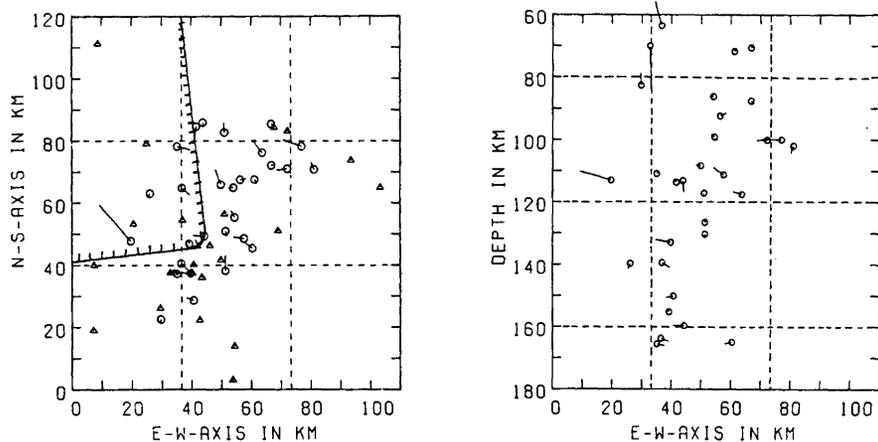


Fig. 15. Similar to Fig. 13, but now showing the shifts of the hypocentres obtained simultaneously with the 3D structure of the model M_K in Fig. 14

sizes in Fig. 13. Thus, there is substantial leaking of the block velocities into the adjacent blocks, prohibiting a precise delineation of the high-velocity subduction zone.

Finally, Fig. 15 shows the shift of the SSH hypocentres relative to the original ones. One can observe, in most cases, a shift in westward direction and also to lower depths, similar to the results of Fig. 13. Thus, from the hypocentral distribution alone, the inclination of the subducting slab appears to be less steep than the one inferred by Fuchs et al. (1979) from the distribution of the hypocentres located with a laterally homogeneous velocity model (see also Engdahl [1973]). These results demonstrate, once more, the systematic errors which are made when one uses a laterally homogeneous earth model for the relocation of the Vrancea earthquakes instead of 3D ray tracing, taking the strong lateral velocity heterogeneities of this region into account.

Summary

In the present paper, a new nonlinear 3D SSH method has been developed which overcomes the systematic approximations inherent in the original SSH method of Aki and Lee (1976), i.e. simplified block-sampling and no ray-tracing. To remedy these approximations, an exact 3D ray tracing has been developed. The method, which was conceived as a shooting method, evaluates the travel times along the true ray paths in the heterogeneous medium specified by rectangular blocks. Snell's law is applied at the block interfaces. With the help of the ray-tracing procedure, the full nonlinear SSH problem is solved. The method developed here is based on the Levenberg-Marquardt algorithm (Levenberg, 1944; Marquardt, 1963) which is famous for its excellent stability and convergence properties. The inverse method is an iterative one where, in each iteration step, a linear inverse problem has to be solved.

A substantial decrease of the degree of nonuniqueness and of the instabilities in the geophysical inverse problem results from the use of *a priori* information on the model space (Jackson, 1979). Thus, the present SSH method has been conceived to handle efficiently *a priori* constraints in the model, e.g. crustal structure known from explosion experiments.

The SSH method has been tested on various 1D and 3D test models which were essentially established in such a way as to represent fairly well the seismic conditions of

the Vrancea region, Romania. The results are encouraging, but clearly demonstrate the intricacies in the inverse solution due to the trade-off between hypocentre depths, origin times and seismic velocities. This is especially true for a station array with small horizontal extensions relative to the depths of the seismic events, as is the case for the Vrancea region.

With 3D ray tracing we investigate the effect of the often hypothesized subduction plate (McKenzie, 1972; Fuchs et al., 1979) on the relocation of the seismic events when these are relocated with a standard location procedure and a laterally homogeneous earth model. The results demonstrate a systematic mislocation which, on the other hand, can be substantially remedied with the 3D SSH method and tracing rays through the true heterogeneous seismic structure. However, due to the trade-off between hypocentral depths and origin times it is not possible to determine the hypocentres without bias.

The SSH procedure has been applied to real crustal- and intermediate-depth events of the Vrancea region recorded during two campaigns, 1977 (Fuchs et al., 1979) and 1979 (Jung, 1983), and to seismic events recorded with the fixed Romanian seismic network during 1964–1981 (Koch, 1982). Because of the limited computer storage capacity, the crustal- and intermediate-depth events were inverted separately. First, the crustal events were inverted for the retrieval of the 3D-crustal structure. This information was then used as an *a priori* constraint in the inversion of the intermediate-depth events, thus reducing the degree of non-uniqueness in the lithospheric structure and achieving better vertical resolution.

The 3D seismic structure of the lithosphere of the Vrancea region inferred in this way decreases the original RMS value for the travel-time residuals in the laterally homogeneous earth model from 0.5 s to about 0.3 s. In spite of this statistically significant improvement, the computed standard errors and the widths of the resolution kernels are still too large to uniquely corroborate the 3D-velocity model retrieved. Its essential feature is a zone of high velocity with 8.4–8.6 km/s, extending from 80 to 160 km depth under the eastern Carpathian foredeeps. This is also in agreement with results of inversion of teleseismic travel-time data of Hovland and Husebye (1982) and of Oncescu et al. (1984) and could substantiate the plate tectonic concepts for the Vrancea region (McKenzie, 1972; Fuchs et al., 1979). Moreover, the model shows a slight decrease of the

seismic velocity to 7.9–7.8 km/s between 40 and 80 km depth, which could support the idea of a low-viscosity zone proposed by Fuchs et al. (1979) to account for the aseismic nature of this depth range.

In conclusion, further investigations are needed for an unambiguous substantiation of the results obtained in the present paper for the velocity structure of the seismic zone of Vrancea. Future recordings from the new Romanian seismic station network, which gives a better coverage of the important eastern Carpathian foredeeps, will certainly give further insight into the seismic structure of this complicated region (e.g. Oncescu et al. 1984).

Acknowledgements. I thank Prof. Dr. K. Fuchs and Prof. Dr. H. Wilhelm for many valuable discussions during the course of this work. I am very indebted to Derrick Balsiger for reading the manuscript. Thanks are due to Prof. G. Mueller for reviewing the manuscript and for numerous helpful comments. The present study has been performed within the framework of the SFB 108 'Stress and stress relaxation in the lithosphere' at the University of Karlsruhe.

References

- Aki, K., Lee, W.H.K.: Determination of three-dimensional velocity anomalies under a seismic array using first *P* arrival times from local earthquakes 1. A homogeneous initial model, *J. Geophys. Res.* **81**, 4381–4399, 1976
- Aki, K., Richards, P.G.: *Quantitative seismology*, San Francisco: W.H. Freeman 1980
- Aki, K., Christoffersson, A., Huseybye, E.S.: Determination of the three-dimensional seismic structure of the lithosphere. *J. Geophys. Res.* **82**, 277–296, 1977
- Backus, G.E., Gilbert, F.: Numerical applications of a formalism for geophysical inverse problems. *Geophys. J.R. Astron. Soc.* **13**, 247–276, 1967
- Backus, G.E., Gilbert, F.: The resolving power of gross earth data. *Geophys. J.R. Astron. Soc.* **16**, 169–205, 1968
- Backus, G.E., Gilbert, F.: Uniqueness in the inversion of inaccurate gross earth data. *Phil. Trans. R. Soc. Lond. A* **266**, 123–192, 1970
- Beck, J.V., Arnold, K.J.: *Parameter estimation in science and engineering*. New York, London: John Wiley & Sons 1977
- Buland, R.: The mechanics of locating earthquakes, *Bull. Seismol. Soc. Am.* **66**, 173–187, 1976
- Cěrvěný, V., Molotkov, I.A., Psencik, I.: *Ray method in seismology*. Prag: Univ. Karlova Press 1977
- Crosson, R.S.: Crustal structure modelling of earthquake data, 1) simultaneous least squares estimation of hypocenter and velocity parameters. *J. Geophys. Res.* **81**, 3036–3046, 1976a
- Crosson, R.S.: Crustal structure modelling of earthquake data, 2) velocity structure of the Pudget Sound region, Washington. *J. Geophys. Res.* **81**, 3047–3054, 1976b
- Draper, N.R., Smith, H.: *Applied regression analysis*. New York, London, Sydney: John Wiley & Sons, Inc., 1966
- Engdahl, E.R.: Relocation of intermediate depth earthquakes in the central Aleutians by seismic ray tracing. *Nature Phys. Sci.* **245**, 23–25, 1973
- Engdahl, E.R., Lee, W.H.K.: Relocation of local earthquakes by ray tracing. *J. Geophys. Res.* **81**, 4300–4306, 1976
- Fuchs, K., Bonjer, K.P., Bock, G., Cornea, I., Radu, C., Enescu, D., Jianu, D., Nourescu, A., Merkle, G., Moldoveanu, T., Tudorache, G.: The Romanian earthquake of March 4, 1977. II. Aftershocks and migration of seismic activity. *Tectonophysics* **53**, 235–247, 1979
- Gill, P.E., Murray, W., Wright, M.H.: *Practical optimization*. London, New York: Academic Press 1981
- Hawley, P.W., Zandt, G., Smith, R.B.: Simultaneous inversion for hypocenters and lateral velocity variations: an iterative solution with a layered model. *J. Geophys. Res.* **86**, 7073–7083, 1981
- Hoerl, H.A., Kennard, R.W.: Ridge regression: biased estimation for nonorthogonal problems, *Technometrics* **12**, 55–82, 1970
- Horie, A., Aki, K.: Three-dimensional velocity structure beneath the Kanto district, Japan. *Journ. Phys. Earth* **30**, 255–282, 1982
- Hovland, J., Husebye, E.S.: Upper mantle heterogeneities beneath eastern Europe. *Tectonophysics* **90**, 137–151, 1982
- Huppert, L.N., Froehlich, C.: The *P*-velocity within the Tonga Benioff zone determined from traced rays and observations. *J. Geophys. Res.* **86**, 3771–3782, 1981
- Jackson, D.D.: Interpretation of inaccurate, insufficient and inconsistent data. *Geophys. J. R. Astron. Soc.* **28**, 97–109, 1972
- Jackson, D.D.: The use of a priori data to resolve non-uniqueness in linear inversion. *Geophys. J.R. Astron. Soc.* **57**, 137–157, 1979
- Jung, P.: Herdparameter und Ausbreitungseffekte krustaler und mitteltiefer Erdbeben in den rumänischen Karpaten unter besonderer Berücksichtigung der Vrancea-Region. Diplomathesis, Univ. Karlsruhe, 1983
- Koch, M.: Seismicity and structural investigations of the Romanian Vrancea region: evidence for azimuthal variations of *P*-wave velocity and Poisson's ratio. *Tectonophysics* **90**, 91–115, 1982
- Koch, M.: The determination of lateral velocity inhomogeneities by linear and nonlinear inversion of teleseismic and local seismic travel times – application to the seismic zone Vrancea, Romania (in German). Ph.D. Thesis, University of Karlsruhe, 1983a
- Koch, M.: A theoretical and numerical study on the determination of the 3D-structure of the lithosphere by linear and nonlinear inversion of teleseismic travel times. *Geophys. J.R. Astron. Soc.*, 1983b
- Lanczos, C.: *Linear differential operators*. London: D. Van Nostrand Co. 1961
- Lawson, C.L., Hanson, R.J.: *Solving least squares problems*. Englewood Cliffs, New Jersey: Prentice Hall 1974
- Lee, W.H.K., Lahr, J.C.: HYPO 71, a computer program for determining hypocenter, magnitude and first motion pattern of local earthquakes. U.S. Geological Survey, Open File Report, 1972
- Levenberg, K.: A method for the solution of certain nonlinear problems in least squares. *Quart. Appl. Math.* **2**, 164–168, 1944
- Marquardt, D.W.: An algorithm for least squares of nonlinear parameters. *J. Soc. Industr. Appl. Math.* **11**, 431–441, 1963
- Marquardt, D.W.: Generalized inverses, ridge regression, biased linear estimation and nonlinear estimation, *Technometrics* **12**, 591–612, 1970
- McKenzie, D.: Active tectonics of the Mediterranean region. *Geophys. J.R. Astron. Soc.* **30**, 109–185, 1972
- Oncescu, M.C., Burlacu, V., Anghel, M., Smalbergher, V.: Three-dimensional *P*-wave velocity image under the Carpathian arc. *Tectonophysics* **106**, 305–320, 1984
- Pavlis, G.L., Booker, J.R.: The mixed discrete-continuous inverse problem: application to the simultaneous determination of earthquake hypocenters and velocity structure. *J. Geophys. Res.* **85**, 4801–4810, 1980
- Pavlis, G.L., Booker, J.R.: A study of nonlinearity in the inversion of earthquake arrival time data for velocity structure. *J. Geophys. Res.* **88**, 5047–5055, 1983
- Roecker, S.: The velocity structure of the Pamir-Hindu Kush region: possible evidence of subducted crust. *J. Geophys. Res.* **87**, 945–959, 1982
- Sollugub, V.B.: Seismic crustal studies in southeastern Europe. In: *The earth's crust and upper mantle*, pp 189–195, Washington, 1969
- Spencer, C., Gubbins, D.: Travel-time inversion for simultaneous earthquake location and velocity structure determination in laterally varying media. *Geophys. J.R. Astron. Soc.* **63**, 95–116, 1980

- Takanami, T.: Three-dimensional seismic structure of the crust and upper mantle beneath the Orogenic belts in Southern Hokkaido, Japan. *Journ. Phys. Earth* **30**, 87–104, 1982
- Thurber, C.H., Ellsworth, W.L.: Rapid solution of ray tracing problems in heterogeneous media. *Bull. Seismol. Soc. Am.* **70**, 1137–1148, 1980
- Tykhonov, A.N., Arsenine, V.: *Méthodes de résolution de problèmes mal posés*. Editions MIR, Moscow, 1976
- Vinnik, L.P., Lenartovich, E.: Horizontal inhomogeneities in the upper mantle of the Carpathians and Caucasus. *Tectonophysics* **28**, 275–291, 1975
- Wiggins, R.A.: The general linear inverse problem: Implications of surface waves and free oscillations for earth structure. *Rev. Geophys. Space Phys.* **10**, 251–285, 1972

Received January 20, 1984; Revised version September 20, 1984

Accepted November 5, 1984

The relationship between the magnetic anisotropy and the *c*-axis fabric in a massive hematite ore

F. Hrouda¹, H. Siemes², N. Herres^{2*}, and C. Hennig-Michaeli²

¹ Geofyzika, národní podnik, P.O. Box 62, 612 46 Brno, Czechoslovakia

² Institut für Mineralogie und Lagerstättenlehre, Rheinisch-Westfälische Technische Hochschule, D-5100 Aachen, Federal Republic of Germany

Abstract. Preferred orientation of hematite ore from Minas Gerais, Brazil, was investigated by reflected-light microscopy, X-ray structural goniometry and magnetic anisotropy. A close relationship was found between *c*-axis fabrics determined by magnetic and non-magnetic methods; experiments confirmed the results of the theoretical treatment. For routine work it is advantageous to use both types of methods, profiting from rapidity of measurement of magnetic anisotropy and from detailed *c*-axis pole figures of pilot specimens provided by X-ray goniometry.

Key words: Hematite ore – *c*-axis fabric – Magnetic anisotropy

Introduction

The lattice-preferred orientation of crystals in a hematite ore is usually investigated by special reflected-light microscopy (e.g. Cameron and Green, 1950; Hennig-Michaeli, 1976) and X-ray pole figure goniometry (e.g. Siemes, 1977). In addition, it can also be studied by magnetic anisotropy (Hargraves, 1959; Uyeda et al., 1963; Porath and Chama-laun, 1966; Porath, 1968). As each of these methods has specific merits (cheap instrument for reflected-light microscopy, detailed lattice orientation pattern provided by X-ray goniometry, rapidity of measurement by magnetic anisotropy), it is advantageous to combine them. For this purpose it is desirable to know the interrelationship among the *c*-axis fabric determinations by these methods. To study this relationship we investigated hematite ore from Minas Gerais, Brazil, using all three methods.

Material used

The itabiritic hematite ore studied is a hand specimen from an unknown locality in Minas Gerais, Brazil (coming from the collection of the Department of Mineralogy and Economic Geology, RWTH Aachen). Macroscopically, a rather inhomogeneous distribution of hematite is indicated by high concentrations in massive ore layers and places of slightly lower content of hematite due to sheet silicate inclusions. Cross-sections of the specimen exhibit tight angular

folds in a massive ore layer and elongated aggregations of sheet silicates as well as open elongated pores in the less rich ore (Fig. 1a).

One of the fracture surfaces of the specimen is a cleavage plane (s_1) along a planar fold limb consisting of massive ore. The arrangement of the sheet silicate flakes indicates a foliation (s_2) inclined at about 20° to s_1 (Fig. 1a and b). On s_1 , several slightly scattered lineations are visible.

From the hand specimen, two slices approximately 2.5 cm thick were cut perpendicular to s_1 and to the mean direction of lineation (the shapes of these slices are clear from Fig. 1a and c). One slice was investigated using reflected-light microscopy and X-ray pole figure goniometry and the other using magnetic anisotropy.

From the first slice, three mutually perpendicular sections were cut for microstructural studies and X-ray pole figure goniometry (for the location, see Fig. 1c):

section $A(\perp s_1, \perp$ mean direction of lineations,
plane of cross-section),

section $B(\perp s_1, \perp A)$,

section $C(\parallel s_1, \perp A, \perp B)$.

In addition, six C sections were prepared for investigating the variability in the *c*-axis distribution. Three of them came from the upper part of the hand specimen (*IT-O* in Fig. 1c) and three from the lower part (*IT-U* in Fig. 1c). The sections within each triad were cut 2 mm apart.

From the second slice, 13 test cube specimens for measuring the magnetic anisotropy were cut with their surfaces parallel to A , B and C , respectively, from the left part of the slice (see Fig. 1d).

Reflected-light microscopy

Microstructurally, the ore is characterized by polygonal tabular hematite crystals (diameter 100–200 μm , thickness 20–50 μm) which lack intracrystalline deformation features (Fig. 2a–c). The tabular hematites lie parallel to either s_1 or s_2 . In the A section a strong preferred orientation of elongated grains along the trace of s_2 is observed locally.

The directional distribution of hematite *c*-axes can be determined qualitatively using the reflected-light microscopy. For individual hematite grains the orientation of the trace of the *c*-axis is measured by means of the polarization figure (Cameron and Green, 1950; Hennig-Michaeli, 1976). In section B (197 grains) there is a very sharp maximum of traces (Fig. 3a), whereas in section A (430 grains) a larger scattering of the trace directions is observed (Fig. 3b). The

* Present address: IFF-KFA Jülich, D 5170 Jülich, Federal Republic of Germany

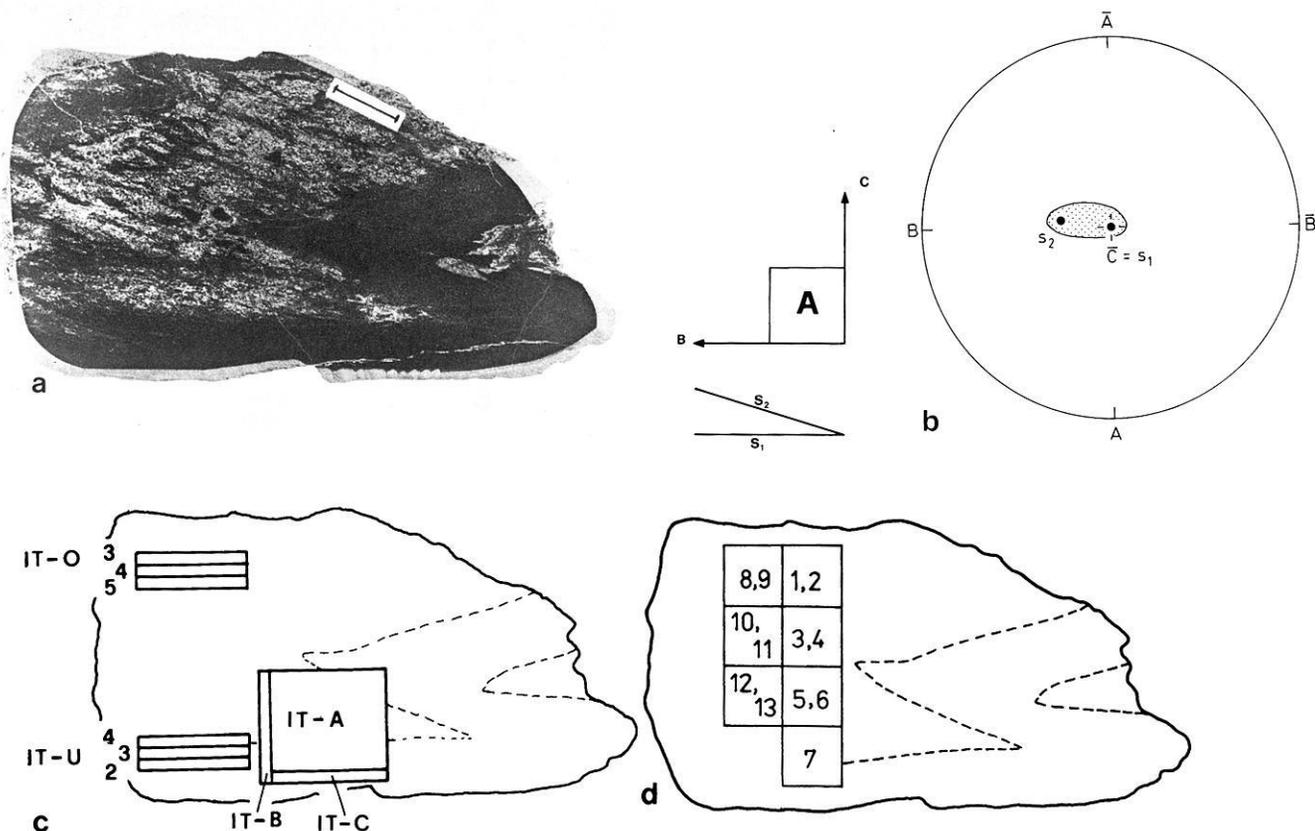


Fig. 1 a-d. Basic information concerning the structure and specimen locations of the sample of hematite ore from Minas Gerais, Brazil. **a** Cross-section of the hematite ore studied. Traces of s_1 (planar fold limb) and s_2 (foliation). Scale bar = 1 cm, **b** Poles of macroscopic structural faces (s_1 , s_2); closed circles denote the mean directions, stippled is the area of occurrence. Equal area projection, lower hemisphere. **c** Locations of the specimens for X-ray pole figure goniometry, **d** Locations of the specimens for magnetic anisotropy.

The reference axes used in all figures are defined as shown in Fig. 1a: A axis lies in the s_1 foliation and is parallel to the mean direction of lineations, B axis also lies in the s_1 foliation and is perpendicular to the A axis and C axis is perpendicular to the s_1 foliation. The notation of sections and planes in this paper is rather unusual from the point of view of physics, but in agreement with that used in X-ray pole figure goniometry – the sections and planes have the same names as the respective axes they are perpendicular to, not the names of the axes lying in them

trace distributions indicate that the c -axes of the hematite grains have a pronounced maximum in an orientation more or less parallel to the mean direction of the normals of s_1 and s_2 (Fig. 1b).

X-ray analysis of preferred orientation

Measurement of preferred orientation by means of an X-ray pole figure goniometer has been performed in the three mutually perpendicular sections A, B and C. The intensity data of the planes A and B were rotated into the plane C and combined with those of plane C to give complete pole figures. Details of the measurements of preferred orientation by means of X-rays and details of the presentation of the data are given by Siemes (1977). The complete pole figures of the (006)-, {104}-, {110}- and {300}-reflections are shown in Fig. 4a-d. If the {104}-reflection has a relative intensity of 100, the (006)-reflection has only a relative intensity of 2. Thus it is possible to measure the former reflection with great accuracy, but the latter with less precision.

Additional, but incomplete, pole figures were prepared from six C sections mentioned earlier to illustrate the variability in the c -axis preferred orientation (see Figs. 5 and

6). The calculated mean intensities of the complete pole figures, which are the intensities of pole figures of hematite without any preferred orientation, were used to establish reference levels for the relative intensities of the incomplete pole figures. Since all measurements were made under the same conditions, all the pole figures are comparable. From Figs. 5 and 6 it follows that in the lower part of the hand specimen (sections U 2, U 3, U 4, see Fig. 5), i.e. in the massive limb of the ore, the preferred orientation is stronger than in the upper part (sections O 3, O 4, O 5, see Fig. 6), as is to be expected from the microscopic study. In the upper part of the hand specimen (Fig. 6), the maximum of the (006)-reflection is split into two areas of high intensity.

Though the hand specimen exhibits monoclinic structural symmetry, the distributions of the intensities in the pole figures show approximately orthorhombic symmetry with identical orientation of the mirror planes in the {104}- as well as in the (006)-pole figures. Only in Fig. 6a and b is there an unresolved discrepancy between the mirror planes in the {104}-pole figure and the (006)-pole figure.

The slightly different orientations of the symmetry planes in the pole figures obtained from different C sections prove that the preferred orientation is varying from place

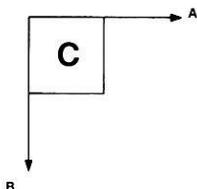
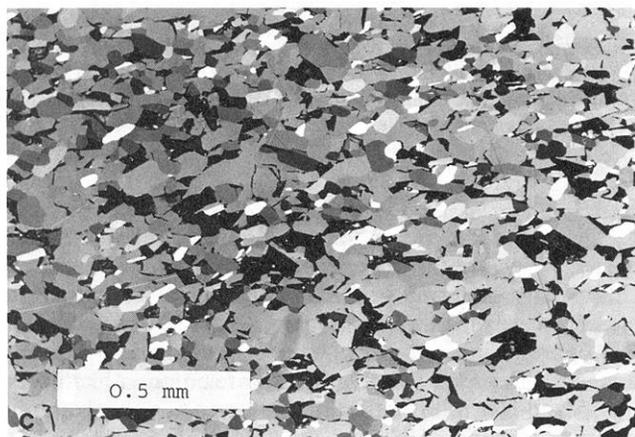
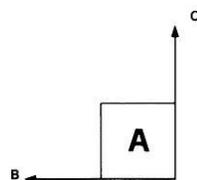
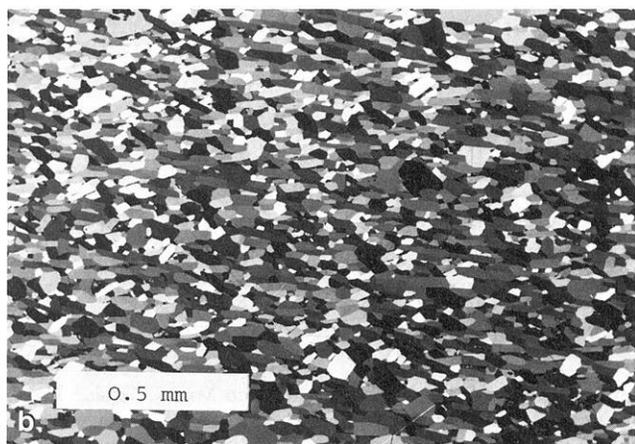
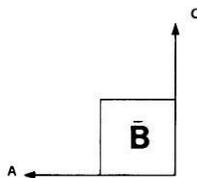
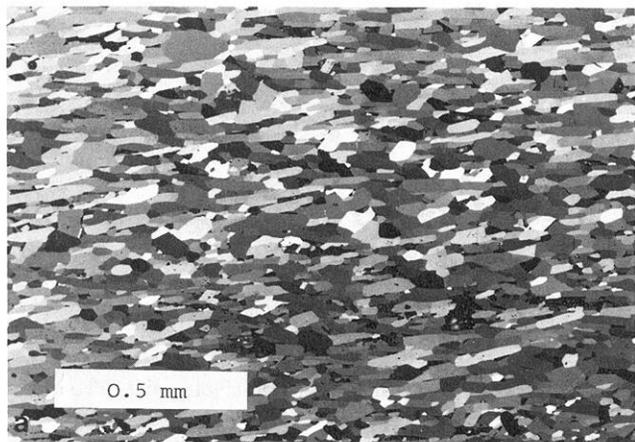


Fig. 2a-c. Microphotographs of the hematite ore. a Section B, b Section A, c Section C

to place in the specimen. This is confirmed by the discrepancy of the orientation of the symmetry planes of the (006)- and {104}-pole figures in Fig. 4a and b and the orientation of the symmetry planes of the {110}- and {300}-pole figures in Fig. 4c and d. The main information of the (006)- and {104}-pole figures comes from sections C of the hand specimen, whereas the main information of {110}- and {300}-pole figures comes from sections A and B.

Magnetic mineralogy of the ore

As shown in the preceding sections, the ore hand specimen investigated is composed virtually of hematite with a small

amount of silicate inclusions. As silicates exhibit magnetic susceptibility of at least one order of magnitude lower than hematite, their contribution to the susceptibility and its anisotropy of the ore can be neglected. On the other hand, small admixtures of magnetite, as low as 0.1%, can influence the measured data considerably. For this reason it is necessary to search for magnetite by the most sensitive methods possible.

Under the ore microscope no magnetite has been identified. In order to confirm this by a magnetic method, we measured the magnetic anisotropy in different fields using the torque magnetometer. As the torque in fields higher than that for saturation of magnetite is independent of the

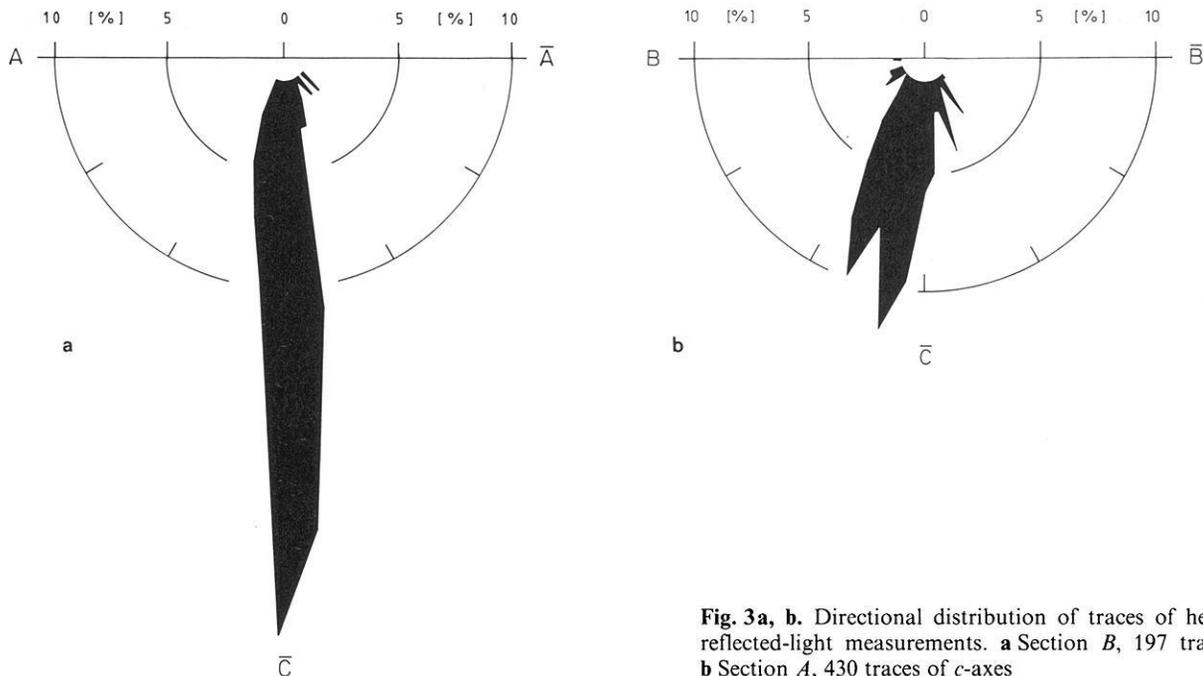


Fig. 3a, b. Directional distribution of traces of hematite c -axes, reflected-light measurements. **a** Section B , 197 traces of c -axes, **b** Section A , 430 traces of c -axes

field for magnetite (Stacey, 1960) and linearly variable for hematite (Porath and Chamalaun, 1966), we can estimate the contribution of both minerals to the anisotropy of the ore.

The magnetic anisotropy in a range of fields from 100–1000 mT was measured on the torque magnetometer developed by Parma (1980) and the measurements were processed by the TOR 1 program developed by Jelínek (unpublished). The results are shown in Fig. 8, where the dependence of the amplitudes of the second harmonic terms of the torque in the B plane of the specimen is presented. It is clear from the figure that the torque is, very closely, linearly dependent on the field. Consequently, the magnetic anisotropy of the ore is controlled exclusively by hematite.

Magnetic fabric and its correlation with optically and X-ray determined c -axis fabric

Let us assume, for the sake of mathematical simplicity, that hematite ore consists only of hematite crystals of the same size and magnetic properties not interacting magnetically and being so numerous that the distribution of their c -axes can be described by a continuous function. Then, the susceptibility tensor of the ore is (see Owens, 1974)

$$\mathbf{k} = \bar{K} \int f(\theta, \phi) \mathbf{K}(\theta, \phi) \sin \theta d\phi d\theta, \quad (1)$$

where \bar{K} is the mean susceptibility (defined as the arithmetical mean of the principal susceptibilities) of a crystal, \mathbf{K} the normed susceptibility tensor of a crystal (the normalizing factor is the mean susceptibility) and $f(\theta, \phi)$ the frequency density function characterizing the distribution of the c -axes of hematite crystals on a surface of a unit hemisphere; θ, ϕ are conventional polar angles specifying the orientation of the c -axes. As the mean susceptibility of hematite is about 5×10^{-3} (Zapletal, 1983), the magnetic interactions in the field of hundreds μT (used in low-field anisotropy meters) can be regarded as negligible (according

to the criteria of Owens and Rutter, 1978). (SI units are used throughout the present paper.)

Introducing the ratio $P_c = K_{ab}/K_c$, where K_{ab} is the susceptibility in the basal plane of hematite crystal and K_c that along the c -axis, the components of the tensor \mathbf{K} are

$$\begin{aligned} K_{11} &= 3[(1 - P_c) \cos^2 \phi \sin^2 \theta + P_c]/(2P_c + 1) \\ K_{22} &= 3[(1 - P_c) \sin^2 \phi \sin^2 \theta + P_c]/(2P_c + 1) \\ K_{33} &= 3[(1 - P_c) \cos^2 \theta + P_c]/(2P_c + 1) \\ K_{12} = K_{21} &= 3(1 - P_c) \sin^2 \theta \cos \phi \sin \phi / (2P_c + 1) \\ K_{23} = K_{32} &= 3(1 - P_c) \sin \theta \cos \theta \sin \phi / (2P_c + 1) \\ K_{13} = K_{31} &= 3(1 - P_c) \sin \theta \cos \theta \cos \phi / (2P_c + 1). \end{aligned} \quad (2)$$

As the tensor representation of magnetic anisotropy is not too illustrative, specially constructed parameters and orientations of principal susceptibilities are used in magnetic anisotropy studies instead of tensors. Among many anisotropy parameters introduced up till now, one pair are of great importance, viz. the P' and T parameters introduced by Jelínek (1981). The P' parameter characterizes the anisotropy degree regardless of the shape of the susceptibility ellipsoid and is defined as follows:

$$P' = \exp \sqrt{\{2[(\eta_1 - \eta)^2 + (\eta_2 - \eta)^2 + (\eta_3 - \eta)^2]\}}, \quad (3)$$

where $\eta_1 = \ln k_1, \eta_2 = \ln k_2, \eta_3 = \ln k_3$,

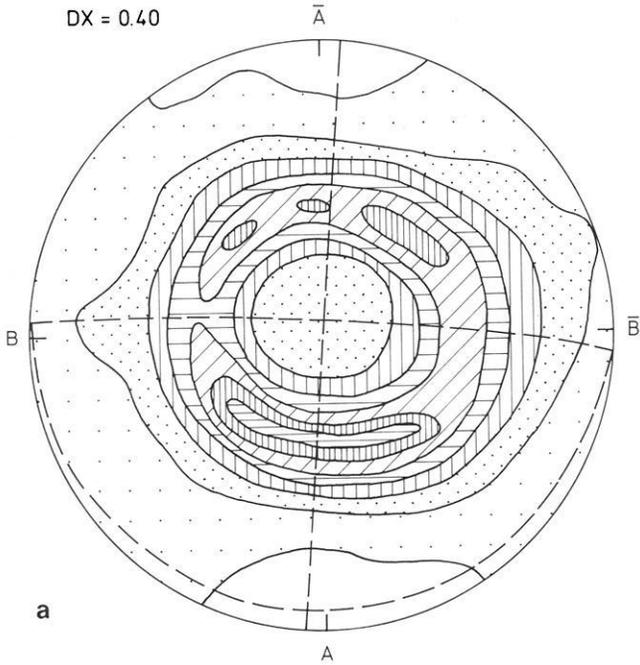
$\eta = (\eta_1 + \eta_2 + \eta_3)/3$ and $k_1 \geq k_2 \geq k_3$ are the principal susceptibilities. The T parameter characterizes the type of the preferred orientation regardless of the anisotropy degree and is defined as

$$T = 2(\eta_1 - \eta_3)/(\eta_2 - \eta_3) - 1. \quad (4)$$

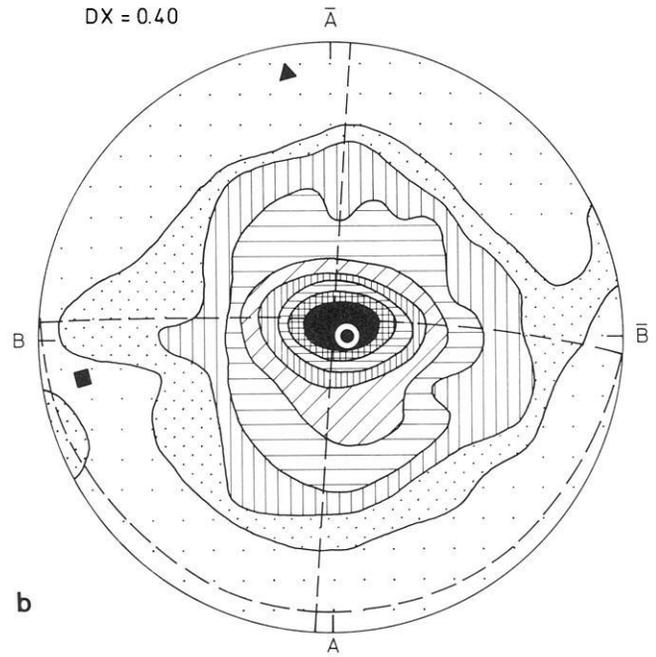
It ranges $0 < T < +1$, if the c -axis pattern in hematite ore is more akin to the cluster pattern than to the girdle one and $-1 < T < 0$, if this pattern is more akin to the girdle than to the cluster.

If the magnetic fabric of an ore is represented by the P' and T parameters and the orientations of principal susceptibilities, the knowledge of \bar{K} , as it follows from Eqs.

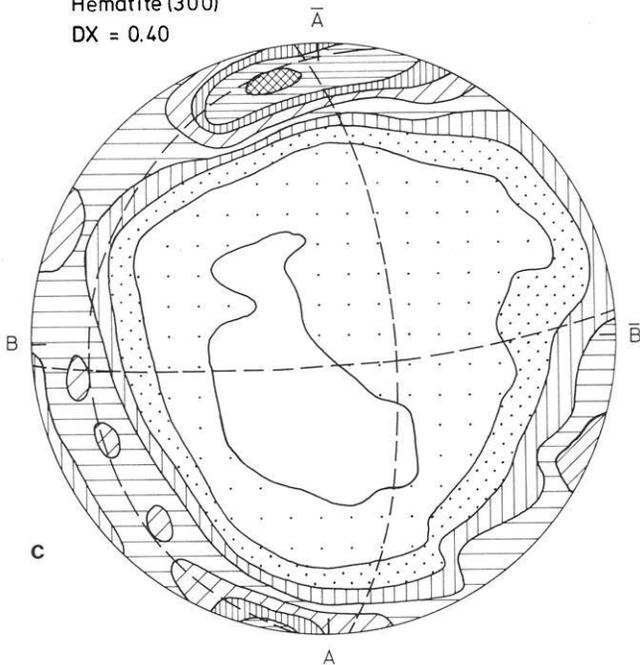
Itabirite Minas Gerais
Hematite (104)
DX = 0.40



Itabirite Minas Gerais
Hematite (006)
DX = 0.40



Itabirite Minas Gerais
Hematite (300)
DX = 0.40



Itabirite Minas Gerais
Hematite (110)
DX = 0.40

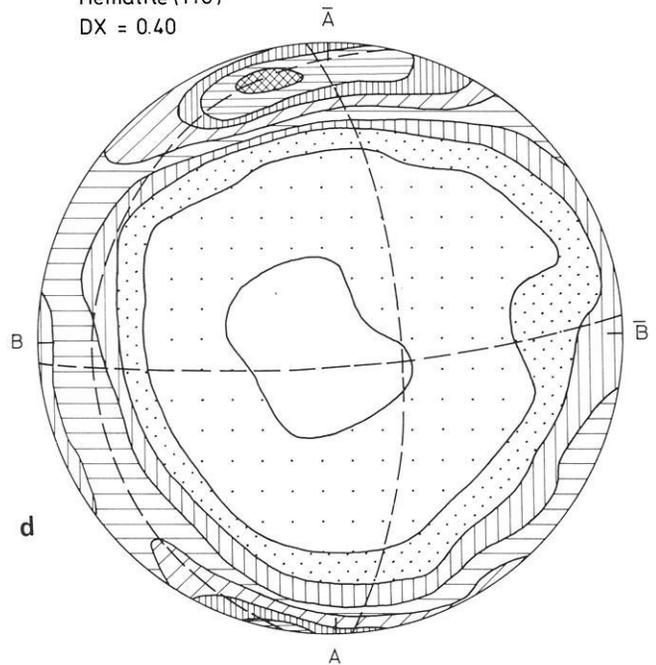
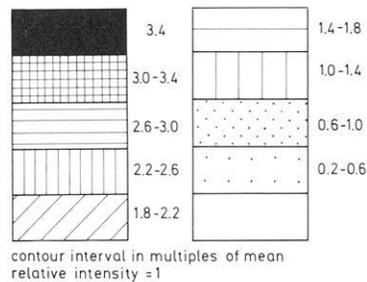


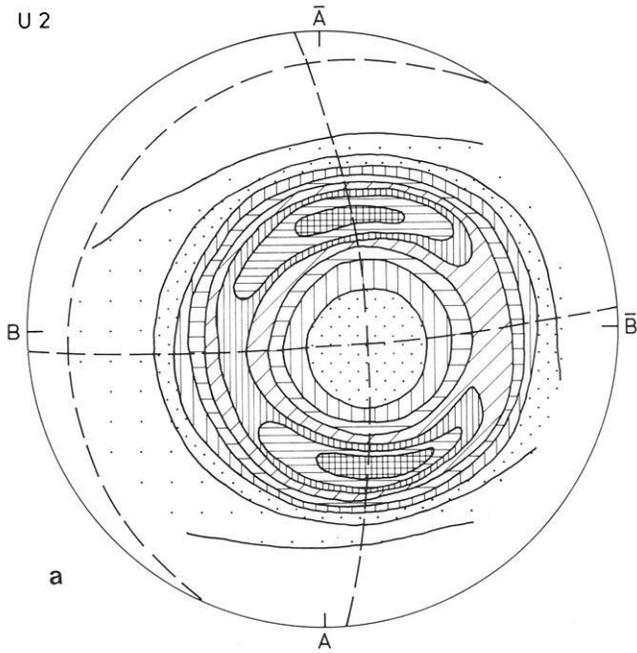
Fig. 4a-d. Complete pole figures of the hematite ore and directions of principal susceptibilities of the theoretical magnetic anisotropy (calculated using Eq. (1)–(4), see the main text). **a** {104}-reflection, **b** {006}-reflection, **c** {300}-reflection, **d** {110}-reflection

Legend for Figs. 4–7: ----- Planes of orthorhombic symmetry in X-ray pole figure diagrams

▲, ■, ● directions of the maximum, intermediate and minimum susceptibilities of the theoretical magnetic anisotropy; △, □, ○ directions of the maximum, intermediate and minimum susceptibilities of the theoretical magnetic anisotropy. Equal-area projection on lower hemisphere



Itabirite Minas Gerais
Hematite (104)
DX = 0.40
U 2



Itabirite Minas Gerais
Hematite (006)
DX = 0.40
U 2

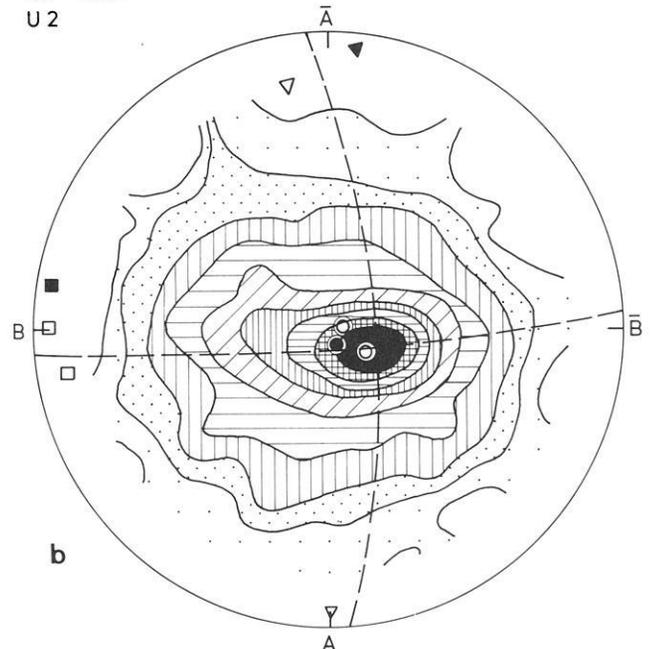
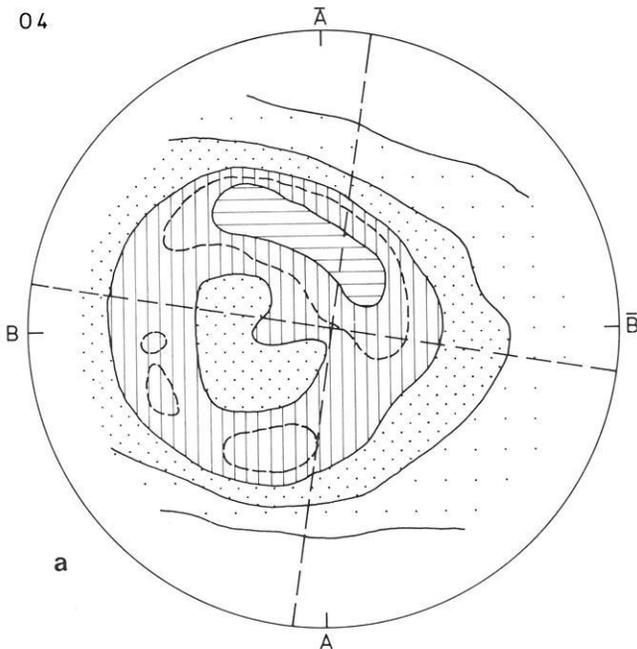


Fig. 5a, b. Incomplete pole figures from the lower part of the hematite ore. **a** {104}-reflection of section *U 2*, **b** (006)-reflection of section *U 2* and orientations of principal susceptibilities of theoretical magnetic anisotropy: section *U 2* (closed symbols), sections *U 3*, *U 4* (open symbols)

Itabirite Minas Gerais
Hematite (104)
DX = 0.40
O 4



Itabirite Minas Gerais
Hematite (006)
DX = 0.40
O 4

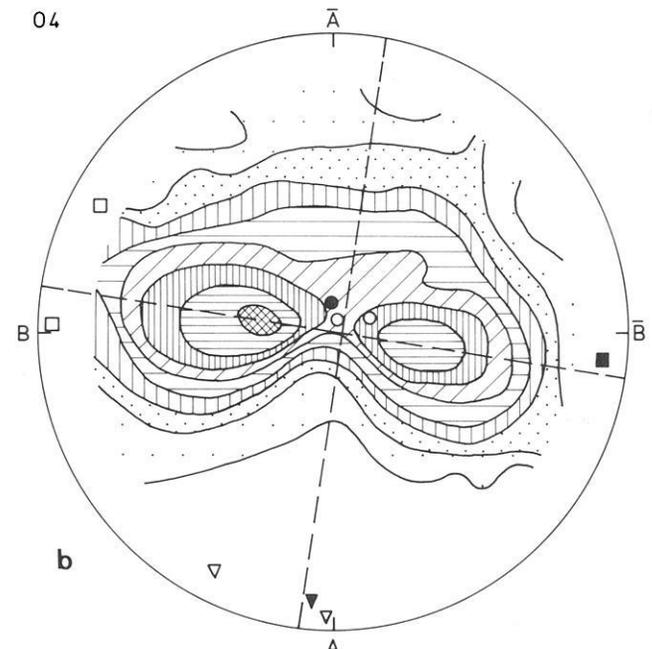


Fig. 6a, b. Incomplete pole figures from the upper part of the hematite ore. **a** {104}-reflection of section *O 4*, **b** (006)-reflection of section *O 4* and orientations of principal susceptibilities of theoretical magnetic anisotropy: section *O 4* (closed symbols), sections *O 3*, *O 5* (open symbols)

Itabirit Minas Gerais
Hematite(006)
DX = 0.40
IT-C

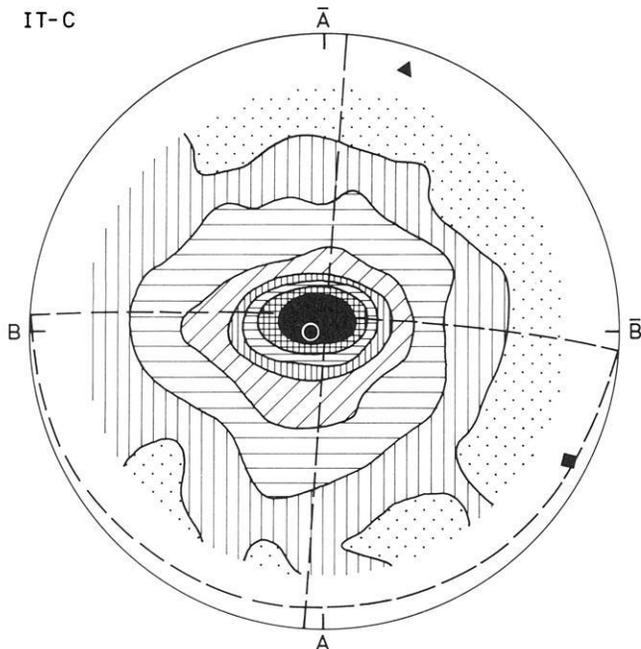


Fig. 7. Incomplete pole figure of the data of the C section of Fig. 4b. Equal area projection of the lower hemisphere of the (006)-reflection

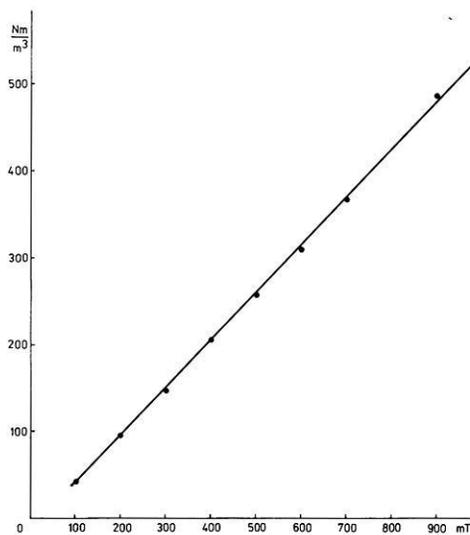


Fig. 8. Diagram to show the dependence of the amplitude of the second harmonic term of the torque on the magnetic field for the hematite ore from Minas Gerais, Brazil

(1)–(4), is not necessary for calculating these characteristics. In addition, Hrouda (1980) showed that in ores with $P_c > 100$ the anisotropy is controlled only by the intensity of c -axis orientation of the highly anisotropic crystals, while the particular value of P_c is unimportant; for example, at constant c -axis concentration, $P_c = 100$, $P_c = 1000$, and $P_c = 10000$ give rise to virtually the same values of ore anisotropy degree. As hematite single crystals exhibit high and virtually isotropic susceptibility in the basal plane and

Table 1.. Theoretical magnetic anisotropy parameters (for definition, see the text) calculated from the (006)-pole figures of hematite ore from Minas Gerais, Brazil, using Eqs. (1)–(4)

Specimen No.	P'	T
IT-CBA	2.30	0.81
IT-C	2.46	0.97
IT-U2	3.15	0.89
IT-U3	3.22	0.94
IT-U4	3.12	0.95
IT-03	2.93	0.87
IT-04	2.75	0.44
IT-05	2.88	0.83
Arithmetical mean	2.85	0.84
Standard deviation	0.33	0.17

very low susceptibility along the c -axis, with $P_c > 100$ (Uyeda et al. 1963; Porath and Chamalaun, 1966), any value of P_c higher than 100 may be used for calculation of the theoretical anisotropy.

Using these factors we calculated the theoretical anisotropy of our hematite ore according to Eqs. (1)–(4) and employing the (006)-pole figure for the function $f(\theta, \phi)$. The data for X-ray measurements were stored in digital form and subsequently used in calculation. The integration over the $f(\theta, \phi)$ function was made numerically by computer. The results are summarized in Figs. 4–7 showing the orientation of principal susceptibilities and in Table 1 giving the values of the P' and T parameters.

It can be seen in Figs. 4–7 that the minimum susceptibility directions mostly lie in the maxima of the c -axis concentrations. Only in the specimen *O 4*, whose c -axis cluster is doubled, does the minimum susceptibility direction lie between the two partial maxima (Fig. 6b). All the minimum susceptibility directions are near the C coordinate, i.e. roughly perpendicular to the s_1 foliation. The maximum susceptibility directions in all specimens are roughly perpendicular to the elongations of the c -axis maxima, lie near the A axis and the mean lineation. The orientations of all the principal susceptibilities are in agreement with the qualitative theoretical predictions following from the properties of hematite single crystals and the c -axis distribution patterns. The P' values in Table 1 are relatively high and correspond to strong c -axis orientation. T values $T > 0$ are in qualitative agreement with the c -axis pattern (cluster type).

The magnetic anisotropy was measured by the KLY-2 Kappabridge (Jelínek, 1980) in a field of $380 \mu\text{T}$ (300 A/m) and calculated using the ANISO 11 program (Jelínek, 1977). The results are summarized in Fig. 9 and Table 2. In Fig. 9 the orientations of principal susceptibilities and of the main fabric elements with respect to the A , B , C coordinates are presented, while in Table 2 the values of the \bar{k} ($\bar{k} = (k_1 + k_2 + k_3)/3$), P' and T parameters are given.

From Fig. 9 it is clear that the minimum susceptibility directions (poles to the magnetic foliation) are very well concentrated in space near the C coordinate and lie in a slightly elongated group in the area between the poles of the s_1 and s_2 planes. The directions of the maximum (magnetic lineation) and intermediate susceptibilities are scattered more than those of the minimum susceptibility and create partial girdles.

As obvious from Table 2, the anisotropy degree of our ore is very high, which indicates an intense preferred orien-

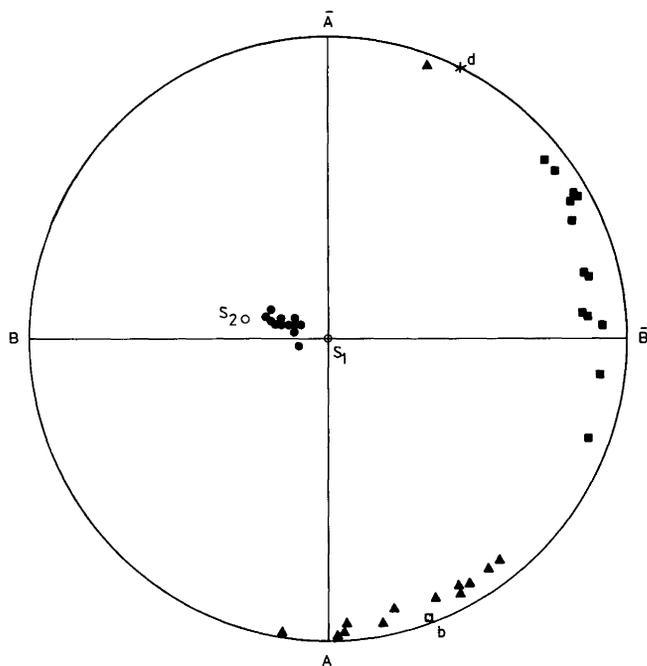


Fig. 9. Orientations of principal susceptibilities and main structural elements with reference to the A , B , C coordinate system for the hematite ore from Minas Gerais, Brazil. ▲ maximum susceptibility, ■ intermediate susceptibility, ● minimum susceptibility, ○ foliation pole, □ lineation (presumably parallel to the fold axis), * intersection line of s_1 and s_2 foliations. Equal area projection of the lower hemisphere

Table 2. Magnetic anisotropy parameters (for definition, see the text) for the hematite ore from Minas Gerais, Brazil

Specimen No.	$\bar{k}(10^{-6})$	P'	T
1	4823	2.222	0.49
2	4823	2.055	0.51
3	4929	1.990	0.38
4	5000	1.977	0.54
5	4730	1.914	0.45
6	5068	2.095	0.47
7	5253	2.901	0.74
8	4959	2.068	0.31
9	4753	1.951	0.36
10	4817	2.161	0.53
11	4950	2.129	0.68
12	5222	2.105	0.59
13	4811	2.132	0.55
Arithmetical mean	4934	2.131	0.51
Standard deviation	167	0.248	0.12

tation of the hematite by crystal lattice in the ore. The T values, ranging from 0.31 to 0.74, indicate that the hematite c -axes create a slightly elongated cluster pattern being in agreement with the X-ray c -axis fabric determination (see Figs. 4–7).

It can be seen in Figs. 4–7 and 9 that both the calculated minimum susceptibilities and the measured ones are oriented very near the C coordinate, i.e. the pole of the s_1 foliation. In this respect, the measured data is in very good agreement with that predicted. The maximum susceptibility directions, both the calculated and the measured, display

relatively large scatters; nevertheless, their mean directions are near the A coordinate. Hence, despite these scatters the calculated and measured maximum susceptibility directions may be regarded as agreeing.

It can be seen in Tables 1 and 2 that the mean value of the anisotropy degree P' calculated is higher than that measured. The scatter of the calculated values is also higher than that of measured values (compare the standard deviations in Tables 1 and 2). The mean value of the shape factor T calculated is higher than that measured. The scatters do not differ too much.

It can be seen in Fig. 1c and d, too, that the anisotropy specimens Nos. 1, 2, 8, 9 are near the X-ray $IT-O$ specimens and Nos. 5, 6, 7 are near the $IT-CBA$ specimen. Specimen No. 7 is also near the $IT-U$ specimens. It is clear from Tables 1 and 2 that the calculated values of the P' and T parameters are higher than those measured in all the three respective groups. The differences may have these explanations:

a) The measurement of the magnetic anisotropy is performed on the total volume of a specimen whereas the pole figures only represent thin mutually perpendicular specimen layers. The preferred orientation of the total specimen volume can be measured by means of neutron radiation texture goniometry (see e.g. Esling et al., 1978). Unfortunately, this method was not at our disposal.

b) Use of incomplete pole figures introduces additional errors into the calculation of the magnetic anisotropy. This can be demonstrated by an example. Figure 7 shows the incomplete (006)-pole figure which was taken from the C section of the hand specimen ($IT-C$). It represents the central part of the complete pole figure, Fig. 4b, but the plotted directions of the principal susceptibilities were calculated from the data of the incomplete pole figure. The maximum and intermediate susceptibilities of the incomplete pole figure deviate noticeably from those of the complete pole figure ($IT-CBA$, Fig. 4b). This discrepancy can be avoided if the complete pole figure is calculated from the incomplete data by means of the vector method (Ruer, 1976; Ruer and Baro, 1977). Complete pole figures may be obtained from only two incomplete measurements by means of the computer program developed by Vadon (1981).

c) All the assumptions for Eq. (1) need not be fulfilled. For example, the grains are obviously not of the same size (see Fig. 2).

Conclusions

The c -axis fabric of massive hematite ore from Minas Gerais, Brazil, was investigated by reflected-light microscopy, X-ray pole figure goniometry and magnetic anisotropy. From this work, the following results have emerged:

1. The orientations of principal susceptibilities measured are essentially the same as those predicted from the c -axis fabric determined by reflected-light microscopy and X-ray pole figure goniometry.

2. The measured values of the anisotropy degree and shape factor are slightly higher than those predicted from the c -axis fabric. These differences may result from the fact that the magnetic anisotropy and the X-ray measurements were not executed on exactly the same specimens, that the volumes of the ore measured by the respective methods were different (1 cm³ cubes in magnetic anisotropy, thin

layers in X-ray pole figure goniometry), that for some predictions incomplete pole figures were used and that, for predictions, a very simple theory was used which may not have fully corresponded to reality. Nevertheless, the differences are small and the data obtained by both methods may be regarded as compatible.

3. For routine work it is advantageous to combine these different types of methods. Large collections of specimens should be investigated by magnetic anisotropy because of the rapidity of this method. Specially selected pilot specimens should subsequently be investigated by reflected-light microscopy and X-ray pole figure goniometry because these methods enable various fabric elements to be investigated and provide detailed pole figures with all partial maxima and minima.

Acknowledgements. We wish to thank Drs. W.H. Owens and L. Kübler as well as an anonymous reviewer for constructive criticism and comments that helped us to improve the paper.

The Deutsche Forschungsgemeinschaft provided the X-ray equipment. The computer programs were run on the Cyber 1975 of the Computer Centre of the RWTH Aachen. Thanks are due to the staff of the Mineralogical Institute, Aachen for their technical assistance.

The Technical Development Foundation of the Geofyzika Brno supported the magnetic anisotropy part of the work. Part of the work was made during the stay of F.H. at the RWTH Aachen.

References

- Cameron, E.N., Green, L.H.: Polarisation figures and rotation properties in reflected light and their application of ore minerals. *Econ. Geol.* **45**, 719–754, 1950
- Esling, C., Wagner, F., Baro, R., Englander, M.: Textures of iron oxides and topotactical relationships. In: Textures of materials, G. Gottstein and K. Lücke, eds.: Proceedings of the Fifth International Conference on Textures of Materials, pp 221–230, Vol. II, 1978
- Hargraves, R.B.: Magnetic anisotropy and remanent magnetism in hemo-ilmenite from ore deposits at Allard Lake, Quebec, *Jour. Geophys. Res.* **64**, 1565–1578, 1959
- Hennig-Michaeli, Ch.: Mikroskopische Gefügeuntersuchungen an experimentell und natürlich verformten Hämatiterzen. Diss. RWTH Aachen, 225 pp, 1976
- Hrouda, F.: Magnetocrystalline anisotropy of rocks and massive ores: a mathematical model study and its fabric implications. *J. Struct. Geol.* **2**, 459–462, 1980
- Jelínek, V.: The statistical theory of measuring anisotropy of magnetic susceptibility of rocks and its application. *Geofyzika Brno*, 88 pp, 1977
- Jelínek, V.: Kappabridge KLY-2. A precision laboratory bridge for measuring magnetic susceptibility of rocks (including anisotropy). Leaflet, *Geofyzika Brno* 1980
- Jelínek, V.: Characterization of magnetic fabric of rocks. *Tectonophysics* **79**, 563–567, 1981
- Owens, W.H.: Mathematical model studies on factors affecting the magnetic anisotropy of deformed rocks. *Tectonophysics* **24**, 115–131, 1974
- Owens, W.H., Rutter, E.H.: The development of magnetic susceptibility anisotropy through crystallographic preferred orientation in a calcite rock. *Phys. Earth Planet. Inter.* **16**, 215–222, 1978
- Parma, J.: Torque magnetometer (in Czech). Report of *Geofyzika Brno*, 25 pp, 1980
- Porath, H.: The magnetic anisotropy of Yampi Sound hematite ores. *Pure Appl. Geophys.* **69**, 168–178, 1968
- Porath, H., Chamalaun, F.H.: The magnetic anisotropy of hematite bearing rocks. *Pure Appl. Geophys.* **64**, 81–88, 1966
- Ruer, D.: Méthode vectorielle d'analyse de la texture. Thèse, Univ. de Metz, 1976
- Ruer, D., Baro, R.: A new method for the determination of the texture of materials of cubic structure from incomplete reflection pole figures. *Adv. X-Ray Analys.* **20**, 187–200, 1977
- Siemes, H.: Fabric analysis and fabric development in ores. *Geol. Fören. Stockh. Förh.* **99**, 172–185, 1977
- Stacey, F.D.: Magnetic anisotropy of igneous rocks. *J. Geophys. Res.* **65**, 2429–2442, 1960
- Uyeda, S., Fuller, M.D., Belshé, J.C., Girdler, R.W.: Anisotropy of magnetic susceptibility of rocks and minerals. *J. Geophys. Res.* **68**, 279–291, 1963
- Vadon, A.: Généralisation et optimisation de la méthode vectorielle d'analyse de la texture. Thèse, Univ. de Metz, 1981
- Zapletal, K.: Some magnetic properties of magnetite and hematite measured in weak magnetic fields (in Czech). Unpublished report of *Geofyzika Brno*, 1983

Received April 17, 1984; Revised December 27, 1984

Accepted February 4, 1985

The inverse scattering problem for reflection of electromagnetic dipole radiation from Earth with vertical variation

H.J. Vidberg and D.O. Riska

Department of Physics, University of Helsinki, Siltavuorenpenger 20 D, SF-00170 Helsinki, Finland

Abstract. The inverse scattering problem of determining the electromagnetic profiles of layers with vertical variation from surface measurements of the electromagnetic field for electric and magnetic dipole radiation is formulated and solved. It is shown that the surface data for dipole radiation, if known at two fixed frequencies and at all transmitter-receiver separations, contain sufficient information for the complete determination of the permittivity, permeability and conductivity distributions of the subterranean layers. The problem is solved by an integral equation formulation and the solution is illustrated with some analytical and numerical examples.

Key words: Inverse scattering method - Electromagnetic dipole radiation

Introduction

The fundamental problem of electromagnetic sounding methods in prospecting work is the determination of the subterranean distribution of the electromagnetic parameters from surface reflection data. This problem is quite complex even in the simplest situation where the subterranean strata have only vertical variation. In the case of reflection of vertically incident plane wave radiation, the inversion problem can, in principle, be solved by the coupled integral equation method of Jaulent and Jean (1972) and Jaulent (1976) or by the less powerful but simpler iterative method of Riska and Vidberg (1983). A drawback of this solution is the fact that the reflection coefficient for plane wave radiation does not, even when known at all frequencies, contain sufficient information for the determination of the permittivity, permeability and conductivity profiles of the subterranean structure. Furthermore, the solution of the inverse scattering problem for plane wave radiation is of limited practical usefulness as most sounding work is carried out with finite-size radiation sources - e.g. dipole antennas. It is, therefore, well worth investigating the corresponding inverse scattering problem for incident dipole radiation.

In this paper we formulate the inverse scattering problem for reflection of electric and magnetic dipole radiation from layers with purely vertical variation. We consider in detail the cases of vertical and horizontal electric and magnetic dipole sources above ground which approximate typical radiation sources used in electrodynamic prospecting work. In contrast to the situation for incident plane wave radiation, the electromagnetic field for dipole radiation depends on position at the surface as well as the radiation frequency. This position- and frequency-dependent electromagnetic field contains sufficient information for the complete determination of the subterranean permittivity, permeability and conductivity profiles. Furthermore, the determination of these profiles is possible from a knowledge of the surface field at only two fixed, but arbitrary, frequencies. Hence no extrapolation into the unmeasurable high-frequency regime is required.

The formulation of the inversion problem for reflection of dipole radiation is mathematically more complicated than that for plane wave radiation, the essential difference being caused by the finite distance to the radiation source. The solution of this inversion problem and the solution algorithms developed in this paper are, on the other hand, simpler. While the inverse scattering problem for plane wave radiation involves a Sturm-Liouville equation with a complex eigenvalue-dependent potential, the corresponding problem for dipole radiation can be formulated in a way that avoids explicit dependence on the eigenvalue in the potential.

In this paper we reduce the inversion problem for dipole radiation to the determination of the complex potential in a Sturm-Liouville equation from the spectrum. The problem is solved by the usual methods of inverse scattering theory. We give a solution based on a Marchenko-type integral equation (Agranovich and Marchenko, 1964), relying on the derivation of the formalism developed by Weidelt (1972). The resulting inverse scattering transform is mathematically very similar to that developed in the work of Coen (1981) on the inverse scattering problem for reflection of elastic waves.

This paper falls into eight sections. Firstly we formulate the inverse scattering problem for a vertical magnetic dipole antenna above ground. Then we solve the problem by means of an inverse scattering integral

* Research supported in part by the Institute of Physics, Outokumpu Oy

Offprint requests to: D.O. Riska

equation. In the subsequent section we discuss the numerical implementation of the solution and then generalize the treatment to the case of a horizontal magnetic dipole source and to the cases of horizontal and vertical electric dipole sources. Following this, we illustrate the solution algorithm for the case of a step function discontinuity which can be handled analytically, and then demonstrate the utility of the solution by some numerical examples. The final section contains a summarizing discussion.

The inversion problem for magnetic dipole radiation

We first consider reflection of the radiation from a vertical magnetic dipole source of strength $me^{-i\omega t}$ at height h from the flat ground surface (Fig. 1). The conductivity, permittivity and permeability distributions below ground are assumed to depend only on depth (z). The situation is then axially symmetric around a vertical axis through the dipole which we choose as the z -axis (we take the downward direction as positive and the position of the dipole as the origin). The electric field is purely azimuthal and depends only on depth (z) and horizontal distance from the dipole (ρ).

The azimuthal electric field amplitude $E(\rho, z)$ satisfies the wave equation

$$\frac{\partial}{\partial z} \frac{1}{\mu(z)} \frac{\partial E}{\partial z} + \frac{1}{\mu(z)} \frac{\partial}{\partial \rho} \frac{1}{\rho} \frac{\partial}{\partial \rho} (\rho E) + \omega^2 K^2(z) E = 0, \quad (1)$$

where we have defined the quantity $K(z)$ as

$$K^2(z) = \left[\varepsilon(z) + i \frac{\sigma(z)}{\omega \varepsilon_0} \right] \varepsilon_0 \mu_0. \quad (2)$$

Here ω is the angular frequency of the dipolar radiation source and σ , ε and μ the vertically varying conductivity, relative permittivity and permeability parameters. This wave equation may be simplified by introducing the Hankel transform of the electric field:

$$\tilde{E}(z, \lambda) = \int_0^{\infty} d\rho \rho J_1(\lambda \rho) E(\rho, z). \quad (3)$$

The transformed field \tilde{E} then satisfies the differential equation

$$\frac{\partial}{\partial z} \frac{1}{\mu(z)} \frac{\partial \tilde{E}}{\partial z} - \frac{\lambda^2}{\mu(z)} \tilde{E} + \omega^2 K^2(z) \tilde{E} = 0. \quad (4)$$

Finally, the first-order derivative of \tilde{E} in Eq. (4) may be removed by the field transform

$$\tilde{E}(z, \lambda) = \sqrt{\mu(z)} Z(z, \lambda). \quad (5)$$

The transformed field Z then satisfies the generalized Sturm-Liouville equation

$$\frac{d^2 Z}{dz^2} + \{k^2 - \lambda^2 - V(z, k)\} Z = 0, \quad (6)$$

in which k is defined as

$$k = \frac{\omega}{c}, \quad (7)$$

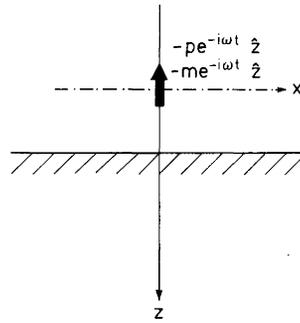


Fig. 1. Vertical dipole source above vertically layered earth

with c being the velocity of light in vacuum. The potential function V in Eq. (6) is defined as

$$V(z, k) = k^2 \{1 - \varepsilon(z) \mu(z)\} + \sqrt{\mu(z)} \frac{d^2}{dz^2} \frac{1}{\sqrt{\mu(z)}} - ik \frac{\sigma(z) \mu(z)}{\varepsilon_0 c}. \quad (8)$$

The potential function thus defined has support in the region $z > h$, where h is the position of the ground surface, and it is well defined for continuously differentiable permeability profiles.

The differential equation, Eq. (6), has the continuously distributed eigenvalues

$$v = \begin{cases} \sqrt{\lambda^2 - k^2}, & k < \lambda \\ -i\sqrt{k^2 - \lambda^2}, & k > \lambda. \end{cases} \quad (9)$$

In the following we shall treat the frequency variable k as a fixed parameter and only consider the real branch of the eigenvalue ($\lambda > k$), i.e. only the components of the field with non-oscillatory dependence on depth. By Eq. (3), this implies that the near-field is used in the inversion.

Once k is taken to be fixed, the potential V has no dependence on the variable eigenvalue v and then the differential equation, Eq. (6), may be treated as a standard Sturm-Liouville equation with a complex potential:

$$\frac{d^2 Z}{dz^2} - [v^2 + V(z)] Z = 0. \quad (10)$$

As the potential V vanishes above ground ($z \leq h$), the general solution to Eq. (10) above ground is a linear combination of $\exp(vz)$ and $\exp(-vz)$. Below ground ($z > h$), the general solution to Eq. (10) may be formed as linear combinations of two linearly independent solutions with exponential behaviour above ground. Such solutions are given by the Jost solutions

$$f_{\pm}(z) = e^{\pm vz} + \theta(z-h) \int_h^z dz' \frac{\sin h v(z-z')}{v} V(z') f_{\pm}(z'). \quad (11)$$

These solutions are well defined provided the potential V is constant or decreasing beyond some arbitrary finite depth d . The boundary condition relevant to the geophysical case is that in which the potential becomes constant at large depths beyond the region of prospecting interest.

It is easy to see by forming the Wronskian of the functions (11) that they form a linearly independent pair of solutions to Eq. (10). They may, therefore, be used to express the electric field due to a vertical magnetic dipole of strength m at height h above ground ($z = 0$) as

$$E(\rho, z) = \int_0^\infty d\lambda \lambda J_1(\lambda \rho) \tilde{E}(z, \lambda), \quad (12)$$

with

$$\tilde{E}(z, \lambda) = i \frac{\mu_0 m \omega}{4\pi} \sqrt{\mu(z)} \frac{\lambda}{v} \{f_-(z) + S(v)f_+(z)\}. \quad (13)$$

This expression is valid for $z > 0$ (below the dipole). The function $S(v)$ in Eq. (13) can be viewed as the surface reflection coefficient.

The surface reflection coefficient $S(v)$ may be calculated from the electric field at the surface using the form of the Jost solutions above earth as

$$S(v) = -e^{-2vh} - i \frac{4\pi v}{\mu_0 m \omega \lambda} e^{-vh} \tilde{E}(h, \lambda). \quad (14)$$

Knowledge of the electric field at the surface by Eq. (3) thus determines the reflection coefficient for all positive values of λ . The reflection coefficient of course also depends on frequency (k).

The inverse scattering problem is the problem of determining the complex potential function V , Eq. (8), from the known reflection coefficient. The solution of this problem is based on the properties of the Jost functions (11) which carry the connection between the reflection coefficient and the potential function.

A convenient integral expression for the reflection coefficient S can be obtained by considering the following alternative expression for the physical solution to the differential Eq. (10):

$$\xi(z) = e^{-vz} - \frac{1}{2v} \int_h^\infty dz' e^{-v|z-z'|} V(z') \xi(z'). \quad (15)$$

Comparing the behaviour of this solution to the combination of Jost functions in Eq. (13) for $0 < z < h$ one obtains the result

$$S(v) = -\frac{1}{2v} \int_h^\infty dz' e^{-vz'} V(z') \xi(z'). \quad (16)$$

In a subsequent section we develop an iterative algorithm for the determination of the potential from S based on this expression and the iterative solution of Eq. (15). In the following section we develop a more powerful integral equation method for the determination of V from S .

If the inverse scattering problem is solved for two frequencies, the frequency-dependent components of the potential (8) may be determined separately from the frequency-independent component. The frequency-dependent terms involve the products $\varepsilon\mu$ and $\sigma\mu$, whereas the frequency-independent component involves only μ . Denoting the frequency-independent (real) component of V as $W(z)$ we have

$$W(z) = \sqrt{\mu(z)} \frac{d^2}{dz^2} \frac{1}{\sqrt{\mu(z)}}. \quad (17)$$

The permeability distribution $\mu(z)$ may be determined from $W(z)$ by iterative solution of the integral equation

$$[\mu(z)]^{-1/2} = 1 + \int_h^z dx (z-x) W(x) [\mu(x)]^{-1/2}. \quad (18)$$

Here we have assumed that both the permeability and its gradient are continuous on the surface. A discontinuity in μ or $d\mu/dz$ at the surface would lead to a singular potential function $V(z)$. The singular part would have to be handled analytically by extracting the starting values $\mu(h)$ and $\mu'(h)$ from the asymptotic expansion

$$S(v) e^{2vh} \rightarrow \frac{\mu(h) - 1}{\mu(h) + 1} + \frac{\mu'(h)}{v[\mu(h) + 1]^2},$$

$$|v| \rightarrow \infty, \quad \text{Re } v > 0. \quad (19)$$

Once $\mu(z)$ has been determined from Eq. (18), $\varepsilon(z)$ may be determined from the frequency-dependent real part of $V(z)$ and $\sigma(z)$ from the imaginary part of $V(z)$.

Integral equation for the inversion problem

The determination of the potential (8) from the reflection coefficient S , Eq. (14), may be carried out by means of a linear integral equation similar to that considered by Weidelt (1972). In order to derive this integral equation we write the Jost solutions, Eq. (11), in the alternative form (Weidelt, 1972)

$$f_\pm(z) = e^{\pm vz} + \theta(z-h) \int_{2h-z}^z dz' e^{\pm vz'} A(z, z'), \quad (20)$$

where the function $A(z, z')$ is a function independent of v with support in the region $2h-z < z' < z$.

Substituting this representation for the fundamental solutions into the differential Eq. (10) and exploiting the v -independence of $A(z, z')$ gives the conditions

$$A_{zz}(z, z') - A_{z'z'}(z, z') = V(z) A(z, z'), \quad (21a)$$

$$2A_z(z, z) = V(z), \quad (21b)$$

$$A(z, 2h-z) = 0, \quad (21c)$$

for the function $A(z, z')$. Equation (21b) can be used to determine the potential $V(z)$ once $A(z, z')$ has been determined.

Consider now the function ξ , Eq. (15), which, by Eq. (11), may be expressed as

$$\xi(z) = f_-(z) + S(v)f_+(z). \quad (22)$$

Here $S(v)$ is the reflection coefficient, Eq. (14). From Eq. (20) one then obtains

$$\xi(z) - e^{-vz} = S(v) e^{vz} + \theta(z-h) \int_{2h-z}^z dz' A(z, z') e^{-vz'} \\ + \theta(z-h) \int_{2h-z}^z dz' A(z, z') e^{vz'} S(v). \quad (23)$$

Taking the inverse Laplace transform of this equation with respect to v gives

$$\begin{aligned} & \frac{1}{2\pi i} \int_{-i\infty+\gamma}^{i\infty+\gamma} dv e^{vt} \{\xi(z) - e^{-vz}\} \\ & = \tilde{S}(z+t) + \theta(z-h) A(z, t) \\ & + \theta(z-h) \int_{2h-z}^z dz' A(z, z') \tilde{S}(t+z'). \end{aligned} \quad (24)$$

Here \tilde{S} is the inverse Laplace transform of the reflection coefficient $S(v)$:

$$S(v) = \int_0^{\infty} dz e^{-vz} \tilde{S}(z). \quad (25)$$

For $t < z$ the left hand side of Eq. (24) vanishes. This can be proven by evaluating the inverse Laplace transform of the successive terms in the iterative expansion of Eq. (15). Note that by the same method one can prove that

$$\tilde{S}(t) = 0 \quad \text{for } t < 2h, \quad (26)$$

since for $t < 2h$ the integral in Eq. (28) may be closed in the right half v -plane and $S(v)$ is analytic for the real part of v sufficiently large and positive. Exploiting the well known analyticity of the Jost solutions, Eq. (20), one can show using Eq. (22) that the iterative expansion of the function ξ , Eq. (15), converges and that hence $S(v)$ is analytic when $\text{Re } v > \sqrt{\sup |V(z)|}$.

For $2h - z < t < z$ one then has the integral equation

$$\tilde{S}(z+t) + A(z, t) + \int_{2h-t}^z dz' A(z, z') \tilde{S}(t+z') = 0, \quad (27)$$

for the determination of $A(z, t)$. Since both \tilde{S} and A are complex functions, Eq. (27) actually represents two coupled integral equations. The solution of the inverse scattering problem thus involves: (1) determination of $S(v)$ from the surface data at a fixed frequency (k) by Eq. (14); (2) Laplace inversion of the thus determined reflection coefficient, Eq. (25); (3) solution of the integral Eq. (27) and, finally, determination of the potential V from the function $A(z, t)$ by Eq. (21 b).

In practice, while the solution of the integral Eq. (27) can be carried out by straightforward numerical techniques the calculation of the inverse Laplace transform of the reflection coefficient can be difficult. Furthermore, the inversion of the Laplace transform is an ill-posed problem and is the origin of the general instability of the solution to the inverse scattering problem.

Numerical implementation of the inverse scattering solution

The integral equation for the kernel function A , Eq. (27), from which the final solution to the inverse scattering problem is obtained, is a Fredholm equation of the second kind and hence may be solved by standard matrix inversion methods. The main problem in the numerical implementation of the inverse scattering solution is associated with the need to evaluate the in-

verse Laplace transform of the reflection coefficient, Eq. (25). The explicit Mellin formula for the inverse Laplace transform:

$$\tilde{S}(z) = \frac{1}{2\pi i} \int_{\gamma-i\infty}^{\gamma+i\infty} dv e^{vz} S(v) \quad (28)$$

requires integration along a line parallel to the imaginary axis in the complex v -plane. As $S(v)$ is calculated from the surface field as a Hankel transform, Eq. (14), that converges only for real λ , use of the Mellin formula, Eq. (28), requires analytic extrapolation of $S(v)$ into the complex v -plane. This is a procedure beset with instability. It should also be noted that the usual methods developed in the literature for numerical solution of the integral transform, Eq. (28), (McWirther and Pike, 1978) only apply for such functions $\tilde{S}(z)$ that vanish as $z \rightarrow \infty$, a condition not satisfied in the geophysical case with absorption.

To determine the function $\tilde{S}(z)$ numerically we consider here the simple approach of parametrizing $S(v)$ by a Padé approximant (Baker, 1975)

$$S(v) = \frac{P_M(v)}{Q_N(v)}, \quad (29)$$

where the degree N of the denominator polynomial Q_N is two units larger than the degree M of the numerator polynomial P_M . With the Padé approximant, Eq. (29), the function $\tilde{S}(z)$ may be calculated directly from Eq. (28). The function $\tilde{S}(z)$ will be of the form

$$\tilde{S}(z) = \sum_i A_i \exp(\gamma_i z), \quad (30)$$

where γ_i are the poles of the rational function P/Q and A_i their corresponding residues. In the cases of geophysical interest there will appear poles with positive and negative real parts.

The errors caused by the Padé approximation in the function $\tilde{S}(z)$, which is the input to the inverse scattering integral Eq. (27), will of course cause distortions in the solution for $V(z)$. Nevertheless, we shall demonstrate by numerical examples in a later section that the Padé approximant method appears to yield reliable results for $V(z)$ down to depths of the order of the average skin depth of the subterranean structure studied. This limitation is of course very natural on physical grounds as one cannot in general expect any method of solving the inverse scattering problem to yield information for depths much beyond this average skin depth.

Given the restriction that the Padé approximant method for calculating the function $\tilde{S}(z)$ is reliable only up to the average skin depth, one may be content to solve the inverse scattering integral Eq. (27) iteratively. The solution for the potential $V(z)$, Eq. (8), will then take the form of a series expansion

$$V(z) = \sum_{n=1}^{\infty} V_n(z) \quad (31)$$

with the terms

$$V_1(z) = -2 \frac{d}{dz} \tilde{S}(2z), \quad (32a)$$

$$V_n(z) = 2 \frac{d}{dz} (-1)^n \int_{2h-z}^z dy_1 \int_{2h-y_1}^z dy_2 \dots \int_{2h-y_{n-2}}^z dy_{n-1} \tilde{S}(z+y_1) \tilde{S}(y_1+y_2) \dots \tilde{S}(y_{n-1}+z). \quad (32b)$$

We have verified up to $n=3$ that this expansion, Eq. (31), is the same as that obtained by using the iterative Jost-Kohn formula (Jost and Kohn, 1952) for solving the inverse scattering problem.

To obtain an estimate of the conditions for convergence of the series (31) we note that the iterative solution $A(z, x)$ of the integral Eq. (27) converges if the norm of the kernel \tilde{S} is less than unity. This condition may be expressed as

$$1 > \|\tilde{S}\| = \int_h^z dy \int_h^z dx |\tilde{S}(x+y)|^2. \quad (33)$$

By Eq. (32a), this condition is satisfied provided

$$\frac{(z-h)^2}{2\sqrt{3}} \sup_{h < u < z} |V_1(u)| < 1. \quad (34)$$

This result shows that there will always exist a region in which the integral Eq. (27) may be solved by iteration. Even when condition (34) is satisfied, the potential series (31) which is derived from the iterative expansion of $A(z, x)$ by differentiation [Eq. (21b)] does not necessarily converge. Following the method by Jost and Kohn (1952) one can prove, however, that the series (31) does converge when the norm of S is sufficiently small [the factor $2\sqrt{3}$ in Eq. (34) replaced by $(2 \log 2 - 1)/2$].

In the analytical example treated later we show that the convergence depth is of the same order of magnitude as the skin depth, which, e.g. for granite with $\sigma \sim 10^{-4}$ S/m, exceeds 600 m for frequencies below 1 kHz. From the point of view of prospecting work this will often be quite sufficient. We shall illustrate the convergence properties and the utility of the iterative solution of the inverse scattering integral Eq. (27) by a set of numerical examples with synthetic data using the Padé approximant method discussed above for deriving the inverse Laplace transform of the reflection coefficient.

The inversion problem for radiation from a horizontal magnetic dipole and vertical or horizontal electric dipoles

The inverse scattering problem for reflection of radiation from a vertical magnetic dipole treated above is particularly simple because of the axial symmetry. In general, one may split the electromagnetic field in a vertically varying medium into a transverse electric ($E_z = 0$) and a transverse magnetic mode ($B_z = 0$). With radiation sources that excite the transverse electric mode one can formulate the inversion problem as an inverse scattering problem with the same potential function as in Eq. (8) by using appropriate field amplitudes.

In the case of a horizontal magnetic dipole source (Fig. 2) the vertical component of the magnetic induction can be written as

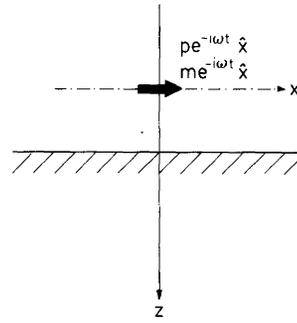


Fig. 2. Horizontal dipole source above vertically layered earth

$$B_z = B(\rho, z) \cos \phi, \quad (35)$$

where the field amplitude $B(\rho, z)$ satisfies the reduced wave equation, Eq. (1). In analogy with the development after Eq. (1), this equation may again be reduced to the Sturm-Liouville form by introducing the transformed field

$$Z(z, \lambda) = \frac{1}{\sqrt{\mu(z)}} \int_0^\infty d\rho \rho J_1(\lambda \rho) B(\rho, z). \quad (36)$$

The transformed field Z then satisfies Eq. (6) with the potential function V given in Eq. (8) and the branches of the eigenvalue ν chosen as in Eq. (9).

The physical solution for Z may be expressed as a linear combination of the Jost solutions (11) as

$$Z(z, \lambda) = \frac{\mu_0 m \lambda}{4\pi} \{f_-(z) + S(\nu) f_+(z)\}, \quad z > 0. \quad (37)$$

The reflection coefficient $S(\nu)$ can finally be calculated from the measured vertical component of the magnetic induction using Eqs. (36) and (37). The recovery of the potential V in this case from the reflection coefficient $S(\nu)$ is performed as before.

In the case of a horizontal electric dipole source, the vertical component of the magnetic induction has the form

$$B_z = B(\rho, z) \sin \phi, \quad (38)$$

where the field amplitude $B(\rho, z)$ again satisfies the reduced wave Eq. (1). Hence, a transformed field Z that satisfies the Sturm-Liouville Eq. (6) with the potential (8) may again be introduced as in Eq. (36). The physical solution for this transformed field is, in this case,

$$Z(z, \lambda) = -i \frac{p \omega \mu_0 \lambda}{4\pi \nu} \{f_-(z) + S(\nu) f_+(z)\}, \quad (39)$$

where p is the strength of the electric dipole. The reflection coefficient $S(\nu)$ can then be calculated from the measured vertical magnetic field component on the surface, and the recovery of the potential function V from $S(\nu)$ can be carried out by the method described earlier.

We finally consider the case of a vertical electric dipole source. As the magnetic field in this case is purely azimuthal, only the transverse magnetic mode is excited. This leads to a Sturm-Liouville equation with a

potential function that is different from that in the previously considered cases.

In the case of a vertical electric dipole source, the azimuthal magnetic field amplitude $H(\rho, z)$ satisfies the reduced wave equation

$$\frac{\partial}{\partial z} \frac{1}{K^2(z)} \frac{\partial H}{\partial z} + \frac{1}{K^2(z)} \frac{\partial}{\partial \rho} \frac{1}{\rho} \frac{\partial}{\partial \rho} (\rho H) + \omega^2 \mu(z) H = 0. \quad (40)$$

The function $K(z)$ is defined in Eq. (2). One may introduce a transformed field Z as

$$Z(z, \lambda) = \sqrt{\frac{\epsilon_0 \mu_0}{K^2(z)}} \int_0^\infty d\rho \rho J_1(\lambda \rho) H(\rho, z), \quad (41)$$

which satisfies the Sturm-Liouville equation

$$\frac{d^2 Z}{dz^2} - [v^2 + \bar{V}(z, k)] Z = 0. \quad (42)$$

Here the potential function \bar{V} has been defined as

$$\bar{V}(z, k) = k^2 [1 - \epsilon(z) \mu(z)] - ik \frac{\mu(z) \sigma(z)}{\epsilon_0 c} + K(z) \frac{d^2}{dz^2} \frac{1}{K(z)}. \quad (43)$$

The branches of the eigenvalue v are chosen as in Eq. (9).

A reflection coefficient $\bar{S}(v)$ for this situation may be introduced by expressing the physical solution for Z in terms of Jost solutions \bar{f}_\pm , defined as in Eq. (11) but with the potential function \bar{V} in place of the function V :

$$Z(z, \lambda) = -i \frac{\omega p \lambda}{4\pi v} \{ \bar{f}_-(z) + \bar{S}(v) \bar{f}_+(z) \}. \quad (44)$$

This reflection coefficient $\bar{S}(v)$ may be calculated from the measured azimuthal magnetic field component on the surface using Eqs. (41) and (44). The potential function $\bar{V}(z, k)$ may then be determined from $\bar{S}(v)$ by the method described earlier. In this case one has, however, to require that both the conductivity and the permittivity profiles be continuously differentiable in order that the potential function \bar{V} , Eq. (43), be well defined. As these conditions are not usually satisfied, use of a vertical electric dipole radiation source is unattractive in comparison to other orientations. The complex structure of the potential function $\bar{V}(z, k)$, Eq. (43), also makes the disentangling of the conductivity, permittivity and permeability distributions from \bar{V} a complicated task, requiring information for more than the two frequencies which proved sufficient in the previously considered cases.

Analytic treatment of a step discontinuity

In order to demonstrate the utility of the integral equation method for solving the inverse scattering problem and the convergence properties of its iterative solution, we treat the case of a step discontinuity at the surface analytically. More precisely, we consider the case of a vertical magnetic dipole source at height h above ground which is assumed to have a constant conduc-

tivity σ and relative permittivity $\epsilon \neq 1$ and permeability $\mu = 1$. Assuming that the electromagnetic properties of air are the same as in vacuum, the potential function V , Eq. (8), will be

$$V(z) = k^2 [1 - \epsilon] \theta(z - h) - ik \frac{\sigma}{\epsilon_0 c} \theta(z - h) \equiv U \theta(z - h). \quad (45)$$

As the frequency (k) is taken to be fixed, U can be treated as a complex constant.

The Jost solutions, Eq. (11), for the step potential (45) are readily obtained by matching the solutions of Eq. (6) for $z \geq h$ using the continuity of the field Z and its vertical derivative at $z = h$:

$$f_\pm(z) = \theta(h - z) e^{\pm vz} + \theta(z - h) \frac{e^{\pm vh}}{2v'} \cdot \{ (v' + v) e^{\pm v'(z-h)} + (v' - v) e^{\mp v'(z-h)} \}. \quad (46)$$

Here we have used the notation

$$v' = \sqrt{v^2 + U}. \quad (47)$$

From Eq. (13) one obtains the reflection coefficient $S(v)$ as

$$S(v) = e^{-2vh} \frac{v - v'}{v + v'}. \quad (48)$$

The inverse Laplace transform of $S(v)$ is, Eq. (25),

$$\bar{S}(z) = -\frac{2}{z - 2h} J_2[(z - 2h)\sqrt{U}] \theta(z - 2h). \quad (49)$$

The kernel function $A(z, t)$ in the representation for the Jost functions, Eq. (20), may be obtained by taking an inverse Laplace transform with respect to v of Eq. (46):

$$A(z, t) = \frac{(t + z - 2h)\sqrt{U}}{2\sqrt{(z - h)^2 - (t - h)^2}} I_1 \{ \sqrt{[(z - h)^2 - (t - h)^2] U} \}. \quad (50)$$

With the help of the power series expansions of the Bessel functions J_2 and I_1 in the expressions for \bar{S} and A , it is then possible to demonstrate explicitly that the integral Eq. (27) is satisfied by the expressions (49) and (50). One can also easily verify that $A(z, t)$ does produce the original step potential $V(z)$ in Eq. (45) through Eq. (21 b).

We turn then to the illustration of the utility of the iterative algorithm described earlier. By the convergence criterion (34) one expects that the step potential (45) can be recovered from the reflection coefficient, Eq. (48), by the iterative algorithm for depth values H roughly satisfying the inequality

$$H^2 |U| \leq 1. \quad (51)$$

If the imaginary part of the potential U (the conductive term) is the one of main importance, which is the case in typical low-frequency prospecting work, it is easy to see that condition (51) is equivalent to

$$H \ll \delta = \sqrt{\frac{2}{\mu_0 \sigma \omega}}, \quad (52)$$

where δ is the skin depth. This is a very natural result as one ought not to expect any solution algorithm for the inverse scattering problem to yield a reliable answer for depths much beyond the skin depth where the radiation does not penetrate.

The lowest-order potential (Born approximation) V_1 is readily obtained from the reflection coefficient, Eq. (48) or (49) using Eq. (32a) as

$$V_1(z) = U \{J_0(\rho) - J_4(\rho)\} = U \left\{ 1 - \frac{\rho^2}{4} + \frac{5}{384} \rho^4 + O(\rho^6) \right\}, \quad (53)$$

where the variable ρ has been defined as

$$\rho = 2(z-h)\sqrt{U}. \quad (54)$$

As the first term in the power series expansion in Eq. (53) gives the exact result, Eq. (45), the higher terms represent the error of the Born approximation. The Born approximation is obviously reliable for $|\rho| \ll 1$, which is equivalent to the skin depth condition, Eq. (51).

The second- and third-order terms in the iterative expansion of the solution may be obtained from Eq. (32b) as

$$V_2(z) = \frac{16U}{\rho^2} J_2^2(\rho) = \frac{U}{4} \rho^2 \left[1 - \frac{\rho^2}{6} + O(\rho^4) \right] \quad (55)$$

$$V_3(z) = \frac{8U}{\rho} J_2(\rho)(D_1 + 3D_3) - \frac{4U}{\rho} \sum_{k=1}^{\infty} [(k+2)J_{k+2}(\rho) - (k-2)J_{k-2}(\rho)](D_{1+k} - D_{3+k}) = \frac{11U}{384} \rho^4 [1 + O(\rho^2)]. \quad (56)$$

We have used the notation

$$D_n = \begin{cases} \frac{\rho}{4-n^2} (J_1 J_n - J_2 J_{n-1}) - \frac{J_2 J_n}{n+2} & (n \neq 2) \\ \frac{1}{4}(1 - J_0^2) - \frac{1}{2} J_1^2 - \frac{1}{4} J_2^2 & (n=2), \end{cases} \quad (57)$$

where the argument of the Bessel function is ρ . By adding V_1 , V_2 and V_3 as given by these expressions, we obtain a result for the potential with an error that is proportional to ρ^6 .

In order to illustrate the convergence of the iterative expansion of the potential we have plotted the first, second and third iterative results for the potential as a function of the dimensionless depth variable $|\rho|$ in Fig. 3. We have chosen U to be purely imaginary, which corresponds to a conductivity step and makes

$$|\rho| = (z-h)\sqrt{f\sigma\{\text{Hz Sm}^{-1}\}}/178 \text{ m.}$$

The iterative method is seen to converge well up to depth $|\rho| \leq 1$. With a typical conductivity value $\sigma = 10^{-4}$ S/m corresponding to granite and frequency values $f = 100$ Hz and 1 kHz, the convergence depths with three terms in V are of the order of 1800 and 600 m, respectively.

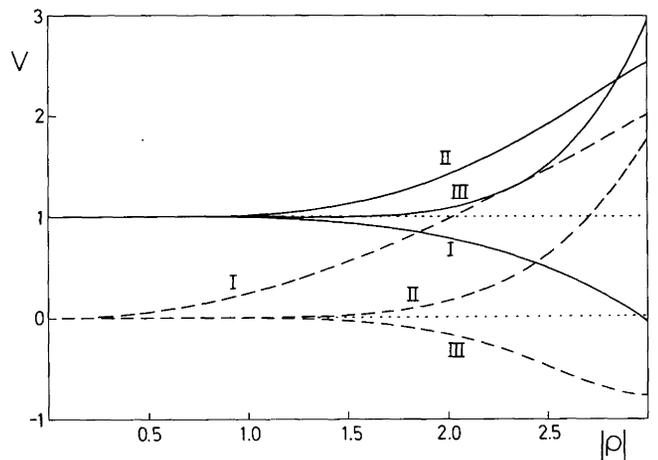


Fig. 3. First (I), second (II) and third (III) iterative results for a conductivity step potential as functions of the depth variable $|\rho|$. The solid lines denote the real and the dashed lines the imaginary parts of the iterated potential divided by the exact value. The deviation from 1 gives the error in the real and the deviation from 0 the error in the imaginary parts

Numerical examples

To examine the numerical quality of the solution to the inverse scattering problem using the Padé approximant method for obtaining the inverse Laplace transform of the reflection coefficient $S(v)$ [Eqs. (29) and (30)] we consider the following examples. We first study numerically the step discontinuity problem considered analytically in the previous section. In the second example we consider a step discontinuity, Eq. (45), with an additional layer of large conductivity, which corresponds to the geophysical situation of a high-conductivity layer in a background of low-conductivity rock. In the third example we consider the case of a staircase conductivity profile chosen to illustrate the solution in the case of an approximately continuously varying conductivity profile.

In all three examples above one can express the reflection coefficient $S(v)$, Eq. (16), as a combination of elementary functions. To generate synthetic data for the surface field one can then use these expressions for $S(v)$ in Eqs. (12) and (13) and the code of Anderson (1979) for the numerical evaluation of the Hankel transform. Going backwards from the calculated surface field using the same computer code, one can then obtain numerical values for the reflection coefficient that will have errors associated with the double numerical quadrature. In this way we obtain values for $S(v)$ at discrete real v -values (equidistant on a logarithmic scale to save computer time). A small number of these numerical values for $S(v)$ are then used to generate Padé approximants to $S(v)$ of the form of Eq. (29) by pointwise fitting (Baker, 1975). The inverse Laplace transform of the Padé approximant is finally evaluated analytically as described earlier. Finally, the inverse scattering integral equation, Eq. (27), is solved iteratively using Eqs. (31) and (32) including terms up to third order ($V_1 + V_2 + V_3$).

In Fig. 4 we display the result thus obtained for the potential step, Eq. (45), considered analytically in the

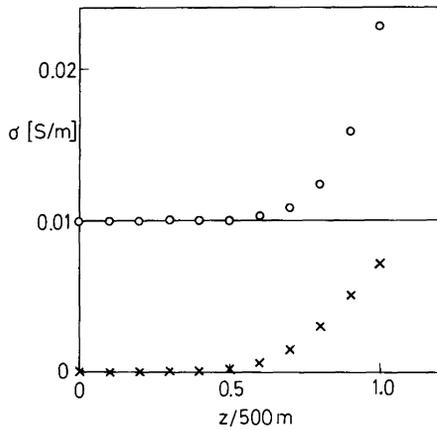


Fig. 4. Real (\circ) and imaginary (\times) parts of the reconstructed uniform conductivity profile as a function of depth, in units of the penetration depth for the frequency 100 Hz

previous section. In the figure the depth coordinate has been divided by the skin depth for a radiation frequency $f=100$ Hz. In this case a 7-point Padé approximant was used. The numerical results in this case are excellent up to half the skin depth.

In Fig. 5 we show the result of the numerical recovery of the conductivity profile in which the conductivity below ground is taken to be 10^{-3} S/m except in the layer $50\text{ m} < z < 70\text{ m}$ where it is taken to be 10^{-1} S/m. The depth scale in the figure has been divided by the skin depth in the high-conductivity layer, which is 160 m at the frequency 100 Hz. Here a 13-point Padé approximant was used. In this case the quality of the numerical recovery is only average: the region of high conductivity is recovered as a smooth bell-shaped distribution. In any case, the average strength of the region of high conductivity is satisfactorily recovered.

In Fig. 6 we show the result of the numerical recovery of a multi-step conductivity profile. The multi-step conductivity profile is superimposed on a uniform background conductivity of 0.025 S/m. The profile contains four conductivity increases of 0.025 S/m and four corresponding decreases of equal magnitude. Again a 13-point Padé approximant was used. The depth scale has again been divided by the minimum skin depth, which in this case is 140 m at 100 Hz. The results show that the quality of the inverse scattering solution is very satisfactory.

The conclusion from these results is that the numerical method based on the Padé approximant is most reliable when the conductivity profile is smooth. When there is a large conductivity contrast, as in the case considered in Fig. 5, the recovery of the original profile by the inverse scattering method is only qualitative, the reason being the inability of a finite Fourier series, Eq. (30), to represent sharp discontinuities. That this is indeed the source of the error can be verified by exact integration based on the Mellin formula, Eq. (28). In that case the reconstructed conductivity profile is in excellent agreement with the original one, in agreement with the convergence criterion, Eq. (34).

The results in Figs. 4–6 also show that the permittivity profile is poorly reconstructed. Although the

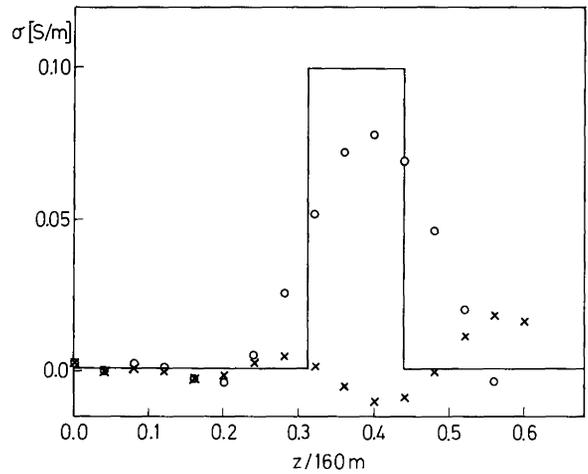


Fig. 5. Reconstructed conductivity profile (\circ =real part, \times =imaginary part) compared with the original profile (solid). The length unit is the skin depth in the high-conductivity layer for the frequency 100 Hz

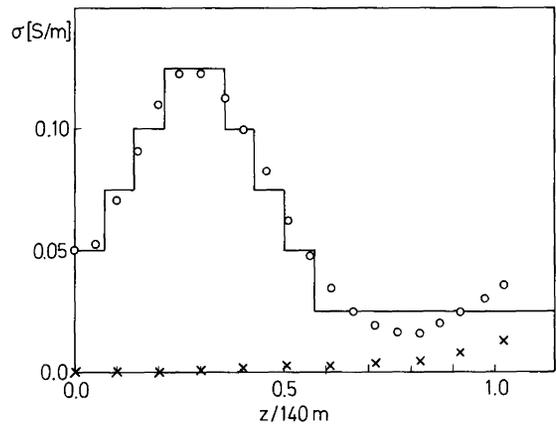


Fig. 6. Reconstruction of a staircase conductivity profile (\circ =real part, \times =imaginary part, solid=original profile). The length unit is the skin depth in the high-conductivity layer at 100 Hz

relative permittivity in all the models considered was kept at unity, the numerical results give a non-zero if small imaginary part for the conductivity, which by Eq. (8), corresponds to a non-unitary permittivity value. For a reliable recovery of the permittivity profile, the frequency has to be high enough to make the real and imaginary parts of the potential (8) of the same order of magnitude. This means, in practice, that the permittivity profile can be determined reliably by the inverse scattering method only in strata of very low conductivity as, otherwise, the skin depth will be too small to allow any useful information to be obtained.

Because a staircase variation of the permeability by Eq. (8) would make the potential singular and thus numerically difficult to recover, the permeability was not allowed to vary in these examples. A non-vanishing imaginary conductivity, however, might also be interpreted as a permeability profile with a non-vanishing second derivative. In order to separate the permeability, permittivity and conductivity profiles using the fre-

quency dependence as described at the end of the second section, one might be forced to use impracticably high frequencies. In practice, however, it might be a very good approximation to neglect the permeability variations completely.

Conclusions

The main qualitative result of the present development of the inverse scattering solution to the problem of determining the electromagnetic parameter profiles of a subterranean structure with only vertical variation from the surface reflection coefficient for electromagnetic dipole radiation is that a complete determination may be achieved by measurements at two fixed frequencies. This result substantially increases the utility of the inverse scattering method in interpretation of electromagnetic prospecting data as such data are invariably obtained with field apparatus with a restricted low-frequency range. Previous investigations of the inverse scattering method for electromagnetic prospecting work (Weidelt, 1972; Riska and Vidberg, 1983) dealt with reflection of plane wave radiation for which the solution to the inverse scattering problem requires knowledge of the reflection coefficient over the whole frequency spectrum $0 \leq \omega \leq \infty$, and thus had to confront the question of extrapolating the data into the unmeasurable high-frequency range.

The solution to the inverse scattering problem for dipole radiation at fixed frequency is obtained by solving a linear integral equation of the Marchenko-type (Agranovich and Marchenko, 1964), Eq. (27), based on Weidelt's (1972) representation for the Jost solutions, Eq. (20). This integral equation has the inverse Laplace transform of the reflection coefficient as a driving term. While the solution to the integral equation can be obtained by straightforward methods, the determination of the inverse Laplace transform requires a numerical inversion of Eq. (25). To solve the inverse scattering problem numerically we use the method of Padé approximants to determine the inverse Laplace transform of the reflection coefficient and then solve the inverse scattering integral equation by iteration. This method is shown to be adequate for the recovery of the potential function that contains the electromagnetic parameter profiles for depths less than the penetration depth. On physical grounds it should also be

obvious that the solution should be unstable for depths beyond the average skin depth of the subterranean structure which are essentially unprobed by the radiation field.

The solution to the inverse scattering problem has been developed here for the cases of horizontal and vertical electric and magnetic dipole radiation sources. The solutions in the case of the magnetic dipole sources and the horizontal electric dipole source are formally similar. In the case of a vertical electric dipole source the potential function containing the electromagnetic parameter profiles is more complicated than in the other cases and the disentangling of these profiles is far more complicated. The simpler cases of a vertical magnetic and horizontal electric dipole source are fortunately those that correspond to practical field equipment used in prospecting work.

References

- Agranovich, Z.S., Marchenko, V.A.: The inverse scattering theory. New York: Gordon and Breach, 1964
- Anderson, W.L.: Computer program. Numerical integration of related Hankel transforms of order 0 and 1 by adaptive digital filtering. *Geophysics* **44**, 1287-1305, 1979
- Baker, G.A. Jr.: Essentials of Padé approximants. New York: Academic Press, 1975
- Coen, S.: The inverse problem of the shear modulus and density profiles of a layered earth - torsional vibration data. *J. Math. Phys.* **22**, 2338-2341, 1981
- Jaulent, M.: Inverse scattering in absorbing media. *J. Math. Phys.* **17**, 1351-1360, 1976
- Jaulent, M., Jean, C.: The inverse s-wave scattering problem for a class of potentials depending on energy. *Commun. Math. Phys.* **28**, 177-220, 1972
- Jost, R., Kohn, W.: Construction of a potential from a phase shift. *Phys. Rev.* **87**, 977-992, 1952
- McWirther, J.G., Pike, E.R.: On the numerical inversion of the Laplace transform and similar Fredholm equations of the first kind. *J. Phys. A: Math. Gen.* **11**, 1729-1745, 1978
- Riska, D.O., Vidberg, H.J.: Iterative solution of the electromagnetic inversion problem for vertically varying earth. *Geophysica* **19**, 137-156, 1983
- Weidelt, P.: The inverse problem of geomagnetic induction. *Z. Geophys.* **38**, 257-289, 1972

Received November 14, 1983; Revised December 27, 1984
Accepted January 11, 1985

The method of integral equation in the direct current resistivity method and its accuracy

R. Schulz

Niedersächsisches Landesamt für Bodenforschung, Postfach 510153, D-3000 Hannover 51, Federal Republic of Germany

Abstract. Model calculations of potential are necessary to interpret quantitatively d-c geoelectrical measurements over three-dimensional bodies. The method of integral equation is very suitable for a body buried in a horizontally stratified half-space. In this approach, the integral representation for the surface potential is not valid for completely conducting bodies. On the other hand, the integral representation for the so-called surface charge density is generally valid.

Solving the integral equation numerically, discretization errors appear. These errors are not negligible because the numerical procedure has only a linear order of convergence. To improve the numerical results, an extrapolation can be used.

Model calculations demonstrate the importance of the ratio of the resistivities of the overburden and the substratum. The more conducting the substratum in relation to the overburden, the more significant is the effect which is caused by a conducting body.

Key words: D.c. resistivity method - Integral equation - 3D modelling - Boundary element method - Numerical accuracy

Introduction

In the direct current resistivity method, information about the real distribution of resistivity of the underground is gained by measurements of apparent resistivity on the surface of the earth. The apparent resistivity is the quotient of measured potential and introduced current, multiplied by a geometry factor. This factor is defined so that the apparent resistivity is equal to the real resistivity in the case of a homogeneous isotropic ground. The apparent resistivity of the frequently used Schlumberger array (Fig. 1) is formulated as follows:

$$\rho_a = \pi \frac{(AB/2)^2 - (MN/2)^2}{MN} \cdot \{U(\mathbf{r}_M, \mathbf{r}_A) - U(\mathbf{r}_N, \mathbf{r}_A) - U(\mathbf{r}_M, \mathbf{r}_B) + U(\mathbf{r}_N, \mathbf{r}_B)\} / I \quad (1)$$

and analogously for the Schlumberger half-array (Fig. 1)

$$\rho_a = 2\pi \frac{AO^2 - (MN/2)^2}{MN} \{U(\mathbf{r}_M, \mathbf{r}_A) - U(\mathbf{r}_N, \mathbf{r}_A)\} / I. \quad (2)$$

Theoretical values of potential distribution are necessary to interpret field measurements. There are model calculations (e.g. Koefoed, 1979) and inversion procedures (Mundry and Dennert, 1980) for the case of a horizontally stratified earth. The interpretation for more complicated geological structures is more difficult. Generally, potential distribution can be calculated only by numerical methods, e.g. by finite differences (Dey and Morrison, 1979; Mundry, 1984), finite elements (Coggon, 1971) or by the method of integral equations (Dieter et al., 1969; Okabe, 1981; 1982).

The integral equation approach is especially suitable in the case of three-dimensional problems. In this approach, the calculation can be reduced to a two-dimensional problem and its numerical solution can be carried out on a computer with an acceptable computation time.

Although the integral equation approach has been well known for a long time and is also used for elec-

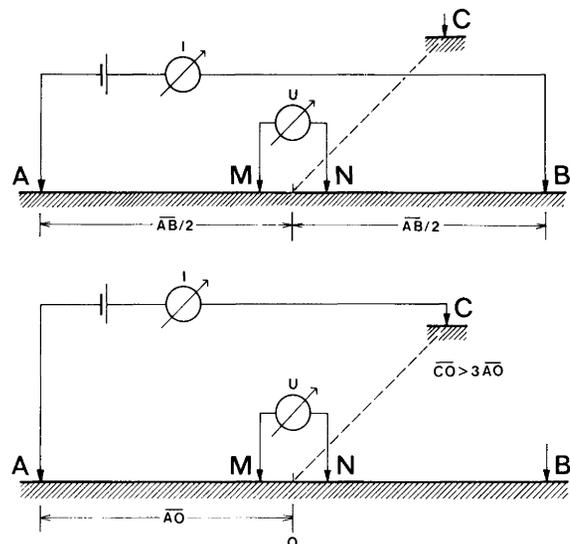


Fig. 1. Schlumberger array (above) and Schlumberger half-array (below). $AB/2 = |\mathbf{r}_A - \mathbf{r}_B|/2 = AO$; $MN = |\mathbf{r}_M - \mathbf{r}_N|$; $MN \ll AB$

tromagnetics, there are no papers about the numerical accuracy of the method.

Fundamental solutions of the potential equation

Suppose that a point source of current density I is located at \mathbf{r}_A , then the potential equation for an isotropic medium of conductivity σ can be written as

$$-\nabla \cdot \{\sigma(\mathbf{r}) \nabla U(\mathbf{r}, \mathbf{r}_A)\} = I \delta(\mathbf{r} - \mathbf{r}_A), \quad (3)$$

where δ is the Dirac function. For a homogeneous space of resistivity ρ_1 , Eq.(3) is reduced to Poisson's equation

$$\Delta U(\mathbf{r}, \mathbf{r}_A) = -I \rho_1 \delta(\mathbf{r} - \mathbf{r}_A). \quad (4)$$

Its unique solution is given by

$$U(\mathbf{r}, \mathbf{r}_A) = \frac{I \rho_1}{4\pi} \frac{1}{|\mathbf{r} - \mathbf{r}_A|}. \quad (5)$$

For a homogeneous half-space the potential can be derived by the method of images

$$U(\mathbf{r}, \mathbf{r}_A) = \frac{I \rho_1}{4\pi} \left(\frac{1}{|\mathbf{r} - \mathbf{r}_A|} + \frac{1}{|\mathbf{r} - \tilde{\mathbf{r}}_A|} \right) \quad (6)$$

with $\tilde{\mathbf{r}}_A = (x_A, y_A, -z_A)$ as the image of the source point.

On boundary interfaces S , at which the conductivity is discontinuously changing, the normal current density and the potential itself are continuous:

$$\sigma(\mathbf{r}_0^+) \mathbf{n} \cdot \nabla U(\mathbf{r}_0^+, \mathbf{r}_A) - \sigma(\mathbf{r}_0^-) \mathbf{n} \cdot \nabla U(\mathbf{r}_0^-, \mathbf{r}_A) = 0, \quad (7)$$

$$U(\mathbf{r}_0^+, \mathbf{r}_A) - U(\mathbf{r}_0^-, \mathbf{r}_A) = 0, \quad (8)$$

where \mathbf{r}_0 is a point on S ; \mathbf{n} is the normal unit vector of S , which is outwardly directed in the case of closed surfaces or has the direction of the positive coordinates in the case of a plane. The superscripts $+$ and $-$ at \mathbf{r}_0 denote the direction of limiting.

In electrostatics, charges are influenced at an uncharged body by a point source. The sum of the influenced charges over the whole body vanishes. The charge influenced at the surface S of the body is determined by the Gauss theorem

$$\hat{\mu}(\mathbf{r}_0) = \varepsilon_0 [\mathbf{n} \cdot \mathbf{E}(\mathbf{r}_0^+, \mathbf{r}_A) - \mathbf{n} \cdot \mathbf{E}(\mathbf{r}_0^-, \mathbf{r}_A)],$$

where \mathbf{r}_0 is a point on S , $\hat{\mu}$ is the surface charge density (As/m^2) and ε_0 is the dielectric constant (e.g. Smythe, 1968). The analogue of the Gauss theorem is formulated for stationary currents as follows:

$$\tilde{\mu}(\mathbf{r}_0) = \sigma [\mathbf{n} \cdot \mathbf{E}(\mathbf{r}_0^+, \mathbf{r}_A) - \mathbf{n} \cdot \mathbf{E}(\mathbf{r}_0^-, \mathbf{r}_A)].$$

$\tilde{\mu}$ is now a surface current density (A/m^2). In the original paper (Dieter et al., 1969) and in papers following it, the value $\mu = \tilde{\mu}/\sigma$ is used for the derivation of the integral equation; μ (V/m) was denoted as surface charge density. The same notation will be used in this paper.

The normal derivative of potential is discontinuous at interface boundaries, Eq. (7), and the discontinuity is defined as

$$\mu(\mathbf{r}_0) = \mathbf{n} \cdot \nabla U(\mathbf{r}_0^-, \mathbf{r}_A) - \mathbf{n} \cdot \nabla U(\mathbf{r}_0^+, \mathbf{r}_A). \quad (9)$$

In the case of a horizontally stratified half-space the conductivity only depends on the depth z and is constant in each layer:

$$\sigma(z) = \sigma_i; \quad z_{i-1} \leq z < z_i, \quad i = 1, \dots, n; \quad z_0 = 0, \quad z_n = \infty.$$

The thickness of the i -th layer is denoted by

$$h_i = z_i - z_{i-1}.$$

In order to derive the integral representation and to solve the integral equation, the potential has to be formulated for any arbitrary point in space. The potential can be calculated by a recursive procedure, used also in electromagnetics (Weidelt, 1978). Exponential expressions with positive exponents are avoided in this recursive procedure. In doing this, good stability in numerical calculation is obtained.

The differential equation, Eq. (3), is solved by separation into cylindrical coordinates (e.g. Morse and Feshbach, 1953). The complete solution is given by

$$U(\mathbf{r}, \mathbf{r}_A) = \frac{I}{2\pi} \int_0^\infty f(z, z_A, \lambda) J_0(\lambda \bar{r}) d\lambda, \quad (10)$$

with $\bar{r} = \sqrt{(x - x_A)^2 + (y - y_A)^2}$. J_0 is the Bessel function of order zero. The unknown function f has to be derived from the differential equation, Eq. (3), and the boundary conditions, Eqs. (7) and (8). Assuming a fictitious boundary plane z_k for the plane of the point source ($z_k = z_A$) and assuming that the point \mathbf{r} is situated below r_A in the j -th layer ($z_A \leq z_{j-1} \leq z < z_j$), function f can be given at the boundary z_{j-1} by (Schulz, 1983)

$$f(z_{j-1}, z_A, \lambda) = \frac{e^{-\lambda(z_{j-1} - z_A)}}{\sigma(z_A)(T_{k+1} + R_{k+1})} \prod_{i=k+1}^{j-1} \frac{1 + T_i}{1 + \frac{\sigma_{i+1}}{\sigma_i} T_{i+1}}. \quad (11)$$

Expressions T_i and R_i are calculated by the following recursion formulae:

$$T_i = \begin{cases} 1; & i = n \\ \frac{\sigma_i \tanh(\lambda h_i) + \sigma_{i+1} T_{i+1}}{\sigma_i + \sigma_{i+1} T_{i+1} \tanh(\lambda h_i)}; & i = n-1, \dots, k+1 \end{cases} \quad (12a)$$

$$R_i = \begin{cases} 0; & i = 1 \\ \frac{\sigma_{i-1} \tanh(\lambda h_{i-1}) + R_{i-1}}{\sigma_i R_{i-1} \tanh(\lambda h_{i-1}) + 1}; & i = 2, \dots, k+1. \end{cases} \quad (12b)$$

Now the function f , Eq. (11), can be continued for an arbitrary point in the j -th layer. If $j = n$, i.e. \mathbf{r} is a point in the last layer, then

$$f(z, z_A, \lambda) = f(z_{n-1}, z_A, \lambda) e^{-\lambda(z - z_{n-1})}. \quad (13a)$$

Otherwise, if $j < n$, then

$$f(z, z_A, \lambda) = f(z_{j-1}, z_A, \lambda) \{ \cosh[\lambda(z - z_{j-1})] - T_j \sinh[\lambda(z - z_{j-1})] \}. \quad (14)$$

In this equation, a difference of high values can appear. For stability, it is useful to reconstruct Eq. (14) in such

a manner that only negative exponents will appear. Equation (12a) yields

$$\frac{\sigma_i + \sigma_{i+1} T_{i+1}}{\sigma_i - \sigma_{i+1} T_{i+1}} e^{\lambda h_i} = \frac{1 + T_i}{1 - T_i} e^{-\lambda h_i}.$$

Inserting this into Eq. (14) we get, for $z_{j-1} \leq z < z_j$ with $j < n$,

$$f(z, z_A, \lambda) = 1/2 f(z_{j-1}, z_A, \lambda) (1 + T_j) \cdot \left[e^{-\lambda(z-z_{j-1})} + e^{-\lambda(z_j-z+h_j)} \frac{\sigma_j - \sigma_{j+1} T_{j+1}}{\sigma_j + \sigma_{j+1} T_{j+1}} \right]. \quad (13b)$$

Equation (13), connected with Eq. (11), is valid for $z \geq z_A$; while reciprocity can be used for $z < z_A$.

In the special case, of source and measuring point lying at the surface of the earth ($z = z_A = 0$), Eq. (11) is reduced to

$$f(0, 0, \lambda) = 1/\sigma_1 T_1,$$

where T_1 is calculated by Eq. (12a). The potential, Eq. (10), becomes

$$U(\mathbf{r}, \mathbf{r}_A) = \frac{I}{2\pi\sigma_1} \int_0^\infty \frac{1}{T_1(\lambda)} J_0(\lambda \bar{r}) d\lambda.$$

This representation is identical with the formula of Pekeris cited by Koefoed (1979), considering that the kernel K_i of Koefoed is identical to the expression $1/T_i$ of this paper.

The normal derivatives $\mathbf{n} \cdot \nabla U$ and, therefore, the partial derivatives of U are needed for the derivation of the integral representation of the potential of a buried body. Considering $J_0(r) = -J_1(r)$, we obtain

$$\frac{\partial}{\partial x} U(\mathbf{r}, \mathbf{r}_A) = -\frac{I}{2\pi} \frac{x - x_A}{\bar{r}} \int_0^\infty f(z, z_A, \lambda) \lambda J_1(\lambda \bar{r}) d\lambda,$$

$$\frac{\partial}{\partial y} U(\mathbf{r}, \mathbf{r}_A) = -\frac{I}{2\pi} \frac{y - y_A}{\bar{r}} \int_0^\infty f(z, z_A, \lambda) \lambda J_1(\lambda \bar{r}) d\lambda,$$

$$\frac{\partial}{\partial z} U(\mathbf{r}, \mathbf{r}_A) = \frac{I}{2\pi} \int_0^\infty \frac{\partial}{\partial z} f(z, z_A, \lambda) J_0(\lambda \bar{r}) d\lambda.$$

All the integrals above and the integral of Eq. (10) have the form

$$h(\bar{r}) = \int_0^\infty g(\lambda) J_v(\lambda \bar{r}) d\lambda; \quad v = 0, 1.$$

Such integrals can be calculated by the theory of linear filters (Ghosh, 1971a; b). Filter coefficients, upper and lower limits of summation depend on the discretization width and the required accuracy. The coefficients in this paper are computed by the method of fast Hankel transforms published by Johansen and Sørensen (1979).

The method of integral equation

The potential of a buried body (Fig. 2) can be given by an integral representation. The potential outside the body will be denoted by $U^{(1)}$, the potential within the body by $U^{(2)}$. The body K with surface S has con-

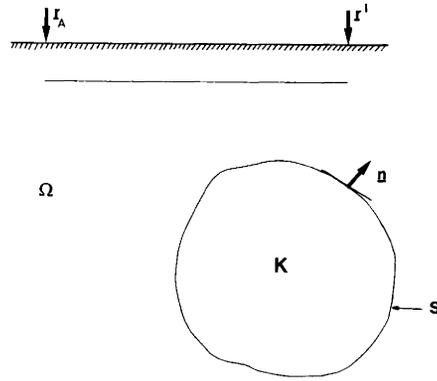


Fig. 2. Body buried in an inhomogeneous half-space

ductivity $\sigma_K(\mathbf{r})$; the space outside the body is denoted by Ω . The surface S is formed so that Green's theorems can be applied. The extended derivation of the following formulae can be found in Okabe (1981; 1982) and Schulz (1983).

It is supposed that fundamental solutions $\Psi^{(1)}$ in the space and $\Psi^{(2)}$ in the body exist, governed by:

$$\begin{aligned} -\nabla \cdot \{\sigma(\mathbf{r}) \nabla \psi^{(1)}(\mathbf{r}, \mathbf{r}')\} &= \delta(\mathbf{r} - \mathbf{r}'); & \mathbf{r} \in \Omega \\ -\nabla \cdot \{\sigma_K(\mathbf{r}) \nabla \psi^{(2)}(\mathbf{r}, \mathbf{r}')\} &= \delta(\mathbf{r} - \mathbf{r}'); & \mathbf{r} \in K. \end{aligned} \quad (15)$$

Fundamental solutions are, for example

$$\Psi(\mathbf{r}, \mathbf{r}') = U(\mathbf{r}, \mathbf{r}')/I$$

with the potential U of

- Eq. (5) for a homogeneous space,
- Eq. (6) for a half-space,
- Eq. (10) for a layered half-space.

To apply Green's formulae, fundamental solutions, Eq. (15), and differential expressions of the potential, Eq. (3), are multiplied and integrated over corresponding domains.

Assuming the source point \mathbf{r}_A and the point \mathbf{r} outside the body, the potential can be formulated as follows (differentiations and integrations - *do* - are taken with respect to the variable \mathbf{r}):

$$\begin{aligned} U^{(1)}(\mathbf{r}', \mathbf{r}_A) &= I \psi^{(1)}(\mathbf{r}_A, \mathbf{r}') \\ &+ \int_S \{ U^{(1)}(\mathbf{r}, \mathbf{r}_A) [\sigma(\mathbf{r}) \mathbf{n} \cdot \nabla \psi^{(1)}(\mathbf{r}, \mathbf{r}') - \sigma_K(\mathbf{r}) \mathbf{n} \cdot \nabla \psi^{(2)}(\mathbf{r}, \mathbf{r}')] \\ &+ \sigma(\mathbf{r}) \mathbf{n} \cdot \nabla U^{(1)}(\mathbf{r}, \mathbf{r}_A) [\psi^{(2)}(\mathbf{r}, \mathbf{r}') - \psi^{(1)}(\mathbf{r}, \mathbf{r}')] \} d\sigma. \end{aligned} \quad (16)$$

The potential of a point source for a buried body consists of the usual potential and a disturbed potential depending on the unknown surface potential and its derivative. Now this integral representation has to be reduced by Green's formula in such a way that the integral will be carried out over the potential itself or over its normal derivative.

The second term of Eq. (16)

$$\int_S \sigma(\mathbf{r}) \mathbf{n} \cdot \nabla U^{(1)}(\mathbf{r}, \mathbf{r}_A) \{ \psi^{(2)}(\mathbf{r}, \mathbf{r}') - \psi^{(1)}(\mathbf{r}, \mathbf{r}') \} d\sigma \quad (16a)$$

can be rebuilt by boundary condition, Eq. (7), to form the expression

$$\int_S \sigma_K(\mathbf{r}) \mathbf{n} \cdot \nabla U^{(2)}(\mathbf{r}, \mathbf{r}_A) \{ \psi^{(2)}(\mathbf{r}, \mathbf{r}') - \psi^{(1)}(\mathbf{r}, \mathbf{r}') \} d\mathbf{o},$$

to which Green's second formula can be applied. This procedure is only valid if the conductivity of the body is constant: $\sigma_K(\mathbf{r}) = \sigma_K$. As the potential $U^{(2)}$ and the fundamental solutions $\Psi^{(1)}$ and $\psi^{(2)}$ satisfy the potential equations, Eqs. (3) and (15), within the body, the volume integral vanishes. From Eq. (16a) we obtain

$$\int_S \sigma_K U^{(2)}(\mathbf{r}, \mathbf{r}_A) \mathbf{n} \cdot \nabla \{ \psi^{(2)}(\mathbf{r}, \mathbf{r}') - \psi^{(1)}(\mathbf{r}, \mathbf{r}') \} d\mathbf{o}.$$

This term is inserted into Eq. (16) instead of the second integral term and with respect to the continuity of the potential, Eq. (8), the following integral representation is obtained

$$U^{(1)}(\mathbf{r}', \mathbf{r}_A) = I \psi^{(1)}(\mathbf{r}_A, \mathbf{r}') + \int_S U^{(1)}(\mathbf{r}, \mathbf{r}_A) [\sigma(\mathbf{r}) - \sigma_K] \mathbf{n} \cdot \nabla \psi^{(1)}(\mathbf{r}, \mathbf{r}') d\mathbf{o}. \quad (17)$$

This representation is only valid if

$$\sigma_K < \infty, \quad \text{i.e. } \rho_K \neq 0. \quad (18)$$

The unknown surface potential is determined by an integral equation. If the point \mathbf{r}' tends to a point \mathbf{r}_0 of the surface S , we obtain

$$U(\mathbf{r}_0^+, \mathbf{r}_A) = I \psi(\mathbf{r}_A, \mathbf{r}_0) + \frac{\sigma(\mathbf{r}_0) - \sigma_K}{2\sigma(\mathbf{r}_0)} U(\mathbf{r}_0, \mathbf{r}_A) + \int_S U(\mathbf{r}, \mathbf{r}_A) [\sigma(\mathbf{r}) - \sigma_K] \mathbf{n} \cdot \nabla \psi(\mathbf{r}, \mathbf{r}_0) d\mathbf{o}.$$

The integral becomes singular, but convergent; the integration has to be carried out without the singularity at \mathbf{r}_0 . The indices (1), appearing at potential and fundamental solutions, are omitted.

The equation for the determination of the surface potential is obtained

$$\frac{\sigma(\mathbf{r}_0) + \sigma_K}{2\sigma(\mathbf{r}_0)} U(\mathbf{r}_0, \mathbf{r}_A) = I \psi(\mathbf{r}_A, \mathbf{r}_0) + \int_S U(\mathbf{r}, \mathbf{r}_A) [\sigma(\mathbf{r}) - \sigma_K] \mathbf{n} \cdot \nabla \psi(\mathbf{r}, \mathbf{r}_0) d\mathbf{o}. \quad (19)$$

Every constant potential is an eigensolution of Eq. (19) in the case of σ_K tending to infinity. Therefore, and because of Eq. (18), the case of a highly conducting body, which is very important for the exploration of ore, can not be treated by the integral equation for surface potential. This method produces incorrect results for cases of resistivity ρ_K tending to 0 (see Table 2). It is for this reason that this method, published by Barthes and Vasseur (1978), will not be used further.

The first term in Eq. (16)

$$\int_S U^{(1)}(\mathbf{r}, \mathbf{r}_A) \{ \sigma(\mathbf{r}) \mathbf{n} \cdot \nabla \psi^{(1)}(\mathbf{r}, \mathbf{r}') - \sigma_K(\mathbf{r}) \mathbf{n} \cdot \nabla \psi^{(2)}(\mathbf{r}, \mathbf{r}') \} d\mathbf{o} \quad (16b)$$

will now be rebuilt in such a way that only the derivative of the potential will appear. $U^{(1)}$ is replaced by $U^{(2)}$ because of the continuity of the potential, Eq. (8). In analogy to the derivation above, the following expression is obtained by Green's formula

$$\int_S \mathbf{n} \cdot \nabla U^{(2)}(\mathbf{r}, \mathbf{r}_A) \{ \sigma(\mathbf{r}) \psi^{(1)}(\mathbf{r}, \mathbf{r}') - \sigma_K \psi^{(2)}(\mathbf{r}, \mathbf{r}') \} d\mathbf{o}.$$

According to the boundary condition, Eq. (7), the integral representation, Eq. (16), can be reconstructed to give

$$U(\mathbf{r}', \mathbf{r}_A) = I \psi(\mathbf{r}_A, \mathbf{r}') + \int_S \sigma(\mathbf{r}) \mathbf{n} \cdot \nabla U^{(1)}(\mathbf{r}, \mathbf{r}_A) \frac{\sigma(\mathbf{r}) - \sigma_K}{\sigma_K} \psi^{(1)}(\mathbf{r}, \mathbf{r}') d\mathbf{o}. \quad (20)$$

The unknown normal derivative of potential is replaced by the also unknown surface charge density. Equations (7) and (9) yield

$$\mathbf{n} \cdot \nabla U^{(1)}(\mathbf{r}_0^+, \mathbf{r}_A) = \sigma_K \mu(\mathbf{r}_0) / \{ \sigma(\mathbf{r}_0) - \sigma_K \}.$$

Replacing this in Eq. (20), the integral representation over the surface charge density is obtained

$$U(\mathbf{r}', \mathbf{r}_A) = I \psi(\mathbf{r}_A, \mathbf{r}') + \int_S \sigma(\mathbf{r}) \mu(\mathbf{r}) \psi(\mathbf{r}, \mathbf{r}') d\mathbf{o}. \quad (21)$$

In this representation the surface charge density μ is unknown. Differentiating this equation with respect to \mathbf{r}' and letting \mathbf{r}' tend to a point \mathbf{r}_0 on the surface of the body, an integral equation for the surface charge density is obtained:

$$\frac{1}{2} \frac{\sigma(\mathbf{r}_0) + \sigma_K}{\sigma(\mathbf{r}_0) - \sigma_K} \mu(\mathbf{r}_0) = I \mathbf{n} \cdot \nabla \psi(\mathbf{r}_A, \mathbf{r}_0) + \int_S \mu(\mathbf{r}) \sigma(\mathbf{r}) \mathbf{n} \cdot \nabla \psi(\mathbf{r}, \mathbf{r}_0) d\mathbf{o}. \quad (22)$$

Inserting the solution of this integral equation into Eq. (21), the potential U can be calculated for every point \mathbf{r}' outside the body. In the special case of a homogeneous half-space, Eqs. (21) and (22) with the corresponding fundamental solution yield the well known integral representation and equation (Dieter et al., 1969).

Without entering into the particulars, it should be noted that there are other integral representations and methods to solve the problem, see e.g. Lee (1975).

The numerical treatment of the integral equation

Integral equation, Eq. (22), can generally be solved only by numerical methods. The unknown surface charge density can be approximated by piecewise defined interpolation polynoms, the coefficients of which have to be determined (boundary element method). In this approach, the surface S of the body is completely divided into N boundary elements S_q (Fig. 3). For practical treatment, the boundary element method is used in the simplest way (Riehle, 1979): the unknown function μ is approximated by a function μ_N , which is constant on every boundary element S_q .

$$\mu_N(\mathbf{r}) = \sum_{q=1}^N d_q(\mathbf{r}) \mu_q \quad \text{with } d_q(\mathbf{r}) = \begin{cases} 1; & \mathbf{r} \in S_q \\ 0; & \mathbf{r} \notin S_q \end{cases}$$

The integral equation, Eq. (22), is transformed by this approach to a system of linear equations

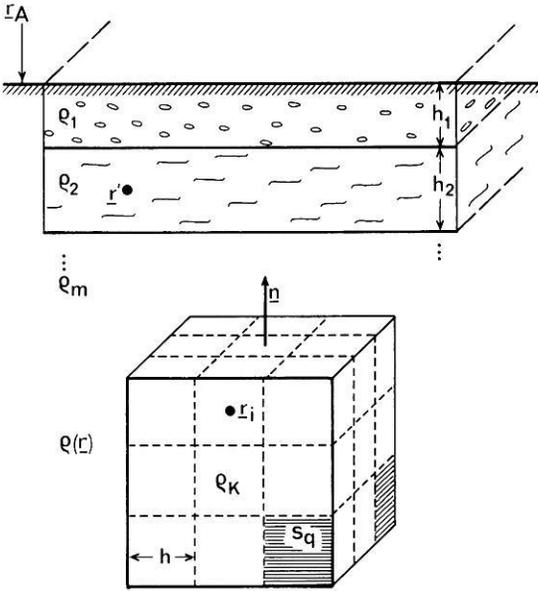


Fig. 3. Three-dimensional body (cube) in a horizontally stratified half-space; dissection of the surface into boundary elements

$$\frac{1}{2} \frac{\sigma(\mathbf{r}_i) + \sigma_K}{\sigma(\mathbf{r}_i) - \sigma_K} \mu_i = \mathbf{I} \mathbf{n} \cdot \nabla \psi(\mathbf{r}_A, \mathbf{r}_i) + \sum_{q=1}^N \mu_q \int_{S_q} \sigma(\mathbf{r}) \mathbf{n} \cdot \nabla \psi(\mathbf{r}, \mathbf{r}_i) d\sigma; \quad i = 1, \dots, N. \quad (23)$$

The potential U , Eq. (21), is approximately represented, with the solution of Eq. (23), by

$$U_N(\mathbf{r}', \mathbf{r}_A) = I \psi(\mathbf{r}_A, \mathbf{r}') + \sum_{q=1}^N \mu_q \int_{S_q} \sigma(\mathbf{r}) \psi(\mathbf{r}, \mathbf{r}') d\sigma. \quad (24)$$

The integrals over the fundamental solutions can be exactly calculated only in special cases, e.g. if Ω is a homogeneous half-space and the boundary elements S_q are planar facets (see Barnett, 1972). But it is sufficient to calculate the integrals numerically. The quadrature formula should be chosen in such a way that the integrals over the approximating functions are calculated exactly. In our case it is sufficient to use the simplest Gauss quadrature formula. Then the linear equations, Eq. (23), have the form

$$\frac{1}{2} \frac{\sigma(\mathbf{r}_i) + \sigma_K}{\sigma(\mathbf{r}_i) - \sigma_K} \bar{\mu}_i = \mathbf{I} \mathbf{n} \cdot \nabla \psi(\mathbf{r}_A, \mathbf{r}_i) + \sum_{q=1}^N \bar{\mu}_q \sigma(\mathbf{r}_q) \mathbf{n} \cdot \nabla \psi(\mathbf{r}_q, \mathbf{r}_i) |S_q|; \quad i = 1, \dots, N, \quad (25)$$

where \mathbf{r}_q is the centre of gravity and $|S_q|$ is the area of the boundary element S_q . The potential, Eq. (24), runs as follows

$$\bar{U}_N(\mathbf{r}', \mathbf{r}_A) = I \psi(\mathbf{r}_A, \mathbf{r}') + \sum_{q=1}^N \bar{\mu}_q \sigma(\mathbf{r}_q) \psi(\mathbf{r}_q, \mathbf{r}') |S_q| \quad (26).$$

The same system of equations and expression for the potential can be obtained by using Gauss simplest

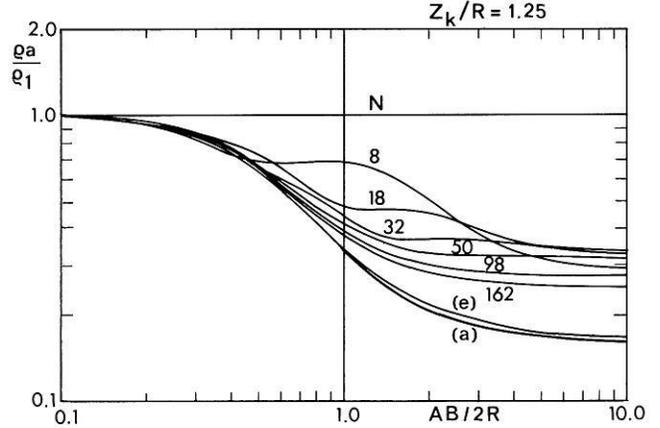


Fig. 4. Influence of the numerical discretization, represented by model curves of apparent resistivity of a Schlumberger sounding array for a sphere (radius R , depth of the centre $1.25 R$, resistivity $\rho_k/\rho_1=0$) buried in a homogeneous half-space. Point of measurement: epicentre of the sphere. N : number of boundary elements; (e): extrapolated values, Eq. (28); (a): analytical (exact) values

Table 1. Extrapolation table for the numerical solution of the integral equation, calculated for one model value of the apparent resistivity ρ_a/ρ_1 of a Schlumberger array ($AB/2 = 100 R$) for a sphere (radius R , depth of the centre $1.25 R$, resistivity $\rho_k/\rho_1=0$) buried in a homogeneous half-space (resistivity ρ_1). Point of measurement: epicentre of the sphere. N : number of boundary elements, h_N : width of the elements, which are quadratic in bi-spherical coordinates, $\rho_a^{(N)}$: value calculated by Eq. (26), $\rho_a^{(a)}$: analytical value (Snyder and Merkel, 1973), V_{ik}, W_{ik} : extrapolated values (Stoer, 1972) with

$$V_{ik} = \sum_{j=i-k}^i c_{kj}^{(i)} V_{j0} \quad \text{with} \quad c_{kj}^{(i)} = \prod_{l=i-k}^i \frac{h_l}{(h_l - h_j)}$$

$$W_{ik} = 2V_{i+1,k} - V_{ik}$$

($i = 2, \dots, 12$; $k = 1, 2, 3$)

N	h_N	$i = \pi/h_N$	$\rho_a^{(N)} = V_{i0}$	V_{i1}	V_{i2}	V_{i3}	$\rho_a^{(a)}$
8	1.57	2	0.2810				
18	1.05	3	0.3205	0.3995			
32	0.78	4	0.3267	0.3454	0.2914		
50	0.63	5	0.3115	0.2508	0.1087	-0.0131	
72	0.52	6	0.2923	0.1962	0.0871	0.0654	
98	0.45	7	0.2754	0.1739	0.1182	0.1596	
128	0.39	8	0.2618	0.1668	0.1455	0.1911	
162	0.35	9	0.2511	0.1651	0.1592	0.1866	
200	0.31	10	0.2425	0.1649	0.1638	0.1745	
242	0.29	11	0.2354	0.1648	0.1643	0.1657	
288	0.26	12	0.2295	0.1646	0.1636	0.1612	
		11	0.2235	0.1644	0.1628	0.1568	0.1596
		10	0.2283	0.1647	0.1649	0.1570	
		9	0.2338	0.1646	0.1684	0.1624	
		8	0.2403	0.1634	0.1729	0.1822	
		7	0.2482	0.1597	0.1728	0.2250	
		6	0.2585	0.1516	0.1493	0.2539	
		5	0.2731	0.1416	0.0654	0.1438	
		4	0.2963	0.1560	-0.0740		
		3	0.3330	0.2914			
		2	0.3600				
		i	W_{i0}	W_{i1}	W_{i2}	W_{i3}	$\rho_a^{(a)}$

Table 2. Comparison of different methods of calculation of apparent resistivity for a buried body. Model: Sphere (radius R , resistivity ρ_K , depth of the centre $Z_K/R=2$) in a homogeneous half-space (resistivity ρ_1). Array: Schlumberger array with $AB/2 = 100R$. Parameters of the model: X/R , Y/R centre of the Schlumberger array; Z_K/R depth of the centre of the sphere; ρ_K/ρ_1 ratio of the resistivities. Analytical method: see Brass et al. (1981). Integral equation: surface charge density, Eq. (21) and Eq. (26); surface potential, Eq. (17); N =number of boundary elements; Extrap=extrapolated values, Eq. (28); *=values which are not correct, Eq. (18)

Parameter				Analytical method	Integral equation					
X/R	Y/R	Z_K/R	ρ_K/ρ_1		Surface charge density			Surface potential		
					$N=98$	$N=162$	Extrap.	$N=98$	$N=162$	Extrap.
0.0	0.0	2.0	0.0	0.7560	0.7938	0.7864	0.7605	0.7897	0.7830	0.7595
0.0	0.0	2.0	0.1	0.8160	0.8383	0.8338	0.8179	0.8357	0.8316	0.8174
0.0	0.0	2.0	0.2	0.8592	0.8727	0.8699	0.8600	0.8710	0.8685	0.8597
0.0	0.0	2.0	0.5	0.9379	0.9407	0.9401	0.9379	0.9403	0.9398	0.9379
0.0	0.0	2.0	2.0	1.0502	1.0520	1.0516	1.0502	1.0517	1.0514	1.0502
0.0	0.0	2.0	10.0	1.1082	1.1174	1.1152	1.1078	1.1157	1.1139	1.1079
0.0	0.0	2.0	9999.0	1.1264	1.1392	1.1362	1.1257	1.1367	1.1343	1.1259
0.5	0.0	2.0	0.0	0.8160	0.8452	0.8395	0.8196	0.7662 *	0.7692 *	0.7794 *
0.5	0.0	2.0	0.1	0.8613	0.8785	0.8751	0.8629	0.8566	0.8583	0.8644
0.5	0.0	2.0	0.2	0.8939	0.9043	0.9022	0.8946	0.8938	0.8945	0.8968
0.5	0.0	2.0	0.5	0.9532	0.9554	0.9549	0.9532	0.9533	0.9535	0.9542
0.5	0.0	2.0	2.0	1.0377	1.0392	1.0388	1.0377	1.0392	1.0387	1.0372
0.5	0.0	2.0	10.0	1.0813	1.0884	1.0868	1.0809	1.0869	1.0854	1.0802
0.5	0.0	2.0	9999.0	1.0950	1.1049	1.1026	1.0944	1.1025	1.1006	1.0937

quadrature formula in Eqs. (21) and (22). Then $\bar{\mu}_q$ has the meaning of the value of μ at the centre of gravity, \mathbf{r}_q , but $\mu(\mathbf{r}_q)$ is generally not equal to $\bar{\mu}_q$.

Convergence of the numerical solution

The problem of how the numerical solution $\bar{\mu}_N$ of Eq. (25) converges to the exact solution μ of Eq. (22) for $N \rightarrow \infty$ has not been analysed yet. The difference between numerical and exact solution can be significant, as Fig. 4 shows. The problem can be solved when theorems of convergence for singular integral equations (Michlin and Pröbldorf, 1980) are applied.

The solution of the last section can be interpreted as a special case of the Ritz-Galerkin method (Michlin and Pröbldorf, 1980, Chapter 18). In the Appendix, Eq. (33), the convergence of the numerical solution to the exact one is proved; the procedure is stable, and the discretization error is estimated by Eq. (35).

In order to get an impression of the dimensions of this error, the estimation is analysed for a special case: it is assumed that the surface S of the body can be divided into boundary elements S_q , each of which is a square of length h (see Fig. 3). The error between the numerically calculated potential U_N , Eq. (24), and the exact potential U , Eq. (21), can be estimated by Eq. (40):

$$|U_N - U| \leq c \cdot h. \quad (27)$$

c is a constant, which depends on the continuity of the function and generally can not be more exactly determined. If a quadrature formula is used for solving the integral equation, the estimation of the approximation error is not essentially changed, Eq. (41). The error between the exact potential, Eq. (21), and the potential

calculated completely numerically, Eq. (26), is estimated by

$$|\bar{U}_N - U| \leq c \cdot h + d \cdot h^2.$$

The quadrature formula is less important with regard to convergence. The order of convergence is not improved by exact integration. Equation (27) shows that the considered numerical procedure for solving the integral equation has only a linear order of convergence. This is verified by model calculations. Figure 4 shows model curves for a Schlumberger array (see Eq. (1)) for a sphere ($\rho_K/\rho_1=0$) buried in a homogeneous half-space. The curves are calculated by solving the integral equation numerically, Eqs. (25) and (26). Beforehand, the Cartesian coordinates must have been transformed into bi-spherical coordinates and the integral equation must have been discretized in this coordinate system (Schulz, 1983). So the surface of the sphere has not been approximated by a polygon, as in Barnett (1972). Nevertheless, it is obvious that the numerical solution significantly differs from the analytical solution even if the number of boundary elements is relatively high. The analytical solution is given by a sum of Legendre, trigonometric and exponential functions (e.g. Snyder and Merkel, 1973).

The order of convergence, Eq. (27), can be used to increase the accuracy by an extrapolation technique. Table 1 shows extrapolated values, which are calculated by Lagrange's interpolation formula (Stoer, 1972). It is difficult to give an exact mathematical proof of the validity of this extrapolation, but in all cases the extrapolated values approximate the exact ones very well.

The computation time for producing such tables is very high, but it has turned out that it is sufficient to obtain the extrapolated value by two computations only (see also Fig. 4). If $\mu_{(1)}$ and $\mu_{(2)}$ are the solutions of

two systems of equations which are obtained by two different discretization widths h_1 and h_2 , then the extrapolation yields

$$\bar{\mu}_{(e)} = (1/h_1 \bar{\mu}_{(1)} - 1/h_2 \bar{\mu}_{(2)}) / (1/h_1 - 1/h_2). \quad (28)$$

This extrapolation corresponds to the second column of Table 1. Of course, it must be assumed that the discretization steps h_i are sufficiently small, i.e. the number of boundary elements is sufficiently high. Especially in the case of shallow bodies, for example see Fig. 4, should a close-meshed discretization be chosen.

Table 2 shows that the extrapolation, Eq. (28), generally yields good results. In Table 2 some results of model calculations (Schulz, 1983) are summarized. The model body is a sphere of variable resistivity and constant depth (depth of the centre $z_K = 2R$, R is the radius), buried in a homogeneous half-space. It is given the apparent resistivity for a Schlumberger array, Eq. (1), with $AB/2 = 100R$; centre of coordinates is the epicentre of the sphere. For $\rho_K = 0$ the analytical solution has a closed form (Snyder and Merkel, 1973) and for $\rho_K > 0$ it is computed by a system of equations (Brass et al., 1981).

The results of Table 2 show significantly:

- The method of integral equation for the surface potential is not valid for a completely conducting body.
- The numerical results of the method of integral equations are different from the exact results, even if the number of boundary elements is high.
- The extrapolation, Eq. (28), yields a good correspondence between numerical and analytical results.

Model curves

Model curves for a body buried in a stratified half-space can be calculated by the described procedure. Studies have shown that the shape of the body influences the model curve only to a low extent. On the other hand, resistivity of the overburden produces significant effects (Schulz, 1983).

As the number of parameters are high, only the fundamental effect is presented by one example (Fig. 5). A completely conducting cube ($\rho_K = \rho_3 = 0$) of width b , buried in a two-layer earth, is chosen as the model. The overburden has thickness $d/b = 0.1$ and resistivity ρ_1 ; the substratum has resistivity ρ_2 . All dimensions of length are normalized by the width b of the cube, all resistivities are normalized by the resistivity ρ_1 of the overburden.

In Fig. 5 the depth of the top of the body ($t/b = 0.25, 0.5, 0.75, 1.0$) and the resistivity of the substratum ($\rho_2/\rho_1 = 10$ - well-conducting overburden; $\rho_2/\rho_1 = 1$ - homogeneous half-space; $\rho_2/\rho_1 = 0.1$ - poorly conducting overburden) are varied.

The Schlumberger half-array, Eq. (2), is chosen as the electrode array. This array has been used in field measurements, with good results (Brass et al., 1981).

The shape of the curves is changed according to the point of measurement. If this point is situated before the epicentre of the body, as in Fig. 5, the curves can have slopes of more than 45° .

It is worth noting that the asymptotic value of apparent resistivity for great distances can be higher than

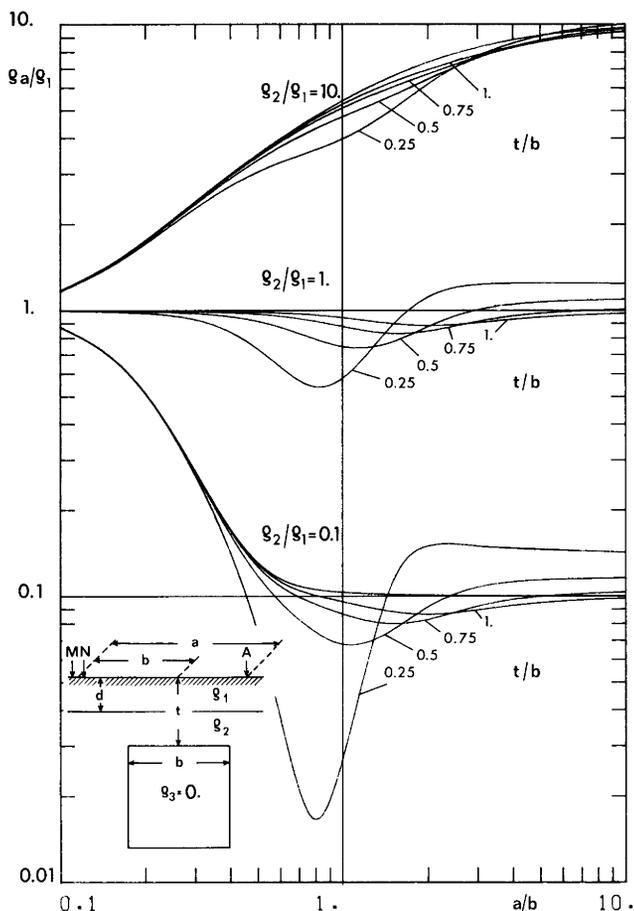


Fig. 5. Model curves of apparent resistivity of Schlumberger sounding half-array for a cube (width b , $\rho_3 = 0$) buried in a two-layer earth (thickness of the overburden $d/b = 0.1$, resistivity of the overburden ρ_1). The curves are plotted for different ratios of resistivity (ρ_2/ρ_1) and different depths of the top of the cube (t/b); the curve on top of each set represents the two-layer case without a buried body. Point of measurement: b before the epicentre of the cube

the resistivity of the substratum, although the buried body has a vanishing resistivity.

The curves are distinctly marked by the resistivity of the overburden. The lower the ratio ρ_2/ρ_1 , the more distinct is the shape of the curve. Therefore, the ratio of resistivities of overburden and substratum will influence of depth of investigation for buried bodies.

Field measurements can be interpreted by comparison with such model curves. This method has been applied, with success, in the exploration of graphite (Schulz, 1983).

Acknowledgements. The author is greatly indebted to Professor Mundry, Geological Survey of Lower Saxony (NLfB), Hannover, and Professor Weidelt, Federal Institute for Geoscience and Natural Resources (BGR), Hannover, for important suggestions and helpful discussions. The author also wants to thank Professor Rosenbach, Technical University of Clausthal, for many useful hints.

The model calculations have been carried out on Cyber 76 at the Regional Computer Centre of Lower Saxony (RRZN), Hannover. This paper has been supported by the Federal Ministry of Research and Technology (BMFT), Bonn, under project no. NTS 1052-R201 and no. RG 81043.

Appendix: proof of convergence

The function space, in which the integral equation, Eq. (22), has to be solved, is the space $L_2(S)$ of functions for which the Lebesgue integral of their square exists. The norm of space $L_2(S)$ is defined by

$$\|s\| := \left\{ \int_S [s(\mathbf{r})]^2 d\mathbf{o} \right\}^{1/2}; \quad s \in L_2(S). \quad (29)$$

A complete sequence of orthogonal coordinate functions is defined by the functions $d_q(\mathbf{r})$, i.e.

$$(d_p, d_q) = \int_S d_p(\mathbf{r}) d_q(\mathbf{r}) d\mathbf{o} = \begin{cases} |S_q|; & p = q \\ 0; & p \neq q \end{cases} \quad (30)$$

and

$$\lim_{N \rightarrow \infty} \int \left\{ s(\mathbf{r}) - \sum_{q=1}^N s_q d_q(\mathbf{r}) \right\}^2 d\mathbf{o} = 0. \quad (31)$$

$|S_q|$ is the area of the boundary element S_q .

The orthogonal projection from the space $L_2(S)$ into the space of piecewise constant functions is noted by O_N :

$$O_N(s) = \sum_{q=1}^N s(\mathbf{r}_q) d_q(\mathbf{r}), \quad (32)$$

where it is assumed that \mathbf{r}_q is not a singularity of s .

With these notes, the following estimation between the exact solution μ of the integral equation, Eq. (22), and the solution μ_N of the system of linear equations, Eq. (23), is valid (Michlin and Prößdorf, 1980, Chapter 18)

$$\|\mu_N - \mu\| \leq (1 + \varepsilon_N) \|(I - O_N)\mu\| \quad (33)$$

with $\lim_{N \rightarrow \infty} \varepsilon_N = 0$.

μ_N converges to the exact solution μ for $N \rightarrow \infty$; also, the procedure is stable.

In the system of equations, Eq. (25), the integrals are numerically calculated. If it is assumed that the functions are differentiable, it can be shown, by expansion in a Taylor series, that

$$\max_{\mathbf{r} \in S} |\bar{\mu}_N(\mathbf{r}) - \mu_N(\mathbf{r})| \leq c_1 \max |R(\mu, \sigma \mathbf{n} \cdot \nabla \psi)| \quad (34)$$

is valid, where c_1 is a constant independent of \mathbf{r} and R is the residual term of the quadrature formula used. Combining Eqs. (33) and (34), we get the solution of the system of equations, Eq. (25), to be:

$$\|\bar{\mu}_N - \mu\| \leq (1 + \varepsilon_N) \|(I - O_N)\mu\| + \hat{c}_1 \max |R(\mu, \sigma \mathbf{n} \cdot \nabla \psi)|. \quad (35)$$

Applying Schwarz's inequality, it ensues that, for the potential U_N , Eq. (24),

$$|U_N - U| \leq \|\mu_N - \mu\| \|\sigma \psi\|. \quad (36)$$

In analogy to Eq. (34), it can be shown that

$$|\bar{U}_N - U_N| \leq c_2 \max |R(\mu, \sigma \psi)| \quad (37)$$

is valid and, therefore,

$$|\bar{U}_N - U| \leq \|\mu_N - \mu\| \|\sigma \psi\| + \hat{c}_2 \max |R(\mu, \sigma \psi)|. \quad (38)$$

This estimation is analysed for the following special case: the surface S is divided into boundary elements S_q , each of which is a square of the same width h . It is assumed that all functions of the integral equation and representation are twice continuously differentiable. Then the residual term of the simplest Gauss quadrature formula used can be estimated by Taylor expansion

$$\max_{\mathbf{r} \in S_q} |R| \leq c_3 h^4.$$

Since the number N of boundary elements is reciprocally proportional to h^2 , it ensues that

$$\max_{\mathbf{r} \in S} |R| \leq c_4 h^2 \quad (39)$$

for the whole surface S .

The term $\|\mu_N - \mu\|$ is additionally significant for convergence because of Eqs. (38) and (33).

It ensues, from Eq. (32), that

$$\|(I - O_N)\mu\| = \left(\sum_{q=1}^N \int_{S_q} (\mu(\mathbf{r}) - \mu(\mathbf{r}_q))^2 d\mathbf{o}_r \right)^{1/2}.$$

Assuming the differentiability of μ , one gets from Eq. (36)

$$|U_N - U| \leq c_6 h \quad (40)$$

and from Eq. (37)

$$|\bar{U}_N - U| \leq c_6 h + c_7 h^2. \quad (41)$$

References

- Barnett, C.T.: Theoretical modelling of induced polarization effects due to arbitrarily shaped bodies. Dissertation, Colorado School of Mines, Golden, Col., 1972
- Barthes, V., Vasseur, G.: Use of d.c. electrical sounding for the detection of a conducting heterogeneity buried in a stratified medium. Luxembourg: Seminar on geothermal energy, Brussels, 6.-8.12.1977, Commission of the European Communities, EUR S 920, Vol. I, pp. 309-319, 1978
- Brass, G., Flathe, H., Schulz, R.: Resistivity profiling with different electrode arrays over a graphite deposit. *Geophys. Prosp.* **29**, 589-600, 1981
- Coggon, J.H.: Electromagnetic and electrical modeling by the finite element method. *Geophysics* **36**, 132-155, 1971
- Dey, A., Morrison, H.F.: Resistivity modeling for arbitrarily shaped three-dimensional structures. *Geophysics* **44**, 753-780, 1979
- Dieter, K., Paterson, N.R., Grant, F.S.: IP and resistivity type curves for three-dimensional bodies. *Geophysics* **34**, 615-632, 1969
- Gosh, D.P.: The application of linear filter theory to the direct interpretation of geoelectrical resistivity sounding measurements. *Geophys. Prosp.* **19**, 197-217, 1971a
- Gosh, D.P.: Inverse filter coefficients for the computation of apparent resistivity standard curves for a horizontally stratified earth. *Geophys. Prosp.* **19**, 769-775, 1971b
- Johansen, H.K., Sørensen, K.: Fast Hankel transforms. *Geophys. Prosp.* **27**, 876-901, 1979
- Koefoed, O.: *Geosounding principles*, Vol. I. Resistivity sounding measurements. Amsterdam: Elsevier 1979
- Lee, R.: An integral equation and its solution for some two-

- and three-dimensional problems in resistivity and induced polarisation. *Geophys. J.R. Astron. Soc.* **42**, 81–95, 1975
- Michlin, S.G., Prößdorf, S.: *Singuläre Integraloperatoren*. Berlin: Akademischer Verlag 1980
- Morse, P.M., Feshbach, H.: *Methods of theoretical physics*. New York: McGraw-Hill 1953
- Mundry, E.: Geoelectrical model calculations for two-dimensional resistivity distributions. *Geophys. Prosp.* **32**, 124–131, 1984
- Mundry, E., Dennert, U.: Das Umkehrproblem in der Geoelektrik. *Geol. Jb.* **E 19**, 19–38, 1980
- Okabe, M.: Boundary element method for the arbitrary inhomogeneities problem in electrical prospecting. *Geophys. Prosp.* **29**, 39–59, 1981
- Okabe, M.: Reciprocal averaging techniques in the geoelectrical boundary element approach. *Geophys. Prosp.* **30**, 653–672, 1982
- Riehle, H.J.J. (ed.): *Colloquium numerical treatment of integral equations*. Amsterdam: Mathematisch Centrum 1979
- Schulz, R.: *Potentialberechnungen zur Interpretation von gleichstromgeoelektrischen Messungen über dreidimensionalen Störkörpern*. Dissertation, Technische Universität Clausthal, 1983
- Smythe, W.R.: *Static and dynamic electricity*. New York: McGraw-Hill 1968
- Snyder, D.D., Merkel, R.M.: Analytic models for the interpretation of electrical surveys using buried current electrodes. *Geophysics* **38**, 513–529, 1973
- Stoer, J.: *Einführung in die Numerische Mathematik I*. Berlin: Springer 1972
- Weidelt, P.: *Entwicklung und Erprobung eines Verfahrens zur Interpretation zweidimensionaler Leitfähigkeitsstrukturen in E-Polarisation*. Habilitationsschrift, Göttingen, 1978

Received October 10, 1984; Revised December 21, 1984

Accepted December 21, 1984

Stochastic ion acceleration by coherent electrostatic waves

H. Varvoglis* and K. Papadopoulos

Astronomy Program, University of Maryland

Abstract. The limits of test ion acceleration by electrostatic, coherent, low frequency waves, are examined using the Poincaré surface of section method. The analysis indicates that Hydrogen cyclotron waves preferentially accelerate O^+ , while lower hybrid waves favor H^+ . This process can be described by a diffusion equation within some phase space boundaries, even for infinite autocorrelation times. Consequences of the results to auroral acceleration are presented.

Key words: Preferential ion acceleration – Auroral Physics – Ion conics formation

Introduction

Particle acceleration in the presence of low frequency electrostatic waves is of paramount importance in the understanding of the spectra of energetic particles in space. A large amount of work has been reported on the various mechanisms for particle acceleration by MHD waves or discontinuities such as the Fermi and shock acceleration processes. It was concluded, in general, that MHD acceleration processes are efficient only for particles above a certain rather large energy threshold, thereby requiring a first stage acceleration. This fact and the observation of energetic particles in regions where MHD waves are absent, motivated the work on the limits of particle acceleration in the presence of short wavelength electrostatic waves (i.e. microturbulent acceleration). The traditional approach follows a quasilinear theory, and computes the energy transfer from the waves to the particles on the basis of an equation of the form (Kennel and Engelman, 1966)

$$\frac{\partial}{\partial t} f = \frac{\pi}{2} \frac{e^2}{M^2} \sum_k \sum_l \frac{1}{k^2} |E_k|^2 \cdot \left\{ \frac{\Omega l}{v_{\perp}} \frac{\partial}{\partial v_{\perp}} + k_z \frac{\partial}{\partial v_z} \right\} \cdot J_l^2 \left(\frac{k_{\perp} v_{\perp}}{\Omega} \right) \delta(\omega - k_z v_z - l\Omega) \left\{ \frac{\Omega l}{v_{\perp}} \frac{\partial}{\partial v_{\perp}} + k_z \frac{\partial}{\partial v_z} \right\} f. \quad (1)$$

This equation can be generalized to include resonance broadening (Dum and Dupree, 1970; Davidson, 1972;

Palmadesso et al., 1974), by replacing the delta function $\delta(\omega - k_z v_z - l\Omega)$ by a resonance function $R(\omega - k_z v_z - l\Omega)$. The use of Eq. (1) is subject to the constraints of broadband spectra and small amplitudes, in the sense that $\tau_{AC} \ll \tau_B$, where τ_{AC} is the autocorrelation time of the spectrum and τ_B the bounce time of the resonant particles in the wave. In using quasi-linear theory to describe particle acceleration, for the case where $\omega > \Omega$, we can distinguish two cases. For particles that satisfy the resonance condition $\omega - k_z v_z - l\Omega_j = 0$, we have the usual resonance at the appropriate harmonic number l . The other possibility is to use the unmagnetized theory which requires that

$$\omega = \mathbf{k} \cdot \mathbf{v} \quad (2)$$

be satisfied.

Since (Davison 1972) $\tau_B = \left(\frac{eE}{kM} \right)^{-1/2}$, the autocorrelation time restriction implies that

$$E \ll \frac{Mk}{e} \frac{1}{\tau_{AC}} \quad (3)$$

where M is the ion mass. Therefore Eq. (1) cannot be applied for particle energization by coherent waves (i.e. $\tau_{AC} \rightarrow \infty$). For the ion cyclotron wave observations the autocorrelation time $\tau_{AC} \gg \frac{1}{\Omega_i}$. E.g. for the case discussed by Kintner et al. (1979) where $\tau_{AC} > 10f_H^+$ $\simeq 60 \Omega_H^{-1}$, and by taking $kR_H \approx 1$ where R_H is the Hydrogen gyroradius, we find $E \ll 10^{-3} \frac{mV}{m}$ resulting in exceedingly small wave amplitudes, which will be completely inefficient for acceleration. Therefore neither the magnetized theory nor the unmagnetized theory can be used to describe particle energization by coherent waves.

The observation of perpendicularly accelerated ions in the presence of coherent Hydrogen cyclotron waves (HCW) in the auroral zones (Kintner et al., 1978, 1979; Lysak et al., 1980) is a typical case where a diffusion formalism should not be applied because $\tau_{AC} \gg \tau_B$. An additional problem in attempting to attribute the auroral ion energization to the observed HCW is the presence of energetic O^+ with conic distributions. Since HCW have frequencies $\omega > \Omega_H$, it is not easy to satisfy

* Permanent address: Astronomy Department University of Thessaloniki, Greece

Offprint requests to: H. Varvoglis

the resonance condition for any appreciable number of particles of the heavier species, such as O^+ . The limitations on the resonant heating of species with $\omega - \Omega_j \gg k_z v_z$ forced Fisk (1978), and Ashour-Abdalla et al. (1981) to postulate, in addition to Hydrogen, excitation of ${}^4\text{He}^+$ and O^+ cyclotron waves respectively.

The limitations and restrictions of the resonant acceleration motivated the work of Papadopoulos et al. (1980), who demonstrated that, in the presence of electrostatic Hydrogen cyclotron waves exceeding a certain amplitude threshold, the maximum non-resonant energization for a subset of particles with a large harmonic number l (i.e. $|\omega - l\Omega_j| \gg k_z v_z$) exceeds by orders of magnitude the resonant one. An additional feature of the mechanism is that stochastic behavior occurs even for infinite autocorrelation time (i.e. monochromatic waves).

The work of Papadopoulos et al. (1980) was based on the intrinsic stochasticity properties of conservative Hamiltonian systems (Chirikov, 1979; Smith and Kaufmann, 1978; Karney, 1979) and was successful in accounting for the preferential perpendicular acceleration of O^+ in the auroral regions. An important corollary was the demonstration that the maximum stochastic acceleration due to Hydrogen cyclotron waves (HCW) scales as $M^{5/2} q^{-3/2}$ with the mass (M) and the charge (q) of the accelerated ions. This paper presents a more detailed account of the process than the space limited short letter of Papadopoulos et al. (1980) and extends

the work to include mass ratios $\frac{M}{m_H} < 3$ (where m_H is the Hydrogen mass). A generalized diffusion theory (Karney, 1979), not subject to small amplitude, broad bandwidth constraints [i.e. Eq. (2)], is then applied to examine the observable consequences of the combined action of lower hybrid waves (LHW) and HCW on the auroral particles. A word of caution to the reader is that the emphasis in this paper is on exploring the fundamental physics of the stochastic process under consideration by applying the unified diffusion theory (which is irrespective of the value of the autocorrelation time), rather than trying to do detailed modelling. For this reason we have confined ourselves to the completely non-resonant case by examining waves with $k_z = 0$. This is a good approximation for both the LHW and the HCW, which have $\frac{k_z}{k} \ll 1$.

Finite k_z does not affect the non-resonant processes discussed here to a great extent, but could play an important role in the resonant acceleration or heating of ions (i.e. small harmonic number l) (Abe et al., 1980; Singh et al., 1981; Varvoglis and Papadopoulos, 1984). We will comment on this below in section Non-resonant ion acceleration by electrostatic HCW of this paper.

Basic Problem Formulation

Consider a monochromatic low frequency electrostatic wave

$$\underline{E}(x, t) = \hat{e}_x E_0 \cos(k_\perp x - \omega t) \quad (4)$$

with ω above the Hydrogen cyclotron frequency Ω_H , propagating perpendicular to the ambient magnetic field $\underline{B} = B_0 \hat{e}_z$. Using *cgs-esu* units the phase space motion of an ion with mass M and charge q , treated as a test particle, is given by the Hamiltonian

$$H = \frac{1}{2M} [P_x^2 + (P_y - qA_y)^2 + P_z^2] - \frac{qE_0}{k_\perp} \sin(k_\perp x - \omega t) \quad (5)$$

where $\underline{A} = A_y \hat{e}_y$ is the vector potential of the ambient magnetic field $\underline{B} = B_0 \hat{e}_z$. We observe that y and z are cyclic coordinates of Eq. (5), so that P_y and P_z are integrals of motion. Therefore, without loss of generality, we may take $P_y = P_z = 0$. Using then the fact that $A_y = xB_0$, Eq. (5) can be written in dimensionless (barred) variables

$$\bar{H} = \frac{1}{2} \bar{P}_x^2 + \frac{1}{2} \bar{x}^2 - \alpha \sin(\bar{x} - \nu t) = \bar{h}. \quad (6)$$

In Eq. (6) we have normalized time to Ω_i^{-1} and length to k_\perp^{-1} . In these units $\bar{P}_x = \bar{v}_x$, $\bar{x} = \bar{v}_y$, $\nu = \frac{\omega}{\Omega_i}$ and the wave amplitude is given by

$$\alpha = \frac{k_\perp E_0}{\Omega_i B_0}. \quad (7)$$

If we perform the canonical transformation

$$\bar{x} = (2I_1)^{1/2} \sin \theta_1 \quad \bar{P}_x = (2I_1)^{1/2} \cos \theta_1$$

$$\bar{t} = \frac{\theta_2}{\nu} \quad \text{and} \quad \bar{h} = -\nu I_2$$

where \bar{h} is the numerical value of the Hamiltonian function given by Eq. (6), our Hamiltonian will be given in the action (I_1, I_2) and angle (θ_1, θ_2) variables as

$$H = I_1 + \nu I_2 - \alpha \sin[(2I_1)^{1/2} \sin \theta_1 - \theta_2]. \quad (8a)$$

Using the Bessel function identity Eq. (8a) becomes

$$H = I_1 + \nu I_2 - \alpha \sum_{l=-\infty}^{\infty} J_l(r) \sin(l\theta_1 - \theta_2) \quad (8b)$$

where $\frac{1}{2}r^2 = I_1$. Notice that r represents the perpendicular momentum of the particle normalized to $M \frac{\Omega_i}{k_\perp}$. It is beyond the purpose of our paper to discuss the enormous amount of work on non-integrable Hamiltonians of the type (8) over the last few years, referring the interested reader to Jorna (1978) and Helleman (1980) for a detailed exposition. We focus here on the theory of non-integrable Hamiltonian systems as applied to wave-particle interactions.

From Eq. (6) we see that for small α the particle motion is that of a reversible slightly perturbed harmonic oscillator, whose trajectory lies close to a constant energy surface. Above some $\alpha = \alpha_{\text{thr}}$, a dramatic transition occurs (often called stochastic instability), at which the particle trajectory in at least a region of phase space becomes chaotic. This allows the particles to cross the surfaces of constant energy (i.e. $I_1 = \text{const.}$) in phase space and thus be accelerated or decelerated. Our problem is thus reduced into finding the threshold

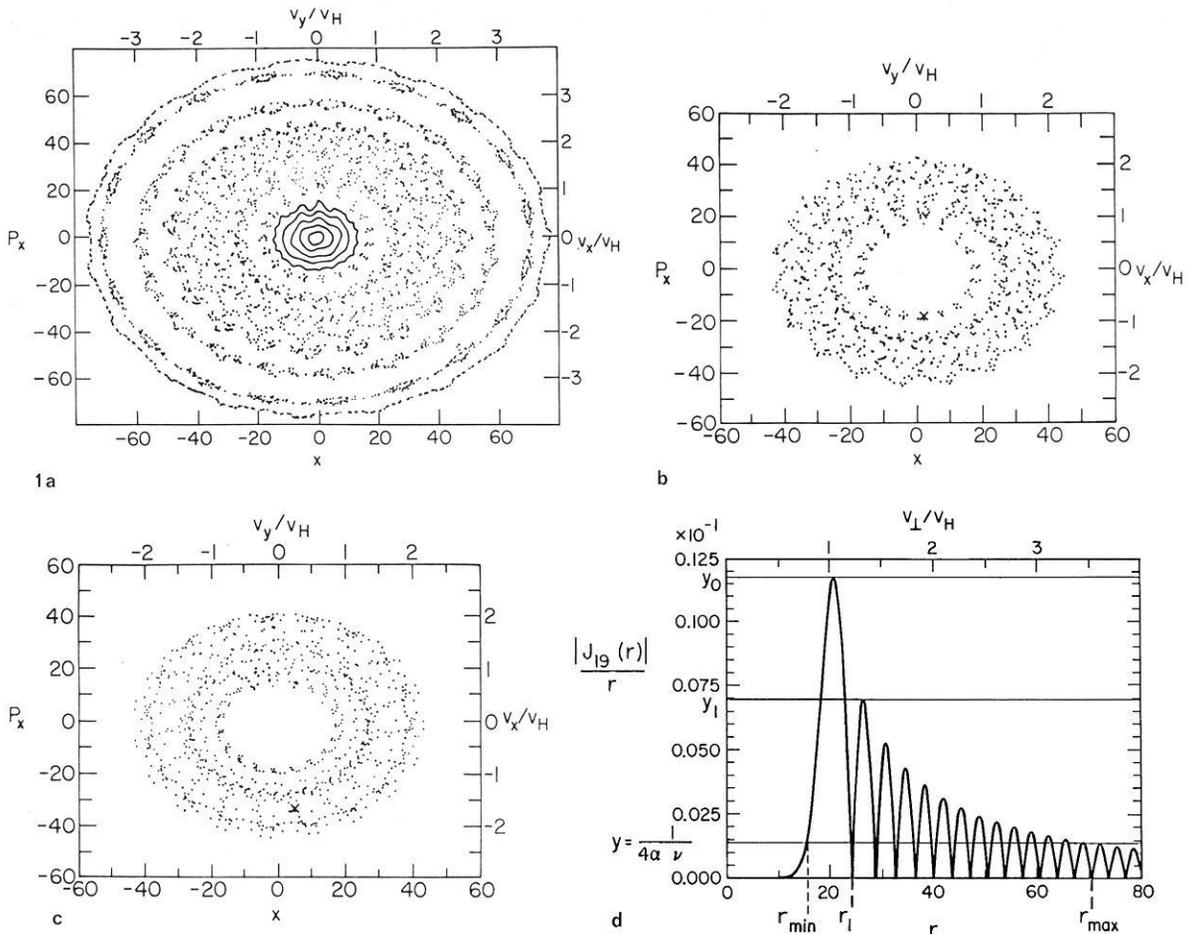


Fig. 1. **a** Surface of section of the system (8) for $\nu=19.2$, $\alpha=9$ and for various initial velocities r_0 , **b, c** Surface of section of the system (8) for $\nu=19.2$, $\alpha=9$ and $r_0=18.55$ and 29.15 respectively (r_0 is marked by an x), **d** The function $\left| \frac{J_{19}(r)}{r} \right|$ versus r and graphical calculation of α_{thr} , r_{min} (approximately) and r_{max} (upper bound)

value of α , and the upper and lower bounds of the ergodic region in phase space. Before doing this we note an interesting scaling property derived from the definition of α given by Eq. (7). For constant values of B_0 and the wave (i.e. \bar{E}_0, k_0), $\alpha \sim \frac{1}{\Omega_i} \sim M_i$, thereby increasing with mass. We expect therefore a mass selectivity, i.e. heavy particles will enter the stochastic region easier. This has been the premise of the Papadopoulos et al. (1980) paper. The value of the threshold α and the bounds of the stochastic region can be found either by a numerical or an approximate analytic method.

The numerical method is based on the concept of surface of section (Birkhoff, 1927; Poincaré, 1957; Berry, 1978). The trajectories describing solutions of the Hamiltonian with $\alpha=0$ lie on a torus in phase space with major radius $I_2 = \text{const.}$ and minor $I_1 = \text{const.}$ called invariant torus. The angles θ_1 and θ_2 describe the system location on the surface of the invariant torus. A good visualization of the type of trajectories followed by the system can be found by looking at the intersection of the invariant tori with a plane $\theta_2 = \text{const.}$ (we usually take $\theta_2 = \pi$). This is called the surface of section plane. For $\alpha=0$, the intersections of all the trajectories with the surface of section are points (consequents), lying on concentric circles (invariant curves), each curve

corresponding to one trajectory. For $\alpha \neq 0$ and beyond a value of α (threshold value) some trajectories are destabilized and the respective consequents can fill up the space between surviving invariant curves, indicating random motion. By numerically examining the surface of section plots we can find the threshold value (Karney, 1978). This picture is shown in Fig. 1a and will be discussed in the next section.

The analytic method, usually called criterion of overlapping resonances (Rosenbluth et al., 1966; Contopoulos, 1967; Chirikov, 1969, 1979) is based on the following concept. For $\alpha \neq 0$ a canonical perturbation theory of the Hamiltonian given in Eq. (8b) results in a series of resonances (Fukuyama et al., 1977; Lichtenberg, 1979). For finite α each of these resonances cause the appearance of families of islands on the surface of section, with a finite width, monotonically increasing with α . When the size of two neighboring islands belonging to different families becomes such that they touch each other (e.g. overlap), the particles can move across the $I_1 = \text{const.}$ lines and be accelerated. If we apply this criterion to Eq. (8b), we find that the "stochastic" transition occurs for

$$\frac{|J_{[v]}(r)|}{r} > \frac{1}{4\nu\alpha} \quad (9)$$

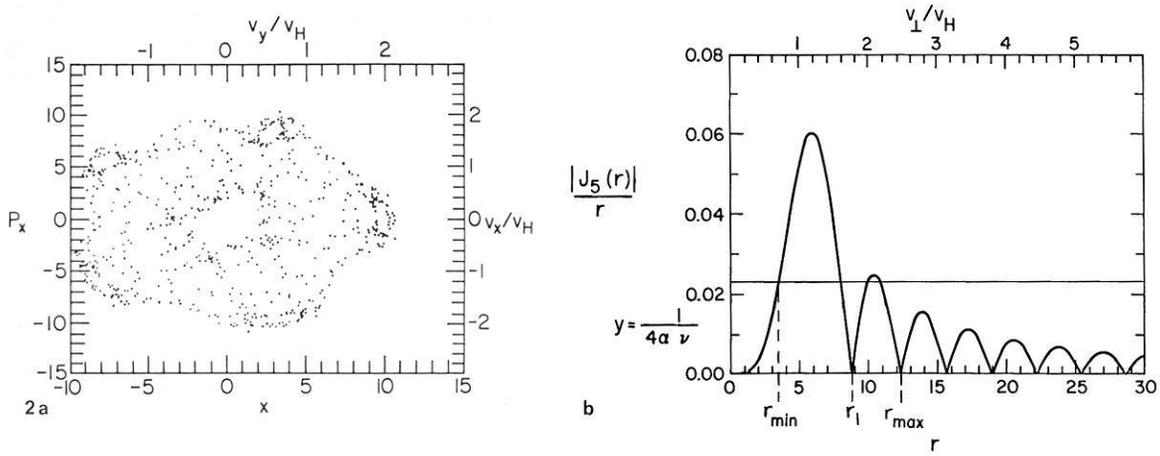


Fig. 2. **a** Surface of section of the system (8) for $v=4.8$, $\alpha=2.25$ and $r_0 \approx r_{\min}$, **b** The function $\frac{|J_5(r)|}{r}$ versus r and graphical calculation of r_{\min} (approximately) and r_{\max} (upper bound)

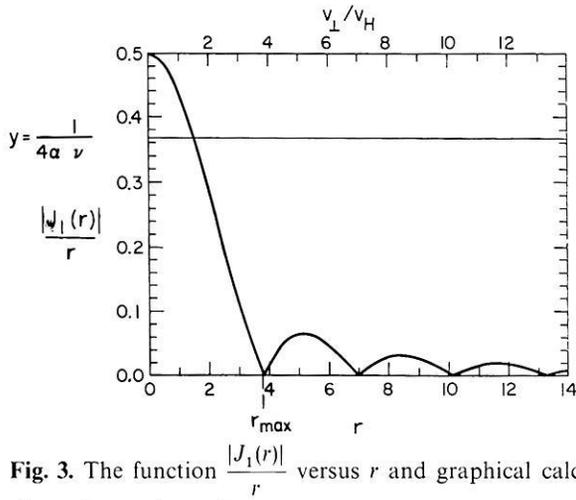


Fig. 3. The function $\frac{|J_1(r)|}{r}$ versus r and graphical calculation of r_{\max} (upper bound)

where $[v]$ is the integer nearest v (Fukuyama et al., 1977; Karney, 1978; Lichtenberg, 1979; Abe et al., 1980; Hsu, 1982). Eq. (9) is strictly valid for $r > v \geq 3$ and therefore can give (a) the threshold value of α and (b) the upper bound of the ergodic region in velocity space only in cases when the above condition is satisfied. However a comparison with surface of section plots shows that it gives relatively correct values of α_{thr} and r_{\max} when $3 > [v] > 1$ (Hsu, 1982) and even acceptable rough estimates of r_{\min} when $[v] \geq 1$ (see Figs. 1d, 2b and 3).

Non-Resonant Ion Acceleration by Electrostatic HCW

We proceed next to examine the non-resonant test particle acceleration in the presence of an electrostatic HCW. For the sake of definiteness we use the following parameters:

$$\omega = 1.2 \Omega_H, \quad k_{\perp} = (1.5)^{1/2} R_H^{-1} \quad \text{and} \quad \frac{e\phi}{T} \approx 0.4.$$

These parameters are consistent with the linear theory of the HCW instability (Kadomtsev, 1965) and the observations of the saturated amplitude (Kintner et al., 1978; Böhmer, 1976). There is nothing special about

these numbers and can be varied to conform to any other situations under study. We use them here only as basic guidelines for our numerical work. Before exploring the Hamiltonian

$$H = I_1 + v I_2 - \alpha \sin(r \sin \theta_1 - \theta_2)$$

using the surface of section method, it is important to stress the strong dependence of our results on the mass ($M = AM_H$) and charge ($q = Qq_H$) of the test ion species under consideration. For a test ion (M, q), and for the parameters mentioned above, the velocity v corresponding to the dimensionless velocity r is given by

$$v(A, Q) = \frac{Q}{A} \frac{r}{(1.5)^{1/2}} v_H \quad (10a)$$

where v_H is the thermal velocity of Hydrogen. The parameters v and α upon which our Hamiltonian depends are similarly given by

$$\alpha(A, Q) = 1.5 \frac{e\phi}{T} \frac{A}{Q} \quad (10b)$$

and

$$v(A, Q) = 1.2 \frac{A}{Q}. \quad (10c)$$

Oxygen ($A=16, Q=1$): We examine first the acceleration of O^+ . In this case, $v=19.2$ and $\alpha=9$. Figure 1a shows the surface of section $x-P_x$ for a variety of initial velocities r_0 . It is essentially the $I_1 \theta_1$ surface in cartesian coordinates. Note that, as was mentioned earlier, $x=v_y$ and $r^2=x^2+P_x^2$. For values of $r_0 < 11.5$, corresponding to particle velocities $v < 1/2 v_H$, the consequents lie on smooth curves indicating that the corresponding trajectories are ordered. For $r_0=11.5$ (i.e. $v \approx 1/2 v_H$), we have an onset of "stochasticity". The connected stochastic region extends to a maximum $r_{\max}=51$ corresponding to $2.6 v_H$. The ergodic properties of this region are depicted in an illustrative way in Fig. 1b, c. These are surface of section plots for two trajectories: one with initial velocity $r_0=18.55$ and the other with $r_0=29.15$. Notice how the two trajectories are wandering in the same region of phase space, due to the fact that both are in the connected stochastic

region ($11.5 < r < 51$). If we proceed to increase r_0 past 51, we discover a series of additional small stochastic bands, which are disconnected from the first one and thus we do not expect any particles to reach them, at least on a time scale physically significant to our problem, i.e. $t \approx 10^3 - 10^4 \Omega_0^{-1}$.

We also note that the lower stochasticity threshold $r_0 = 11.5$ corresponds to 2.3 times the thermal velocity of O^+ , and thus only a small fraction (0.2% for a Maxwellian) of O^+ ions can enter the acceleration mechanism. This fraction can be larger if the distribution is non-Maxwellian, e.g. because of resonant acceleration from non-coherent electrostatic noise (Whalen et al., 1978; Klumpar, 1979), but is always small, ensuring our theoretical treatment of O^+ ions as test particles.

We can also find the stochasticity regions for O^+ , in the above example, by using the resonance overlap criterion to our problem as given in Eq. (9). The graphical solutions of the inequality is shown in Fig. 1d

for the case $[v] = 19$. For values of α such that $\frac{1}{4\alpha v} > y_0 = \frac{1}{4\alpha_{\text{thr}} v}$, the wave amplitude is so small that the particle behaves adiabatically. For values of α such that

$$y_1 < \frac{1}{4\alpha v} < y_0$$

only small scale acceleration takes place in the sense that the stochastic region is restricted in the neighborhood of $r \approx v$ (notice that $r_{\text{min}} \approx \frac{\omega}{k_{\perp}} = v$, the phase velocity of the wave; this corresponds to the standard quasi-linear result for the unmagnetized case). The maximum velocity that can be achieved is given by r_1 , the first zero of the Bessel function (Abe et al., 1980). Finally for

$$\frac{1}{4\alpha v} < y_1$$

large scale non-resonant acceleration is possible. In this case the particle moves towards higher velocities by jumping from one "stochastic cell" to the next across the zeros of the Bessel function J_n (e.g. see MacKay et al., 1984). An upper bound of the maximum velocity r_{max} will be given by the first zero of J_n after the last intersection of the line

$$y = \frac{1}{4\alpha v}, \quad \alpha > \alpha_{\text{thr}}$$

with $|J_{[v]}(r)|$ (Abe et al., 1980). We emphasize that the so calculated r_{max} is an upper bound, because the last cell satisfying Eq. (9) may not be stochastic over all its width or even it may not be connected to the previous one. For the case of large v and α an analytic expression for inequality (9) can be derived by using the asymptotic expansion for $J_{[v]}(r)$, giving

$$r_{\text{max}} = \left[\frac{32\alpha^2 v^2}{\pi} \right]^{1/3}. \quad (11)$$

This is the result used by Papadopoulos et al. (1980) in their discussion of charge to mass ratio scaling in the auroral acceleration of heavy ions.

Helium ($A=4, Q=1$): The He^+ acceleration can be examined along the lines of O^+ . In Fig. 2a we give the surface of section for a He^+ trajectory with $r_0 \approx v_{\text{ph}} \approx v_H$. Clearly the diffuse points are wandering inside an annulus whose inner and outer radii give the "barriers" r_{min} and r_{max} . In Fig. 2b we give a graph of the inequality (9) for $n=5$ [as given from Eq. (10c)] and $\alpha = 2.25$ [as given from Eq. (10b) for $\frac{e\phi}{T_H} \approx 0.4$]. It can be seen that the efficiency of the acceleration is substantially smaller in this case (the line $y = \frac{1}{4\alpha v}$ hardly intersects the second bump of $\frac{|J_5(r)|}{r}$). Numerically the lower threshold r_{min} is found to be $r_{\text{min}} = 2.4 \approx 0.5 v_{\text{ph}} \approx 0.5 v_H$, while the maximum velocity is found to be $r_{\text{max}} = 10.5 \approx 2.1 v_H$, corresponding to an energy of $18 T_H$. Notice how the estimates from the stochasticity criterion (9) compare with the numerically (from the surface of section) determined values of r_{min} and r_{max} , i.e. the inner and outer radii of the stochastic annulus in Fig. 2a.

Hydrogen ($A=1, Q=1$): In this case we expect our criterion (11) to give only a qualitative picture of the behavior of the system, because in deriving (9) terms of the same order as $\frac{J_1(r)}{r}$ were neglected. In Fig. 3 we have plotted the function $\frac{|J_1(r)|}{r}$ versus r as well as the line $y = \frac{1}{4\alpha v} = 0.34$. This figure shows that stochastic protons should not have a lower limit in velocity (there are stochastic protons all the way down to $r \approx 0$), while the maximum velocity should be less than 3.8. These results were confirmed by the numerical method: a surface of section plot shows that the area covered with irregular trajectories is small compared to the cases of Oxygen and Helium, does not have a lower limit in velocity and extends up to $r_{\text{max}} \approx 2.0 (\approx 2 v_H)$. It is obvious that protons cannot be accelerated to high perpendicular energies by the HCW even for $\frac{e\phi}{T} \approx 1$. The amplitude required for the line y to intersect the second bump corresponds to $\frac{e\phi}{T} \gtrsim 2$. We expect therefore that the $\omega - \Omega_i = k_z v_i$ resonant acceleration dominates.

In summarizing the results of this section we note the strong dependence of the acceleration on the mass. This was the main result of Papadopoulos et al. (1980). It is interesting to examine the scaling of the maximum velocity for a wave with $\omega = \bar{v} \Omega_H$, and $k_{\perp} = \frac{\bar{k}}{R_H}$. Using the asymptotic formula given by Eq. (11), we find for the maximum energy

$$\frac{E_{\text{max}}}{T_H} \approx 5 \left(\frac{e\phi}{T_H} \right)^{4/3} (\bar{k} \bar{v}^2)^{2/3} \frac{A^{5/3}}{Q^{2/3}}. \quad (12)$$

In using this formula one should be careful, to be consistent with its validity region for $r > v > 3$. Eq. (12) exhibits a strong dependence on A , thereby predicting

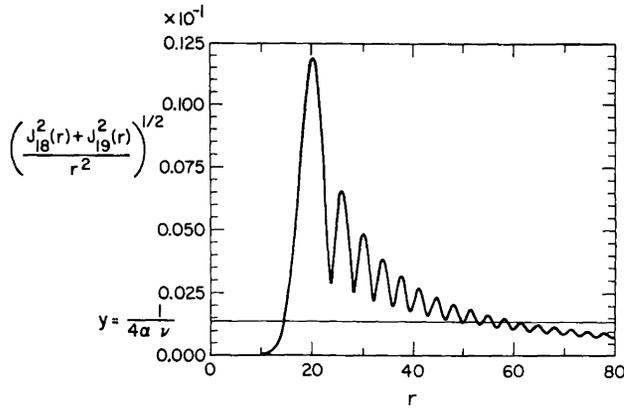


Fig. 4. The function $\frac{[J_{18}^2(r) + J_{19}^2(r)]^{1/2}}{r}$ versus r and graphical calculation of r_{\max} for the case of O^+ acceleration, following Lysak et al. (1980)

preferential acceleration of heavy ions. Another interesting feature is the strong dependence on the frequency $\bar{\nu}$ and wavelength \bar{k} . For HCW, the value of \bar{k} is of order unity and therefore not an important parameter. However, $\bar{\nu}$ can take values near 2, 3 etc. due to excitation of harmonics. The acceleration due to the first harmonic will dominate only if

$$\frac{\phi_1}{\phi_n} > n$$

where ϕ_n is the wave energy of the n^{th} harmonic. In general the presence of the second harmonic will increase the Helium energy found above, Eq. (12), by a factor

$$\left(\frac{2\phi_2}{\phi_1}\right)^{4/3} = 2.5 \left(\frac{\phi_2}{\phi_1}\right)^{4/3}$$

while not affecting the H^+ . The H^+ can however enter the stochastic regime by the third harmonic, and reach energies of the order of

$$E_{\max} \simeq 40 \left(\frac{e\Phi_3}{T_H}\right)^{4/3}$$

Before closing this section we should comment on the effect of a finite Δk bandwidth in the spectrum, instead of a single wave, and of a finite k_z .

Finite Δk bandwidth

In this case the Hamiltonian of Eq. (5) will be

$$H = \frac{1}{2M} [P_x^2 + (P_y - qA_y)^2 + P_z^2] - q \sum_k \frac{E_k}{k_{\perp}} \sin(k_{\perp} x - \omega t). \quad (13)$$

As long as the wave number $\Delta k < k_{\perp}$, where Δk is the spectrum bandwidth, the analysis can be carried out in a similar fashion as above for each wave-number. The resulting stochasticity criterion will then look like Fig. 4, from which we see that the acceleration limits are not affected.

Finite k_z

The case of a finite k_z has been considered by Abe et al. (1980) and Singh et al. (1981, 1982) who found that finite k_z effects are important for small values of ν (and thus small A/Q) but cannot be neglected even for large ν (and A/Q). As discussed by Karney (1978) and recently by Varvoglis and Papadopoulos (1984) for small $\xi = k_z/k_{\perp}$ (e.g. $k_z/k_{\perp} < \frac{1}{3}$) and for $A/Q \gg 1$ the reduced frequency ν in Eq. (9) should be replaced by the Doppler shifted frequency $\nu^* = \nu - \frac{k_z}{k_{\perp}} p_z$. As a result then Eq. (11) is replaced by the relation

$$r_{\max} = \left[\frac{32}{\pi} (\nu^2 - \xi^2 r_{\max}^2) \alpha^2 \right]^{1/3} \quad \alpha \geq \alpha_{\text{thr}}$$

which gives estimates for the maximum energization of the ions very close to the numerical results of Singh et al. (1981). Namely for the O^+ energization case considered by Singh et al. (i.e. $\alpha = 18.5$, $\nu = 19.2$) this last equation gives $r_{\max} = 84.5$ while Singh et al. find 88.5. This shows that the dominant character of the interaction remains non-resonant, at least for the case of the O^+ ions. However this is not true for H^+ and generally for ions with small A/Q (e.g. $A/Q \lesssim 3$). Note that Eq. (11) and therefore its corrected (for the finite k_z) form cannot be used for the computation of the energization of He^+ , since the $A/Q \gg 1$ condition is violated.

Diffusion from Monochromatic Waves

In the previous sections we established the fact that even a coherent wave with amplitude above threshold can cause particle diffusion between r_{\min} and r_{\max} as given before. The basic physical picture is that particles jump from one resonant island to the next. A diffusion coefficient should then be given by $\frac{\langle (\Delta r)^2 \rangle}{2\Delta t}$, where Δr is

the typical step size and Δt is the corresponding time interval. The change $\langle \Delta r^2 \rangle$ in r , between two consecutive "kicks" the particle gets from the wave, is given by $\langle (\Delta r)^2 \rangle = \frac{1}{4} \frac{\alpha^2 n^2}{r^2 (r^2 - n^2)^{1/2}}$ (Appendix); this occurs

twice at every gyration, that is over time $\Delta t = (2\pi\Omega_i^{-1})/2 = \pi$. The so calculated diffusion coefficient is by a factor of two smaller than the results of Karney (1979) and Antonsen and Ott (1981)

$$D(r) = \frac{1}{2} \frac{\alpha^2 \nu^2}{r^2 (r^2 - \nu^2)^{1/2}} \quad (14)$$

that was derived by using standard mapping techniques.

The diffusion equation

$$\frac{\partial f(r, t)}{\partial t} = \frac{1}{r} \frac{\partial}{\partial r} r D(r) \frac{\partial}{\partial r} f(r, t) \quad (15)$$

can be solved exactly if the diffusion coefficient (14) is approximated (for $r \gg \nu$) by

$$D(r) \simeq \frac{1}{2} \frac{\alpha^2 \nu^2}{r^3}.$$

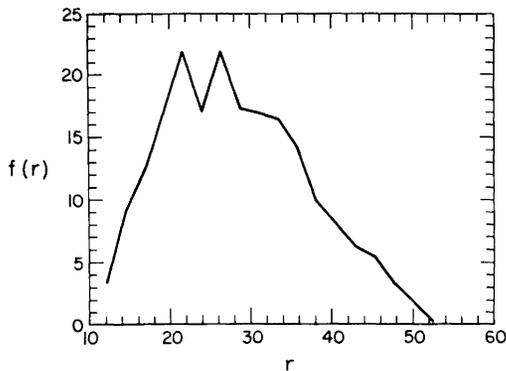


Fig. 5. Distribution function (in arbitrary units) at $t=265$ calculated numerically from a very narrow initial distribution (a “numerical” delta function) at $r_0=14.5$ for $\alpha=9$ and $\nu=19.2$

If moreover we take $f(r,0)=f_0(r)=n_0\delta(r-r_0)$, then the solution can be written as (Dum, 1978; Wu et al., 1981)

$$\begin{aligned} f(r,t) &= Ct^{-2/5} \exp\left(-\frac{2(r_0^5+r^5)}{25\alpha^2\nu^2t}\right) \\ &\simeq Ct^{-2/5} \exp\left(-\frac{2r^5}{25\alpha^2\nu^2t}\right) \end{aligned} \quad (16)$$

where we have assumed in the last approximation $r_0 \lesssim \nu \ll r$. The asymptotic expansions leading to the simple result in Eq. (16) are valid for

$$\frac{4\nu^5}{25\alpha^2\nu^2} = \frac{4\nu^3}{25\alpha^2} < t \ll \frac{2r_{\max}^5}{25\alpha^2\nu^2} \simeq 3.82(\alpha\nu)^{4/3}.$$

For $\alpha=9$, $\nu=19.2$ the range of validity is $15 < t < 3.5 \cdot 10^3$, while the typical interaction time in the case of Oxygen acceleration by HCW is of the order of 10^2 .

To check how accurately Eq. (16) describes the actual situation, we have performed the following numerical calculations. The trajectories of 100 particles with initial velocities given by a delta function distribution and identical initial phases $\theta_1 = \frac{-\pi}{2}$ were followed in time. Their velocities were recorded at time intervals $\Delta t = 2\pi$ and the resulting distribution function was calculated. Figure 5 shows the distribution function calculated in this way for $t=265$ in a case with $\alpha=9$, $\nu=19.2$, and with a very narrow initial distribution function (a “numerical” delta function) at $r_0=14.5$. To suppress fluctuations due to the small number of particles, Fig. 5, we average $f(r,t)$ over $\Delta t=8\pi$ (i.e. over four recording time intervals).

The distribution function of Fig. 5 should be compared with the solution of the diffusion equation [Eq. (16)], plotted for the same parameter and time values, shown in Fig. 6. Note that this curve starts at $r=24$, because the diffusion coefficient (14) is valid for $r \gg \nu$ only, and in Fig. 6 we have picked $r > j_{n,1} \simeq n + 1.8n^{1/3}$, which for $n=19$ gives $r > 24$. We see that the numerical (Fig. 5) and the analytic (Fig. 6) results are in good agreement. The two spikes at $r=22$ and $r=26$ in Fig. 5 are probably due to the presence of an island family at $r \simeq 24$, which has been ignored in the derivation of Eq. (14).

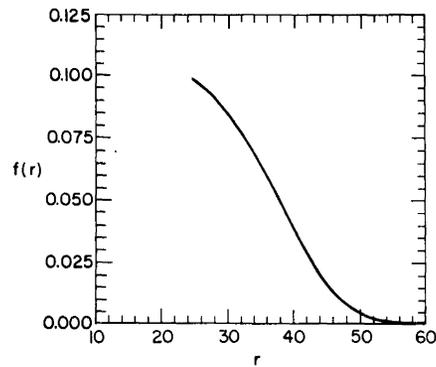


Fig. 6. Solution of the diffusion Eq. (15) (in arbitrary units) at $t=265$ for the same parameter values as Fig. 5

Some Implications of Non-Resonant Acceleration to Auroral Physics

Although the results presented above are very general, with many potential applications to magnetospheric and solar physics, we attempt here to examine some of the implications related to auroral acceleration. In the auroral zones, in addition to HCW, there is substantial wave activity near the lower hybrid frequency. The occurrence of lower hybrid waves (LHW) in the auroral zones was discussed by Horita and Watanabe (1969) and later by Papadopoulos and Palmadesso (1975). They usually occur at altitudes lower than the HCW (Mozer et al. 1980). For a strongly magnetized plasma (i.e. $\frac{\Omega_e}{\omega_e} > 1$), the lower hybrid frequency Ω_{LH} equals approximately to the Hydrogen plasma frequency ω_H . Since the analysis presented in the previous two sections is valid for waves with frequency $\omega \gtrsim \Omega_i$ it is interesting to examine its implications to the acceleration by LHW.

In the previous section we established the fact that for waves with amplitude above threshold the magnetized and unmagnetized diffusion equation that describes heating or acceleration can be used without the amplitude constraints imposed by the autocorrelation time. In assessing the effectiveness of acceleration of ionospheric ions by LHW we can use either a diffusion approach or a nonlinear test particle analysis. Chang and Coppi (1981) examined the ion energization by LHW using the first approach and assuming a broadband Δk spectrum. For the wave amplitude considered (i.e. $E_{\text{rms}} \simeq 50 \frac{\text{mV}}{\text{m}}$), the above assumption is unnecessary since the LHW amplitude is above threshold for non-resonant acceleration of Hydrogen ions by a single wave. We will follow the nonlinear test particle method here, because it allows a direct acceleration efficiency comparison with the HCW and in addition allows us to determine the maximum allowable acceleration (i.e. the velocity at which $D \approx 0$), which is not possible from the diffusion approach. The following standard parameters were taken for the LHW.

$$\omega_{LH} = 20\Omega_H, \quad k_{\perp}R_H = 6, \quad \frac{e\phi}{T_H} \simeq 0.4.$$

Table 1.

Species	E_{\min} (ev)		E_{\max} (keV)		Relevant time scales in $t\Omega_H$			
	HCW	LHW	HCW	LHW	HCW		LHW	
					Flight	Acceleration	Flight	Acceleration
H^+	0	30	0.05	2.5	2.5×10^3	2.5	3.7×10^4	7×10^3
He^+	10	110	0.25	25	5×10^3	400	7.5×10^4	10^6
O^+	55	450	2.5	250	10×10^3	65×10^3	15×10^4	1.8×10^8

HCW: $\omega = 1.2\Omega_H$, $k_{\perp}R_H = (1.5)^{1/2}$, LHW: $\omega = 20\Omega_H$, $k_{\perp}R_H = 6$, $T_H = 10$ ev

These conform with the observed frequency and the expected wave number estimates (Temerin, 1979). No assumption about the bandwidth is necessary since the diffusion process will operate even for coherent waves ($\Delta k = 0$). The results are shown in Table 1 where they are compared with the minimum and maximum acceleration limits for HCW. It is obvious from this table that Hydrogen enters easily in the acceleration region for both HCW and LHW, but achieves high energies by non-resonant acceleration only by the LHW. Of course this is in addition to the always present resonant HCW acceleration. On the other extreme reasonable fluxes of He^+ and O^+ can enter the acceleration region for HCW, but only O^+ can be accelerated to substantial energies. Finally acceleration of He^+ and O^+ by LHW is not favored for two reasons. First because of a rather high minimum threshold, and second because of an inefficient acceleration rate. Namely even if some fluxes of He^+ and O^+ enter the stochastic regime, they will drift out of the turbulent region faster than the time scale in which they can gain substantial energy. For example, as it is shown in Table 1, the parallel motion will carry an H^+ ion outside the length $L \approx 1,000$ km of the LHW turbulent region in a time $t_{\text{flight}} \approx 3.7 \times 10^4 \Omega_H^{-1}$, an He^+ ion in $7.5 \times 10^4 \Omega_H^{-1}$ and an O^+ ion in $15 \times 10^4 \Omega_H^{-1}$. This should be compared to the acceleration time scale [the time needed to make the exponent in Eq. (16) unity] which is $t_{\text{scale}} \approx 7 \times 10^3 \Omega_H^{-1}$ for H^+ , $10^6 \Omega_H^{-1}$ for He^+ and $1.8 \times 10^8 \Omega_H^{-1}$ for O^+ ions respectively.

Before closing we discuss some implications of our results to auroral physics. We limit ourselves to the high energy particles, which cannot be accounted by resonant heating. These are the particles that S3-3 can measure (i.e. >100 eV). Consider a portion of the auroral field lines around S3-3 altitudes as shown in Fig. 7, namely a region I with LHW, followed by a region II with HCW at higher altitude. According to our previous results (Table 1) Hydrogen will be selectively accelerated in region I forming high energy conics but will be basically unaffected by HCW in region II, except of course for parallel electric field and mirroring effects which can simply be superimposed in our results. On the other hand O^+ will be unaffected by LHW in region I but accelerated in II. It will be interesting to search in the S3-3 data for evidence of different regions of conic formation for H^+ and O^+ .

We can also speculate with respect to the source of the LHW and HCW. We feel in agreement with Chang and Coppi (1981) and Rowland et al. (1981) that the LHW are excited by the energetic (i.e. keV) electron be-

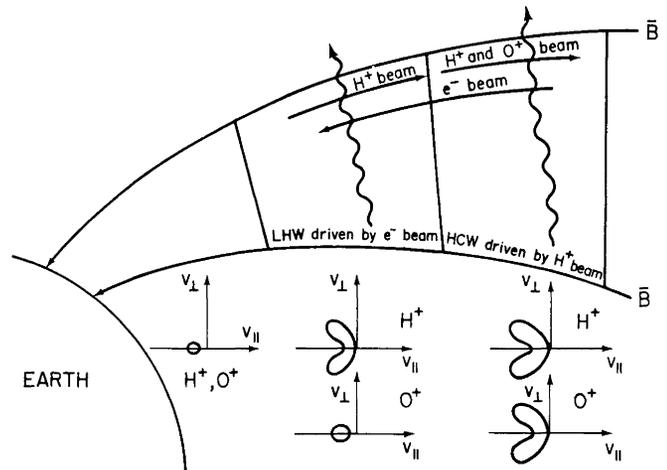


Fig. 7. Proposed scenario for LHW and HCW excitation and H^+ , O^+ conics formation in auroral regions

ams. The LHW waves could create conics in the energetic Hydrogen distribution functions. If these conics move to higher altitudes, they can excite HCW (Mikhailovskii, 1974), which in their turn can form O^+ conics. If the proton conics move downwards an equivalent scenario can be followed. If this is the case the HCW will not be driven by the cold electron current (Kindel and Kennel, 1971) as it is the usual assumption. The above comments are of course very speculative and should be taken as such. We are currently examining this possibility on a detailed basis including non-local effects such as mirror forces and parallel electric fields (Chiu et al., 1981). However, the main point of this section is the fact that the understanding of the non-resonant particle acceleration by low frequency electrostatic turbulence in conjunction with S3-3 type observations can be very fruitful in resolving basic auroral physics questions, such as the one posed by Kintner et al. (1979): "Did the HCW modes accelerate the ions or the ions excited the waves?" We feel that a closer examination of the S3-3 data in combination with the physics discussed here can actually help to resolve this ambiguity.

Acknowledgements. This work was supported by NASA NAGW #81 and ONR N00014-79-C0665 grants. Our numerical computations were partially supported by the Computer Science Center of the University of Maryland.

Appendix

Assuming that the chaotic motion of the consequents in velocity space is a Markovian process (Chandrasekhar, 1943) we

can compute a diffusion coefficient for the diffusion of ions in phase (and velocity) space in the following way. We find the change Δr in the velocity of an ion that happens in a time interval Δt , and then take the ratio

$$D = \frac{\langle (\Delta r)^2 \rangle}{2\Delta t} \quad (\text{A-1})$$

where the average is taken on the surface of section plane $r\theta_1$. As appropriate time interval Δt we take the time needed for an ion to complete one-half gyration around its guiding center, because in one full gyration the particle receives two "kicks" from the wave (one when $p_x \approx 0$, $x > 0$ and one when $p_x \approx 0$, $x < 0$). Accordingly $\Delta t \equiv \frac{2\pi\Omega_i^{-1}}{2} = \pi$ in our system units.

The change in velocity in this interval can be calculated from the equations of motion derived from the "resonant" Hamiltonian (Greene, 1980)

$$H_R = I_1 + \nu I_2 - \alpha J_n(r) \sin(n\theta_1 - \theta_2).$$

We find

$$\frac{dI_1}{dt} = r \frac{dr}{dr} = \alpha n J_n(r) \cos(n\theta_1 - \theta_2) \quad (\text{A-2})$$

$$\frac{d\theta_1}{dt} = 1 - \frac{\alpha}{r} J_n'(r) \sin(n\theta_1 - \theta_2) \approx 1 \text{ (since } r \gg \alpha \text{)}.$$

Then $\frac{dr}{d\theta_1} \approx \frac{\alpha n J_n(r)}{r} \cos(n\theta_1 - \theta_2)$ and

$$\Delta r \approx \frac{\alpha n J_n(r)}{r} \cos(n\theta_1 - \theta_2) \Delta\theta_1$$

$$\langle (\Delta r)^2 \rangle = \frac{\alpha^2 n^2 J_n^2(r)}{r} \langle \cos^2(n\theta_1 - \theta_2) \rangle \langle (\Delta\theta_1)^2 \rangle. \quad (\text{A-3})$$

Because of the already mentioned facts that the particle receives two "kicks" per revolution and $\Delta\theta_1 \approx \Delta t$, we take $\langle (\Delta\theta_1)^2 \rangle = \pi^2$. If moreover we take the average value of (A-3) over a distance $\Delta r = \pi$, we find (assuming that in this distance

$\frac{J_n^2(r)}{r^2}$ remains practically constant)

$$D \approx \frac{\pi}{4} \frac{\alpha^2 n^2}{r^2} \frac{2}{\pi} \frac{1}{(r^2 - n^2)^{1/2}} \left\langle \cos^2 \left(r - \frac{n\pi}{2} - \frac{\pi}{4} \right) \right\rangle \\ = \frac{1}{4} \frac{\alpha^2 n^2}{r^2 (r^2 - n^2)^{1/2}}.$$

This is by a factor of two smaller than the result obtained by Karney (1979) and by Antonsen and Ott (1981), which we used in solving analytically the diffusion equation (15).

References

- Abe, H., Momota, H., Itatani, R.: High-energy tail formation by a monochromatic wave in the magnetized plasma. *Phys. Fluids* **23**, 2417–2424, 1980
- Antonsen, T.M., Jr., Ott, E.: Diffusion coefficient for ions in the presence of a coherent lower hybrid wave. *Phys. Fluids* **24**, 1635–1640, 1981
- Ashour-Abdalla, M., Okuda, H., Cheng, C.Z.: Acceleration of heavy ions in auroral fields lines. *Geophys. Res. Lett.* **8**, 795–798, 1981
- Barbanis, B.: On the Isolating Character of the "Third" integral in a resonance case. *Astron. J.* **71**, 415–424, 1966
- Berry, M.V.: Regular and irregular motion. In: *Topics in nonlinear dynamics*, S. Jorna, ed. pp. 16–120. New York: American Institute of Physics, 1978

- Birkhoff, G.D.: *Dynamical systems*. New York: American Mathematical Society Colloquium Publications 1927, pp. 143, 152
- Böhmer, H.: Excitation of ion cyclotron harmonic waves with an ion beam of high perpendicular energy. *Phys. Fluids* **19**, 1371–1374, 1976
- Chandrasekhar, S., *Stochastic Problems in Physics and Astronomy*, Rev. Mod. Physics **15**, 1–89, 1943
- Chang, T., Coppi, B.: Lower hybrid acceleration and ion evolution in the suprathermal region. *Geophys. Res. Lett.* **8**, 1253–1256, 1981
- Chirikov, B.V.: Research concerning the Theory of Nonlinear Resonance and Stochasticity, Preprint 267 (Institute of Nuclear Physics, Novosibirsk, 1969 (Engl. Transl. CERN Transl. 71-40, 1971))
- Chirikov, B.V.: A universal instability of many-dimensional oscillator systems. *Phys. Rep.* **52**, 264–379, 1979
- 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
- Contopoulos, G.: Resonance phenomena and the non-applicability of the "third" integral, *Bull. Astron. 3e ser.* **2**, Fasc. 1, 223–241, 1967
- Davidson, R.C.: *Methods in nonlinear plasma theory*. New York: Academic Press, 1972
- Dum, C.T., Dupree, T.D.: Nonlinear stabilization of high frequency instabilities in a magnetic field. *Phys. Fluid* **13**, 2064–2081, 1970
- Dum, C.T.: Anomalous heating by ion sound turbulence. *Phys. Fluids* **21**, 945–955, 1978
- Fisk, L.A.: ³He-rich flares, a possible explanation. *Astrophys. J.* **224**, 1049–1055, 1978
- Ford, J.: A picture book in stochasticity. In: *Topics in nonlinear dynamics*. S. Jorna: ed. pp. 121–146 New York: American Institute of physics, 1978
- Fukuyama, A., Momota, H., Itatani, R., Takizuka, T.: Stochastic acceleration by an electrostatic wave near ion cyclotron harmonics. *Phys. Rev. Lett.* **38**, 701–704, 1977
- Green, J.M.: The calculation of K.A.M. surfaces. In: *Proc. Int. Conf. on nonlinear dynamics*, R. Helleman ed., Ann. N.Y. Acad. Sci. **357**, 80–89, 1980
- Helleman, R.H.G.: Self-generated chaotic behavior in nonlinear mechanics. In: *Fundamental Problems in Statistical Mechanics Vol. 5*, E.G.D. Cohen ed.: pp. 165–233, New York: North Holland 1980
- Horita, R.E., Watanabe, T.: Electrostatic waves in the ionosphere excited around the lower hybrid resonance frequency. *Planet. Space. Sci.* **17**, 61–74, 1969
- Hsu, J.Y.: Frequency mismatch and stochastic heating in the cyclotron frequency range. *Phys. Fluids.* **25**, 159–163, 1982
- Jorna, S. (ed): *Topics in non-linear dynamics*. New York: American Institute of Physics, 1978
- Kadomtsev, B.B.: *Plasma Turbulence*, p. 73. London: Academic Press, 1965
- Karney, C.F.F.: Stochastic ion heating by a lower hybrid wave. *Phys. Fluids* **21**, 1584–1599, 1978
- Karney, C.F.F.: Stochastic ion heating by a lower hybrid wave: II *Phys. Fluids* **22**, 2188–2209, 1979
- Kennel, C.F., Engelman, F.: Velocity space diffusion from plasma turbulence in a magnetic field. *Phys. Fluids.* **9**, 2377–2388, 1966
- Kindel, J.M., Kennel, C.F.: Topside current instabilities. *J. Geophys. Res.* **76**, 3055–3078, 1971
- Kintner, P.M., Kelley, M.C., Sharp, R.D., Ghielmetti, A.G., Temerin, M., Cattell, C., Mizera, P.F., Fennell, J.F.: Simultaneous observations of energetic (keV) upstreaming ions and electrostatic hydrogen cyclotron waves. *J. Geophys. Res.* **84**, 7201–7212, 1979
- Kintner, P.M., Kelley, M.C., Mozer, F.S.: Electrostatic hydrogen cyclotron waves near one earth radius altitude in

- the polar magnetosphere. *Geophys. Res. Letters*, **5**, 139–142, 1978
- Klumpar, D.M.: Transversely accelerated ions: an ionospheric source of hot magnetospheric ions. *J. Geophys. Res.* **84**, 4229–4237, 1979
- Lichtenberg, A.: Adiabatic and stochastic motion of charged particles in the field of a single wave. In: *Stochastic behavior in classical and quantum Hamiltonian systems*, G. Casati and J. Ford, (eds): pp. 18–34, New York: Springer, 1979
- Lysak, R.L., Hudson, M.K., Temerin, M.: Ion heating by strong electrostatic ion cyclotron turbulence. *J. Geophys. Res.* **85**, 678–686, 1980
- Mackay, R.S., Meiss, J.D., Percival, I.C.: Transport in Hamiltonian systems. *Physica* **13D**, 55–81, 1984
- Mikhailovskii, A.B.: *Theory of plasma instabilities*, New York: Consultants Bureau, 1974
- 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
- Palmadesso, P.J., Coffey, T.P., Ossakov, S.L., Papadopoulos, K.: Topside ionosphere ion heating due to electrostatic ion cyclotron turbulence. *Geophys. Res. Lett.* **1**, 105–108, 1974
- Papadopoulos, K., Gaffey, Jr., J.D., Palmadesso, P.J.: Stochastic acceleration of large m/q ions by hydrogen cyclotron waves in the magnetosphere. *Geophys. Res. Lett.* **7**, 1014–1016, 1980
- Papadopoulos, K., Palmadesso, P.J.: Excitation of lower hybrid waves in a plasma by electron beams. *Phys. Fluids* **19**, 605–606, 1975
- Poincaré, H.: *Les méthodes nouvelles de la mécanique céleste*, Vol. III, p. 75. New York: Dover, 1957
- Rosenbluth, M.N., Sagdeev, R.Z., Taylor, J.B., Zaslavski, G.M.: Destruction of magnetic surfaces by magnetic field irregularities. *Nucl. Fusion* **6**, 297–300, 1966
- Rowland, H.L., Palmadesso, P.J., Papadopoulos, K.: Anomalous resistivity of auroral field lines. *Geophys. Res. Lett.* **8**, 1257–1260, 1981
- Singh, N., Schunk, R.W., Sojka, J.J.: Energization of ionospheric ions by electrostatic hydrogen cyclotron waves. *Geophys. Res. Lett.* **8**, 1249–1252, 1981
- Singh, N., Schunk, R.W., Sojka, J.J.: Cyclotron resonance effects on stochastic acceleration of light ionospheric ions. *Geophys. Res. Lett.* **9**, 1053–1056, 1982
- Smith, G.R., Kaufman, A.N.: Stochastic acceleration by an obliquely propagating wave—an example of overlapping resonances. *Phys. Fluids* **21**, 2230–2241, 1978
- Temerin, M.A.: Doppler shift effects on double-probe-measured electric field power spectra. *J. Geophys. Res.* **84**, 5929–5934, 1979
- Treve, Y.M.: Theory of chaotic motion with application to controlled fusion research. In: *Topics in nonlinear dynamics*. S. Jorna: ed., pp. 147–220. New York: AIP, 1978
- Varvoglis, H., Papadopoulos, K.: Large scale chaotic motion of charged particles in a longitudinal electrostatic wave. *J. Phys. A* **17**, 311–321, 1984
- Whalen, B.A., Bernstein, W., Daly, P.W.: Low altitude acceleration of ionospheric ions. *Geophys. Res. Lett.* **5**, 55–58, 1978
- Wu, C.S., Gaffey, Jr. J.D., Liberman, B.: Statistical acceleration of electrons by lower hybrid turbulence. *J. Plasma Phys.* **25**, 391–402, 1981

Received November 7, 1984; Revised version January 23, 1985
Accepted January 30, 1985

Book Reviews

Wilfried Schröder: Das Phänomen des Polarlichtes. Wissenschaftliche Buchgesellschaft Darmstadt, 156 S., 1984

Die Geschichte der Geophysik und der Meteorologie wurde in der Vergangenheit speziell in Deutschland ziemlich vernachlässigt. Es ist eines der Ziele des Autors dieses Bandes, diese Lücke für die Geschichte des „Polarlichtes“ zu füllen. Der Autor ist Cochairman der Interdivisional Commission of the International Association of Geomagnetism and Aeronomy (IAGA) und Sekretär des Arbeitskreises Geophysik der Deutschen Geophysikalischen Gesellschaft. Er hat mehrere Arbeiten und einige Bücher über einschlägige Themen veröffentlicht.

Das Büchlein konzentriert sich auf die Geschichte der Polarlichtforschung und nicht etwa auf die Physik dieser Erscheinungen. Diese werden nur kurz im ersten Kapitel gestreift, das den Titel „Polarlicht in der Geschichte und in der Gegenwart“ trägt. Das 2. Kapitel faßt Hinweise auf Polarlicht in der alten, der mittelalterlichen und der neueren Literatur zusammen. Im 3. Kapitel wird dargestellt, wie sich im 18. Jahrhundert das Verständnis der Naturerscheinungen grundsätzlich geändert hat. Dieser Prozeß wurde zweifellos eingeleitet durch Beobachtungen des Polarlichtes vom 17. März 1716, die an vielen Stellen in Europa gemacht wurden. Im Anschluß daran traten physikalische anstatt metaphysische Deutungen in den Vordergrund. Mit zunehmender Zahl der Beobachtungen und verbesserten Untersuchungsmethoden entwickelte sich das Verständnis der Zusammenhänge zwischen optischen Beobachtungen des Himmels, solaren Ereignissen und erdmagnetischen Variationen, kurz gesagt solar-terrestrische Physik, im 19. Jahrhundert rapide.

Gleichzeitige Fortschritte in der Spektroskopie ermöglichten die Identifizierung des Polarlichtes als Strahlung atmosphärischer Gase angeregt durch solare Korpuskeln. Im Laufe mehrerer Expeditionen in Polargebiete und vor allem durch die vereinten Anstrengungen während des 1. Internationalen Polarjahres 1882–83 wurden praktisch alle wichtigen Eigenschaften des Polarlichtes entdeckt, mit Ausnahme der Höhe, in der es entsteht. Erst als es norwegischen Forschern gelang, Polarlichter gleichzeitig von zwei 30 km voneinander entfernten Punkten zu fotografieren, wurde diese Frage gelöst.

Der derzeitige Stand des Wissens vom Polarlicht wird im letzten Abschnitt des Kapitels 3 zusammengefaßt. Das 4. Kapitel ist einigen theoretischen Gesichtspunkten des Gebietes und einem Abriß der Polarlichtforschung im Zusammenhang mit dem allgemeinen Fortschritt der Naturwissenschaft gewidmet. Dabei wird der Beitrag von Amateurbesachtern auf diesem Gebiet deutlich.

Der Druck des Bandes ist klar, die Abbildungen, darunter viele historische, sind ausgezeichnet. Eine Liste von ca. 200 Referenzen und ein Personenregister schließt sich an. Das Buch ist allen zu empfehlen, die sich für einen allgemeinen Abriß der Geschichte der Geophysik interessieren.

W. Dieminger

Bruce A. Bolt: Erdbeben. Eine Einführung. Übersetzt aus dem Englischen von R. Gutdeutsch. Springer Verlag, Berlin, 236 S., 1984

„Seldom does a seismologist happen upon a well written, lucid book on earthquakes that one can hand to one's spouse for reading“ sagt D.M. Boore in einer Buchbesprechung des englischen Originals: Earthquakes, A primer by B.A. Bolt, W.H. Freeman and Co., San Francisco, 1978.

Diesem Buch gelingt es in der Tat, das Thema „Erdbeben“ fesselnd und anschaulich darzustellen. Ohne die effekthaschenden Züge von Sensationsreportagen, sondern mit der realistischen Beschreibung von Erdbebenzenarios ist es auf natürliche Weise spannend. Bolt hat es verstanden, durch lebensnahe Schilderung von aktuellen und historischen Begebenheiten, durch Anekdoten und

persönliche Erlebnisse das Thema aufzulockern. Ohne die verfälschende Simplifikation mancher allgemeinverständlicher Schriften, aber auch ohne die den Laien abschreckende mathematische Formulierung, sondern mit einer guten Kombination leicht fasslicher physikalischer Erklärungen, Plausibilitätsbetrachtungen und anschaulicher Vergleiche werden die geophysikalischen Zusammenhänge erläutert. Bolts Prinzip ist nicht die abstrakte Darstellung, sondern die Erklärung anhand genau untersuchter Erdbebenbeispiele. Für manchen Geschmack sind diese Fallschilderungen vielleicht sogar etwas zu zahlreich und zu detailliert. In jeder Hinsicht begrüßenswert ist jedoch das sehr umfangreiche und großenteils neue Bildmaterial, was dieses Buch in besonderer Weise auszeichnet. Charakteristisch sind auch die an den Leser gerichteten Anregungen zu eigenen Überlegungen und Schlußfolgerungen.

Dieses Buch kann als – im besten Sinne verstanden – populärwissenschaftlich bezeichnet werden: Es eröffnet einem breiten Leserkreis ein wissenschaftliches Verständnis für das Phänomen „Erdbeben“. Das Buch wendet sich ausdrücklich an Leser ohne mathematische und technische Vorbildung, allerdings, so meine ich, sollte eine ausgeprägte naturwissenschaftliche Neigung und ein sehr spezielles Interesse, auch für Detailfragen im Zusammenhang mit Erdbeben, vorhanden sein. Besonders angesprochen werden etwa Studienanfänger der Geowissenschaften oder Personen, die im weiteren Sinne beruflich mit Erdbeben befaßt sind. So gesehen ist Bolts Buch kein „Erdbebenbuch für jedermann“, füllt jedoch in seiner Art eine Lücke zwischen allgemeinverständlicher Vermittlung von Basiswissen und studienorientiertem Lehrbuch.

Die Inhaltsübersicht zeigt im wesentlichen die Fragestellungen, auf die der Leser Antwort bekommt und auf die er später selbst Antwort geben kann:

Wo Erdbeben vorkommen – Was wir bei einem Erdbeben wahrnehmen – Störungszonen in der Erde – Die Ursachen von Erdbeben – Erdbeben, Vulkane und Tsunamis – Erdbebenmessungen – Die Größe eines Erdbebens – Anregung von Erdbeben durch Wasser – Ereignisse, die einem Erdbeben vorangehen – Selbstschutz bei Erdbeben – Umweltstudien für erdbebensichere Bauplanung.

Der Anhang enthält neben einem Verzeichnis wichtiger Erdbeben, neben Intensitätsskalen und Erläuterungen der verwendeten Fachwörter sowie physikalischer Zusammenhänge als Besonderheit ein Erdbeben-Quiz.

Was das Buch weiterhin auszeichnet und es auch für den Geophysiker und Seismologen interessant macht, ist die aktuelle und kritische Darstellung der Themen: seismisches Risiko, Erdbebenvorläuferphänomene und Erdbebenvorhersage. In der Besprechung der infrastrukturellen, wirtschaftlichen, gesellschaftlichen und letztlich soziologisch-psychologischen Hintergründe einer Erdbebenvorhersage und in der Darstellung der Schutzmöglichkeiten, wie erdbebensichere Bauweise, Verhaltensmaßregeln, Versicherungsschutz etc. findet auch der fachkundige Leser noch Neues (Kapitel 9, 10 und 11). Insbesondere die Erörterung der induzierten Seismizität und der seismischen Instrumentierung von wichtigen großen Bauten (Anhang E) wendet sich an einen sehr speziell interessierten Leser.

Durch die weitgehend originalgetreue Übersetzung ins Deutsche bleibt der persönliche Stil des Autors erhalten. Einige Passagen sind dagegen redigiert und für den deutschsprachigen Leser verständlich gemacht. Im Anhang wurde die in Europa verwendete MSK-Intensitätsskala und eine Liste mitteleuropäischer Schadenbeben beigefügt. Eine Anpassung des Buches an mitteleuropäische Verhältnisse ist dadurch sicher nicht erfolgt und wäre meiner Ansicht nach auch nicht wünschenswert. Bolts Buch stammt aus dem „Erdbebenland Kalifornien“ und spiegelt, abgesehen von der Beschreibung einiger weltweiter Schadenbeben und der Erdbebenvorhersage in China, in Text und Bild die Verhältnisse in den U.S.A. wieder. Dieses Charakteristikum hat die deutsche Ausgabe erhalten.

Wolfgang Brüstle